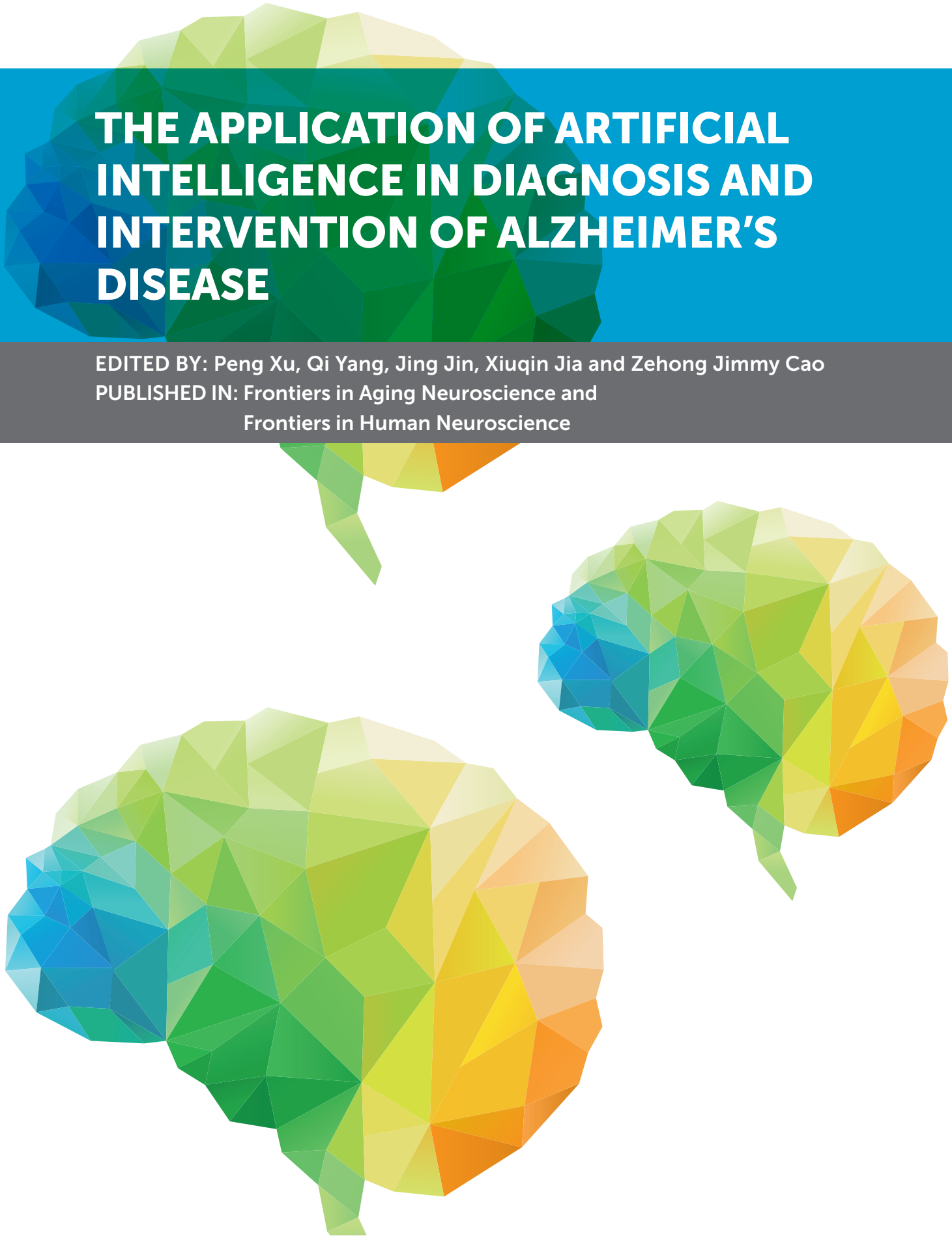
THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN DIAGNOSIS AND INTERVENTION OF ALZHEIMER'S DISEASE

EDITED BY: Peng Xu, Qi Yang, Jing Jin, Xiuqin Jia and Zehong Jimmy Cao

PUBLISHED IN: Frontiers in Aging Neuroscience and
Frontiers in Human Neuroscience



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-83250-575-5
DOI 10.3389/978-2-83250-575-5

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN DIAGNOSIS AND INTERVENTION OF ALZHEIMER'S DISEASE

Topic Editors:
Peng Xu, University of Electronic Science and Technology of China, China
Qi Yang, Capital Medical University, China
Jing Jin, East China University of Science and Technology, China
Xiuqin Jia, Capital Medical University, China
Zehong Jimmy Cao, University of South Australia, Australia

Citation: Xu, P., Yang, Q., Jin, J., Jia, X., Cao, Z. J., eds. (2022). The Application of Artificial Intelligence in Diagnosis and Intervention of Alzheimer's Disease. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-575-5

Table of Contents

frontiers in Aging Neuroscience

OPINION
published: 12 November 2021
doi: 10.3389/fnagi.2021.789547

Electronic Science Games Used to Enhance Cognitive Ability: Opinion of Design From Personalization and Adaptation

*Dong Wen[1], Wang Yao[2], Jian Xu[2], Shaochang Wang[2], Yingzhu Zhong[2], Hongqian Chen[2], Xianling Dong[3], M. Iqbal Saripan[4] and Yanhong Zhou[5]**

[1] *Institute of Artificial Intelligence, University of Science and Technology Beijing, Beijing, China,* [2] *School of Information Science and Engineering, Yanshan University, Qinhuangdao, China,* [3] *Department of Biomedical Engineering, Chengde Medical University, Chengde, China,* [4] *Faculty of Engineering, Universiti Putra Malaysia, Serdang, Malaysia,* [5] *School of Mathematics and Information Science and Technology, Hebei Normal University of Science and Technology, Qinhuangdao, China*

Keywords: electronic science game, cognitive ability, personalization, adaptation, enhance

INTRODUCTION

In recent years, more and more cognitive scientists have incorporated personalized and adaptive elements into the design of cognitive training games. From the perspective of personalization, some researchers allow game players to choose roles, scenes or personalized cognitive training tasks according to their own preferences to achieve personalized game design (Reategui et al., 2006; Wei, 2009; Hollingdale and Greitemeyer, 2013; Lin et al., 2013; Li et al., 2016; Nagle et al., 2016; Orji et al., 2017; Orji and Moffatt, 2018; Soares et al., 2018; Waltemate et al., 2018; Zibrek et al., 2018; González et al., 2019; Knutas et al., 2019; Troussas et al., 2019). There are also some researchers from the adaptive point of view, by dynamically changing the parameters in the game, automatically adapting to the player's game difficulty, dynamically generating new content and other methods to achieve the adaptive design of the game (Carro et al., 2002; Hunicke, 2005; Togelius et al., 2010; Johnson et al., 2014; Li et al., 2014; Yannakakis and Togelius, 2015; Shaker et al., 2016; Schadenberg et al., 2017; Soler-Dominguez et al., 2017; Ashish et al., 2018; Lopes et al., 2018; Shi and Chen, 2018; Souza et al., 2018; Denisova and Cairns, 2019; Dey et al., 2019; Hendrix et al., 2019; Liang et al., 2019; Pan et al., 2019; Papadimitriou et al., 2019; Peng et al., 2019; Plass et al., 2019; Sepulveda et al., 2019). Relevant research showed that adding personalized design to electronic science games for improving cognitive abilities could enhance the cognitive training experience of gamers, stimulate their interest in cognitive training, and better enhance the training experience and cognition ability of gamers (Reategui et al., 2006; Wei, 2009; Hollingdale and Greitemeyer, 2013; Lin et al., 2013; Li et al., 2016; Nagle et al., 2016; Orji et al., 2017; Orji and Moffatt, 2018; Soares et al., 2018; Waltemate et al., 2018; Zibrek et al., 2018; González et al., 2019; Knutas et al., 2019; Troussas et al., 2019); adding adaptive design to electronic science games used to improve cognitive ability, which can match the player's level with the difficulty of the game, so that gamers can obtain the best training effect (Carro et al., 2002; Hunicke, 2005; Togelius et al., 2010; Johnson et al., 2014; Li et al., 2014; Yannakakis and Togelius, 2015; Shaker et al., 2016; Schadenberg et al., 2017; Soler-Dominguez et al., 2017; Ashish et al., 2018; Lopes et al., 2018; Shi and Chen, 2018; Souza et al., 2018; Denisova and Cairns, 2019; Dey et al., 2019; Hendrix et al., 2019; Liang et al., 2019; Pan et al., 2019; Papadimitriou et al., 2019; Peng et al., 2019; Plass et al., 2019; Sepulveda et al., 2019). From the perspective of personalization and adaptability, this article systematically discusses the research status and design methods of electronic science games to enhance cognitive ability, as well as the advantages and challenges of personalized and adaptive design in electronic science games.

OPEN ACCESS

Edited by:
Peng Xu,
University of Electronic Science and Technology of China, China

Reviewed by:
Fangzhou Xu,
Qilu University of Technology, China
Minpeng Xu,
Tianjin University, China

****Correspondence:***
Yanhong Zhou
yhzhou168@163.com

Received: *05 October 2021*
Accepted: *20 October 2021*
Published: *12 November 2021*

Citation:
Wen D, Yao W, Xu J, Wang S, Zhong Y, Chen H, Dong X, Saripan MI and Zhou Y (2021) Electronic Science Games Used to Enhance Cognitive Ability: Opinion of Design From Personalization and Adaptation. Front. Aging Neurosci. 13:789547. doi: 10.3389/fnagi.2021.789547

THE STUDY VALUE OF ELECTRONIC SCIENCE GAMES WITH PERSONALIZATION AND ADAPTATION

In recent years, a series of studies based on brain science methods have shown that electronic science games have the effect of promoting cognitive enhancement and that individual differences in game players will make them obtain different effects in cognitive training (Shang and Zhang, 2017). Therefore, in order to allow gamers to obtain the best training results, more and more researchers are beginning to design games from the perspective of personalization and adaptation, and prove the advantages of personalized and adaptive games through experiments.

In 2019, Ahmed Tlili et al. divided 51 learners into a control group and an experimental group, and they learned through the non-personalized version and the personalized version of the game, respectively (Tlili et al., 2019). The experimental results showed that personalized educational games not only reduced the cognitive burden of learners, but also allowed learners to show higher perceived usefulness and willingness to use. In the same year, Troussas et al. conducted research on personalized games and showed that integrating personalization and collaboration into mobile game learning could further help higher education students to improve their knowledge and cognition level (Troussas et al., 2019).

In 2013, Sampayo-Vargas et al. began to explore adaptive design. They compared adaptive difficulty adjustment games and non-adaptive difficulty adjustment games. The results showed that players who played adaptive games had achieved better results (Sampayo-Vargas et al., 2013).

In 2014, Montani developed a new adaptive game, and the results showed that adaptive games can improve the cognitive ability of game players (Montani et al., 2014). Recently, Daghestani et al. developed a gamification learning system that combines gamification, classification and adaptive technologies. The results showed that the adaptive gamification learning system had a positive impact on students' learning participation and academic performance (Daghestani et al., 2020).

In summary, electronic science games incorporating personalized and adaptive elements have a significant positive effect on increasing player participation, reducing the burden on players in the game, and achieving game training goals.

PERSONALIZATION AND ADAPTATION DESIGN OF ELECTRONIC SCIENCE GAMES

With the continuous development of cognitive science and the video game industry, electronic science games for cognitive enhancement have been sought after. However, how to design such games is the core issue of this research. The following will introduce how such games are designed from the perspective of personalization and adaptation.

Personalized Design in Games

Personalized game design refers to games tailored for gamers based on their personality characteristics, game abilities, game styles, preferences, etc. (Sedleniece and Cakula, 2012; Aljabali and Ahmad, 2018). The personalized design of the game is mainly divided into four aspects.

1) Personalized virtual characters. Many researchers use different methods to personalize the design of virtual characters in the game, which can significantly increase the immersion and real experience of game players, and stimulate their own cognitive abilities (Reategui et al., 2006; Hollingdale and Greitemeyer, 2013; Zibrek et al., 2018; Peng et al., 2019).
2) Personalized game mode. The first personalization mode is to divide players into different types, and different types of players play different games (Nagle et al., 2016; Orji et al., 2017). The second type of personalization involves recommending different games according to the game players' gaming preferences without having to divide them into different types of players (González et al., 2019).
3) Personalized scene effects and sound effects. We can use visual effects and sound effects to design the personalize features of game. Separate design from the sequence of game scenes and the background music of the game to make the game more attractive and retention rate (Lin et al., 2013; Li et al., 2016).
4) Personalized game content. Personalized game content means that the game content is tailor-made for gamers, is "personalized" and adjusted according to the needs and preferences of specific gamers. Kucirkova et al. personalize the game content in different ways, so that gamers have an excellent gaming experience (Kucirkova and Flewitt, 2020).

Adaptive Design in Games

Most of the research on adaptive design is mainly reflected in the difficulty adaptation of the game (Mishra et al., 2016). The adaptively designed game will provide game players with continuous and appropriate level challenges, thereby enhancing the player's participation experience in the game (Orvis et al., 2008; Belanich et al., 2013; Csikszentmihalyi, 2014). The adaptive design of the game is mainly divided into three aspects.

1) Change the parameters in the game in real time. Real-time change the number of static objects in the scene, the speed of the character, the attack value of the character, and so on. Studies have shown that the difficulty of the game is balanced with the player's ability, which can stimulate the enthusiasm of the player, effectively master various skills in the game, and enable the player to successfully complete the challenge (Denisova and Cairns, 2019; Peng et al., 2019; Plass et al., 2019).
2) Automatically adjust the game difficulty level for game players. Before the game starts, each game has a corresponding difficulty level. In the game, according to the data generated by the player during the game, the player is automatically matched with a game of suitable difficulty.

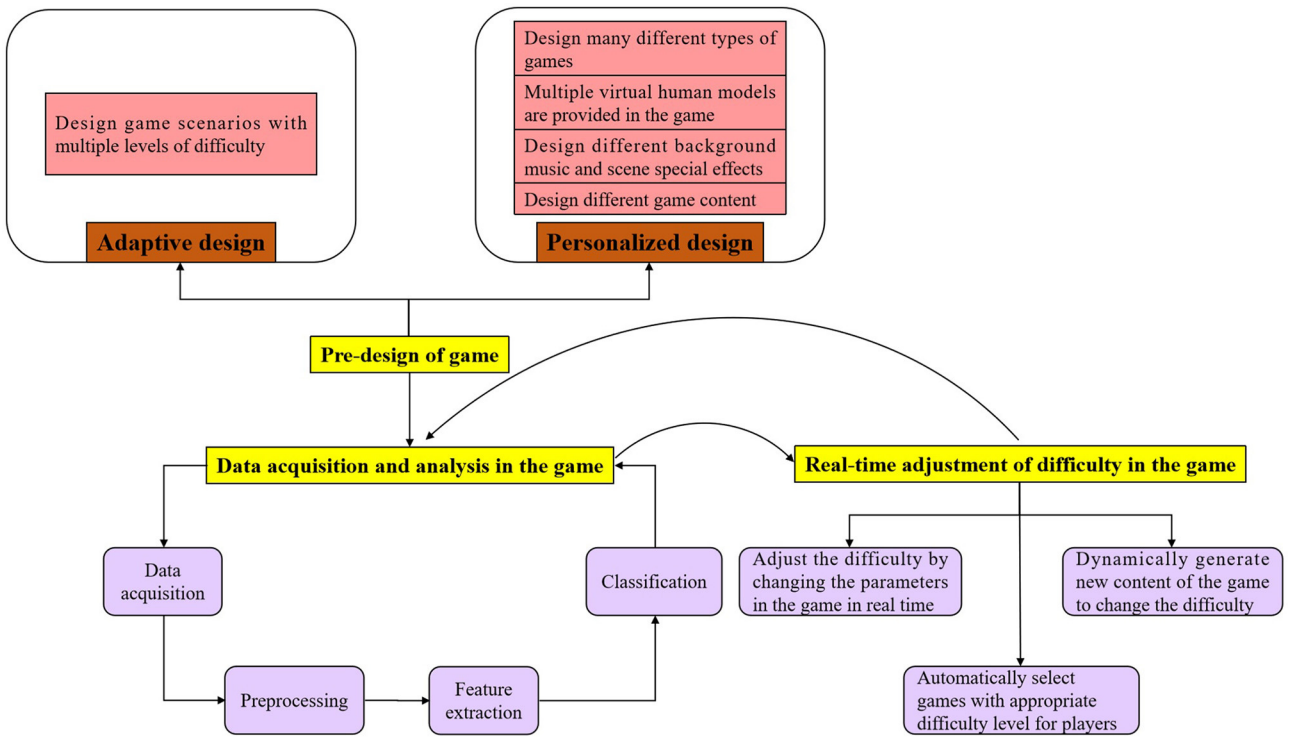


FIGURE 1 | Flow chart of adaptive and personalized game design. "Pre-game design" includes personalized design and adaptive design of the game. After the completion of the pre-game design, we can start "data collection and analysis in the game," and then adjust the difficulty of the game in real time according to the analysis results. Among them, "data acquisition and analysis in the game" and "real-time adjustment of difficulty in the game" feedback each other.

3) Dynamically generate new game content. The new content of the game is not generated manually, but automatically generated by the computer through the monitoring of the game data during the game. The game contains help, hints, storyline, etc. Different from personalized game content, adaptively and dynamically generated game content does not completely depend on the player's needs and interests, but is adjusted according to the player's real-time performance in the game.

Suggestions in Games Design

Combining the above-mentioned personalization and adaptive design of current electronic science games, we put forward the following suggestions for integrating personalization and adaptive design into electronic science games.

1) Early stage of game design. According to personalized design methods, many different types of games can be designed. The game can provide a variety of virtual character models, or design different content for the game (such as game tasks, help content, etc.). According to the adaptive design method, scenes with multiple difficulty levels can be designed for the same game.
2) Data collection and analysis in the game. Collect behavioral data of game players. When necessary, collect physiological signals such as EEG and ECG of game players, and perform real-time pre-processing, feature extraction and classification operations after acquiring the data. The game state is dynamically adjusted according to the data information of the game player.
3) Real-time difficulty adjustment design in the game. According to the previously suggested analysis results, determine whether the current game difficulty matches the player's ability, and then combine the adaptive design method to change the game difficulty so that the game player can get the best gaming experience.

After completing the pre-design of the game, we can play the game, collect and analyze the physiological signals of the subjects in real time, and then adjust the difficulty of the game in real time according to the analysis results. Among them, the data collection and analysis in the game and the real-time adjustment of the difficulty in the game feedback each other. The specific design process is shown in **Figure 1**.

DISCUSSION

In short, electronic science games that integrate personalization and adaptive design to improve cognitive ability can not only stimulate the interest of gamers, but also give gamers a stronger sense of substitution, and can also cultivate their concentration and execution. In terms of cognition, it has higher practicality (Guangxin and Fei, 2006).

However, there are still some problems to be solved in the personalized and adaptive design of electronic science games at this stage.

1) The ability of personalized recommendation needs to be improved. According to existing research, the personalized design concept of games is reflected in the fact that gamers can choose their favorite virtual avatars, game modes, scene effects, and game content, but most of the time gamers cannot distinguish which is their favorite, the result of the choice will cause the training results may not be the most effective.
2) The performance of data analysis algorithms needs to be improved. In the game, when the difficulty of the game is dynamically adjusted according to the player's state, it is necessary to collect player behavior data, and dynamically change the difficulty of the game according to the results of data analysis. In order to obtain the best adaptation effect, better data analysis algorithms are needed.
3) The player's immersive experience needs to be improved. Although the personalized and adaptive design of the game has been achieved at this stage, some players still have problems such as inattention, so next you can consider integrating the game into virtual reality to enhance their immersive experience.

Therefore, with the continuous development of electronic science games, we need to overcome the above-mentioned challenges and contribute to their further application in cognitive enhancement. In addition, considering that Alzheimer's Disease patients have obvious cognitive decline, it is a good choice to improve their cognitive ability to use the electronic science games that embody personalization and adaptability in the future.

AUTHOR CONTRIBUTIONS

YZhou, DW, and WY contributed to conception and design of the study. WY, YZhon, and HC searched the database. DW, WY, YZhon, and HC performed the analysis of literatures. WY, JX, and DW wrote the first draft of the manuscript. SW, YZhon, and HC wrote sections of the manuscript. XD and MS revised this paper and analyzed the literatures. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported in part by National Natural Science Foundation of China (61876165 and 61503326), Natural Science Foundation of Hebei Province in China (F2016203343).

REFERENCES

Aljabali, R. N., and Ahmad, N. (2018). "A review on adopting personalized gamified experience in the learning context," in 2018 *IEEE Conference on e-Learning, e-Management and e-Services* (Johor: IEEE). doi: 10.1109/IC3e.2018.8632635

Ashish, D., Dhaval, S., Bhasin, A., Das, A., Lahiri, U. (2018). An intelligent, adaptive, performance-sensitive, and virtual reality-based gaming platform for the upper limb. *Comput. Anim. Virt. Worlds*. 29, 155–169. doi: 10.1002/cav.1800

Belanich, J., Orvis, K. L., and Sibley, D. E. (2013). Pc-based game features that influence instruction and learner motivation. *Milit. Psychol*. 25, 206–217. doi: 10.1037/h0094963

Carro, R. M., Breda, A. M., Castillo, G., and Bajuelos, A. L. (2002). "A methodology for developing adaptive educational-game environments," in *International Conference on Adaptive Hypermedia and Adaptive Web-based Systems* (Aveiro: Springer-Verlag). doi: 10.1007/3-540-47952-X_11

Csikszentmihalyi, M. (2014). "Toward a psychology of optimal experience," in: *Flow and the Foundations of Positive Psychology*. Dordrecht: Springer.

Daghestani, L. F., Ibrahim, L. F., Reem, S., Al-Towirgi, R. S., and Salman, H. A. (2020). Adapting gamified learning systems using educational data mining techniques. *Comput. Appl. Eng. Educ*. 28, 568–589. doi: 10.1002/cae.22227

Denisova, A., and Cairns, P. A. (2019). Player experience and deceptive expectations of difficulty adaptation in digital games. *Enter. Comput*. 29, 56–68. doi: 10.1016/j.entcom.2018.12.001

Dey, A., Chatourn, A., and Billinghurst, M. (2019). "Exploration of an EEG-based cognitively adaptive training system in virtual reality," in 2019 *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Osaka: IEEE). doi: 10.1109/VR.2019.8797840

González, C., Toledo, P., Muñoz, V., and Torres-Carrion, P.V. (2019). Serious games for rehabilitation: gestural interaction in personalized gamified exercises through a recommender system. *J. Biomed. Inform*. 97:266. doi: 10.1016/j.jbi.2019.103266

Guangxin, W., and Fei, D. (2006). Summary of educational games' basic structure and character. *China Educ. Technol*. 7, 77–80. doi: 10.3969/j.issn.1006-9860.2006.07.022

Hendrix, M., Bellamy-Wood, T., Mckay, S., Bloom, V., and Dunwell, I. (2019). Implementing adaptive game difficulty balancing in serious games. *IEEE Trans. Games*. 11, 320–327. doi: 10.1109/TG.2018.2791019

Hollingdale, J., and Greitemeyer, T. (2013). The changing face of aggression: the effect of personalized avatars in a violent video game on levels of aggressive behavior. *J. Appl. Soc. Psychol*. 43, 1862–1868. doi: 10.1111/jasp.12148

Hunicke, R. (2005). The case for dynamic difficulty adjustment in games. *Int. Conf. Adv. Comput. Enter. Technol*. 265, 429-433. doi: 10.1145/1178477.1178573

Johnson, A., Tang, Y., and Franzwa, C. (2014). "kNN-based adaptive virtual reality game system," in *IEEE International Conference on Networking* (Miami, FL: IEEE), 457–462. doi: 10.1109/ICNSC.2014.6819669

Knutas, A., Roy, R. V., Hynninen, T., Granato, M., Kasurinen, J., and Ikonen, J. (2019). A process for designing algorithm-based personalized gamification. *Multimed. Tools Appl*. 78, 13593–13612. doi: 10.1007/s11042-018-6913-5

Kucirkova, N., and Flewitt, R. (2020). The future-gazing potential of digital personalization in young children's reading: views from education professionals and app designers. *Early Child Dev. Care*. 190, 135–149. doi: 10.1080/03004430.2018.1458718

Li, S., Yuan, X., and Kai, C. (2016). "Design and implementation of personalized interface of Chinese army chess," in *2016 Chinese Control and Decision Conference (CCDC)*. (Yinchuan: IEEE). doi: 10.1109/CCDC.2016.7531732

Li, Y. N., Yao, C., Li, D. J., and Kang, Z. (2014). Adaptive difficulty scales for parkour games. *J. Visual Lang. Comput*. 25, 868–878. doi: 10.1016/j.jvlc.2014.09.003

Liang, Z., Zhou, K., and Gao,. K. (2019). Development of virtual reality serious game for underground rock-related hazards safety training. *IEEE Access*. 7, 118639–118649. doi: 10.1109/ACCESS.2019.2934990

Lin, C., Yeh, Y., Hung, Y., and Chang, R. I. (2013). Data mining for providing a personalized learning path in creativity. *Comput. Educ*. 68, 199–210. doi: 10.1016/j.compedu.2013.05.009

Lopes, R., Eisemann, E., and Bidarra, R. (2018). Authoring adaptive game world generation. *IEEE Trans. Comput. Intellig. AI Games*. 10, 42–55. doi: 10.1109/TCIAIG.2017.2678759

Mishra, J., Anguera, J. A., and Gazzaley, A. (2016). Video games for neuro-cognitive optimization. *Neuron*. 90, 214–218. doi: 10.1016/j.neuron.2016.04.010

Montani, V., Michele, D., and Zorzi, M. (2014). A new adaptive videogame for training attention and executive functions: design principles and initial validation. *Front. Psychol.* 5:e00409. doi: 10.3389/fpsyg.2014.00409

Nagle, A., Wolf, P., and Riener, R. (2016). Towards a system of customized video game mechanics based on player personality: relating the big five personality traits with difficulty adaptation in a first-person shooter game. *Enter. Comp*. 13, 10–24. doi: 10.1016/j.entcom.2016.01.002

Orji, R., Mandryk, R. L., and Vassileva, J. (2017). Improving the efficacy of games for change using personalization models. *ACM Trans. Comp. Hum. Interac. (TOCHI)*. 24, 1–22. doi: 10.1145/3119929

Orji, R., and Moffatt, K. (2018). Persuasive technology for health and wellness: state-of-the-art and emerging trends. *Health Inform. J*. 24, 66–91. doi: 10.1177/1460458216650979

Orvis, K. A., Horn, D. B., and Belanich, J. (2008). The roles of task difficulty and prior videogame experience on performance and motivation in instructional videogames. *Comput. Hum. Behav*. 24, 2415–2433. doi: 10.1016/j.chb.2008.02.016

Pan, L., Song, A., Wang, S., and Duan, S. (2019). Experimental study on upper-limb rehabilitation training of stroke patients based on adaptive task level: a preliminary study. *Biomed Res. Int*. 2019, 1–9. doi: 10.1155/2019/2742595

Papadimitriou, S., Chrysafiadi, K., and Virvou, M. (2019). Fuzzeg: fuzzy logic for adaptive scenarios in an educational adventure game. *Multimedia Tools Appl*. 77, 32023–32053. doi: 10.1007/s11042-019-07955-w

Peng, C., Wang, D., Zhang, Y., and Xiao, J. (2019). A visuo-haptic attention training game with dynamic adjustment of difficulty. *IEEE Access* 7, 68878-68891. doi: 10.1109/ACCESS.2019.2918846

Plass, J. L., Homer, B. D., Pawar, S., Brenner, C., and Macnamara, A. P. (2019). The effect of adaptive difficulty adjustment on the effectiveness of a game to develop executive function skills for learners of different ages. *Cogn. Dev*. 49, 56–67. doi: 10.1016/j.cogdev.2018.11.006

Reategui, E., Boff, E., and Campbell, J. A. (2006). "Endowing a virtual character with personalization capabilities," in *IEEE International Conference on Tools With Artificial Intelligence* (Washington, DC: IEEE Computer Society). doi: 10.1109/ICTAI.2006.57

Sampayo-Vargas, S., Cope, C. J., Zhen, H., and Byrne, G. J. (2013). The effectiveness of adaptive difficulty adjustments on students' motivation and learning in an educational computer game. *Comput. Educ*. 69, 452–462. doi: 10.1016/j.compedu.2013.07.004

Schadenberg, B. R., Neerincx, M. A., Cnossen, F., and Looije, R. (2017). Personalising game difficulty to keep children motivated to play with a social robot: a bayesian approach. *Cogn Syst Res*. 43, 222–231. doi: 10.1016/j.cogsys.2016.08.003

Sedleniece, M., and Cakula, S. (2012). "Framework for personalized E-learning model," in *Wseas International Conference on Communications* (Kos Island)

Sepulveda, G. K., Besoain, F., and Barriga, N. A. (2019). "Exploring dynamic difficulty adjustment in videogames," in 2019 *IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)* (Valparaso: IEEE). doi: 10.1109/CHILECON47746.2019.8988068

Shaker, N., Togelius, J., and Nelson, M. J. (2016). *Procedural Content Generation in Games. Springer International Publishing.*

Shang, J., and Zhang, L. (2017). A review of game-based learning based on cognitive neuroscience. *E-educ. Res*. 38, 104–111. doi: 10.13811/j.cnki.eer.2017.02.014

Shi, P., and Chen, K. (2018). Learning constructive primitives for real-time dynamic difficulty adjustment in super mario bros. *IEEE Trans. Games*. 10, 155–169. doi: 10.1109/TCIAIG.2017.2740210

Soares, D., Bruno, F., and Furtado, A. L. (2018). Player behavior and personality modeling for interactive storytelling in games. *Entertain. Comp*. 28, 32–48. doi: 10.1016/j.entcom.2018.08.003

Soler-Dominguez, J. L., Camba, J. D., Contero, M., and Alcaiz, M. (2017). "A proposal for the selection of eye-tracking metrics for the implementation of adaptive gameplay in virtual reality based games," in *Virtual, Augmented and Mixed Reality, 10280*, eds S. Lackey and J. Chen (Valencia; Cham: Springer), 369–380.

Souza, G., Deus, D., Tadaiesky, V., Araújo, I., Monteiro, D. C., and Lima de Santana, A. (2018). Optimizing tasks generation for children in the early stages of literacy teaching: a study using bio-inspired metaheuristics. *Soft Comput*. 22, 6811–6824. doi: 10.1007/s00500-018-3409-1

Tlili, A., Denden, M., Essalmi, F., Jemni, M., Kinshuk, Chen, N., et al. (2019). Does providing a personalized educational game based on personality matter? A case study. *IEEE Access* 7, 119566–119575. doi: 10.1109/ACCESS.2019.2936384

Togelius, J., Yannakakis, G. N., Stanley, K. O., and Browne, C. (2010). "Searchbased procedural content generation," in *Proceedings of the 2010 International Conference on Applications of Evolutionary Computation* (Istanbul). doi: 10.1007/978-3-642-12239-2_15

Troussas, C., Krouska, A., and Sgouropoulou, C. (2019). Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education. *Comput. Educ*. 144:103698. doi: 10.1016/j.compedu.2019.103698

Waltemate, T., Gall, D., Roth, D., Botsch, M., and Latoschik, M. E. (2018). The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Trans. Vis. Comput. Graph* 24, 1643–1652. doi: 10.1109/TVCG.2018.2794629

Wei, P. (2009). Design and evaluation of a computer game to promote a healthy diet for young adults. *Health Commun*. 24, 115–127. doi: 10.1080/10410230802676490

Yannakakis, G. N., and Togelius, J. (2015). "Experience-driven procedural content generation (extended abstract)," in *International Conference on Affective Computing and Intelligent Interaction* (Xi'an: IEEE). doi: 10.1109/ACII.2015.7344619

Zibrek, K., Kokkinara, E., and Mcdonnell, R. (2018). The effect of realistic appearance of virtual characters in immersive environments - does the character's personality play a role? *IEEE Trans. Visual. Comput. Graph*. 24:1681. doi: 10.1109/TVCG.2018.2794638

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 18 November 2021
doi: 10.3389/fnagi.2021.766884

Pathological Gait Signatures of Post-stroke Dementia With Toe-Off and Heel-to-Ground Angles Discriminate From Alzheimer's Disease

Linhui Ni[1], Wen Lv[1], Di Sun[1], Yi Sun[1,2], Yu Sun[1,3,4], Xinxin Xu[1], Mengyue Chang[5], Xing Han[5], Shuai Tao[5], Xingyue Hu[1] and Huaying Cai[1]**

[1] Department of Neurology, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou, China, [2] Electroencephalogram Unit, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou, China, [3] Key Laboratory for Biomedical Engineering of Ministry of Education, Department of Biomedical Engineering, Zhejiang University, Hangzhou, China, [4] Zhejiang Lab, Hangzhou, China, [5] Dalian Key Laboratory of Smart Medical and Health, Dalian University, Dalian, China

OPEN ACCESS

Edited by:
Jing Jin,
East China University of Science and Technology, China

Reviewed by:
Fan Gao,
University of Kentucky, United States
Zeping Lv,
National Research Center for Rehabilitation Technical Aids, China

***Correspondence:**
Xingyue Hu
huxingyue2003@zju.edu.cn
Huaying Cai
caihuaying2004@zju.edu.cn

Received: 30 August 2021
Accepted: 13 October 2021
Published: 18 November 2021

Citation:
Ni L, Lv W, Sun D, Sun Y, Sun Y, Xu X, Chang M, Han X, Tao S, Hu X and Cai H (2021) Pathological Gait Signatures of Post-stroke Dementia With Toe-Off and Heel-to-Ground Angles Discriminate From Alzheimer's Disease.
Front. Aging Neurosci. 13:766884.
doi: 10.3389/fnagi.2021.766884

Given the limited power of neuropsychological tests, there is a need for a simple, reliable means, such as gait, to identify mild dementia and its subtypes. However, gait characteristics of patients with post-stroke dementia (PSD) and Alzheimer's disease (AD) are unclear. We sought to describe their gait signatures and to explore gait parameters distinguishing PSD from post-stroke non-dementia (PSND) and patients with AD. We divided 3-month post-stroke patients into PSND and PSD groups based on the Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), and the activity of daily living (ADL). Thirty-one patients with AD and thirty-two healthy controls (HCs) were also recruited. Ten gait parameters in one single and two dual-task gait tests (counting-backward or naming-animals while walking) were compared among the groups, with adjustment for baseline demographic covariates and the MMSE score. The area under the receiver operating characteristic curve (AUC) was used to identify parameters discriminating PSD from individuals with PSND and AD. Patients with PSD and patients with AD showed impaired stride length, velocity, stride time, and cadence while patients with PSD had altered stance and swing phase proportions (all $p \leq 0.01$, *post hoc*). Patients with AD had smaller toe-off (ToA) and heel-to-ground angles (HtA) ($p \leq 0.01$) than HCs in dual-task gait tests. Individuals with PSD had a shorter stride length, slower velocity, and altered stance and swing phase percentages in all tests ($p \leq 0.01$), but a higher coefficient of variation of stride length (CoV_{SL}) and time (CoV_{ST}) only in the naming animals-task gait test ($p \leq 0.001$) than individuals with PSND. ToA and HtA in the naming animals-task gait test were smaller in individuals with AD than those with PSD ($p \leq 0.01$). Statistical significance persisted after adjusting for demographic covariates, but not for MMSE. The pace and the percentage of stance or swing phase in all tests, CoV_{ST} in the dual-task paradigm, and CoV_{SL} only in the naming animals-task gait test (moderate accuracy, AUC > 0.700, $p \leq 0.01$) could

distinguish PSD from PSND. Furthermore, the ToA and HtA in the naming animals-task gait paradigm discriminated AD from PSD (moderate accuracy, AUC > 0.700, $p \leq 0.01$). Thus, specific gait characteristics could allow early identification of PSD and may allow non-invasive discrimination between PSD and AD, or even other subtypes of dementia.

Keywords: gait, cognition, post-stroke dementia, Alzheimer's disease, dual-task gait

INTRODUCTION

Approximately, 50 million people suffer from dementia (WHO, 2018). It is one of the major causes of disability and mortality among aging adults (WHO, 2018). The two most common types of dementia are Alzheimer's disease (AD) and vascular dementia (VaD) caused by stroke (WHO, 2018; Emrani et al., 2020). Dementia is characterized by amnesia, impaired executive function, visuospatial capacity, and attention. Currently, diagnosis depends on the temporal relationship of the disease symptoms with imaging examination findings and neuropsychological tests (Emrani et al., 2020; Ismail et al., 2020). However, early diagnosis of post-stroke dementia (PSD) and mild AD is difficult because of the limited sensitivity of the cognitive function scales, particularly under repetitive interview conditions. Additionally, educational attainment, cultural background, and even hearing or speaking abilities may reduce the specificity of the neuropsychological tests. Overlapping symptoms and imaging manifestations, multifactorial causes, and homogeneity of the histopathology limit the accuracy of distinguishing among dementia subtypes. Pathological biopsy of brain tissue is the gold standard for dementia classification, but this is not generally applicable due to its invasive nature (Schott et al., 2010). There is no specific biomarker that can robustly identify vulnerable patients with PSD from patients with post-ischemic stroke or non-PSD dementia subtypes. Thus, there is a need to identify safe, reliable, and effective clinical markers to enhance diagnostic accuracy.

The human gait is remarkably complex. Gait in older people is divided into five primary modal domains: pace, rhythm, variability, asymmetry, and postural control (Lord et al., 2013). An integrated gait reflects the health of individuals, particularly in compensating for changes in postural balance and preventing falls. This is controlled by well-balanced neural circuits and specific brain structures involving the frontal and limbic regions, basal ganglia, cerebellum, and optical, vestibular, sensory, and motor systems (Takakusaki, 2013; Tian et al., 2017; Allali et al., 2019). Memory, attention, executive function, and visual-spatial capacity share some overlapping brain regions related to gait (Morris et al., 2016). Therefore, gait is no longer regarded as a purely autonomic movement. A healthy integrated gait requires attention, executive function, and visual and auditory capacities. Spatiotemporal gait characteristics in the single-gait test of cerebrovascular disease and neurodegenerative diseases have been described, particularly for cases with mild cognitive impairment (MCI), AD, and Parkinson's disease (PD) (Mc Ardle et al., 2019; Sanders et al., 2020). In recent years, studies have increasingly implemented the dual-task gait paradigm, which requires subjects to walk while accomplishing an additional cognitive task, to reflect the cognitive challenges at the cognitive-motor interface, increasing the sensitivity for discovering occult cognitive deterioration (Bayot et al., 2018).

Gait stride length and velocity, belonging to the pace domain of gait, have been assessed most commonly in this field, because of the ease of acquisition. The decreased pace and increased instability have been detected in the general older population (Cohen et al., 2016; Noce Kirkwood et al., 2018; Rasmussen et al., 2019). It is controversial whether individuals with MCI show gait dysfunction as compared to matched healthy aging adults. In addition, MCI sufferers who walk slower and who demonstrate higher dual-task costs have been shown to be at risk of progression to dementia (Montero-Odasso et al., 2017). Recently, numerous cohort studies have shown weaker gaits in patients with AD, manifested as decreased pace, greater variability, and worse rhythm in the normal gait test, and pathological gait parameters would be more sensitive measured during dual-task gait measurements (Mc Ardle et al., 2017, 2019). Moreover, evidence suggests that the asymmetry increases with cognitive decline (Ghoraani et al., 2021). There were rarely differences found in postural control during walking between patients with AD and age-matched healthy adults or those with other cognitive impairments (Gillain et al., 2009; Maquet et al., 2010).

In terms of discrete gait characteristic comparisons among dementia subtypes, reports have outlined distinctive patterns of gait damage under a few dual-task gait measurements. Differences in gait damage in individuals with AD and those with non-AD dementia had frequently been reported, mostly in the late stage of AD. Patients with AD showed less impairment in pace, rhythm, and variability than those with non-AD dementia, such as fronto-temporal dementia and Lewy body dementia (Beauchet et al., 2016). People with VaD showed a poorer pace than patients with AD (Allan et al., 2005). However, there are few studies available on domains of gait other than pace for distinguishing between VaD and AD. There have been rare descriptions of differences in gait between individuals with PSD, a subtype of VaD, and those with AD, and there is no specific gait-based predictor that can identify early dementia in post-stroke patients. Early recognition of PSD in patients with ischemic stroke, at 3 months from stroke initiation, is crucial for the follow-up treatment strategy and predicting prognosis, because of

Abbreviations: PSD, post-stroke dementia, PSND, post-stroke non-dementia, AD, Alzheimer's disease, HCs, healthy control subjects, CoV, coefficient of variation, ToA, toe-off angle, HtA, heel-to-ground angle, MMSE, Mini-Mental State Examination, MoCA, Montreal Cognitive Assessment, HAMD, 17-item Hamilton Depression Rating Scale, HAMA, Hamilton Anxiety Rating Scale, ADL, activity of daily living, ROC, receiver operating characteristic, AUC, area under the ROC curve, NIHSS, National Institution Health of Stroke Scale, mRS, Modified Rankin Scale, VaD, vascular dementia, MCI, mild cognitive impairment, PD, Parkinson's disease.

the probable remission and even reversion of PSD (van der Flier et al., 2018). On the other hand, specific subtypes of dementia require targeted therapy to control disease progression. Thus, the early and definite typology of dementia is of great clinical significance. In recent years, studies of machine learning, based on using data from wearable devices to detect spatiotemporal gait parameters, have increased and have been confirmed as an effective method for measuring the relationship between gait parameters and neurological functions (Cheng et al., 2020). The inconsistent cognitive condition of post-stroke patients would place varying degrees of cognitive load on gait performance. We hypothesized that this might be differences in spatiotemporal gait parameters between states of dementia (PSD) and non-dementia post-stroke (PSND). On the other hand, discrete pathologies and cognition formation may result in unique gait patterns between individuals with PSD and those with AD.

The goal of this study was to elucidate a motion marker that could identify PSD in ischemic stroke patients, and to outline the typical gait features in individuals with PSD and those with AD, providing a low-cost, feasible, and effective means for earlier detection of PSD or non-PSD subtypes of dementia in order to implement interventions for cognitive impairment as early as possible. Thus, in the present study, we compared spatiotemporal gait patterns and neuropsychiatric parameters among healthy older individuals (HCs) and age-matched individuals with PSD, PSND, or AD. To this end, we introduced parameters, i.e., the toe-off angle (ToA) and heel-to-ground angle (HtA), respectively, measured at the moment of initiation or end of the swing phase, which has not been reported in previous studies of PSD and AD, to the best of our knowledge. Respective comparisons of PSD with PSND or AD were performed to clarify gait differences that allowed the distinction of these conditions.

PARTICIPANTS AND METHODS

Study Design and Participants

This clinical study is a monocentric cross-sectional study. Sixty-six outpatients at 3-month post ischemic stroke and 31 patients with mild to moderate AD were recruited from the Clinic of the neurology department, Sir Run Run Shaw Hospital of Zhejiang University in China, and 32 matched healthy control subjects (HCs) were recruited from the Clinic or physical examination center. The stroke patients were first-episode with definitive acute ischemic based on MRI, and in normal cognition before the stroke. They were then categorized into the group of PSD or PSND depending on the cognitive and activity of daily living assessments, and related clinical presentations. The stroke related scales of recruited samples were defined as follows: the National Institution Health of Stroke Scale (NIHSS) $\leq$4, the Modified Rankin Scale (mRS) $\leq$2, and the muscle strength $\geq$ 4+ grade.

All participants had to be aged over 55 years old and can walk at least 10 m without any assistant. We will exclude adults who: can't speak fluently; with Parkinsonism symptoms or other neurological diseases influencing cognition or gait (such as PD dementia with Lewy bodies, frontal–temporal dementia, and dystonia); with osteoarticular diseases which might influence on walking; and/or with severe mental illness, such as major depression (total score > 7 on the 17-item Hamilton Depression Rating Scale), anxiety (total score > 7 on Hamilton Anxiety Rating Scale), bipolar disorder, and schizophrenia, or any psychotropic drugs taken.

Clinical and Cognitive Assessment

At all follow-up visits, participants had an interview with the same neurology doctor to complete the demography baseline information collection including age, gender, education level, height, weight, comorbidities, and habits of smoking and drinking. The interviewers combined the results of medical history and imaging reports, with neurological examination to confirm whether the subject enrolled. The depression condition was evaluated using the 17-item Hamilton Depression Rating Scale (HAMD), and the anxiety condition was measured by Hamilton Anxiety Rating Scale (HAMA). The Activity of Daily Living Scale (ADL) was used to assess the self-care ability of patients in daily life. The baseline and 3 months post stroke of NIHSS and mRS of ischemic stroke patients were assessed, with the muscle strength and lesion side reported by MRI were also collected.

Post-stroke dementia was diagnosed by the two same neurology doctors according to the 2019 Chinese Vascular Cognitive Impairment Guideline, which defined PSD as a status with cognitive impairment and impaired activity of daily living lasting for 3 months after stroke onset (Cognitive Impairment Committee Nb, Chinese Medical Doctor Association, 2019). We, therefore, determined the enrolled stroke patients in 3 months post stroke, which is also consistent with the international consensus (within 6 months) (Skrobot et al., 2018). All patients with AD met the 2011 revised criteria for AD diagnosis of National Institute on Aging-Alzheimer's Association diagnostic guidelines (Khachaturian, 2011). The clinical and MRI data of all patients are available.

The cognitive assessments included the Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), and Mini-Cog. On the basis of Chinese national conditions and education levels of the general elderly adults, we defined the cognitive impairment assessed by MMSE to illiteracy $\leq$19, primary school $\leq$22, middle school and above $\leq$26 (Li et al., 2016), MoCA to illiteracy $\leq$13, primary school $\leq$19, middle school and above $\leq$24 (Lu et al., 2011), and Mini-Cog $\leq$3 (McCarten et al., 2011).

Gait Testing Procedure

All gait assessments were performed in a spacious hallway outside the clinic room using wearable motion sensors (JiBuEn^{R} gait analysis system, version 2.3). Patients walked at least 10 m at their comfortable pace with or without a cognitive task. Five steps of each start and end of the pathway were deleted to ensure the acceleration and deceleration phases were not recorded. All participants were asked to accomplish three gait trials. First, participants were asked to walk at their normal, everyday walking speed. The next two dual-task gait tests comprised walking while counting backward, or naming animals, which have been validated in previous clinical trials that robustly increase

cognitive demand of adults (Montero-Odasso et al., 2017). Only participants who completed all three trials were included in the analysis. Spatiotemporal gait parameters representing pace (stride length and velocity), rhythm (stride time, cadence, percentages of stance and swing phase), variability [coefficient of variation (CoV) of stride length (CoV_{SL}) and stride time (CoV_{ST})], and postural control involving ToA and HtA were collected. The CoV was calculated as follows:

$$\text{CoV}(\%) = \text{SD of parameter}/\text{Mean of parameter} \times 100\%$$

Toe-off angle was defined as the angle of toe-off the ground measured at the moment of initiation of the swing phase. HtA was defined as the angle of initial heel stride to the ground measured at the moment of initiation of the stance phase (**Figure 1**).

Covariates

Analyses were adjusted for covariates that included demographics (age, gender, education levels, and height), numbers of comorbidity, and baseline cognition (MMSE). In a comparison of PSD and PSND, muscle strength was added to the covariates.

Statistical Analysis

The baseline demography and clinical information, cognitive assessments, and gait parameters of four groups were displayed through descriptive analysis. We evaluated the data distribution of all quantitative variables in each group using the one-sample Kolmogorov–Smirnov test and histograms. Variables that had normal distribution will be described using means and SD, while variables with non-normal distribution were presented using median and inter-quartile range (IQR). Categorical variables were described by frequencies and percentages.

Physical information, cognitive scores, and gait parameters comparisons among four groups were analyzed using a one-way ANOVA test when quantitative variables were in normal distribution and homoscedasticity, otherwise, they were analyzed using the Kruskal–Wallis *H* test. The chi-square test was used to determine the categorical data. The LSD analysis was used for *post hoc* analysis, of which, significance was defined as $p \leq 0.01$. Student *t*-test was assessed to compare the gait parameters of PSD with PSND or AD group, respectively, yet the pairwise analysis was using Mann–Whitney *U* test if data in non-normal distribution or non-homoscedasticity. General linear regression was used to control for primary demography covariates of age, gender, education levels, height, and numbers of comorbidity with or without MMSE when data in normal distribution and homoscedasticity. Otherwise, generalized linear regression was used. The parameter under $p < 0.05$ was determined as the potential predictor of PSD from AD. To more precisely identify the PSD from patients with stroke, the significance was shrinking to a more conservative threshold of $P \leq 0.01$. Moreover, stepwise logistic regression was used to verify the superiority of the above gait predictor for identifying PSD. The receiver operating characteristic curve analysis (ROC) and area under the curve (AUC) determined the overall accuracy of possible distinguisher for PSD from individuals with ischemic stroke or AD.

RESULTS

Baseline Information

A total of 127 subjects (32 HCs, 32 PSND, 32 PSD, and 31 AD participants) were included in this longitudinal study. The baseline characteristics of participants, including medical and cognitive conditions, are summarized in **Table 1**. The numbers of comorbidities ($\chi^2 = 27.22$, $p \leq 0.001$) and ratios of hypertension ($\chi^2 = 26.8$, $p \leq 0.001$) were significant among the four groups. The age, female proportion, education level, height, and proportion of those with diabetes, smoking and drinking habits, and depression and anxiety conditions were not significantly different ($p > 0.05$). The activities of daily living of PSD and AD were significantly different from HCs and PSND, respectively ($p \leq 0.001$, *post hoc*, **Table 1**). The NIHSS, mRS, muscle strength, and infarcted lesion side were not significantly different between PSD and PSND patients ($p > 0.05$, **Supplementary Table 1**).

The MMSE, MoCA, and Mini-Cog scores were significantly different among the four groups after controlling for education level (MMSE: Wald $\chi^2 = 180.686$, $p \leq 0.001$; MoCA: Wald $\chi^2 = 365.823$, $p \leq 0.001$; Mini-Cog: Wald $\chi^2 = 220.423$, $p \leq 0.001$, **Table 1**). PSD and AD groups showed lower MMSE, MoCA, and Mini-Cog scores than the HCs and PSND groups ($p \leq 0.01$, *post hoc*, **Table 1**), and the AD group had lower Mini-Cog scores than the PSD group ($p \leq 0.01$, *post hoc*, **Table 1**). The gait characteristics of the four gait domains among the four groups are presented in **Table 2**.

Gait Impairment in the Dementia Subtypes (Alzheimer's Disease and Post-stroke Dementia) Compared With Healthy Controls

Compared to HCs, patients with PSD and AD had a shorter stride length, slower gait velocity, decreased cadence, and longer stride time in single or dual-task gait tests ($p \leq 0.01$, *post hoc*, **Table 2**). In the PSD group, the percentage of time spent in the stance phase was longer and that spent in the swing phase was shorter in all gait tests, while disturbed CoV_{ST} and HtA were observed only in the naming animals-task gait test ($p \leq 0.01$, *post hoc*, **Table 2**). We observed greater CoV_{SL} and smaller ToA and HtA in the AD than in the HCs group in all gait paradigms.

Gait Impairment in Post-stroke Dementia Compared to Post-stroke Non-dementia Individuals

Patients with ischemic stroke showed no difference in NIHSS score, mRS score, muscle strength, or lesion side when they returned to the clinic at the third month post-stroke ($p > 0.05$, **Supplementary Table 1**). The spatiotemporal gait parameters of individuals in the PSND group were not different from those of HCs ($p > 0.05$, *post hoc*, **Table 2**). However, a worsening pace and disturbed gait phase of individuals with PSD were noted in the

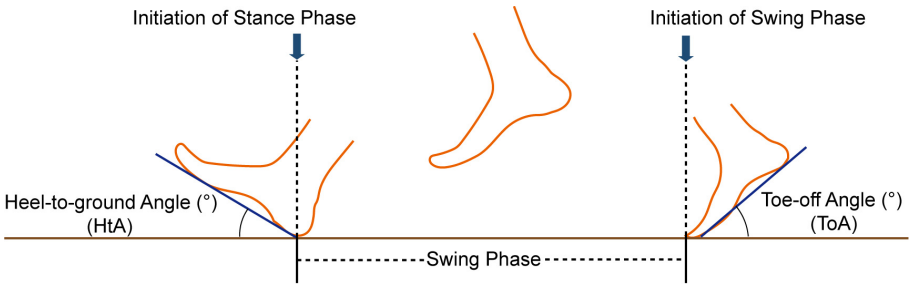


FIGURE 1 | The toe-off angle (ToA) and heel-to-ground angle (HtA) in the gait cycle.

TABLE 1 | Baseline characteristics of participants stratified by included samples.

	HCs	PSND	PSD	AD	F/χ2	*P*
Characteristics	*n* = 32	*n* = 32	*n* = 32	*n* = 31		
Age, mean (SD)	67.7 (5.13)	67.8 (6.53)	68.4 (8.44)	71.8 (8.23)	2.08	0.106
Female, no. (%)	19 (59.4)	11 (34.4)	12 (37.5)	18 (58.1)	6.69	0.082
Education level, no. (%)					10.46	0.106
Illiterate	2 (6.3)	8 (25.0)	10 (31.3)	5 (16.1)		
Primary school	7 (21.9)	8 (25.0)	9 (28.1)	5 (16.1)		
Middle school and above	23 (71.9)	16 (50.0)	13 (40.6)	21 (67.7)		
Height, mean (SD)	161.6 (5.65)	166.0 (7.27)	164.3 (6.71)	161.5 (5.88)	2.60	0.056
No. of comorbidities, median (IQR)	1.0 (2.0)	2.0 (1.0)	2.0 (2.0)	1.5 (2.0)	27.22	**<0.001**
Comorbidities, no. (%)						
Hypertension	11 (34.4)	25 (78.1)	29 (90.6)	16 (51.6)	26.80	**<0.001**
Diabetes	6 (18.8)	10 (31.3)	9 (28.1)	6 (19.4)	2.04	0.565
Smoking, no. (%)	10 (31.3)	15 (46.9)	14 (43.8)	8 (25.8)	4.08	0.253
Drinking, no. (%)	4 (12.5)	8 (25.0)	10 (31.3)	6 (19.4)	3.57	0.312
HAMD, median (IQR)	3.0 (3.0)	2.0 (4.0)	3.0 (5.0)	4.0 (3.0)	1.66	0.645
HAMA, median (IQR)	2.0 (4.0)	3.0 (4.0)	3.0 (5.0)	3.0 (5.0)	5.78	0.123
ADL, median (IQR)	14.0 (0)^{S,A}	14.19 (0)	20.5 (7.0)^{N}	20.3 (12.0)	103.03	**<0.001**
Cognition assessment					***P***	**Adj. *P***
MMSE, median (IQR)	27.0 (2.0)^{S,A}	27.0 (4.0)	22.5 (10.0)^{N}	17.0 (11.0)	**<0.001**	**<0.001**
MoCA, median (IQR)	25.0 (3.0)^{S,A}	25.0 (4.0)	16.5 (7.0)^{N}	13.0 (10.0)	**<0.001**	**<0.001**
Mini-Cog, median (IQR)	5.0 (1.0)^{S,A}	5.0 (1.0)	2.0 (2.0)^{N,A}	1.0 (2.0)	**<0.001**	**<0.001**

Data of continuous variables described as means (SD) were assessed using One-way ANOVA analysis, whereas data displayed as median (IQR) were used Kruskal–Wallis H tests. Data of categorical variables were described by frequencies and percentages using the chi-square test. Bold values highlight the significant difference. Adj. P, P-value when adjusting for education level; S, different to PSD; A, different to AD; N, different to PSND; HCs, healthy controls; PSND, post-stroke non-dementia; PSD, post-stroke dementia; AD, Alzheimer's disease; BMI, body mass index; HAMD, 17-item Hamilton Depression Rating Scale; HAMA, Hamilton Anxiety Rating Scale; ADL, Activity of Daily Living Scale; MMSE, Mini-mental State Examination; MoCA, Montreal Cognitive Assessment.

single and dual-task gait tests ($p \leq 0.01$, **Table 3** and **Figure 2**). Individuals with PSD showed increased CoV_{ST}, stride time, and decreased cadence during dual-tasks, while increased CoV_{SL} was only observed in this group during the naming animals-task gait test ($p \leq 0.01$, **Table 3** and **Figure 2**). The differences remained robust after controlling for primary baseline covariates (age, gender, education levels, height, and numbers of comorbidity, modal 1 of **Table 3**), except for the CoV_{ST} and stride time during the counting backward-task, and cadence in the two dual-task gait paradigms ($p \leq 0.01$, **Table 3**). The differences were no longer significant after additional adjustment for MMSE scores and other primary baseline (age, gender, education levels, height, and numbers of comorbidity, modal 2 of **Table 3**) ($p > 0.01$, **Table 3**).

Moreover, stepwise logistic regression validated the importance of the above parameters to distinguish patients with PSD from that of PSND ($p < 0.05$, **Table 4**), and the AUCs showed that stride length, velocity, and the percentage of time spent in the stance or swing phase in the counting-task gait test showed moderate accuracy for distinguishing PSD from PSND individuals (AUCs ≥ 0.725, **Figure 3**). The CoV_{ST} in the naming animals-gait test might be optimal for recognizing subjects with PSD from individuals with PSND [AUC = 0.800 (0.685–0.915), $p \leq 0.001$, **Figure 3**].

TABLE 2 | Comparison of gait characteristics among controls, non-dementia post-stroke, and dementia subtypes.

	HCs	PSND	PSD	AD	*P*	Adj. *P*
Pace						
Stride length (m)						
Single-task	1.18 (0.15)^{S,A}	1.20 (0.16)	1.05 (0.31)^{N}	1.06 (0.24)	**0.001**	**<0.001**
Counting	1.16 (0.13)^{S,A}	1.24 (0.10)	1.05 (0.26)^{N}	1.08 (0.32)	**<0.001**	**<0.001**
Naming animals	1.12 (0.19)^{S,A}	1.18 (0.19)	1.00 (0.29)^{N}	1.00 (0.35)	**<0.001**	**<0.001**
Velocity (m/s)						
Single-task	1.10 (0.18)^{S,A}	1.05 (0.20)	0.96 (0.28)^{N}	0.91 (0.27)	**<0.001**	**<0.001**
Counting	1.08 (0.18)^{S,A}	1.07 (0.18)	0.92 (0.36)^{N}	0.88 (0.33)	**<0.001**	**<0.001**
Naming animals	0.98 (0.12)^{S,A}	0.92 (0.24)	0.72 (0.23)^{N}	0.70 (0.53)	**<0.001**	**<0.001**
Variability						
CoV_{SL} (%)						
Single-task	3.99 ± 1.35^{S,A}	4.37 ± 1.18	5.26 ± 1.74	5.96 ± 2.00	**<0.001**	**<0.001**
Counting	4.41 (1.81)^{A}	5.35 (1.59)	5.76 (2.61)	5.20 (3.91)	0.066	**0.031**
Naming animals	5.32 (3.96)^{S,A}	5.07 (1.56)	7.16 (4.74)^{N}	8.64 (4.69)	**<0.001**	**<0.001**
CoV_{ST} (%)						
Single-task	2.34 ± 1.06	2.35 ± 0.84	3.10 ± 1.67	3.13 ± 1.19	**0.010**	0.103
Counting	2.22 (2.89)	1.95 (1.15)	3.24 (1.49)^{N}	2.75 (1.80)	**0.005**	0.246
Naming animals	5.00 (5.10)	3.84 (2.19)	6.13 (4.00)^{N}	6.67 (6.73)	**<0.001**	**<0.001**
Rhythm						
Stride time (s)						
Single-task	1.06 (0.07)^{S,A}	1.11 (0.09)	1.16 (0.13)	1.15 (0.21)	**<0.001**	**<0.001**
Counting	1.08 (0.13)^{S,A}	1.23 (0.05)	1.19 (0.18)	1.17 (0.25)	**<0.001**	**<0.001**
Naming animals	1.12 (0.15)^{S,A}	1.22 (0.20)	1.39 (0.35)	1.34 (0.24)	**<0.001**	**<0.001**
Cadence (steps/min)						
Single-task	113.74 (7.56)^{S,A}	108.11 (8.88)	103.00 (12.75)	104.35 (18.09)	**<0.001**	**0.001**
Counting	110.09 (13.25)^{S,A}	106.67 (4.90)	100.84 (16.02)	102.56 (18.56)	**0.001**	**0.014**
Naming animals	101.93 (9.90)^{S,A}	96.00 (16.26)	85.41 (20.83)	86.96 (13.55)	**<0.001**	**<0.001**
Stance phase (%)						
Single-task	62.07 ± 1.70^{S}	62.15 ± 1.77	63.95 ± 2.62^{N}	63.09 ± 2.43	**0.002**	**0.014**
Counting	62.64 ± 2.03^{S}	62.49 ± 1.33	64.50 ± 2.79^{N}	63.56 ± 1.98	**0.001**	**0.009**
Naming animals	64.90 ± 2.49^{S}	64.89 ± 2.34	67.56 ± 4.06^{N}	66.84 ± 3.10	**0.001**	**0.009**
Swing phase (%)						
Single-task	37.93 ± 1.70^{S}	37.86 ± 1.77	36.04 ± 2.62^{N}	36.91 ± 2.44	**0.002**	**0.013**
Counting	37.36 ± 2.03^{S}	37.51 ± 1.33	35.52 ± 2.80^{N}	36.45 ± 1.98	**0.001**	**0.010**
Naming animals	35.12 ± 2.51^{S}	35.12 ± 2.35	32.45 ± 4.06^{N}	33.17 ± 3.11	**0.001**	**0.009**
Postural control						
Toe-off angle (°)						
Single-task	46.05 (5.76)^{A}	43.69 (20.76)	39.10 (25.74)	36.41 (30.27)	**0.002**	**0.009**
Counting	46.05 (7.97)^{A}	43.64 (27.64)	39.35 (26.20)	35.65 (31.38)	**0.007**	0.106
Naming animals	43.98 (7.27)^{S,A}	39.69 (22.91)	35.06 (24.65)	16.30 (25.38)	**<0.001**	**<0.001**
Heel-to-ground angle (°)						
Single-task	34.17 (4.81)	33.90 (23.19)	26.43 (18.36)	29.85 (23.88)	**0.008**	0.051
Counting	33.03 (6.06)	31.08 (25.90)	28.35 (20.41)	30.55 (23.50)	0.176	0.666
Naming animals	30.43 (6.47)^{A}	28.40 (22.01)	25.18 (16.92)	14.17 (8.60)	**<0.001**	**<0.001**

Data of continuous variables described as means ± SD were assessed using One-way ANOVA analysis if normal distributed and homogeneity, whereas displayed as median (IQR) and were used Kruskal–Wallis H test. The post hoc is to compare each group with every other group, respectively. The significant difference assessed by post hoc defined p ≤ 0.01 and was marked in the top right corner of the parameter result. Bold values highlight the significant difference among the four groups. The adjusted modal is controlling for age, gender, education levels, height, and numbers of comorbidity. S, different to PSD; A, different to AD; N, different to PSND; HCs, healthy controls; PSND, post-stroke non-dementia; PSD, post-stroke dementia; AD, Alzheimer's disease; CoV, coefficient of variation.

Differences in Gait Parameters Between Post-stroke Dementia and Alzheimer's Disease Individuals

In the naming animals-task gait test, participants in the AD group demonstrated significantly smaller ToA and HtA than individuals with PSD ($p \leq 0.01$, **Supplementary Table 2** and **Figure 4**), except for the single or counting-task gait tests. These two parameters of the postural control domain were robust, showing significant differences after controlling for age, gender, education level, height, and numbers of comorbidities, with or without MMSE scores (Adjusted modal 1 and Adjusted modal 2 of **Supplementary Table 2**) ($p \leq 0.01$, **Supplementary Table 2**). However, no other significant differences in gait parameters were found in the single or dual-task gait tests between individuals with PSD and AD.

TABLE 3 | Comparison of gait characteristics of individuals with post-stroke dementia (PSD) and individuals with post-stroke non-dementia (PSND).

	Unadjusted model		Adjusted model 1		Adjusted model 2	
	t/U	*P*	F/χ^2	*P*	F/χ^2	*P*
Pace						
Stride length (m)						
Single-task	256.5	**0.001**	14.060	**<0.001**	1.675	0.196
Counting	4.967	**<0.001**	23.278	**<0.001**	6.231	0.013
Naming animals	263.0	**0.004**	7.548	**0.006**	4.595	0.032
Velocity (m/s)						
Single-task	3.769	**<0.001**	10.832	**0.002**	1.939	0.170
Cunting	4.051	**<0.001**	11.016	**0.002**	4.027	0.045
Naming animals	3.875	**<0.001**	12.968	**0.001**	6.032	0.018
Variability						
CoV_{SL} (%)						
Single-task	−2.265	0.027	4.943	0.031	3.234	0.079
Forward counting	−0.914	0.364	0.267	0.608	0.139	0.711
Naming animals	−4.213	**<0.001**	13.161	**<0.001**	1.588	0.208
CoV_{ST} (%)						
Single-task	−2.204	0.033	3.330	0.074	2.668	0.109
Counting	623.0	**0.001**	6.235	0.013	3.867	0.049
Naming animals	−4.369	**<0.001**	16.292	**<0.001**	11.409	**0.001**
Rhythm						
Stride time (s)						
Single-task	541.0	0.266	1.158	0.282	1.228	0.268
Counting	−2.974	**0.005**	5.575	0.018	2.112	0.146
Naming animals	639.0	**0.005**	7.824	**0.005**	3.713	0.054
Cadence (steps/min)						
Single-task	1.699	0.094	2.086	0.155	1.163	0.286
Counting	2.957	**0.005**	5.195	0.023	1.960	0.161
Naming animals	2.752	**0.009**	5.831	0.016	0.957	0.328
Stance phase (%)						
Single-task	−3.163	**0.003**	7.285	**0.009**	3.569	0.065
Counting	−3.649	**0.001**	10.191	**0.001**	4.370	0.037
Naming animals	−3.216	**0.002**	7.455	**0.006**	0.685	0.408
Swing phase (%)						
Single-task	3.198	**0.002**	7.402	**0.009**	3.626	0.063
Counting	3.604	**0.001**	9.912	**0.002**	4.211	0.040
Naming animals	3.212	**0.002**	7.431	**0.006**	0.668	0.414
Postural control						
Toe-off angle (°)						
Single-task	409.0	0.167	0.486	0.486	1.079	0.299
Counting	408.5	0.165	0.077	0.781	0.253	0.615
Naming animals	387.5	0.136	1.315	0.251	1.600	0.206
Heel-to-ground angle (°)						
Single-task	362.5	0.045	2.302	0.129	4.346	0.037
Counting	434.0	0.295	0.157	0.694	0.179	0.674
Naming animals	1.256	0.214	1.135	0.021	1.159	0.287

Normally distributed data used Student's t-test, and control for primary covariates by general linear models, otherwise used Mann–Whitney U test, and control for primary covariates by generalized linear models. The significant difference is confined by $p \leq 0.01$. Bold values highlight the significant differences between the two groups. Adjusted model 1: Controlling for age, gender, education levels, height, muscle strength, and numbers of comorbidity. Adjusted model 2: Controlling for age, gender, education levels, height, muscle strength, numbers of comorbidity, and MMSE.

Furthermore, we modeled the above parameters using stepwise logistic regression (**Supplementary Table 3**). The AUCs showed that ToA and HtA had moderate accuracy for distinguishing AD from patients with PSD (AUC $>$ 0.700, $p \leq 0.01$, **Figure 5**).

DISCUSSION

In the present study, we focused on exploring unique gait markers with high accuracy to distinguish individuals with PSD from those without dementia who had suffered an ischemic stroke

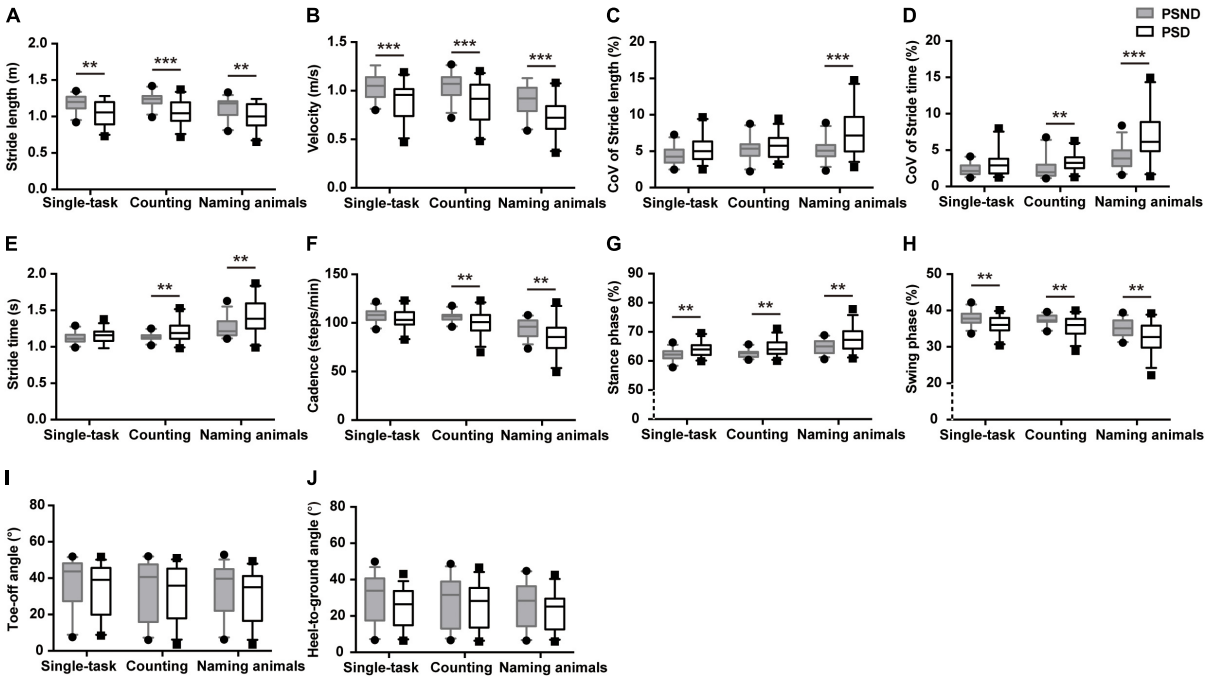


FIGURE 2 | The comparisons of spatiotemporal gait characteristics between individuals with post-stroke non-dementia (PSND) and individuals with post-stroke dementia (PSD) in three gait paradigms. **(A,B)** The pace domain of stride length and velocity. **(C,D)** The variability domain of coefficient of variation (CoV) of stride length (CoV_{SL}) and stride time (CoV_{ST}). **(E–H)** The rhythm domain of stride time, cadence, and percentages of the stance and swing phases. **(I,J)** The ToA and HtA of PSD and PSND individuals. $^{**}p \leq 0.01$, $^{***}p \leq 0.001$.

3 months earlier and compared the gait characteristics of patients with PSD with those of patients with AD to understand the unique signatures of gait that reflect the different pathogeneses and pathologies. First, both individuals with PSD and AD showed an impaired pace domain as compared to HCs, while the PSND group did not differ from HCs. Moreover, patients with PSD walked with an impaired rhythm, while patients with AD demonstrated worse postural control during walking than HCs. One main finding of our study is that individuals with PSD had significantly shorter stride length, slower walking speed, and spent a longer percentage of time in the stance phase than individuals with PSND during the single and dual-task gait tests. Increased CoV_{SL} and CoV_{ST} with longer stride time and worse cadence were found only in individuals with PSD during the naming animals-task gait paradigm. On the other hand, significant differences in ToA and HtA during the naming animals-task gait test might allow distinction of individuals with PSD and AD.

Human gait is typically divided into five domains (Lord et al., 2013), four of which were included in our study. Mobility decline with slowing gait is a continuum that co-exists with or even precedes the decline in cognition, which is pervasive and under-recognized in the majority of cognition-motor studies. A mildly reduced pace in individuals with MCI was detected when using dual-task gait paradigms, even though this finding was controversial in a single task-gait paradigm (Cullen et al., 2019; Latorre Román et al., 2020), indicating that mild cognitive decline might influence gait constitution, particularly in condition of more severe cognitive complaints. In addition, several studies have shown that gait slowing occurred in the early stage of cognitive decline and might be a predictor of the risk of progressing to dementia (Montero-Odasso et al., 2017), demonstrating that gait abnormality occurred before a diagnosis of moderate cognitive impairment. Reduced stride length and walking speed have been reported in subjects with AD, particularly those with moderate to severe AD as compared to aged-matched HCs (Mc Ardle et al., 2017). Gait rhythm was generally impaired, as outlined by a few cross-sectional studies, while findings of increasing variability were inconsistent (Boripuntakul et al., 2014; Mc Ardle et al., 2017; Valkanova and Ebmeier, 2017; Pieruccini-Faria et al., 2021). Our results indicated a longer stride time, with higher CoV_{ST} and CoV_{SL}, and disturbed pace in individuals with AD performing single or dual-task gait tests, as compared to subjects in the HCs group, which supported the previous findings. The ToA and HtA were smaller in patients with AD than in HCs. PSD, which involves a definite stroke event and subsequent cognitive impairment, is a subtype of VaD but is not equal to VaD. Previously, treatments have focused on the recovery of motor, sensory, visual, or articulatory functions. However, about half of patients with stroke suffer from amnesia and decreased executive capacity, which has not received much attention previously. Cognitive degeneration always indicates a poor prognosis, and it predicts the risk of relapse of stroke (van der Flier et al., 2018). However, if this goes undetected, the best period for therapy may be missed. The cognitive

TABLE 4 | Logistic regression of primary gait parameters to identify individuals with PSD from individuals with PSCN.

	Unadjusted model			Adjusted model 1			Adjusted model 2		
	OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*
Single-task									
Stride lengths	0.001	0.0–0.08	**0.002**	0.0	0.0–0.06	**0.003**	0.031	0.0–17.81	0.283
Velocity	0.003	0.0–0.12	**0.002**	0.001	0.0–0.07	**0.002**	0.003	0.0–1.77	0.074
Stance phase	1.46	1.11–1.92	**0.006**	1.95	1.28–2.96	**0.002**	2.33	1.11–4.90	0.026
Swing phase	0.683	0.52–0.90	**0.006**	0.512	0.34–0.80	**0.002**	0.430	0.21–0.90	0.026
Counting									
Stride lengths	0.0	0.0–0.02	**0.001**	0.0	0.0–0.003	**0.001**	0.0	0.0–1.76	0.064
Velocity	0.003	0.0–0.11	**0.001**	0.0	0.0–0.03	**0.001**	0.0	0.0–0.70	0.040
CoV_{ST}	1.62	1.06–2.48	0.027	1.87	1.10–3.19	0.022	3.40	1.12–10.30	0.031
Stride time	1.01	1.00–1.01	0.017	1.01	1.00–1.02	0.021	1.02	0.999–1.04	0.069
Cadence	0.92	0.86–0.99	0.018	0.900	0.83–0.97	**0.009**	0.810	0.66–1.00	0.051
Stance phase	1.57	1.15–2.13	**0.004**	1.91	1.28–2.85	**0.002**	2.44	0.88–6.71	0.085
Swing phase	0.643	0.47–0.87	**0.005**	0.530	0.36–0.79	**0.002**	0.429	0.16–1.14	0.429
Naming animals									
Stride lengths	0.006	0.0–0.30	0.011	0.004	0.0–0.36	0.016	0.003	0.0–10.32	0.164
Velocity	0.003	0.0–0.13	**0.002**	0.002	0.0–0.10	**0.002**	0.0	0.0–0.44	0.032
CoV_{SL}	1.62	1.18–2.21	**0.003**	1.67	1.20–2.34	**0.003**	1.60	0.97–2.64	0.068
CoV_{ST}	1.73	1.24–2.41	**0.001**	1.94	1.27–2.95	**0.002**	4.62	1.34–15.98	0.016
Stride time	1.00	1.001–1.008	**0.007**	1.01	1.001–1.008	**0.008**	1.01	1.001–1.02	0.031
Cadence	0.948	0.91–0.99	0.016	0.947	0.91–0.99	0.019	0.89	0.80–0.99	0.030
Stance phase	1.28	1.07–1.54	**0.008**	1.30	1.07–1.58	**0.010**	1.53	0.99–2.36	0.057
Swing phase	0.781	0.65–0.93	**0.008**	0.771	0.63–0.94	**0.010**	0.662	0.43–1.02	0.058

0.0 means the data is greater than but infinitely close to 0. Bold values highlight the significant differences between the two groups. Adjusted model 1: Controlling for age, gender, education levels, height, muscle strength, and numbers of comorbidity. Adjusted model 2: Controlling for age, gender, education levels, height, muscle strength, numbers of comorbidity, and MMSE.

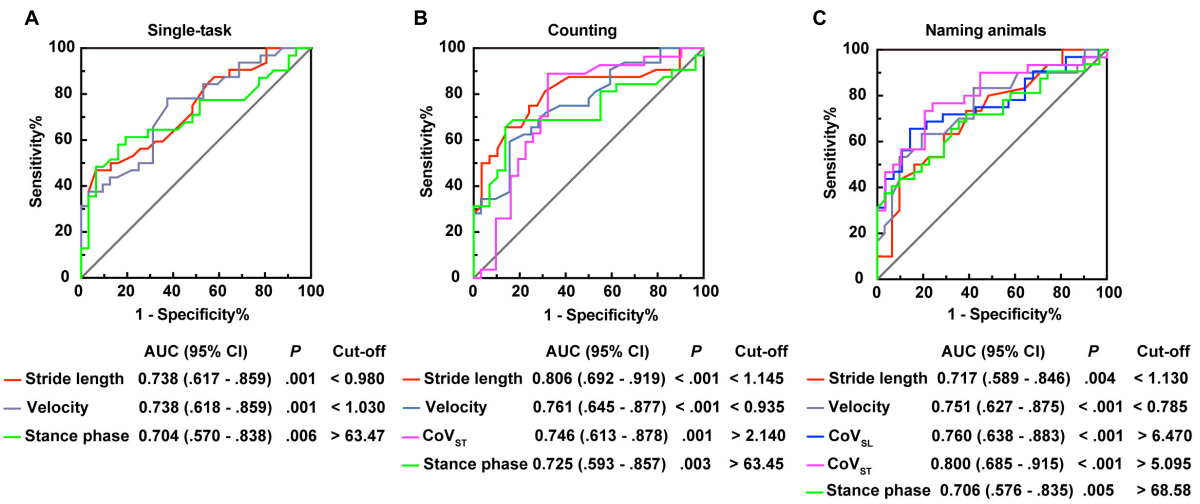


FIGURE 3 | Receiver operating characteristic curve analysis (ROC) plots for gait characteristics identifying the individuals with post-stroke dementia (PSD) from individuals with post-stroke non-dementia (PSND). **(A)** Stride length, velocity, and percentages of the stance phase for identifying patients with PSD in the single-task gait test. **(B)** Stride length, velocity, CoV_{ST}, and percentages of the stance phase for distinguishing patients with PSD in the counting backward-task gait test. **(C)** Stride length, velocity, CoV_{SL}, CoV_{ST}, and percentages of the stance phase for distinguishing patients with PSD in the naming animals-task gait test. AUC, area under the curve; CI, confidence interval; cut-off, cut-off point; CoV, coefficient of variation.

symptoms of PSD, based on one or more infarction lesions in specific brain regions, could sometimes be reversible if diagnosed timely and well-targeted treatment is started early. Thus, the timely distinction of PSD is of great significance. Due to the limited power and accuracy of PSD diagnosis at present, the gait signature, which combines evaluation of cognitive decline

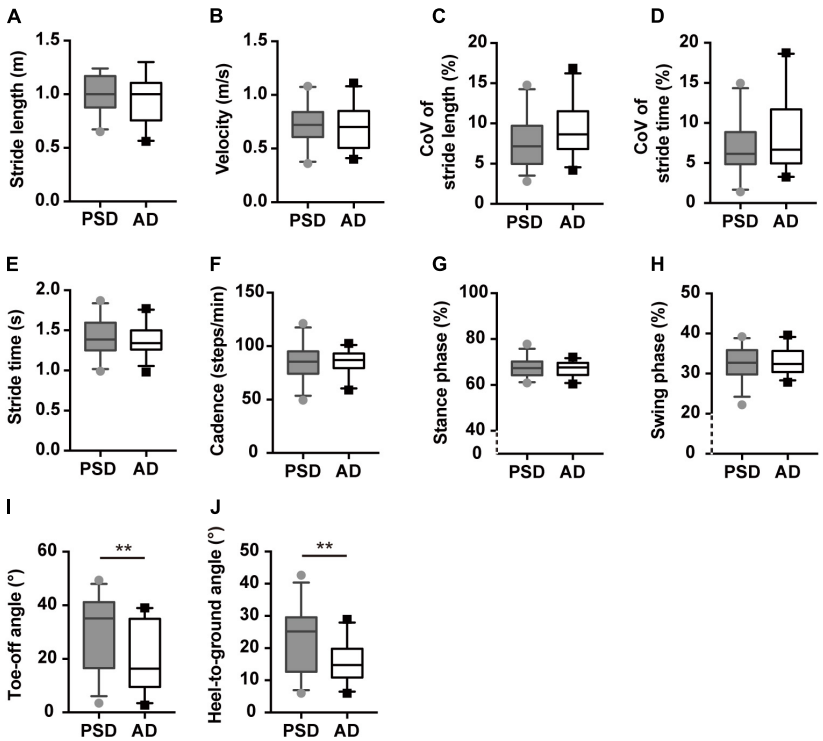


FIGURE 4 | Gait characteristics comparisons of patients with PSD and patients with AD showed in the naming animals-task gait test. **(A,B)** The pace domain of stride length and walking speed. **(C,D)** The variability domain of CoV_{SL} and CoV_{ST}. **(E–H)** The rhythm domain of stride time, cadences, percentages of stance, and swing phases. **(I,J)** The ToA and HtA of patients with PSD and patients with AD showed in the naming animals-task gait test. $^{**}p \leq 0.01$. PSD, post-stroke dementia, AD, Alzheimer's disease, CoV, coefficient of variance.

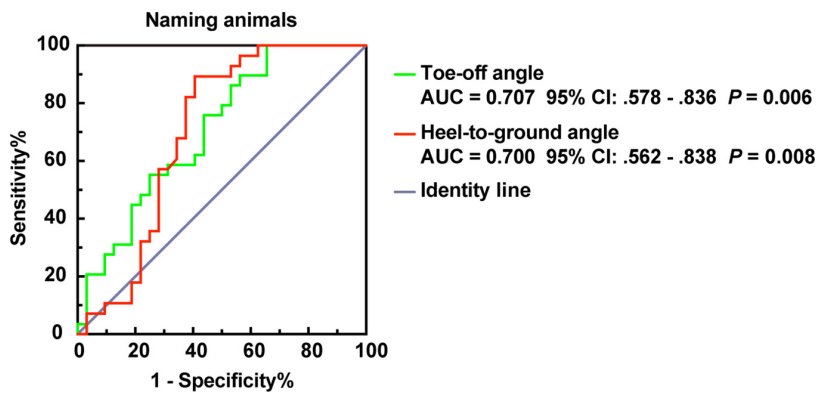


FIGURE 5 | ROC plot for ToA and HtA distinguish PSD and AD of dementia subtypes in the naming animals-task gait test. AUC, area under the curve; CI, confidence interval.

and cognition-motor interaction, may facilitate the distinction of PSD from PSND.

Recently, studies have reported that cognitive decrease influences the movement of the post-stroke population. Two cross-sectional studies have shown that disturbed gait and balance were associated with impaired cognition, particularly executive function deficits (Einstad et al., 2021) in post-stroke patients, while a lesion in the right hemisphere might lead to gait complaints (Ursin et al., 2019). Another prospective study found that gait performance was related to executive function when recall over 1 year after mild or moderate acute supratentorial ischemic stroke (Sagnier et al., 2017). A study by Assayag et al. suggested that gait and balance were predictors of cognitive status within 2 years post-stroke (Ben Assayag et al., 2015). However, little had been reported on the gait characteristics of PSD as compared to PSND or individuals with non-stroke cognitive damage, even though the gait characteristics of patients with stroke with severe motor system damage have been reported

(Dai et al., 2021). To the best of our knowledge, the gait-based discrimination between individuals with PSD and PSND has not been reported previously. Subjects with stroke enrolled in our study included those with left, right, or bilateral cerebral infarctions, and there were no differences in the lesion side distribution, muscle strength, NIHSS, or mRS scores between the PSD and PSND groups. The subjects in the PSD group demonstrated worse gait, with impaired cognition than those in the PSND group when controlling for baseline demographics. However, after adjusting for MMSE scores, the differences were not significant, suggesting that global cognition plays a crucial role in the gait performance of patients with PSD. Additionally, some gait parameters could identify patients with PSD, as evidenced by the moderate AUCs. Previous studies have revealed that patients with VaD walked with a slower velocity and shorter stride length than patients with AD (Tanaka et al., 1995; Sverdrup et al., 2021). However, we did not find a difference in the pace domain between the PSD and AD groups, even though PSD is a subtype of VaD. Additionally, we found that the smaller ToA or HtA might sensitively identify individuals with AD from individuals with PSD during the naming animals-task gait test. This has not been reported to date, to our knowledge. Whether these two parameters can define other phenotypes of dementia requires further study.

The dual-task gait test is the most popular method to investigate gait in individuals challenged by cognitive complaints and it sensitively detects gait perturbation, because it increases the gap between dementia patients and healthy populations (Al-Yahya et al., 2011; Amboni et al., 2013; Matsuura et al., 2019). Walking while performing a cognitive task will divert more attention, executive function, memory, or even visual and aural resources to complete cognitive tasks as a priority. Thus, during dual-task gait tests, people consciously walk using a more cautious gait, to prevent accidents. Under the naming animals-task condition, participants must distribute their visual capacity and attentional and executive functions simultaneously to the extra task along with walking while the counting-backward task is performed without accessory visual-spatial skills but requires memory. In line with previous evidence, overall gait performance in our study is worse as evidenced by increased CoV_{SL}, CoV_{ST}, and stride time but decreased cadence of patients with PSD than those of individuals with PSND only in the dual-task gait tests. In addition, the naming animals-task gait test might be a feasible paradigm for exploring the different gait characteristics of individuals with AD and PSD as patients with AD usually have worse visual-spatial capacity. Even so, more rational dual-task gait tests should be attempted, and uniform standard dual-task paradigms should be defined to enhance detection of occult cognitive impairment by testing cognition–motor interaction.

The mechanisms underlying discrete movement weakening in our study could not be simply explained by aging or motor system damage. A well-balanced dynamic gait is a complex achievement achieved not only by the motor system, such as muscle strength, but also by the cognition of individuals, controlled by widespread brain regions that process sensory, attention, executive, visual, and even memory information (Mc Ardle et al., 2017; Allali et al., 2019). An increasing number of studies have suggested that cognition shares some neural structures and pathology with those by which gait is controlled (Mc Ardle et al., 2017). Thus, higher-order region deterioration leading to cognitive decrease might also result in subtle changes in discrete gait characteristics.

To provide insight into the structural imaging-gait correlations, the relationships between functional structure changes and gait performance have been well-studied. Gait velocity provides an overall view of brain function and connection. Impaired pace, including low speed and short stride length, has been associated with decreased gray matter volume in the cortex (Callisaya et al., 2013), basal ganglia, and caudate nucleus (Dumurgier et al., 2012). White matter hyperintensity is also strongly associated with poor attention and executive processing, and these negative changes in the brain also negatively affect gait pace and variability (Wilson et al., 2019), but no association was found between the reduction in white matter volume and pace disturbance (Ezzati et al., 2015).

Different types of dementia show damage in overlapping brain regions and thus these individuals show roughly similar gait pace impairment as compared to age-matched healthy older individuals. The lower velocity and shorter stride length of subjects with AD and PSD in our study support this view. Individuals with VaD mostly show deficits in basal forebrain cholinergic signaling, i.e., a damaged higher-order region conventionally associated with vasculopathy and amyloid deposition in patients with AD (Kalaria, 2002). On the other hand, different subtypes of dementia may have their own typical pathology in specific brain regions, which can result in unique gait characteristics. Frontal and entorhinal cortex atrophy are typical forms of AD. These areas are deemed to process attention, executive function, and control pace (Wilson et al., 2019). On the other hand, posterior cortical atrophy correlates with the initiation of visual-spatial dysfunction (Spasov et al., 2019). This region mainly processes pace, postural control, and cadence in the rhythm domain (Wilson et al., 2019). Hippocampal atrophy, which induces memory decline, may influence rhythm, variability, and postural control (Zimmerman et al., 2009; Beauchet et al., 2019). However, the infarction lesions of patients with stroke in our study were primarily located from the basal ganglia to the periventricular regions, thalamus, pons, and frontal and parietal lobes. The prefrontal cortex-basal ganglia circuit is responsible for gait velocity and step width, along with executive function, while the limbic regions and thalamus may process stride length, width, and cadence, as indicated in previous studies (Callisaya et al., 2013; Takakusaki, 2013; Wilson et al., 2019).

Compared to HCs, we found that the AD and PSD groups showed some unique parameters, providing evidence that differences in deficits in the respective brain regions could induce diverging cognitive symptoms and disease-related gait performance. Based on the comparison of the AD with the PSD group, no difference in the familiar gait parameters was observed, even with dual-task gait tests, except for the ToA and HtA under the naming animals-task condition. Studies on ToA or HtA measured at the moment of initiation or end of the swing phase of dementia patients are scarce, even though largely reported in the patients with PD. We consider that these two parameters belong to the postural control domain because they

reflect ankle-related muscle strength, lower limb joint excursion, and foot clearance, which are related to postural instability in patients with PD (Killeen et al., 2017). The ToA and HtA were smaller in the AD group than in the PSD group, indicating that walking while conducting a naming animals-task placed a greater attention demand for joint flexion and limb muscle strength on patients with AD. Muscle strength, flexion, and extension involve walking automaticity produced by the spinal cord, brain stem, cerebellum, and the afferent pathway from the cerebral motor cortex to the brain stem and spinal cord (Clark, 2015). During the naming animals-task, higher central region processing of visual afferent information competes for attentional resources on locomotion, resulting in degraded feedback of automated motion (Clark, 2015). The visual capacity deficit, induced by occipital lobe atrophy, and local neural degeneration exerts a cognitive load on patients with AD during the naming animals-task gait test, while severe impairment of the prefrontal cortex leads to worse executive function. Participants might consciously adjust their stride angle to control postural stability once they become aware that the locomotor system was being challenged. After controlling for demographic covariates and the MMSE score, the statistically significant differences in the ToA and HtA between patients with PSD and patients with AD persisted. The differentiated circuit modulating locomotor and automated motion, or specific cognitive domains, other than global cognition, or peripheral function should be further analyzed to explain this phenomenon. These results suggest that ToA and HtA might be key motion parameters for distinguishing AD from PSD using a standard animals naming-task gait test, with modest accuracy (an intermediate AUC). The stride angle has not been routinely examined in dementia subtypes, and research focusing on the relationship between brain structure and stride angle is sparse. Thus, further studies are needed to investigate these issues.

Amyloid deposition and tau hyperphosphorylation are classical theories for the etiology and pathogenesis of AD (Raz et al., 2016). Abnormal production of amyloid and tau causes toxicity to neurons, negatively changes neuronal activity and synaptic plasticity, activates glial cells and neuroinflammation, induces neuronal death, damages neural circuits, and causes cerebrovascular dysfunction (Raz et al., 2016; Charidimou et al., 2017). In individuals with PSD, a specific brain region suffered ischemia, subsequent neuron death, and inflammation around the lesion. Some studies have indicated that deposition of amyloid and tau also occurred in patients with PSD, and the two subtypes of dementia in older individuals always involve some cerebrovascular malfunction, even though VaD *per se* includes a wider range of vesicular pathological changes (Kalaria, 2002; Emrani et al., 2020). Hence, there is clearly a neuropathological overlap of AD and PSD, which could result in difficulty in identifying subtle differences in some common gait parameters. Additionally, pathological changes in the central nervous system affect the cholinergic system, which plays a critical role in motion and cognition. The basal ganglia afferents to the cerebral cortex are mostly cholinergic neurons, which also modulate hippocampal activity and the frontoparietal networks (Tisch et al., 2004). It has been suggested that AD and PSD both involve acetylcholine signaling disruption and that the cholinergic deficits in AD are related to motor disturbances (Emrani et al., 2020). Previous research has shed light on the fact that acetylcholine esterase inhibitors could decrease the variability and fall incidence of people with mild AD (Montero-Odasso et al., 2009). Whether the cholinergic circuit disturbance might affect the ToA or HtA has not been reported. Further studies should focus on the role of the cholinergic system in stride angles.

On the basis of the lack of differences in gait parameters between individuals with PSND and HCs, we compared the gait characteristics of individuals with PSD and individuals with PSND. We found that individuals with PSD showed deficits in the pace domain, demonstrated as markedly shorter stride length and slower velocity, with a disturbed stance/swing phase ratio, despite a lack of difference in the baseline NIHSS, mRS, muscle strength, and infarction side between the PSD and PSND groups. The difference was robust after controlling for major demographic covariates, numbers of comorbidity, and muscle strength, but did not persist after further adjusting for the MMSE score. This indicated that global cognition was the overriding factor accounting for the gait disturbance in individuals with PSD, whose muscle strength did not differ from that of PSND individuals. The counting-task gait test involves people walking while performing serial subtraction of 1, which requires relatively high numeracy, memory, and attention capacity. The stride length and velocity in the counting-task gait better distinguished individuals with PSD from individuals with PSND, as the AUC increased from 0.738 to 0.806 and from 0.738 to 0.761, respectively, for the single-task gait test. Because velocity reflects global brain function, and stride length and velocity are both highly related to attention and executive function, counting backward might be a means for sensitively detecting attention and memory-associated gait disturbances. As the degree of cognitive loading increased, the variability of gait could discriminate individuals with PSD in the naming animals-task, as demonstrated by an AUC of 0.760 for CoV_{SL} and an AUC of 0.800 for CoV_{ST}. The cognitive challenge of the naming animals-task is increased by the additional requirement for visual-spatial skill. However, due to the heterogeneity of the ischemic lesion in the stroke patients in our study, the regional association between specific brain structures and gait characteristics is difficult to depict. Further studies are needed to classify the subtypes of stroke by lesion location and to research the respective gait signature and the interaction of the central structure or function with gait parameters.

Some limitations of our study need to be addressed. First, due to the limited recruitment of patients with AD and patients with PSD, we did not further stratify the patients by the severity of cognitive impairment. Thus, in our study, this population comprised those with mild to moderate dementia. Further studies should focus on more details of gait characteristics in individuals with different levels of cognitive impairment. In addition, the pathological heterogeneity of patients with PSD and patients with AD requires a more rigorous stratification of the underlying

central pathology to study the relationship between specific histomorphology and gait domains further.

CONCLUSION

In this study, we observed gait impairment in patients with PSD and patients with AD, as compared with matched normal aging controls. Some gait parameters, particularly, the stride length in the counting backward-task and the CoV_{ST} in the naming animals-task gait paradigm could allow the distinction of individuals with PSD from 3-month post-stroke patients without dementia. A smaller ToA and HtA might be characteristic gait features distinguishing subjects with AD from subjects with PSD. Overall, our findings suggest that particular gait characteristics could be non-invasive biomarkers facilitating early diagnosis of individuals with PSD, and could support the use of gait for identification of dementia subtypes, to promote appropriate and early intervention.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Ethical Committee of Sir Run Run Shaw Hospital, Zhejiang University School of Medicine. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LN and HC designed the study. WL, YiS, and XX performed the study and acquired the data. LN, MC, and XH executed the statistical analysis and made the corresponding tables and figures. YuS and ST interpreted the raw data. LN wrote the manuscript. HC and XYH reviewed and modified the manuscript. All authors contributed to the article and agreed to be accountable for the content of the work.

FUNDING

This work was supported by the Natural Science Foundation of Zhejiang Province of China (Nos. LY19H090026 and LY19H090027), Medical and Health Research Project of Zhejiang Province of China (No. 2018RC045), and National Natural Science Foundation of China (Nos. 81400926 and 81371258).

ACKNOWLEDGMENTS

The authors thank all patients and healthy participants for participating in this study. The authors also thank Editage (www.editage.cn) for English language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2021.766884/full#supplementary-material

REFERENCES

Allali, G., Montembeault, M., Brambati, S. M., Bherer, L., Blumen, H. M., Launay, C. P., et al. (2019). Brain Structure Covariance Associated With Gait Control in Aging. *J. Gerontol. A Biol. Sci. Med. Sci.* 74, 705–713. doi: 10.1093/gerona/gly123

Allan, L. M., Ballard, C. G., Burn, D. J., and Kenny, R. A. (2005). Prevalence and severity of gait disorders in Alzheimer's and non-Alzheimer's dementias. *J. Am. Geriatr. Soc.* 53, 1681–1687. doi: 10.1111/j.1532-5415.2005.53552.x

Al-Yahya, E., Dawes, H., Smith, L., Dennis, A., Howells, K., and Cockburn, J. (2011). Cognitive motor interference while walking: a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* 35, 715–728. doi: 10.1016/j.neubiorev.2010.08.008

Amboni, M., Barone, P., and Hausdorff, J. M. (2013). Cognitive contributions to gait and falls: evidence and implications. *Mov. Dis.* 28, 1520–1533. doi: 10.1002/mds.25674

Bayot, M., Dujardin, K., Tard, C., Defebvre, L., Bonnet, C. T., Allart, E., et al. (2018). The interaction between cognition and motor control: A theoretical framework for dual-task interference effects on posture, gait initiation, gait and turning. *Neurophysiol. Clin.* 48, 361–375. doi: 10.1016/j.neucli.2018.10.003

Beauchet, O., Annweiler, C., Callisaya, M. L., De Cock, A. M., Helbostad, J. L., Kressig, R. W., et al. (2016). Poor Gait Performance and Prediction of Dementia: Results From a Meta-Analysis. *J. Am. Med. Dir. Assoc.* 17, 482–490. doi: 10.1016/j.jamda.2015.12.092

Beauchet, O., Launay, C. P., Sekhon, H., Montembeault, M., and Allali, G. (2019). Association of hippocampal volume with gait variability in pre-dementia and dementia stages of Alzheimer disease: Results from a cross-sectional study. *Exp. Gerontol.* 115, 55–61. doi: 10.1016/j.exger.2018.11.010

Ben Assayag, E., Shenhar-Tsarfaty, S., Korczyn, A. D., Kliper, E., Hallevi, H., Shopin, L., et al. (2015). Gait measures as predictors of poststroke cognitive function: evidence from the TABASCO study. *Stroke* 46, 1077–1083. doi: 10.1161/strokeaha.114.007346

Boripuntakul, S., Lord, S. R., Brodie, M. A., Smith, S. T., Methapatara, P., Wongpakaran, N., et al. (2014). Spatial variability during gait initiation while dual tasking is increased in individuals with mild cognitive impairment. *J. Nutr. Health Aging* 18, 307–312. doi: 10.1007/s12603-013-0390-3

Callisaya, M. L., Beare, R., Phan, T. G., Blizzard, L., Thrift, A. G., Chen, J., et al. (2013). Brain structural change and gait decline: a longitudinal population-based study. *J. Am. Geriatr. Soc.* 61, 1074–1079. doi: 10.1111/jgs.12331

Charidimou, A., Boulouis, G., Gurol, M. E., Ayata, C., Bacskai, B. J., Frosch, M. P., et al. (2017). Emerging concepts in sporadic cerebral amyloid angiopathy. *Brain* 140, 1829–1850. doi: 10.1093/brain/awx047

Cheng, Q., Wu, M., Wu, Y., Hu, Y., Kwapong, W. R., Shi, X., et al. (2020). Weaker Braking Force, A New Marker of Worse Gait Stability in Alzheimer Disease. *Front. Aging Neurosci.* 12:554168. doi: 10.3389/fnagi.2020.554168

Clark, D. J. (2015). Automaticity of walking: functional significance, mechanisms, measurement and rehabilitation strategies. *Front. Hum. Neurosci.* 9:246. doi: 10.3389/fnhum.2015.00246

Cognitive Impairment Committee Nb, Chinese Medical Doctor Association. (2019). A guideline for the diagnosis and treatment of Chinese vascular cognitive impairment in 2019. *Chin. Med. J.* 99, 2737–2744.

Cohen, J. A., Verghese, J., and Zwerling, J. L. (2016). Cognition and gait in older people. *Maturitas* 93, 73–77. doi: 10.1016/j.maturitas.2016.05.005

Cullen, S., Borrie, M., Carroll, S., Sarquis-Adamson, Y., Pieruccini-Faria, F., McKay, S., et al. (2019). Are Cognitive Subtypes Associated with Dual-Task Gait Performance in a Clinical Setting? *J. Alzheimers Dis.* 71, S57–S64. doi: 10.3233/jad-181196

Dai, S., Piscicelli, C., Clarac, E., Baciu, M., Hommel, M., and Pérennou, D. (2021). Balance, Lateropulsion, and Gait Disorders in Subacute Stroke. *Neurology* 96, e2147–e2159. doi: 10.1212/wnl.0000000000011152

Dumurgier, J., Crivello, F., Mazoyer, B., Ahmed, I., Tavernier, B., Grabli, D., et al. (2012). MRI atrophy of the caudate nucleus and slower walking speed in the elderly. *Neuroimage* 60, 871–878. doi: 10.1016/j.neuroimage.2012.01.102

Einstad, M. S., Saltvedt, I., Lydersen, S., Ursin, M. H., Munthe-Kaas, R., Ihle-Hansen, H., et al. (2021). Associations between post-stroke motor and cognitive function: a cross-sectional study. *BMC Geriatr.* 21:103. doi: 10.1186/s12877-021-02055-7

Emrani, S., Lamar, M., Price, C. C., Wasserman, V., Matusz, E., Au, R., et al. (2020). Alzheimer's/Vascular Spectrum Dementia: Classification in Addition to Diagnosis. *J. Alzheimers Dis.* 73, 63–71. doi: 10.3233/jad-190654

Ezzati, A., Katz, M. J., Lipton, M. L., Lipton, R. B., and Verghese, J. (2015). The association of brain structure with gait velocity in older adults: a quantitative volumetric analysis of brain MRI. *Neuroradiology* 57, 851–861. doi: 10.1007/s00234-015-1536-2

Ghoraani, B., Boettcher, L. N., Hssayeni, M. D., Rosenfeld, A., Tolea, M. I., and Galvin, J. E. (2021). Detection of Mild Cognitive Impairment and Alzheimer's Disease using Dual-task Gait Assessments and Machine Learning. *Biomed. Signal Process Control* 64:102249. doi: 10.1016/j.bspc.2020.102249

Gillain, S., Warzee, E., Lekeu, F., Wojtasik, V., Maquet, D., Croisier, J. L., et al. (2009). The value of instrumental gait analysis in elderly healthy, MCI or Alzheimer's disease subjects and a comparison with other clinical tests used in single and dual-task conditions. *Ann. Phys. Rehabil. Med.* 52, 453–474. doi: 10.1016/j.rehab.2008.10.004

Ismail, Z., Black, S. E., Camicioli, R., Chertkow, H., Herrmann, N., Laforce, R. Jr., et al. (2020). Recommendations of the 5th Canadian Consensus Conference on the diagnosis and treatment of dementia. *Alzheimers Dement.* 16, 1182–1195. doi: 10.1002/alz.12105

Kalaria, R. (2002). Similarities between Alzheimer's disease and vascular dementia. *J. Neurol. Sci.* 203-204, 29–34. doi: 10.1016/s0022-510x(02)00256-3

Khachaturian, Z. S. (2011). Revised criteria for diagnosis of Alzheimer's disease: National Institute on Aging-Alzheimer's Association diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 253–256. doi: 10.1016/j.jalz.2011.04.003

Killeen, T., Easthope, C. S., Demkó, L., Filli, L., Lõrincz, L., Linnebank, M., et al. (2017). Minimum toe clearance: probing the neural control of locomotion. *Sci. Rep.* 7:1922. doi: 10.1038/s41598-017-02189-y

Latorre Román, P., Muñoz Jiménez, M., Salas Sánchez, J., Consuegra González, P., Moreno Del Castillo, R., Herrador Sánchez, J. A., et al. (2020). Complex Gait Is Related to Cognitive Functioning in Older People: A Cross-Sectional Study Providing an Innovative Test. *Gerontology* 66, 401–408. doi: 10.1159/000508245

Li, H., Jia, J., and Yang, Z. (2016). Mini-Mental State Examination in Elderly Chinese: A Population-Based Normative Study. *J. Alzheimers Dis.* 53, 487–496. doi: 10.3233/jad-160119

Lord, S., Galna, B., Verghese, J., Coleman, S., Burn, D., and Rochester, L. (2013). Independent domains of gait in older adults and associated motor and nonmotor attributes: validation of a factor analysis approach. *J. Gerontol. A Biol. Sci. Med. Sci.* 68, 820–827. doi: 10.1093/gerona/gls255

Lu, J., Li, D., Li, F., Zhou, A., Wang, F., Zuo, X., et al. (2011). Montreal cognitive assessment in detecting cognitive impairment in Chinese elderly individuals: a population-based study. *J. Geriatr. Psychiatry Neurol.* 24, 184–190. doi: 10.1177/0891988711422528

Maquet, D., Lekeu, F., Warzee, E., Gillain, S., Wojtasik, V., Salmon, E., et al. (2010). Gait analysis in elderly adult patients with mild cognitive impairment and patients with mild Alzheimer's disease: simple versus dual task: a preliminary report. *Clin. Physiol. Funct. Imaging* 30, 51–56. doi: 10.1111/j.1475-097X.2009.00903.x

Matsuura, T., Sakashita, K., Grushnikov, A., Okura, F., Mitsugami, I., and Yagi, Y. (2019). Statistical Analysis of Dual-task Gait Characteristics for Cognitive Score Estimation. *Sci. Rep.* 9:19927. doi: 10.1038/s41598-019-56485-w

Mc Ardle, R., Galna, B., Donaghy, P., Thomas, A., and Rochester, L. (2019). Do Alzheimer's and Lewy body disease have discrete pathological signatures of gait? *Alzheimers Dement.* 15, 1367–1377. doi: 10.1016/j.jalz.2019.06.4953

Mc Ardle, R., Morris, R., Wilson, J., Galna, B., Thomas, A. J., and Rochester, L. (2017). What Can Quantitative Gait Analysis Tell Us about Dementia and Its Subtypes? A Structured Review. *J. Alzheimers Dis.* 60, 1295–1312. doi: 10.3233/jad-170541

McCarten, J. R., Anderson, P., Kuskowski, M. A., McPherson, S. E., and Borson, S. (2011). Screening for cognitive impairment in an elderly veteran population: acceptability and results using different versions of the Mini-Cog. *J. Am. Geriatr. Soc.* 59, 309–313. doi: 10.1111/j.1532-5415.2010.03249.x

Montero-Odasso, M. M., Sarquis-Adamson, Y., Speechley, M., Borrie, M. J., Hachinski, V. C., Wells, J., et al. (2017). Association of Dual-Task Gait With Incident Dementia in Mild Cognitive Impairment: Results From the Gait and Brain Study. *JAMA Neurol.* 74, 857–865. doi: 10.1001/jamaneurol.2017.0643

Montero-Odasso, M., Wells, J., and Borrie, M. (2009). Can cognitive enhancers reduce the risk of falls in people with dementia? An open-label study with controls. *J. Am. Geriatr. Soc.* 57, 359–360. doi: 10.1111/j.1532-5415.2009.02085.x

Morris, R., Lord, S., Bunce, J., Burn, D., and Rochester, L. (2016). Gait and cognition: Mapping the global and discrete relationships in ageing and neurodegenerative disease. *Neurosci. Biobehav. Rev.* 64, 326–345. doi: 10.1016/j.neubiorev.2016.02.012

Noce Kirkwood, R., de Souza Moreira, B., Mingoti, S. A., Faria, B. F., Sampaio, R. F., and Alves Resende, R. (2018). The slowing down phenomenon: What is the age of major gait velocity decline? *Maturitas* 115, 31–36. doi: 10.1016/j.maturitas.2018.06.005

Pieruccini-Faria, F., Black, S. E., Masellis, M., Smith, E. E., Almeida, Q. J., Li, K. Z. H., et al. (2021). Gait variability across neurodegenerative and cognitive disorders: Results from the Canadian Consortium of Neurodegeneration in Aging (CCNA) and the Gait and Brain Study. *Alzheimers Dement.* 2021:12298. doi: 10.1002/alz.12298

Rasmussen, L. J. H., Caspi, A., Ambler, A., Broadbent, J. M., Cohen, H. J., d'Arbeloff, T., et al. (2019). Association of Neurocognitive and Physical Function With Gait Speed in Midlife. *JAMA Netw. Open* 2:e1913123. doi: 10.1001/jamanetworkopen.2019.13123

Raz, L., Knoefel, J., and Bhaskar, K. (2016). The neuropathology and cerebrovascular mechanisms of dementia. *J. Cereb. Blood Flow Metab.* 36, 172–186. doi: 10.1038/jcbfm.2015.164

Sagnier, S., Renou, P., Olindo, S., Debruxelles, S., Poli, M., Rouanet, F., et al. (2017). Gait Change Is Associated with Cognitive Outcome after an Acute Ischemic Stroke. *Front. Aging Neurosci.* 9:153. doi: 10.3389/fnagi.2017.00153

Sanders, L. M. J., Hortobágyi, T., Karssemeijer, E. G. A., Van der Zee, E. A., Scherder, E. J. A., and van Heuvelen, M. J. G. (2020). Effects of low- and high-intensity physical exercise on physical and cognitive function in older persons with dementia: a randomized controlled trial. *Alzheimers Res. Ther.* 12:28. doi: 10.1186/s13195-020-00597-3

Schott, J. M., Reiniger, L., Thom, M., Holton, J. L., Grieve, J., Brandner, S., et al. (2010). Brain biopsy in dementia: clinical indications and diagnostic approach. *Acta Neuropathol.* 120, 327–341. doi: 10.1007/s00401-010-0721-y

Skrobot, O. A., Black, S. E., Chen, C., DeCarli, C., Erkinjuntti, T., Ford, G. A., et al. (2018). Progress toward standardized diagnosis of vascular cognitive impairment: Guidelines from the Vascular Impairment of Cognition Classification Consensus Study. *Alzheimers Dement.* 14, 280–292. doi: 10.1016/j.jalz.2017.09.007

Spasov, S., Passamonti, L., Duggento, A., Liò, P., and Toschi, N. (2019). A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. *Neuroimage* 189, 276–287. doi: 10.1016/j.neuroimage.2019.01.031

Sverdrup, K., Selbæk, G., Bergh, S., Strand, B. H., Thingstad, P., Skjellegrind, H. K., et al. (2021). Physical performance across the cognitive spectrum and between dementia subtypes in a population-based sample of older adults: The HUNT study. *Arch. Gerontol. Geriatr.* 95:104400. doi: 10.1016/j.archger.2021.104400

Takakusaki, K. (2013). Neurophysiology of gait: from the spinal cord to the frontal lobe. *Mov. Dis.* 28, 1483–1491. doi: 10.1002/mds.25669

Tanaka, A., Okuzumi, H., Kobayashi, I., Murai, N., Meguro, K., and Nakamura, T. (1995). Gait disturbance of patients with vascular and Alzheimer-type

dementias. *Percept. Mot. Skills* 80(3 Pt 1), 735–738. doi: 10.2466/pms.1995.80.3.735

Tian, Q., Chastan, N., Bair, W. N., Resnick, S. M., Ferrucci, L., and Studenski, S. A. (2017). The brain map of gait variability in aging, cognitive impairment and dementia-A systematic review. *Neurosci. Biobehav. Rev.* 74(Pt A), 149–162. doi: 10.1016/j.neubiorev.2017.01.020

Tisch, S., Silberstein, P., Limousin-Dowsey, P., and Jahanshahi, M. (2004). The basal ganglia: anatomy, physiology, and pharmacology. *Psychiatr. Clin. North Am.* 27, 757–799. doi: 10.1016/j.psc.2004.06.004

Ursin, M. H., Bergland, A., Fure, B., Thommessen, B., Hagberg, G., Øksengård, A. R., et al. (2019). Gait and balance one year after stroke; relationships with lesion side, subtypes of cognitive impairment and neuroimaging findings-a longitudinal, cohort study. *Physiotherapy* 105, 254–261. doi: 10.1016/j.physio.2018.07.007

Valkanova, V., and Ebmeier, K. P. (2017). What can gait tell us about dementia? Review of epidemiological and neuropsychological evidence. *Gait Posture* 53, 215–223. doi: 10.1016/j.gaitpost.2017.01.024

van der Flier, W. M., Skoog, I., Schneider, J. A., Pantoni, L., Mok, V., Chen, C. L. H., et al. (2018). Vascular cognitive impairment. *Nat. Rev. Dis. Primers* 4:18003. doi: 10.1038/nrdp.2018.3

WHO (2018). *The Global Dementia Observatory Reference Guide version1.1*. Geneva: WHO.

Wilson, J., Allcock, L., Mc Ardle, R., Taylor, J. P., and Rochester, L. (2019). The neural correlates of discrete gait characteristics in ageing: A structured review. *Neurosci. Biobehav. Rev.* 100, 344–369. doi: 10.1016/j.neubiorev.2018.12.017

Zimmerman, M. E., Lipton, R. B., Pan, J. W., Hetherington, H. P., and Verghese, J. (2009). MRI- and MRS-derived hippocampal correlates of quantitative locomotor function in older adults. *Brain Res.* 1291, 73–81. doi: 10.1016/j.brainres.2009.07.043

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 03 December 2021
doi: 10.3389/fnagi.2021.768229

Application of Artificial Intelligence Modeling Technology Based on Fluid Biopsy to Diagnose Alzheimer's Disease

Yuan Sh[1†], Benliang Liu[2,3†], Jianhu Zhang[1†], Ying Zhou[1†], Zhiyuan Hu[1,4,5,6] and Xiuli Zhang[4*]*

[1] Fujian Provincial Key Laboratory of Brain Aging and Neurodegenerative Diseases, School of Basic Medical Sciences, Fujian Medical University, Fuzhou, China, [2] China National Center for Bioinformation, Beijing, China, [3] Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China, [4] Chinese Academy of Sciences Key Laboratory of Standardization and Measurement for Nanotechnology, Chinese Academy of Sciences Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety, Chinese Academy of Sciences Center for Excellence in Nanoscience, National Center for Nanoscience and Technology of China, Beijing, China, [5] School of Nanoscience and Technology, Sino-Danish College, University of Chinese Academy of Sciences, Beijing, China, [6] School of Chemical Engineering and Pharmacy, Wuhan Institute of Technology, Wuhan, China

OPEN ACCESS

Edited by:
Xiuqin Jia,
Capital Medical University, China

Reviewed by:
Ke Du,
China Medical University, China
Fuzhou Hua,
Second Affiliated Hospital of Nanchang University, China
Kewei Chen,
Banner Alzheimer's Institute, United States

***Correspondence:**
Xiuli Zhang
zhxiuli@gmail.com
Zhiyuan Hu
huzy@nanoctr.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

Received: *31 August 2021*
Accepted: *15 October 2021*
Published: *03 December 2021*

Citation:
Sh Y, Liu B, Zhang J, Zhou Y, Hu Z and Zhang X (2021) Application of Artificial Intelligence Modeling Technology Based on Fluid Biopsy to Diagnose Alzheimer's Disease. Front. Aging Neurosci. 13:768229. doi: 10.3389/fnagi.2021.768229

Background: There are no obvious clinical signs and symptoms in the early stages of Alzheimer's disease (AD), and most patients usually have mild cognitive impairment (MCI) before diagnosis. Therefore, early diagnosis of AD is very critical. This paper mainly discusses the blood biomarkers of AD patients and uses machine learning methods to study the changes of blood transcriptome during the development of AD and to search for potential blood biomarkers for AD.

Methods: Individualized blood mRNA expression data of 711 patients were downloaded from the GEO database, including the control group (CON) (238 patients), MCI (189 patients), and AD (284 patients). Firstly, we analyzed the subcellular localization, protein types and enrichment pathways of the differentially expressed mRNAs in each group, and established an artificial intelligence individualized diagnostic model. Furthermore, the XCell tool was used to analyze the blood mRNA expression data and obtain blood cell composition and quantitative data. Ratio characteristics were established for mRNA and XCell data. Feature engineering operations such as collinearity and importance analysis were performed on all features to obtain the best feature solicitation. Finally, four machine learning algorithms, including linear support vector machine (SVM), Adaboost, random forest and artificial neural network, were used to model the optimal feature combinations and evaluate their classification performance in the test set.

Results: Through feature engineering screening, the best feature collection was obtained. Moreover, the artificial intelligence individualized diagnosis model established

based on this method achieved a classification accuracy of 91.59% in the test set. The area under curve (AUC) of CON, MCI, and AD were 0.9746, 0.9536, and 0.9807, respectively.

Conclusion: The results of cell homeostasis analysis suggested that the homeostasis of Natural killer T cell (NKT) might be related to AD, and the homeostasis of Granulocyte macrophage progenitor (GMP) might be one of the reasons for AD.

Keywords: Alzheimer's disease, mild cognitive impairment, artificial intelligence, predictive diagnostics, blood biomarkers

INTRODUCTION

Alzheimer's disease (AD) is the most common chronic neurodegenerative disease (Burns and Iliffe, 2009). According to the World Health Organization, dementia affects 5–8 percent of people over 60 years. As of September 2020, there were about 50 million people with dementia, with 10 million new cases per year worldwide (Word Health Organization, 2020). Through establishing an individualized diagnosis model for patients with AD in its early onset, it is expected to realize early intervention for patients. At present, some studies have reported artificial intelligence models for AD diagnosis (Lunnon et al., 2013; Li et al., 2018; Way et al., 2018; Ludwig et al., 2019; Stamate et al., 2019). For example, in a European cohort study, a machine-learning approach identified 347 plasma metabolites associated with early diagnosis in AD with an area under curve (AUC) of about 0.85 (Stamate et al., 2019). In a study of circulating non-coding RNA in patients with AD, 21 disease-related features were identified using RT-qPCR, and 18 strongly correlated features were extracted using statistical learning methods to establish a machine learning model, with an AUC of about 0.86 (Herrero-Labrador et al., 2020). In an AD classifier based on texture features, the researchers modeled the high-level semantic features of MRI with an accuracy of about 85% (So et al., 2019).

However, these studies are based on the dichotomous task, ignoring the correlation degree of occurrence and development of the control (CON), mild cognitive impairment (MCI), and AD themselves, and the accuracy is not high. In this study, we incorporated blood mRNA expression data to establish two highly accurate artificial intelligence individualized diagnostic models for CON, MCI, and AD classification problems. Although a few studies analyze and/or predict these three disease states simultaneously, most of these studies are based on medical image data (Rogers et al., 2012). Furthermore, we analyzed the blood cells composition corresponding to blood mRNA profiles. We revealed some of the underlying mechanisms during the early pathogenesis of AD by analyzing the imbalance of five major groups of cells, including Epithelial, Hematopoietic stem cells (HSC), Lymphoid, Myeloid and Stroma. The overall landscape of blood cell imbalance lays a solid foundation for further mechanism research and individualized therapy.

Abbreviations: PD, Parkinson's disease; AD, Alzhaimer's disease; MCI, mild cognitive impairment; GEO, Gene Expression Omnibus; AUC, area under curve; IPA, ingenuity pathway analysis; DEGs, differentially expressed genes.

MATERIALS AND METHODS

Data Source and Preprocessing

We downloaded two sets of peripheral whole blood mRNA expression profiles from the GEO database[1], including GSE63060 and GSE63061 (Sood et al., 2015). These two sets of data were detected by the platforms Illumina HumanHT-12 V3.0 and Illumina HumanHT-12 V4.0, respectively. After deleting fuzzy samples and finally keep 329 samples and 382 samples, a total of 711 samples (**Table 1**). We have carried out standardized processing in the quantity of the data set. The method is as follows: We mark the sample as x, and the expression value of the j_th gene in the sample as. First, calculate the sum of the expression values of all genes in the i_th sample, and then calculate the j_th gene in the i_th sample Divide by the sum in turn $\left(x_{ij}\Big/\sum_{j=1}^{n} x_{ij}\right)$, and finally multiply the obtained value by 10^6. The specific calculation formula is as follows:

$$x_{ij}^{'} = \left(x_{ij}\Big/\sum_{j=1}^{n} x_{ij}\right) * 10^6.$$

Identification of Differentially Expressed Genes

The differential expression genes (DEGs) were recognized with the limma Bioconductor package (limma package v.3.24.15)

[1]https://www.ncbi.nlm.nih.gov/geo/

TABLE 1 | Data distribution diagram.

Datasets	Disease type	Sample numbers	Age [Median (Range)]	Sex (% male)
GSE63061	CON	104	73 (52–87)	40.38
	MCI	80	74 (63–90)	51.25
	AD	145	76 (58–88)	31.72
GSE63060	CON	134	74 (63–91)	39.55
	MCI	109	79 (57–100)	40.37
	AD	139	79 (59–95)	38.85

in R (Smyth, 2005). The limma package use T-statistic as a discriminant that can eliminate the irrelevant genes. Limma package was use to findmarker by each two groups such as AD vs. CON, MCI vs. CON, and AD vs. MCI. We used the FDR-method correction for multiple testing.

Feature Importance Selection

We use recursive feature elimination cross-validation to eliminate low importance. First, we choose a linear model to calculate all feature coefficients. Then we make a loop to eliminate low coefficient features until the number of features meets our requirements. This method is provided in the RFECV function, which is in the scikit-learn module in Python (Pedregosa et al., 2011; Buitinck et al., 2013). We use the default parameters of the RFECV function in the sklearn.feature_selection module, dependent variables are all genes, and independent variables are the results of the numerical transformation of CON, MCI, and AD set to 0, 1, and 2, respectively.

Machine Learning Model

The machine learning models we use include linear models and non-linear models. Among them, the linear model uses a linear support vector machine (SVM) (Cortes and Vapnik, 1995), and the non-linear model includes AdaBoost (Freund and Schapire, 1997), random forest (Liaw and Wiener, 2002), and Artificial neural networks. Among them, linear SVM, random forest, AdaBoost use the function of the scikit-learn module of Python, Artificial neural networks use PyTorch module of Python. In the above four models, we use default parameters for the first three models; for Artificial neural networks, we use a feedforward neural network composed of three hidden layers and one output layer. The number of neurons in each hidden layer is 64, 32, and 16, respectively. The number of neurons in the output layer is 3, and the neurons in the output layer represent the probability values of various samples.

XCell Analysis

XCell is a web analysis tool developed by the University of California, enriched based on gene expression data[2] to obtain the Cell-Type score data (Aran et al., 2017). This method is based on gene signature, which is used to infer 64 types of immune cells and stromal cells.

Disease Ontology Semantic and Enrichment Analysis

R (version 4.0.2) package DOSE (Yu et al., 2015) to analyze which diseases are related to the final features that we found. There are five functions in the DOSE package which we use is enriched function. Using cumulative hypergeometric model to identify which disease ontology that genes are mainly enriched in, where k is the number of genes related to the disease ontology; r is the number of all genes which are involved in all diseases that are collected in the DOSE package, and s is the number of genes which we have identified, the formula is as follows:

$$P = 1 - \sum_{i=0}^{s-1} \frac{\binom{k}{i}\binom{r-k}{k-i}}{\binom{k}{r}}$$

Ingenuity Pathway Analysis

Ingenuity Pathway Analysis (IPA) is a bioinformatics analysis method. We use IPA method to locate features and annotate functions. P-value < 0.05 was considered a statistically significant threshold. Z-value greater than 0 is defined as active, and less than 0 is defined as suppressed. The activation z-score of a hypothesis is calculated from the regulation directions and gene expression changes of the genes in the overlap of data set and hypothesis-regulated genes. It assesses whether there is a significant pattern match between predicted and observed up- and down-regulation, and also predicts the activation state of the regulator ($z > 0$: activating, $z < 0$: inhibiting). The activation z-score is given by:

$$Zscore = \frac{(N^+ + N^-)}{\sqrt{N}}$$

with $N^+(N^-)$ being the number of genes where the product of net-effect and observed direction of gene regulation is greater (less) than zero, and $N = N{+}{+}N^-$ (Krämer et al., 2014).

RESULTS

Establishment and Analysis Process of Individualized Diagnosis Model for Overall Alzheimer's Disease Patients

Based on blood mRNA expression profiles, we analyzed, screened and obtained two sets of potential blood biomarkers for early AD diagnosis and developed two different model frameworks (**Figure 1**). Individualized blood mRNA expression data of 711 patients were downloaded from the GEO database, including 238 CONs, 189 MCIs, and 284 AD patients. Firstly, we analyzed the subcellular localization, protein types and enrichment pathways of the differentially expressed mRNAs in each group, and established an artificial intelligence individualized diagnostic model. Furthermore, the XCell tool was used to analyze the blood mRNA expression data to obtain blood cell composition and quantitative data. New ratio features were established for mRNA and XCell data. Co-linearity and importance analysis of all features were carried out to obtain the optimal feature solicitation. Finally, four machine learning algorithms, including linear SVM, Adaboost, random forest and artificial neural network, were used to established models for the optimal feature set and evaluate their classification performance in test sets.

Next, we analyzed the effects of different mRNAs on the levels of different disease groups (MCI and AD) in terms of subcellular localization, coding protein type and enrichment function. First, we normalized the data and then identified 5,625 differentially expressed genes (DEGs) between

[2]https://xcell.ucsf.edu/

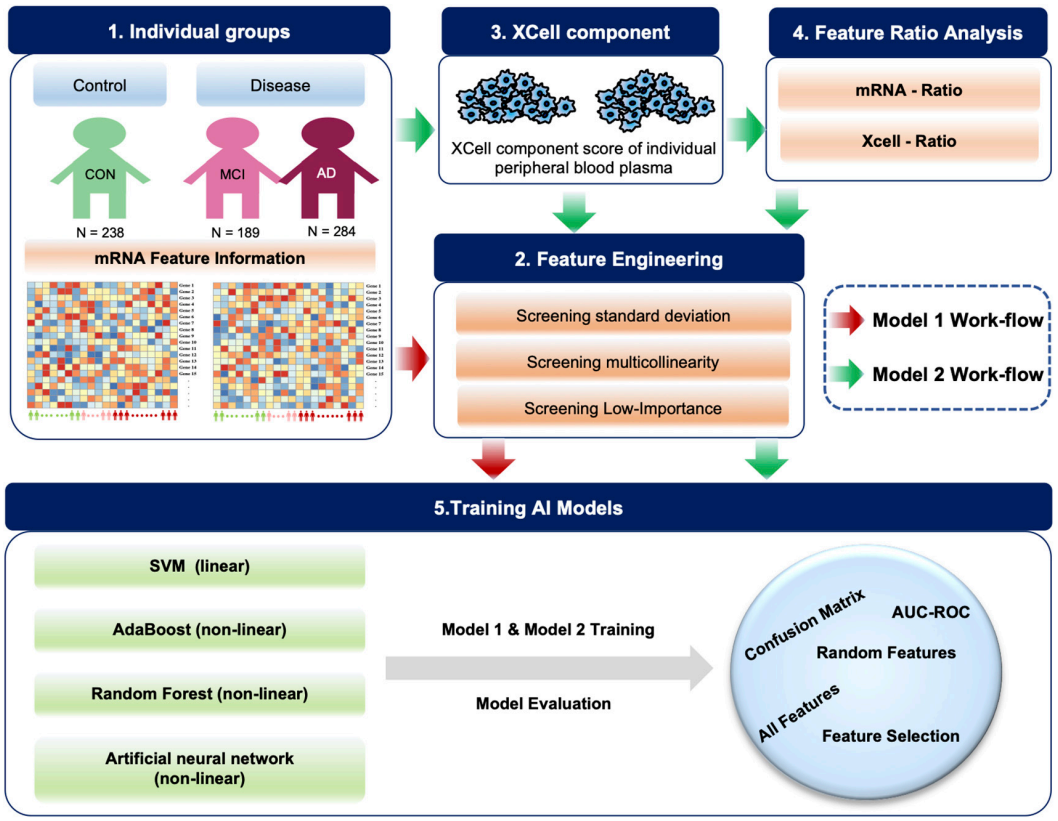


FIGURE 1 | Diagram of the workflow in this study. A total of 711 individual peripheral blood mRNA data were included for modeling and analysis, including 238 control (CON), 189 MCI and 284 AD samples. For Model 1, we performed feature engineering, training and testing. For Model 2, XCell cell composition was first decomposed from blood mRNA data, then feature ratios were calculated using XCell and mRNA features, respectively. We combined mRNA, XCell, mRNA-ratio, and XCell-ratio to perform feature engineering. Four different algorithms were used for AI modeling. The models were evaluated on test data using AUC-ROC and confusion matrix. The red arrows represent Model 1, and the green arrows represent the processes for Model 2.

the two groups (**Supplementary Figure 1**). Meanwhile, we performed cross-evaluation in different datasets. In addition, the results of the IPA show that DEGs in CON and MCI and DEGs in CON and AD differ not only in gene type but also in gene location. In the early stage of AD, the abundance of proteins located in the plasma membrane by DEGs is significantly up-regulated. In contrast, the expression of proteins located in other regions is inhibited. Specifically, the expression levels of Transmembrane receptor, G-protein coupled receptor, Phosphatase, and Kinase in the MCI group are increased. In addition, we also analyzed the same differential genes with disease-related enrichment and enriched and analyzed the up-regulated and down-regulated genes in the disease group (MCI and AD)/normal group, respectively. The upregulated genes are mainly enriched and associated with senile diseases. The down-regulated genes can significantly affect "Parkinson's disease," "Huntington's disease," "AD," and "Oxidative phosphorylation." NDUFA4, NDUFB6, ATP5F1C, CALM2, COX5B, COX4I1, and CYCS are also involved in the three major neurodegenerative diseases, including Parkinson's disease, Huntington's disease, and AD (Stelzer et al., 2016; Adav et al., 2019).

Individualized Diagnostic AI Model Based on Blood mRNA Expression Data

We used standard deviation distribution, Co-linearity analysis (**Figure 2A**) and importance analysis (**Figure 2B**) to perform feature engineering screening on the total mRNA features for screening and to obtain the optimal feature set. In general, starting from 5,625 features, we screened out the features with a standard deviation less than 3 (retaining more than 75% of the features), and then the remaining 4,219 features. After the analysis of Co-linearity (Pearson correlation coefficient between various features is calculated), the features with a correlation greater than 0.9 are filtered out, leaving 1,598 remaining features. Importance analysis results showed that the TOP 5 with the highest contribution to tri-classification modeling are STAT6, KLF6, FCER2, HLA-A, PPBP, etc. (**Figure 2B**). We quantified the three states based on disease progression and assessed the correlation between the selected features and the disease state. Our results showed that SNRPB2, LPP, C7ORF43, HCG27, and RGS14 (Top-5) were positively correlated with the development of CON, MCI, and AD. Negative correlation features included BUD31, GTF2H5, RPS23, MRPS17, and MRPL51 (Top-5) (**Figure 2C**). After importance and correlation analysis (using an

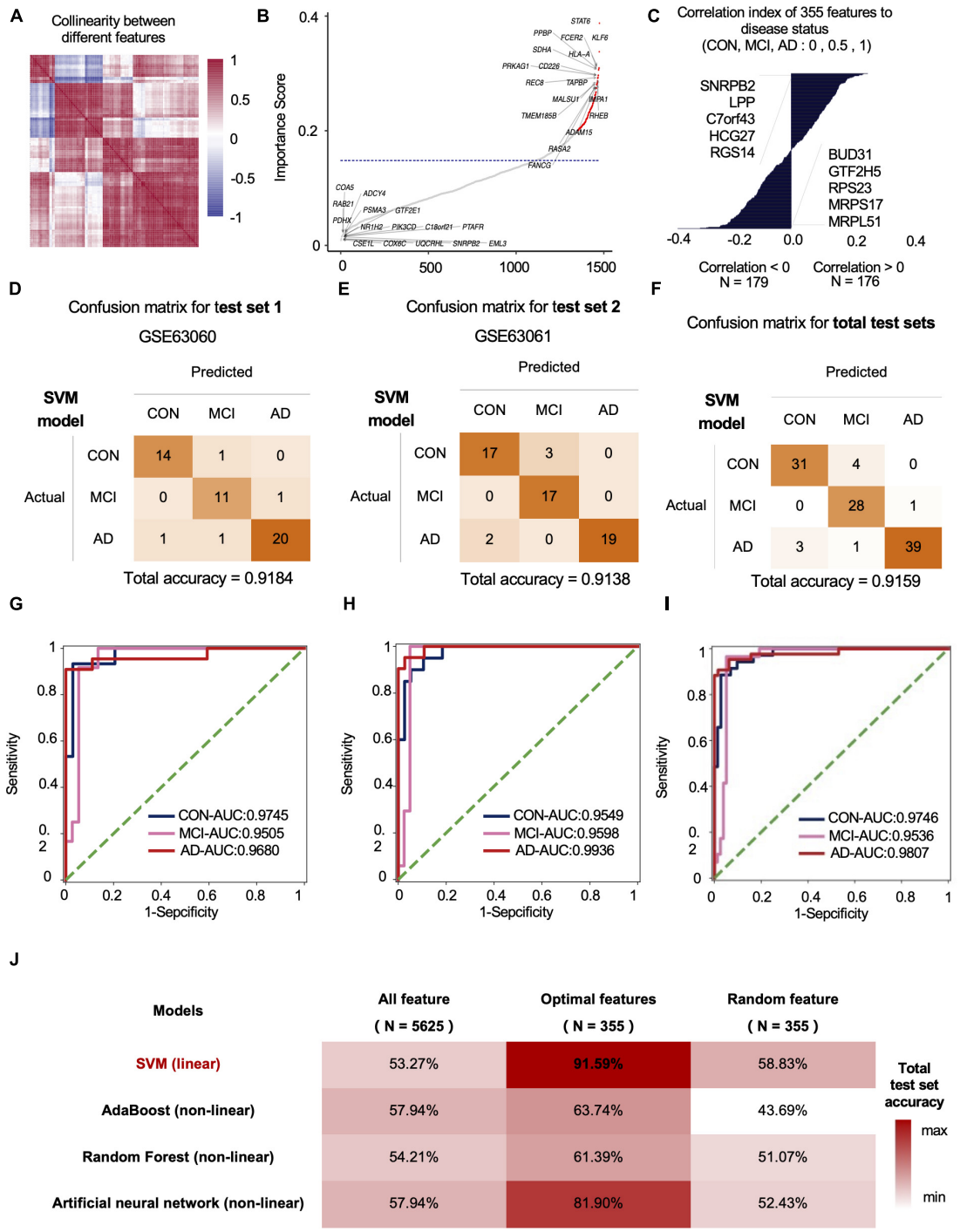


FIGURE 2 | The mRNA data-based feature engineering and AI-modeling of CON, MCI and AD. **(A)** Correlation of each feature, correlation threshold 0.9. **(B)** Filter lower importance features, the y-axis represents the standard deviation, dashed line is the cutting threshold, the x-axis represents the features in order of importance (ascending). **(C)** Correlation between features and labels, show top five features. **(D–F)** Confusion matrix of two independent test sets. **(G–I)** ROC curve of the model developed by optimal feature in two independent test sets. **(J)** Compare the prediction accuracy on test set between different models and features.

iterative method, removing 1% of the features in each iteration), optimal feature set was obtained for 355 mRNA features. We used the SVM algorithm to model the optimal feature set and then tested the model in two independent test sets, with the final test accuracy of 91.84 and 91.38%, respectively, and the average accuracy of 91.59% (**Figures 2D–F**). AUC values for CON, MCI and AD groups were 0.9746, 0.9536, and 0.9807, respectively (**Figures 2G–I**). Compared with the accuracy of

optimal feature set (91.59%), the accuracy of SVM model established by all features (1,920 mRNAs) and random features (355 mRNAs) under SVM algorithm was 53.27 and 58.83%, respectively (**Figure 2J**). In addition to SVM, we also evaluate the classification performance of the models based on other algorithms. The accuracy rates of Adaboost, Random Forest and Artificial Neural Network test sets were 66.36, 62.62, and 80.56%, respectively, lower than the optimal model obtained by the SVM algorithm (**Figure 2J**).

During the Progression of Alzheimer's Disease, Both Blood mRNA and Blood Cells Showed Significant Expression Imbalance

To optimize the classification efficiency of the artificial intelligence diagnostic model, we tried to include more biological information of different mRNAs. With the progression of AD, the composition of the various immune cells in the blood gradually changes. XCell is a method for inferring the quantitative abundance of 64 cell types based on mRNA expression data. We used XCell to analyze the quantitative level and composition of various cells in the blood and their changing trend with the progression of AD. Our results showed that the original blood mRNA expression profiles were mainly composed of five cell categories (41 cell subtypes), including HSC, Lymphoid, Myeloid, Epithelial, and Stroma. With the progression of AD, the proportions of HSC, Lymphoid and Myeloid in blood cells gradually decrease, while the proportions of Epithelial and Stroma gradually increase (**Figure 3A**). **Figure 3B** shows the absolute abundances of some representative cells (from all 41 cell types) and their relative abundances between the CON, MCI, and AD groups. To study the imbalance of cell proportion, we further analyzed the ratio of cell abundance. The top 10 cell ratios in CON, MCI, and AD groups were mainly related to immunity (58.8%). This result further suggest that AD disease is associated with immune dysregulation and can be recognized from the blood (**Figure 3C**; Ikeda et al., 2010). With the progression of the disease, we identified a total of 33 pairs of cells that showed a gradual change in the ratio. Most of them (27 pairs) gradually rise, including melanoma/GMP, melanoma/B-cells and melanoma/pro B-cells (**Figure 3C**, Red, **Figure 3D**, Up). The ratio of six pairs of cells decreased gradually, including $CD8^+$ naive T-cells/Plasma cells and $CD8^+$ naive T-cells/Th1 cells, etc. (**Figure 3D**, Down).

Cell-Related Imbalance Can Be Included in the Feature Set to Participate in the Model Optimization

We fused four types of features, including mRNA, mRNA ratio, XCell, and XCell ratio for feature engineering to obtain the optimal feature set and subsequent AI modeling to diagnose CON, MCI, and AD (**Figure 4A**). Similar to the feature screening method in the first modeling, we recalculated the Co-linearity of each feature in the data, filtered out the remaining 956 features after the Co-linearity was greater than 0.9, and then eliminated the insignificant features by iterative method (each iteration removed 1% of the features), leaving 319. Finally, we selected the features with a cumulative weight greater than 75% to form a new optimal feature set. The optimal feature set contained 119 mRNAs, 56 mRNA-ratio pairs, and 6 XCell-ratio pairs. The mRNA ratio feature accounted for 60% of the top 20 importance rankings of the optimal feature set. Among them, the CFLAR/FCXER2 ratio with high importance was gradually increased in CON, MCI, and AD groups (**Figure 4B**). A previous study reported that the CFLAR is a vital gene encoding apoptosis regulator, and the FCXER2 is an important gene related to immunity (Stelzer et al., 2016).

The inclusion of XCell-ratio features reveals the importance of the imbalance of the proportion between different blood cells in the modeling of CON, MCI, and AD. The most essential XCell-ratio features for modeling included MV colorectal cells/Osteoblast, $CD8^+$ naive T-cells/Mesangial cells and GMP/Osteoblast (**Figure 4C**). The optimal features are closely related to the progression of AD, where SNTB2, ATP6AP1/TRAPPC2L and CFLAR/FCER2 are positively correlated with the progression of AD. In contrast, MRPS17, AIF1 and GTF2H5 are negatively correlated with the progression of AD (**Figure 4D**).

The Introduction of the Concept of Proportion Imbalance Is Beneficial to the Establishment of Artificial Intelligence Individualized Diagnosis Model

The imbalance of mRNA ratio and cell ratio was observed during the progression of AD. To evaluate the impact of the imbalance on AD diagnosis, we incorporated four algorithms, including linear SVM (linear model), Adaboost (non-linear model), random forest (non-linear model) and Artificial neural networks (non-linear model), to established artificial intelligence models. The results show that the accuracy of SVM algorithm is the highest. The accuracy of SVM, Adaboost, Random Forest and artificial neural networks for the test set were 91.59, 66.36, 62.62, and 80.56%, respectively. Compared with the optimal feature set, the accuracy of the model based on the total features and 181 random features was lower, indicating that our method of feature establishment, evaluation and screening is reasonable, effective, and reliable (**Figure 4E**).

At present, the accuracy of the best model obtained by the model in two independent test sets was 93.88 and 89.66%, respectively, with an average accuracy of 91.59% (**Figures 4F–H**). Notably, the average recall rate for AD patients in this set was 93.02%. Further analysis showed that the AUC values of the CON, MCI and AD groups were 0.9524 (CON and other groups), 0.9651 (MCI and other groups) and 0.9807 (AD and other groups), respectively (**Figures 4I–K**).

We also obtained a better tri-classification diagnosis model by including cell ratio and mRNA ratio data. Compared with the previously reported dichotomies, our model is more accurate and stable. Since we have comprehensively considered the changes in the occurrence and development of Con-MCI-AD, the model

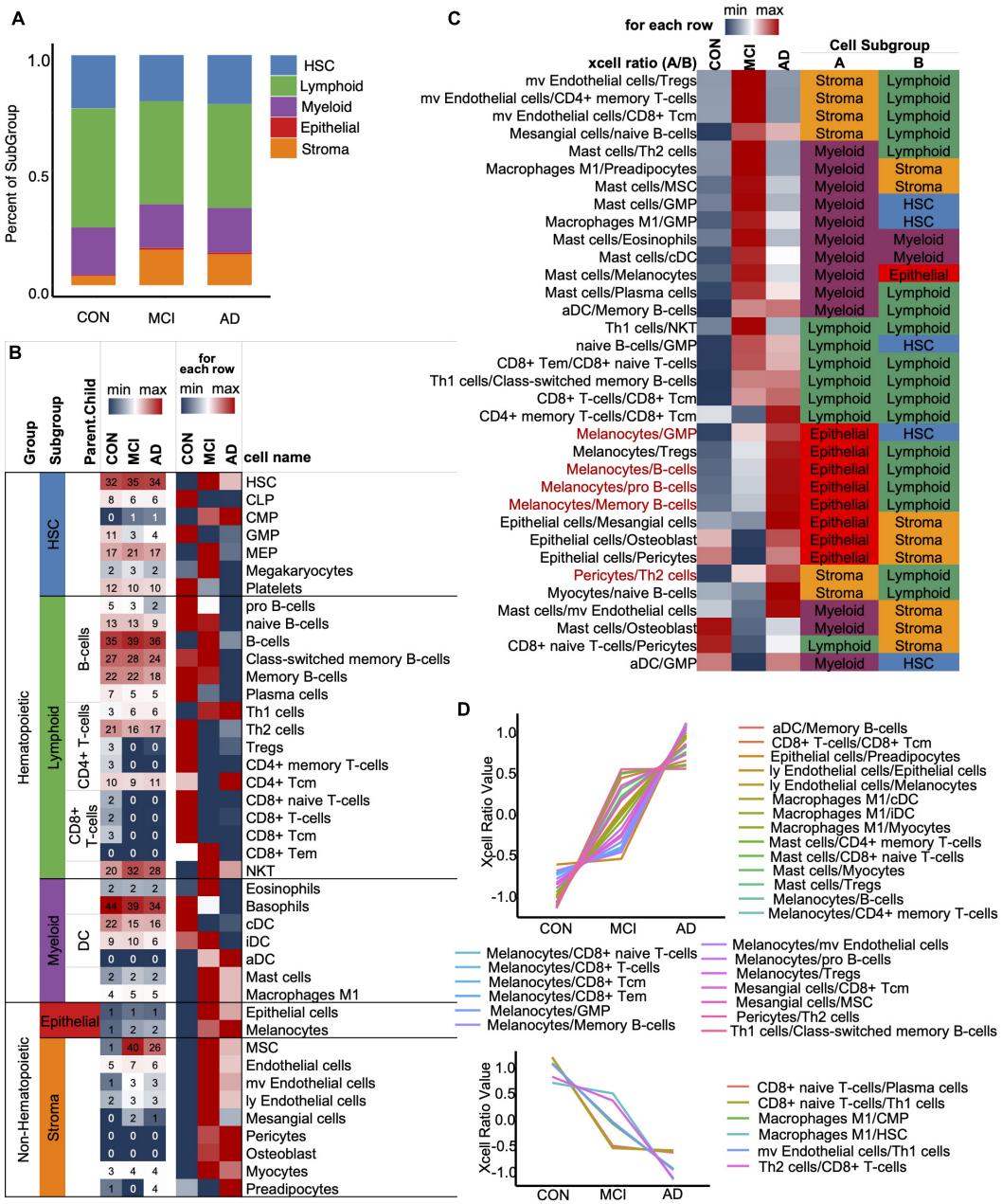


FIGURE 3 | XCell and XCell-Ratio analysis. **(A)** Stacked five cell types in the CON, MCI, and AD groups. **(B)** Representative cells from five cell types. The number is original XCell score. **(C)** The union set of top-10 XCell-ratios in the con, MCI, and AD groups. All values are normalized by row. Red items, non-Hematopoietic/Hematopoietic items show a gradual trend with AD progression. **(D)** Cell ratios show a gradual trend with AD progression.

obtained in this study covers a wider area and applies to more potentially susceptible populations.

Cell Ratio Analysis Showed That There Were Three Aspects of Immune Disorders During the Progression of Alzheimer's Disease

To comprehensively analyze the imbalance of blood cell proportion in the body, we matched the variation trend of each cell with the progression of AD with the differentiation process of pluripotent stem cells. Our results showed three types of imbalances in the blood cells of patients with AD progression (**Figure 5**).

First, the proportion of hematopoietic and non-hematopoietic cells decreased gradually (Abnormal 1). The number of non-hematopoietic cell types in the disease group (MCI and AD) increased significantly compared with CON. MSC cells with multi-organ differentiation potential showed a more substantial increase in MCI patients than the CON. Its downstream, such as

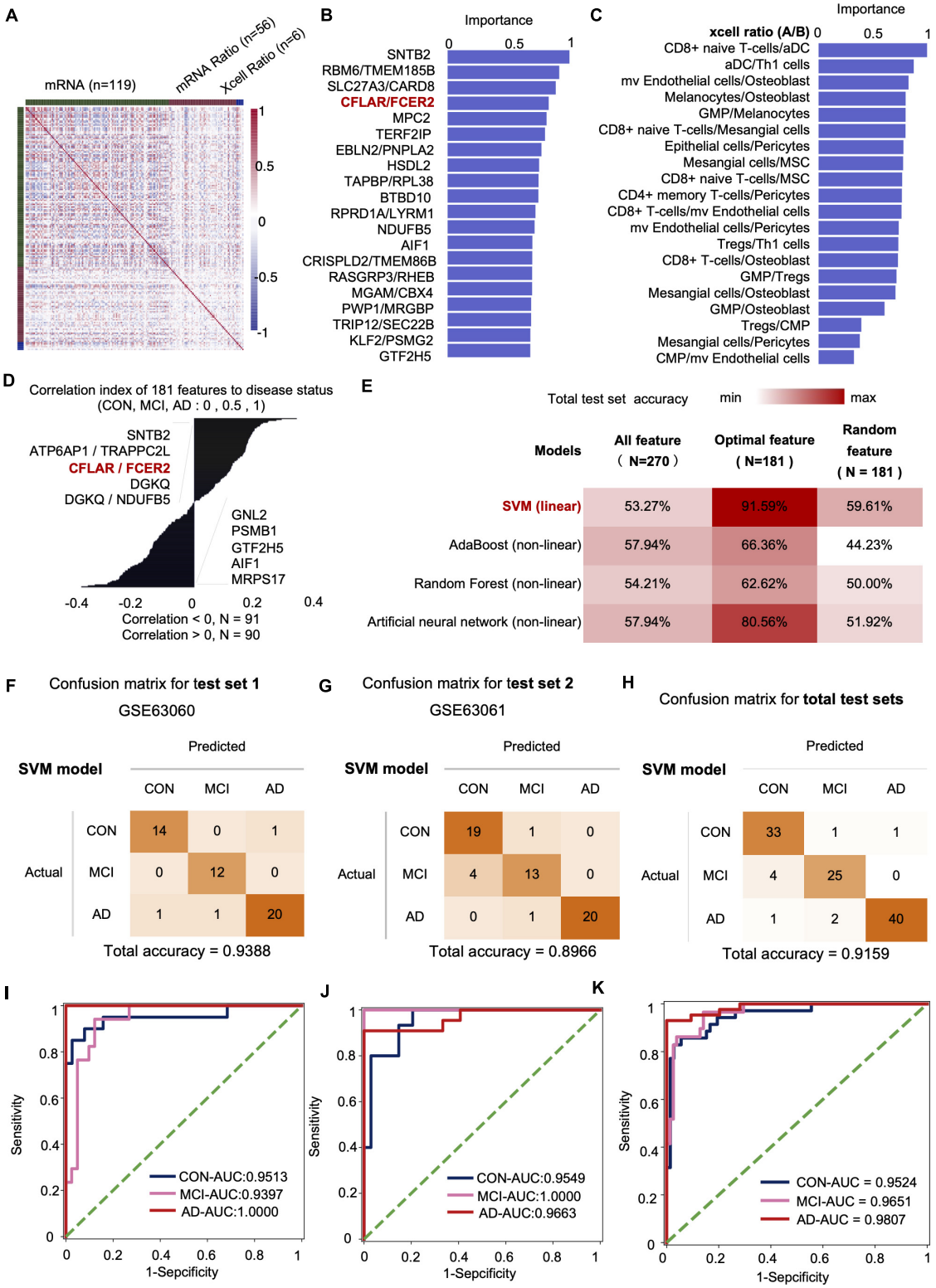


FIGURE 4 | The mRNA and XCell ratio data-based feature engineering and AI-modeling of CON, MCI, and AD. **(A)** Feature type correlation heat map, an overview of the correlations between features after feature engineering. **(B)** Top 20 features importance (from optimal 181 features set). **(C)** XCell-Ratio of the top 20 in importance. The numbers represent the abundance of different cells in CON, MCI and AD. **(D)** The 181 features relevance between features and labels. **(E)** Compare the prediction accuracy on test set between different models and different features. **(F–H)** Confusion matrix of test set. **(I–K)** ROC curve of the SVM model developed by optimal features.

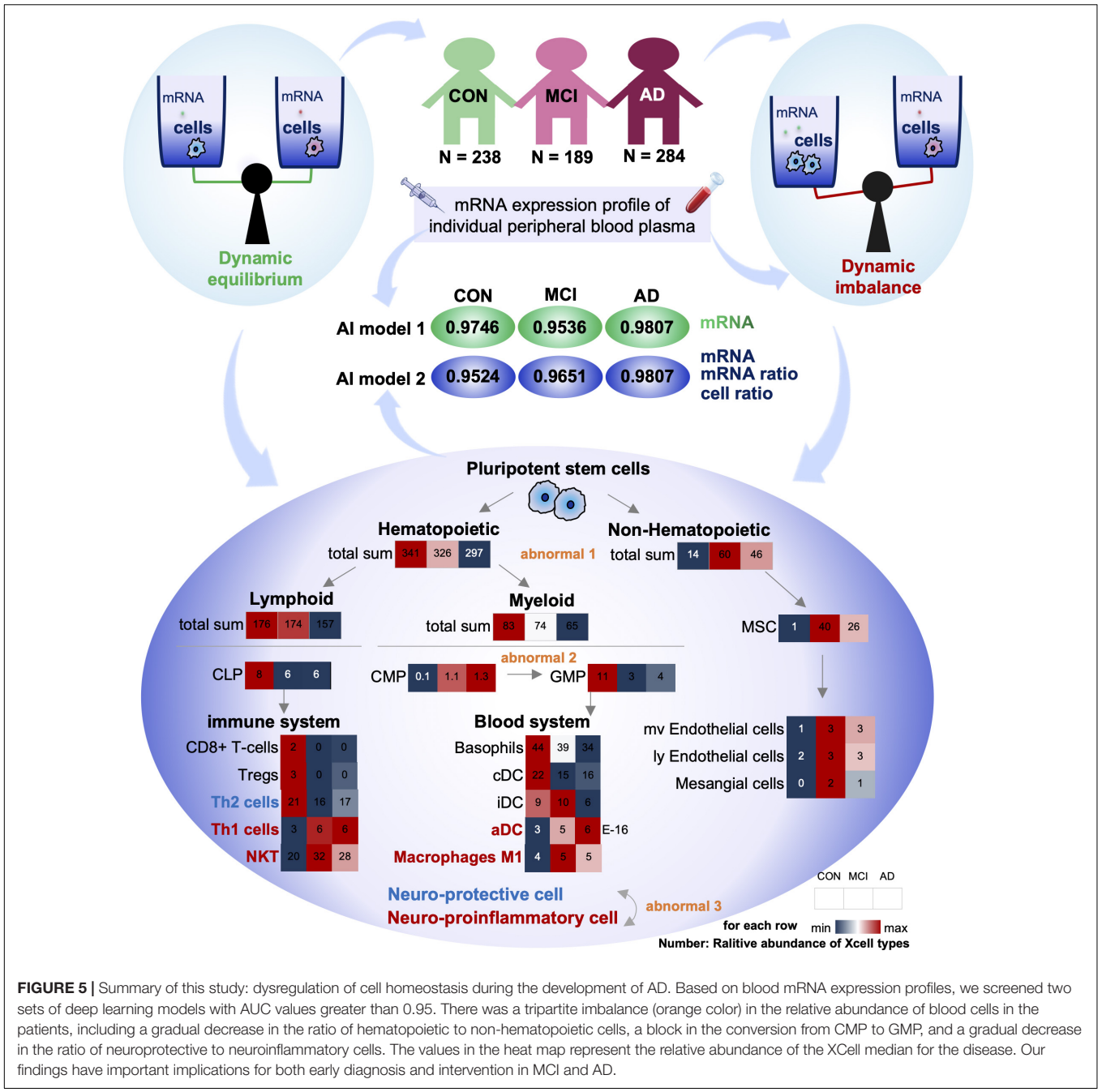


FIGURE 5 | Summary of this study: dysregulation of cell homeostasis during the development of AD. Based on blood mRNA expression profiles, we screened two sets of deep learning models with AUC values greater than 0.95. There was a tripartite imbalance (orange color) in the relative abundance of blood cells in the patients, including a gradual decrease in the ratio of hematopoietic to non-hematopoietic cells, a block in the conversion from CMP to GMP, and a gradual decrease in the ratio of neuroprotective to neuroinflammatory cells. The values in the heat map represent the relative abundance of the XCell median for the disease. Our findings have important implications for both early diagnosis and intervention in MCI and AD.

Endothelial cells and Mesangial cells, also showed similar trends. However, hematopoietic cell-related items were significantly downregulated in the disease group (MCI and AD). It has been reported that hematopoietic microglia can ameliorate the progression of AD by eliminating amyloid deposition through cell-specific phagocytic mechanisms (Lampron et al., 2011). Furthermore, hematopoietic cells and their associated factors have potential therapeutic value in AD, and the downregulation of the relative number of hematopoietic cells may be one of the critical theories (Sanchez-Ramos et al., 2008; Lim et al., 2020).

Secondly, the differentiation of CMP to GMP was resisted (Abnormal 2): in the hematopoietic cells, both Lymphoid and Myeloid cells showed a trend of gradual decrease. The Myeloid is the primary source of cells in the blood system. The common myeloid precursor cell index was significantly higher in the disease group (MCI and AD) compared with CON. As the downstream of its differentiation, the proportion of progenitor cells of granular macrophages decreased gradually. We speculated that blocking the differentiation of CMP to GMP might be closely related to the occurrence of MCI and AD.

Thirdly, the proportion of neuroprotective and neuroinflammatory cells was gradually reduced (Abnormal 3). Next, we analyzed the immune system differentiated by Lymphoid and Blood system differentiated by Myeloid. Our

results showed that with the gradual development of AD, the immune system and circulatory system both changed from increasing to decreasing cell types. There is no apparent difference between the two systems. However, most of the progressively elevated cells are neuroinflammatory, including Th1, NTK, and ADC. The neuroprotective cells declined gradually, including Th2.

Compared with the CON group, the ratio of Th1 cells to Th2 cells in the disease group (MCI and AD) was significantly unbalanced. This conclusion is consistent with the results of previous animal immunotherapy experiments, which showed that Th1 cells decreased and Th2 cells increased in AD mice after immunotherapy (Town et al., 2002; Cao et al., 2009; Marciani, 2016). Notably, various indicators of NKT cells were similar to those of Th1 cells, suggesting that NKT cells also play an important regulatory role in the onset and progression of AD. However, no studies have confirmed that these cells play an important role in the onset and progression of AD. Therefore, NKT cells may be an important feature related to AD that has been recently discovered. In conclusion, we believe that the imbalance of homeostasis in GMP may be one of the important causes of AD, and the imbalance of homeostasis in NKT cells may be closely related to the occurrence and development of AD (**Figure 5**).

DISCUSSION

In this study, we established a complete feature engineering framework and an excellent machine learning model. We identified a set of features with stable classification efficiency based on the multidimensional data of mRNA expression profiles.

We found many mRNA features at the mRNA level that were significantly up- or down-regulated during AD progression. For example, the ATP5F1c gene is significantly down-regulated in the disease group (MCI and AD) compared to the CON group. The ATP5F1c was reported to play an important role in mitochondrial oxidative phosphorylation (Stelzer et al., 2016). The expressed level of mitochondrial electron transport chain complex IV (COX) was significantly reduced in AD patients. A previous study showed that genetic defects in the Cox family might be associated with the genetic risk of AD. In addition to the mRNA features, the mRNA ratio feature accounted for 60% of the top 20 importance rankings of the optimal feature set. Among them, CFLAR is an important gene encoding apoptosis regulator. FCXER2 is an important gene related to immunity (Stelzer et al., 2016). The importance of CFLAR/FCXER2 was higher, and its ratio gradually increased in CON, MCI and AD groups. We speculate that the occurrence and development of AD may be due to neuronal death caused by immune system abnormalities, and this process can be found in the blood.

At the cellular level, we also identified some particularly important features. For example, $CD4^+$ and $CD8^+$ (including $CD8^+$ naive T cells, $CD8^+$ T cells, and $CD8^+$ TCM) cells were significantly lower in AD and patients with mild cognitive impairment. $CD4^+$ T cells are effective mediators of well-known autoimmune diseases in the nervous system, such as multiple sclerosis and narcolepsy, which are involved in developing microglia (Pasciuto et al., 2020). Furthermore, we found a significant increase in the expression of myeloid cells represented by activated dendritic cells (ADC) in the disease group (MCI and AD). In contrast, the expression of lymphocytes represented by B cells was significantly reduced (**Figure 3C**). We found that many cells associated with the immune system showed a gradual increase during the occurrence and development of the disease (**Figure 3D**). Among the cell proportion features with noticeable progressive changes in expression, we found that almost all the proportion features were related to myeloid cells or lymphocytes (88.24%); among them, 58.8% were related to lymphocytes and 38.24% were related to myeloid cells.

We found that mitochondrial dysfunction in the brain tissue of AD patients can be simultaneously detected in the peripheral system (Johri and Beal, 2012), suggesting that AD may be caused by abnormal gene expression or brain damage, as observed in peripheral blood (Johri and Beal, 2012; Leuner et al., 2012; Trushina et al., 2013; Pérez et al., 2017). DEGs in peripheral blood may be one of the important causes of AD. Machine learning is an important branch of artificial intelligence. The main difference between this method and the traditional statistical learning method is that the machine learning method usually does not need a statistical hypothesis, which dramatically improves the accuracy of training results and the adaptability of the model, and is widely used in the study of the pathogenesis of AD (Farran et al., 2013; Goecks et al., 2020).

Compared with traditional feature engineering, this paper not only pays attention to feature selection but also pays attention to the development of new dimension features. In the cell type score obtained based on mRNA expression profile data, we found that Th and NKT cells were different between the disease group (MCI and AD) and the control group, with significantly fewer Th2 cells and significantly more Th1 cells and NKT (**Figure 5**). It is suggested that the occurrence and development of AD are closely related to immune system diseases, consistent with a previous report (Cui and Wan, 2019), we should pay high attention to the homeostatic dysregulation of NKT cells in AD. The features used in our model are highly interpretable.

CONCLUSION

We find 5625 DEGs in CON, MCI, and AD, which are related to the disease. The optimal feature set was obtained through feature engineering screening, and the artificial intelligence individualized diagnosis model established based on this method achieved a classification accuracy of 91.59% in the test set. The AUC of CON, MCI, and AD were 0.9746, 0.9536, and 0.9807, respectively. The relative abundance of five types of cells, including Epithelial, HSC, Lymphoid, Myeloid and Stroma in the blood of CON, MCI and AD patients was obtained by mRNA expression profile analysis. We also included mRNA, cell abundance and ratio information to establish an artificial intelligence model. The diagnostic accuracy of the optimal model in the tri-classification test set was 91.59%, and the

diagnostic AUC of CON, MCI and AD were 0.9524, 0.9651, and 0.9807, respectively. Based on the mRNA profiles, we analyzed the ratio of different cells using XCell. As patients progressively deteriorated from CON, MCI to AD, blood cells displayed three aspects of imbalance, including a progressive decrease in the proportion of hematopoietic cells, a block in the differentiation of CMP to GMP, and a progressive decrease in the proportion of neuroprotective/neuroinflammatory cells. Our findings have important implications for both early diagnosis and intervention in MCI and AD.

In this study, the composition of various cells in the blood of a single patient was analyzed based on the blood mRNA expression profile. Based on this, the balance between different mRNAs and cells in blood was analyzed. For the imbalance of disease and cell proportion in CON, MCI, and AD patients and their contribution to the artificial intelligence model, this study provides new ideas and results for the onset and progression of AD from both basic and application perspectives. The 181 features are composed of four dimensions, which can accurately classify CON, MCI, and AD groups, suggesting that machine learning methods can capture changes in blood biomarkers in AD patients. The results of cell homeostasis analysis suggested that the homeostasis of NKT cells might be related to AD, and the homeostasis of GMP might be one of the possible reasons for AD.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

XZ and ZH designed the study. YS, BL, JZ, and YZ performed the analyses and interpreted the results. YS and XZ wrote the manuscript. XZ conducted this study. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Nos. 32027801, 31870992, and 21775031), the Strategic Priority Research Program of Chinese Academy of Sciences (Grant Nos. XDB36000000 and XDB38010400), CAS-JSPS (Grant No. GJHZ2094), Research Foundation for Advanced Talents of Fujian Medical University (Grant Nos. XRCZX2017020 and XRCZX2019005), Beijing Natural Science Foundation Haidian original innovation joint fund (Grant No. L202023), the China Postdoctoral Science Foundation (Grant No. 2021M690806). The funding body had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2021.768229/full#supplementary-material

Supplementary Figure 1 | The blood mRNA expression of CON, MCI and AD patients showed a different trend. **(A)** Heatmap of GSE63060 and GSE63061, rows represent genes, and columns represent samples. **(B)** DEGs of three groups between each other, bar-plot show the number of all up-down DEGs in different groups, the Venn-plot show that the intersection DEGs, the upset-plot show that all kind of intersections DEGs. The bar graph on the left represents the intersection of up-regulated differential genes between each group. In the middle of the scatter plot, the single point represents the unique gene in the data set, and the point-line-connection represents the intersection of the data and other data. **(C)** IPA localization and types score of DEGs, base-line (gray-line) CON/CON, red represent cell membrane localization. **(D)** Pathway enrichment results, the absolute value of x-axis is log10(*P*-value), the positive and negative signs represent up and down-regulation.

REFERENCES

Adav, S. S., Park, J. E., and Sze, S. K. (2019). Quantitative profiling brain proteomes revealed mitochondrial dysfunction in Alzheimer's disease. *Mol. Brain* 12:8. doi: 10.1186/s13041-019-0430-y

Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18:220. doi: 10.1186/s13059-017-1349-1

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., et al. (2013). API design for machine learning software: experiences from the scikit-learn project. *arXiv* [Preprint]. arXiv:1309.0238

Burns, A., and Iliffe, S. (2009). Alzheimer's disease. *BMJ* 338, b158–b158. doi: 10.1136/bmj.b158

Cao, C., Arendash, G. W., Dickson, A., Mamcarz, M. B., Lin, X., and Ethell, D. W. (2009). Aβ-specific Th2 cells provide cognitive and pathological benefits to Alzheimer's mice without infiltrating the CNS. *Neurobiol. Dis.* 34, 63–70. doi: 10.1016/j.nbd.2008.12.015

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018

Cui, Y., and Wan, Q. (2019). NKT cells in neurological diseases. *Front. Cell. Neurosci.* 13:245. doi: 10.3389/fncel.2019.00245

Farran, B., Channanath, A. M., Behbehani, K., and Thanaraj, T. A. (2013). Predictive models to assess risk of type 2 diabetes, hypertension and comorbidity: machine-learning algorithms and validation using national health data from Kuwait—a cohort study. *BMJ Open* 3:e002457. doi: 10.1136/bmjopen-2012-002457

Freund, Y., and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *J. Computer System Sci.* 55, 119–139. doi: 10.1006/jcss.1997.1504

Goecks, J., Jalili, V., Heiser, L. M., and Gray, J. W. (2020). How machine learning will transform biomedicine. *Cell* 181, 92–101. doi: 10.1016/j.cell.2020.03.022

Herrero-Labrador, R., Trueba-Saiz, A., Martinez-Rachadell, L., Fernandez, de Sevilla, M. E., Zegarra-Valdivia, J. A., et al. (2020). Circulating insulin-like growth factor i is involved in the effect of high fat diet on peripheral amyloid β clearance. *IJMS* 21:9675. doi: 10.3390/ijms21249675

Ikeda, T., Yamamoto, K., Takahashi, K., and Yamada, M. (2010). Immune system-associated antigens on the surface of peripheral blood lymphocytes in patients with Alzheimer's disease. *Acta Psychiatr. Scand.* 83, 444–448. doi: 10.1111/j.1600-0447.1991.tb05573.x

Johri, A., and Beal, M. F. (2012). Mitochondrial dysfunction in neurodegenerative diseases. *J. Pharmacol. Exp. Ther.* 342, 619–630. doi: 10.1124/jpet.112.192138

Krämer, A., Green, J., Pollard, J., and Tugendreich, S. (2014). Causal analysis approaches in ingenuity paway analysis. *Bioinformatics* 30, 523–530. doi: 10.1093/bioinformatics/btt703

Lampron, A., Gosselin, D., and Rivest, S. (2011). Targeting the hematopoietic system for the treatment of Alzheimer's disease. *Brain Behav. Immun.* 25, S71–S79. doi: 10.1016/j.bbi.2010.12.018

Leuner, K., Schulz, K., Schütt, T., Pantel, J., Prvulovic, D., Rhein, V., et al. (2012). Peripheral mitochondrial dysfunction in Alzheimer's disease: focus on lymphocytes. *Mol. Neurobiol.* 46, 194–204. doi: 10.1007/s12035-012-8300-y

Li, X., Wang, H., Long, J., Pan, G., He, T., Anichtchik, O., et al. (2018). Systematic analysis and biomarker study for Alzheimer's disease. *Sci. Rep.* 8:17394. doi: 10.1038/s41598-018-35789-3

Liaw, A., and Wiener, M. (2002). Classification and regression by randomforest. *R. News* 2:23.

Lim, S. L., Tran, D. N., Kieu, Z., Chen, C., Villanueva, E., Ghiaar, S., et al. (2020). Genetic ablation of hematopoietic cell kinase accelerates Alzheimer's disease–like neuropathology in Tg2576 mice. *Mol. Neurobiol.* 57, 2447–2460. doi: 10.1007/s12035-020-01894-6

Ludwig, N., Fehlmann, T., Kern, F., Gogol, M., Maetzler, W., Deutscher, S., et al. (2019). Machine learning to detect Alzheimer's disease from circulating non-coding RNAs. *Genom. Proteomics Bioinform.* 17, 430–440. doi: 10.1016/j.gpb.2019.09.004

Lunnon, K., Sattlecker, M., Furney, S. J., Coppola, G., Simmons, A., Proitsi, P., et al. (2013). A blood gene expression marker of early Alzheimer's disease. *JAD* 33, 737–753. doi: 10.3233/JAD-2012-121363

Marciani, D. J. (2016). A retrospective analysis of the Alzheimer's disease vaccine progress - the critical need for new development strategies. *J. Neurochem.* 137, 687–700. doi: 10.1111/jnc.13608

Pasciuto, E., Burton, O. T., Roca, C. P., Lagou, V., Rajan, W. D., Theys, T., et al. (2020). Microglia require CD4 T cells to complete the fetal-to-adult transition. *Cell* 182, 625–640.e24. doi: 10.1016/j.cell.2020.06.026

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Machine Learn. Res.* 12, 2825–2830.

Pérez, M. J., Ponce, D. P., Osorio-Fuentealba, C., Behrens, M. I., and Quintanilla, R. A. (2017). Mitochondrial bioenergetics is altered in fibroblasts from patients with sporadic Alzheimer's disease. *Front. Neurosci.* 11:553. doi: 10.3389/fnins.2017.00553

Rogers, J. A., Polhamus, D., Gillespie, W. R., Ito, K., Romero, K., Qiu, R., et al. (2012). Combining patient-level and summary-level data for Alzheimer's disease modeling and simulation: a beta regression meta-analysis. *J. Pharmacokinet. Pharmacodyn.* 39, 479–498. doi: 10.1007/s10928-012-9263-3

Sanchez-Ramos, J., Song, S., Cao, C., and Arendash, G. (2008). The potential of hematopoietic growth factors for treatment of Alzheimer's disease: a mini-review. *BMC Neurosci.* 9:S3. doi: 10.1186/1471-2202-9-S2-S3

Smyth, G. K. (2005). "limma: linear models for microarray data," in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor, Statistics for Biology and Health*, eds R. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, and S. Dudoit (New York, NY: Springer-Verlag), 397–420. doi: 10.1007/0-387-29362-0_23

So, J. H., Madusanka, N., Choi, H. K., Choi, B. K., and Park, H. G. (2019). Deep learning for Alzheimer's disease classification using texture features. *Curr. Med. Imaging Rev.* 15, 689–698. doi: 10.2174/1573405615666190404163233

Sood, S., Gallagher, I. J., Lunnon, K., Rullman, E., Keohane, A., Crossland, H., et al. (2015). A novel multi-tissue RNA diagnostic of healthy ageing relates to cognitive health status. *Genome Biol.* 16:185. doi: 10.1186/s13059-015-0750-x

Stamate, D., Kim, M., Proitsi, P., Westwood, S., Baird, A., and Nevado-Holgado, A. (2019). A metabolite-based machine learning approach to diagnose Alzheimer-type dementia in blood: results from the european medical information framework for Alzheimer disease biomarker discovery cohort. *Alzheimer's Dementia: Transl. Res. Clin. Intervent.* 5, 933–938. doi: 10.1016/j.trci.2019.11.001

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., et al. (2016). The genecards suite: from gene data mining to disease genome sequence analyses. *Curr. Protocols Bioinform.* 54:1-1.30.33. doi: 10.1002/cpbi.5

Town, T., Vendrame, M., Patel, A., Poetter, D., DelleDonne, A., Mori, T., et al. (2002). Reduced Th1 and enhanced Th2 immunity after immunization with Alzheimer's h-amyloid1–42. *J. Neuroimmunol.* 132, 49–59. doi: 10.1016/s0165-5728(02)00307-7

Trushina, E., Dutta, T., Persson, X.-M. T., Mielke, M. M., and Petersen, R. C. (2013). Identification of altered metabolic pathways in plasma and CSF in mild cognitive impairment and Alzheimer's disease using metabolomics. *PLoS One* 8:e63644. doi: 10.1371/journal.pone.0063644

Way, G. P., Sanchez-Vega, F., La, K., Armenia, J., Chatila, W. K., Luna, A., et al. (2018). Machine learning detects pan-cancer ras pathway activation in the cancer genome atlas. *Cell Rep.* 23, 172–180.e3. doi: 10.1016/j.celrep.2018.03.046

Word Health Organization (2020). *Word Health Organization*. Geneva: WHO.

Yu, G., Wang, L.-G., Yan, G.-R., and He, Q.-Y. (2015). DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* 31, 608–609. doi: 10.1093/bioinformatics/btu684

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 24 December 2021
doi: 10.3389/fnagi.2021.804384

Classification of Cognitive Impairment and Healthy Controls Based on Transcranial Magnetic Stimulation Evoked Potentials

Jiahao Zhang[1,2†], Haifeng Lu[3†], Lin Zhu[4†], Huixia Ren[5,6], Ge Dang[4], Xiaolin Su[4], Xiaoyong Lan[4], Xin Jiang[7], Xu Zhang[1,2], Jiansong Feng[1,2], Xue Shi[4], Taihong Wang[1,2], Xiping Hu[3,8*] and Yi Guo[4,9*]*

[1] Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China, [2] School of Microelectronics, Southern University of Science and Technology, Shenzhen, China, [3] Gansu Provincial Key Laboratory of Wearable Computing, School of Information Science and Engineering, Lanzhou University, Lanzhou, China, [4] Department of Neurology, Shenzhen People's Hospital (The First Affiliated Hospital of Southern University of Science and Technology, The Second Clinical Medical College of Jinan University), Shenzhen, China, [5] Department of Neurology, The Second Clinical Medical College, Jinan University (Shenzhen People's Hospital), Shenzhen, China, [6] The First Affiliated Hospital, Jinan University, Guangzhou, China, [7] Department of Geratic, Shenzhen People's Hospital (The First Affiliated Hospital of Southern University of Science and Technology, The Second Clinical Medical College of Jinan University), Shenzhen, China, [8] School of Intelligent Systems Engineering, Sun Yat-sen University, Shenzhen, China, [9] Shenzhen Bay Laboratory, Shenzhen, China

Backgrounds: Nowadays, risks of Cognitive Impairment (CI) [highly suspected Alzheimer's disease (AD) in this study] threaten the quality of life for more older adults as the population ages. The emergence of Transcranial Magnetic Stimulation-Electroencephalogram (TMS-EEG) enables noninvasive neurophysiological investi-gation of the human cortex, which might be potentially used for CI detection.

Objectives: The aim of this study is to explore whether the spatiotemporal features of TMS Evoked Potentials (TEPs) could classify CI from healthy controls (HC).

Methods: Twenty-one patients with CI and 22 HC underwent a single-pulse TMS-EEG stimulus in which the pulses were delivered to the left dorsolateral prefrontal cortex (left DLPFC). After preprocessing, seven regions of interest (ROIs) and two most reliable TEPs' components: N100 and P200 were selected. Next, seven simple and interpretable linear features of TEPs were extracted for each region, three common machine learning algorithms including Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbor (KNN) were used to detect CI. Meanwhile, data augmentation and voting strategy were used for a more robust model. Finally, the performance differences of features in classifiers and their contributions were investigated.

Results: 1. In the time domain, the features of N100 had the best performance in the SVM classifier, with an accuracy of 88.37%. 2. In the aspect of spatiality, the features of the right frontal region and left parietal region had the best performance in the SVM classifier, with an accuracy of 83.72%. 3. The Local Mean Field Power (LMFP), Average Value (AVG), Latency and Amplitude contributed most in classification.

OPEN ACCESS

Edited by:
Peng Xu,
University of Electronic Science and Technology of China, China

Reviewed by:
Fangzhou Xu,
Qilu University of Technology, China
Hua Yuan,
Fourth Military Medical University, China

***Correspondence:**
Taihong Wang
wangth@sustech.edu.cn
Xiping Hu
huxp@lzu.edu.cn
Yi Guo
xuanyi_guo@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

Received: *29 October 2021*
Accepted: *25 November 2021*
Published: *24 December 2021*

Citation:
Zhang J, Lu H, Zhu L, Ren H, Dang G, Su X, Lan X, Jiang X, Zhang X, Feng J, Shi X, Wang T, Hu X and Guo Y (2021) Classification of Cognitive Impairment and Healthy Controls Based on Transcranial Magnetic Stimulation Evoked Potentials. Front. Aging Neurosci. 13:804384. doi: 10.3389/fnagi.2021.804384

Conclusions: The TEPs induced by TMS over the left DLPFC has significant differences spatially and temporally between CI and HC. Machine learning based on the spatiotemporal features of TEPs have the ability to separate the CI and HC which suggest that TEPs has potential as non-invasive biomarkers for CI diagnosis.

Keywords: spatiotemporal features, machine learning, cognitive impairment, TEP, TMS-EEG

1. INTRODUCTION

Cognitive impairment (CI) refers to a cognitive function decline beyond typical aging, which is increasingly prevalent in the elderly and loom as a public health issue (Montine et al., 2021).

Clinically, Montreal Cognitive Assessment (MoCA) and Mini-Mental State Examination (MMSE) are commonly used for routine cognitive screening. MoCA is more sensitive than MMSE in detecting mild cognitive impairment (MCI) (Ciesielska et al., 2016). For Alzheimer's disease (AD) diagnosis, there are two main types of biomarkers: biophysiological biomarkers such as β amyloid in cerebrospinal fluid (CSF)/plasma/serum and brain imaging markers (Frisoni et al., 2017; Ng et al., 2019; Cullen et al., 2021). For example, Aβ42/Aβ40 ratio in CSF and blood (Buchhave et al., 2012; Hansson et al., 2018), Positron Emission Tomography (PET) of beta-amyloid and tau proteins (Leuzy et al., 2019; Rabinovici et al., 2019). However, the invasiveness of the collection of body fluids and high cost of PET limit their large-scale use. In recent years, some articles have reported that combining different neuroimaging modalities together can effectively detect CI. For example, combining functional MRI (fMRI) and Diffusion Tracking Imaging (DTI) can reflect functional connectivity changes in neuronal networks between CI and Healthy Controls (HC) (Ye and Bai, 2018). Multimodal fusion combines the advantages of each modality, but it is undeniable that complex data fusion algorithms impose huge challenges for clinical application.

Due to the above problems, some researchers have turned their attention to find a quick, noninvasive, and inexpensive method to detect CI, especially for mild and moderate patients without obvious behavioral symptoms. As a non-invasive, high time resolution method, the electroencephalogram (EEG) has been widely used in clinical examinations. In recent decades, the abnormalities in the resting state EEG of patients with CI have been discovered, such as a shift of the power spectrum to lower frequencies, a decrease in the coherence of fast rhythms, a decreased complexity of EEG patterns (Jeong, 2004).

With the development of non-invasive neuromodulation technology, the combination of Transcranial Magnetic Stimulation and Electroencephalogram (TMS-EEG) allows external input to specific cortical areas of subjects in a controlled and quantitative way for direct functional assessment (Hallett, 2007; Kimiskidis, 2016; Cao et al., 2021). When TMS pulses are applied to the cortex, trans-synaptic activation of local and distal cortical networks is obtained (Tremblay et al., 2019). The sum of synaptic potentials can be recorded simultaneously by high time resolution, multichannel scalp EEG. There are a series of positive and negative deflections after TMS, known as TMS Evoked Potentials (TEPs). The TEPs last 300 ms or more and can be recorded by either local or distal electrodes (Komssi et al., 2002), reflecting the spread of activation over cortical regions that are functionally connected and indicating the state of the brain further (Nikulin et al., 2003). Compared with resting-state EEG, TMS-EEG provides controlled stimulation without the involvement of the participation, which is more stable and objective.

At present, some researchers have used TMS-EEG to assess patients with CI. For example, the prefrontal TMS-evoked activity was able to track disease progression in Alzheimer's Disease (AD) and the P30 amplitude was predictive of the MMSE score in patients with AD (Bagattini et al., 2019). The Motor-Evoked Potentials (MEPs) produced by paired pulses on the primary cortex can be used as indicators in the classification of different subtypes of MCI (Benussi et al., 2021).

In this study, we hypothesized that TEPs resulting from stimulation of the left DLPFC may be associated with the cognitive status, thus, the features of TEPs could further differentiate CI and HC. On the premise of preserving time and space features simultaneously, we extracted some concise, interpretable linear features of TEPs in seven regions of interest (ROIs). We aimed to classify CI and HC automatically through machine learning based on the spatiotemporal features of TEPs and find potential biomarkers for clinical diagnosis.

2. MATERIALS AND METHODS

The framework is shown in **Figure 1**. We removed artifacts of TMS-EEG at first. Then, we divided all the trials of each participant into three segments to the augment dataset. In order to preserve the features of time and space simultaneously, we focused on the TEP's N100 and P200 components in seven ROIs. Next, we explored some concise, interpretable linear features of these two components including Local Mean Field Power (LMFP), Latency, Amplitude, Standard Deviation, Average Value (AVG), Area Under the Curve (AUC), and Range. Finally, we used three common machine learning algorithms: K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Random Forest (RF) to obtain the label of each segment, and voted to get participant's final prediction result (Cover and Hart, 1967; Vapnik, 1999; Liaw and Wiener, 2002).

2.1. TMS-EEG Data Acquisition

2.1.1. Study Participants

All participants in this study were recruited from the Department of Neurology in Shenzhen People's Hospital (The First Affiliated

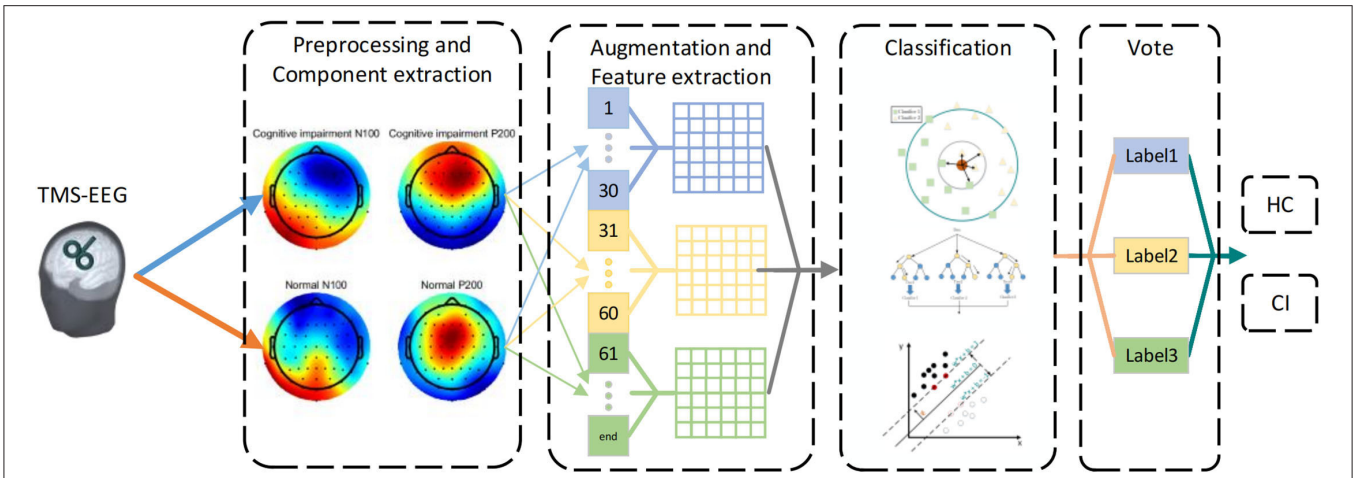


FIGURE 1 | The framework of CI detection based on TEPs. First, the N100 and P200 components of TEPs were selected after removing artifacts. Then, all trials were divided into three segments. Subsequently, seven features were extracted from different segments and regions of interest (ROIs) respectively. Finally, machine learning was used to classify features, and voted on each segment to get the predicted result.

Hospital, Southern University of Science and Technology).This study was approved by the Institutional Review Board of Shenzhen People's Hospital. All participants provided written informed consent.

A total of 43 subjects participated in this study, including 21 patients with CI and 22 HC. The inclusion criteria of CI group were: (a) Clinicians highly suspect the subject has AD based on the clinical diagnostic criteria (National Institute of Neurological and Communicative Disorders and stroke and the AD and Related Disorders Association (NINCDS-ADRDA), the Diagnostic and Statistical Manual of Mental Disorders-V (DSM-V) criteria); (b) Mild and moderate CI diagnosed (10<MoCA<26) (Nasreddine et al., 2005); (c) aged 50–75 years old. The exclusion criteria were: (a) blood vessels and other types of dementia; (b) severe CI (MoCA <10); (c) a history of other psychiatric or other neurological disorders, such as schizophrenia, Parkinson's disease, and multiple sclerosis; (d) any contraindication for TMS, such as a metallic implanted device in or near the head and aneurysms.

All of the subjects in the control group met the following criteria: (a) aged 50–75 years old; (b) never complained of cognition or memory problems; (c) no history of any psychiatric or neurological disorders, brain injury, cranial neurosurgery, alcohol or drug abuse, or any severe chronic systemic illness; (d) no contraindication for TMS.

There was no significant difference in age between the two groups ($p = 0.406$). Demographic information was summarized in **Table 1**.

TABLE 1 | Demographic subjects.

	CI	HC
Subject(s)	21	22
Age (mean ± SD)	61.86 ± 4.77	60.77 ± 3.65
Sex (male/female)	9:12	10:12
MoCA (mean ± SD)	20.33 ± 4.44	/

2.1.2. TMS-EEG Recordings

The dorsolateral prefrontal cortex (DLPFC) is a key node of various cognitive functions such as memory, attention, and execution (Carlén, 2017). As we aimed to research cognition related function, the left DLPFC is also a recommended target for TMS treatment (Ahmed et al., 2012). Therefore, we chose left DLPFC to be the target of stimulation in this study.

All of the subjects in this study underwent a TMS-EEG protocol. A total of 100 single-pulse TMS pulses were delivered using the MagPro X100 with MagOption(MagVenture, Copenhagen, Denmark). The coil (figure-8 coil, Coil B65; external wing diameter, 90 mm) was placed over F3 (International 10/20 EEG system) to target the left DLPFC. The Inter-Stimulus Interval (ISI) was 3s jittered, and the stimulation intensity was 120% Resting Motor Threshold (RMT). The RMT is determined as the minimum stimulus intensity that produces a MEP exceeding 50 μV in a minimum of five out of ten trials in the relaxed right abductor pollicis brevis.

While receiving TMS, EEG signals of subjects were collected by BrainAmp DC amplifier (Brain Products, Munich, Germany) with a 64-channel EEG system. Participants were asked to remain still and relaxed during the EEG recording. The sampling rate was maintained at 5 kHz, and electrode impedances were maintained below 5 kΩ by applying the conductive gel. FCz was used as the reference while AFz was the ground during the EEG recording. All recordings took place in a temperature-controlled and electrically shielded room. Participants were asked to listen to white noise through earphones in order to mask the loud click accompanied by TMS coil discharge. A foam layer was placed under the coil to inhibit bone conduction and scalp sensation caused by the vibration of the coil (Rogasch et al., 2014).

2.2. TMS-EEG Data Preprocessing

2.2.1. Remove Artifacts

The TMS-EEG data in this study were preprocessed offline with TMS-EEG Signal Analyser (TESA) toolbox (Rogasch et al., 2017). TESA is an open source extension for EEGLAB (Delorme and Makeig, 2004), which is used for cleaning and analyzing TMS-EEG data. Both EEGLAB and TESA ran in Matlab (R2020b).

The data were divided into two-second epochs (−1,000 to 1,000 ms, the time of stimulation was marked as 0 s) and then baseline corrected (−500 to −50 ms). In order to remove the huge electromagnetic artifacts associated with TMS, the large amplitude TMS pulse artifact was removed between −5 and 15 ms and cubic interpolation was used to replace the removed data. For more efficient computing, the sampling rate of data was reduced from 5 to 1 kHz. Epochs and channels contaminated seriously were removed during visual inspection.

The first round of Independent Component Analysis (ICA) was performed to remove large value artifacts including TMS-evoked muscle artifacts and decay artifacts. Subsequently, the data were band-pass (1–80 Hz) and band-stop (48–52 Hz) filtered. It was followed by the second round of ICA to remove other relatively small value artifacts including auditory artifacts, blinks, eye movement, persistent scalp muscle activity, and electrode noise. Finally, the missing channels removed in the preprocessing were interpolated using spherical interpolation and all channels were re-referenced to the common average (Rogasch et al., 2014).

2.2.2. Time-Locked Averaging and GMFP

After data preprocessing, we got clean TMS-EEG trials (1s before stimulation, 1s after stimulation). TEPs were computed by averaging selected artifact-free single trial. The grand average TEPs of the two groups were shown in **Figures 2A,B**.

TMS Evoked Potentials can be recorded from the local electrode to the stimulation site, also from the electrode located in the distant cortical region. The Global Mean Field Power (GMFP) is usually calculated as a measure of global cortical excitability if ROIs are not specified. GMFP is the standard deviation (SD) of all channels at a given sampling point (Esser et al., 2006). The calculation formula was shown in Equation (1). The GMFP curve was shown in **Figure 2C**, which also indicated that the two groups may have differences in the two time windows of N100 and P200 (gray areas).

$$\mathrm{GMFP}(i) = \sqrt{\frac{\sum_{j=1}^{K}\left(V_{\mathrm{j}}(i) - V_{\mathrm{mean}}\right)^2}{K}} \tag{1}$$

where K is the number of all channels ($K = 62$ in this study), $V_j(i)$ is defined as the voltage measured with channel j at sampling point i, and V_{mean} represents the average of the voltages across all channels.

2.2.3. Data Augmentation and Gaussian Smoothing

In order to improve the stability of subsequent machine learning, data augmentation technique was applied. We divided all the trials (the trials remaining after artifacts removing) of each subject into three segments (1–30, 31–60, and 61-end), and then averaged the trials in each segment, which meant that each subject had three segments available for training.

The TEPs obtained by a smaller number of trials are not as smooth as all the trials. In order to make the subsequent feature extraction more reliable, especially the Latency and Amplitude, we used Gaussian window method (the length of Gaussian window is set to 20) to smooth the data after averaging (Gwosdek et al., 2011).

2.2.4. TEPs' Time Windows and ROIs Selection

We found four typical characteristic peaks in **Figures 2A,B** which were consistent with previous research (N40, P60, N100, and P200) (Rosanova et al., 2012; Rogasch et al., 2015). In these typical peaks, N100 and P200 are widely regarded as the two most reliable and reproducible peaks (Kerwin et al., 2018). Considering individual variation led to the advancement or delay of the latency of characteristic peaks, we chose two relatively wide time windows (100–160 ms, 180–280 ms) to include peaks in the window to the full extent.

For the spatial features of TEPs, we selected 7 ROIs according to previous research (Kerwin et al., 2018). The seven ROIs are left Frontal (Fl, F1/F3/FC3/FC5), right Frontal (Fr, F2/F4/FC4/FC6), Central (C, Cz/C1/C2), Centroparietal (Cp, CPz/CP1/CP2/Pz/P1/P2), left Parietal (Pl, CP3/CP5/P3/P5), right Parietal (Pr, CP4/CP6/P4/P6), and Occipital (O, Oz/O1/O2), as shown in **Figure 2D**.

2.3. Temporal-Spatial Features Extract

The average TEPs recorded by all channels in each ROI was calculated as the TEP of this ROI, as shown in Equation (2).

$$X(i) = \frac{1}{k}\sum_{j=1}^{k} V_j(i) \tag{2}$$

where X(i) is TEP in the selected ROI, k is the number of channels in this ROI, $V_j(i)$ is defined as the voltage measured with channel j at the sampling point i.

In order to describe the details of the two peaks (N100 and P200) as much as possible, we calculated a series of linear features in the selected time windows and regions. The features we extracted were introduced below:

- Latency and Amplitude. The Latency and Amplitude are the most common approaches for quantifying TEP (Tremblay et al., 2019). That is, the time and amplitude of the largest peak (negative or positive).
- Local Mean Field Power. The LMFP refers to SD across specific channels in the selected ROI (i.e., electrodes of interest, EOI). We calculated at every sampling point in the given time window and then averaged them. The LMFP reflects the dispersion degree of the signals recorded by the electrodes in this region indicating local excitability changes (Pellicciari et al., 2013), as shown in Equation (3).

$$f_{LMFP} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{\Sigma_{j=1}^{k}\left(V_j(i) - V_{\mathrm{mean}}\right)^2}{k}} \tag{3}$$

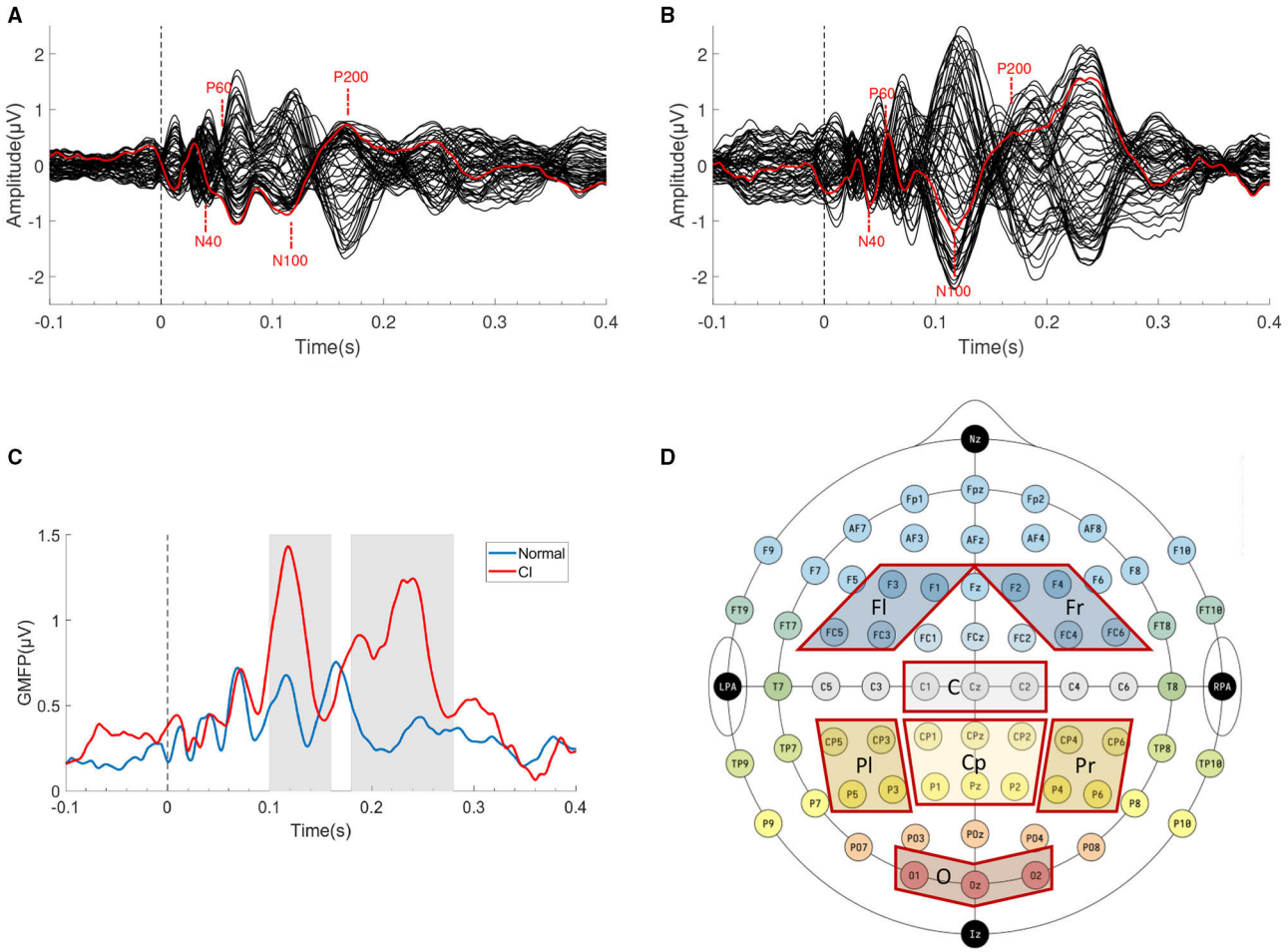


FIGURE 2 | TMS Evoked Potentials (TEPs) components and ROIs selection. The grand average butterfly plot of all channels' TEPs of **(A)** HC group and **(B)** CI group. The electorde (F3) under the TMS coil is indicated in red. **(C)** The GMFP of two groups, the gray areas indicate two time windows of N100 (100–160 ms) and P200 (180–280 ms). **(D)** Schematic diagram of seven ROIs.

where N is the number of all sampling points in the time window.

- Standard Deviation (STD). The STD is the standard deviation of the signal value in the selected time window, reflecting the degree of dispersion of the signal, as shown in Equation (4). This is an important time domain feature in EEG also called activity. The activity is quantified by means of the amplitude variance (Hjorth, 1970).

$$f_{STD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(X(i) - \frac{1}{N}\sum_{i=1}^{N}X(i)\right)^2} \quad (4)$$

- Area Under Curve (AUC). The AUC is the area of the envelope between the signal and the time axis. The upper part of the time axis is positive and the lower part is negative. We used numerical integration to calculate the area, as shown in Equation (5). AUC was also called cortical-evoked activity (CEA) in previous research (Rajji et al., 2013).

$$f_{AUC} = \frac{1}{2f}\sum_{i=1}^{N-1}\left(X(i) + X(i+1)\right) \quad (5)$$

where f is the sampling rate.

- Average Value. The AVG is the average signal value in the selected time window.
- Range. The Range is the difference between the maximum value and the minimum value of the signal in the selected time window, reflecting the fluctuation degree of the signal.

2.4. Machine Learning

In each ROI, we extracted 7 features of N100 and P200 respectively. Finally, 98 (2 time windows*7 ROIs*7 features) features were obtained in each segment of each subject.

Due to the possible correlation of different features, the t-Distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten and Hinton, 2008) was used to reduce the dimension.

After features' dimension reduction, the feature array was normalized to $[-1, 1]$.

In this study, three machine learning algorithms were used. SVM was implemented in the LIBSVM toolbox (Chang and Lin, 2011) with default parameters (linear kernal). Other classifiers [RF($n_{tree} = 7$) and KNN($k = 5$)] were also implemented in Matlab.

Since the features were divided into three segments, three labels that had the same weight were obtained for each participant after the classifier's prediction. We used the voting strategy to fusion three labels. The most pointed category was considered the final label of the subject.

In order to evaluate the performance of the classifiers and to simulate the reality of real CI recognition as much as possible, we adopted a leave-one-out cross-validation (LOOCV) strategy, keeping the minimum subject subset containing all the segment of a subject as the test set and employing all the others for training.

It is necessary to evaluate the classification effect of the model using appropriate indicators. For the binary classification problem, the test set can be divided into: True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). In this study, the subjects with CI were defined as positive samples, HC were defined as negative samples. The several evaluation indicators we used in this section are as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{7}$$

$$Specificity = \frac{TN}{TN + FP} \tag{8}$$

$$F1\text{-}score = \frac{2TP}{2TP + FP + FN} \tag{9}$$

2.5. Statistics

2.5.1. Cluster-Based Permutation Test

Electroencephalogram data has both time and space structure (sampled in multiple channels and multiple time points). Therefore, the difference between CI and HC was the evaluation of a very large number of channel-time pairs, which was a multiple comparisons problem (MCP). For TEPs, we used cluster-based permutation statistics at the whole scalp level to take into account any combination of space and time, while controlling the MCP (Maris and Oostenveld, 2007). We performed an independent *t*-test for the two groups in the selected time windows (100–160 ms, 180–280 ms). If the test statistic value observed in at least two adjacent channels was lower than the threshold value of 0.05, then this value was considered in the cluster arrangement. We performed 5,000 iterations of trial randomization to generate permutation distributions and controlled multiple comparisons across spaces (P <0.025, two-tailed test).

TABLE 2 | Classification results by all classifiers in different components.

Component	Classifier	Accuracy	Sensitivity	Specificity	F1-score
N100	KNN	0.8140	0.7619	0.8636	0.8000
	SVM	**0.8837**	**0.8095**	**0.9545**	**0.8718**
	RF	0.7674	0.7619	0.7727	0.7619
P200	KNN	0.7442	0.7143	0.7727	0.7317
	SVM	0.7907	0.7619	0.8182	0.7805
	RF	0.7442	0.6667	0.8182	0.7179
All components	KNN	0.7907	0.7143	0.8636	0.7692
	SVM	0.8372	0.7619	0.9091	0.8205
	RF	0.7907	0.7619	0.8182	0.7805

The bold values indicate the optimal result under the same index, the same as follow.

3. RESULTS

3.1. Classification Results Based on Different Time Windows

According to the section 2.2.4, we extracted features in different components of TEPs (N100 and P200). Then, we explored the performance of the classifiers by using different components' features. To reduce the dimension of the features matrix, we used t-SNE to reduce the dimension of data.

Table 2 showed the classification results. All components mean merging the features of N100 and P200 components. The best classification performance was achieved by using the features of N100. The highest accuracy of 88.37% was achieved by SVM, with a specificity of 95.45%, the sensitivity of 80.95%, and the F1-score of 87.18%. The features of all components reached slightly weaker but still reasonable classification performance. The classification results of P200 were not satisfactory, the highest accuracy was 79.07%, and other classifiers had even worse results. The sensitivity, specificity, and F1-score were also lower than the N100 in different classifiers.

3.2. Classification Results Based on Different ROIs

Regions of interest were divided according to the section 2.2.4 and we extracted features from each ROI to explore which ROIs are more sensitive to CI. Since there were only 14 features (2 time windows*7 features) in each ROI, the features were directly put into the classifier after normalization without dimensionality reduction.

Table 3 showed the classification results. It could be concluded that different brain regions had great influence on the classification results, and the Fr and Pl region showed the best performance, which achieved 83.72% by SVM. Moreover, when RF was used, the features of Fr were sensitive to the distinction of positive samples, which meant that the patients with CI could be well recognized and the probability of the patients with CI being diagnosed as normal was reduced. The features of the Fl region showed slightly weaker but still reasonable classification performances. The features of the C region can distinguish negative samples well, but it was not good enough to distinguish patients with CI. The features of the Cp region, O

region, and Pr region were basically unable to complete the task of distinguishing normal people from patients with CI.

3.3. Statistical Results

We first performed a cluster-based permutation test on the TEPs of the two groups. The results of statistical analysis were shown in **Figure 3**. The topographic map was generated with Yang's topoplot_bcl function based on EEGLAB's topoplot function (Li et al., 2018). The asterisk indicates that the *p*-value is less than 0.01. In the N100 time window (100–160 ms), right frontal region, left parietal region, and occipital region were significantly different between the two groups ($p < 0.01$). In the P200 time window (180–280 ms), bilateral frontal region, bilateral parietal region, and occipital region were significantly different between the two groups ($p < 0.01$).

Furthermore, we used the violin plot to describe the distribution of 14 features of the right frontal region, as shown in **Figure 4**. We also performed *t*-test on the features in **Figure 4**, the results were shown in the **Table 4**. All the features of N100 in the right frontal region were significantly different between the two groups (p <0.01). There was no difference in the latency of P200 in the right frontal region between the two groups. This may explain why the features of N100 performed better than P200 on classification, and even better than using them simultaneously.

Then, we used the t-SNE to reduce the dimension of the best performing N100 features to 3, and then draw them in three-dimensional space. The result was shown in **Figure 5A**, red represented CI, and blue represented HC. It revealed that the points in CI were more dispersed than normal.

Finally, for the 7 features of N100 in the right frontal region, we used XGBoost to evaluate the importance of features (Chen and Guestrin, 2016). The feature importance ranking was listed in **Figure 5B**. The LMFP, AVG, Latency, and Amplitude were the most important, which exactly were consistent with the most widely used features to measure TEPs in previous studies (Tremblay et al., 2019). Our results verified the conclusions of the predecessors and also instructed doctors to pay more attention to these features in future clinical practice.

TABLE 3 | Classification results by all classifiers in different regions.

Region	Classifier	Accuracy	Sensitivity	Specificity	F1-score
	KNN	0.7209	0.5714	0.8636	0.6667
C	SVM	0.7209	0.6190	0.8182	0.6842
	RF	0.7674	0.7143	0.8182	0.7500
	KNN	0.6977	0.5714	0.8182	0.6486
Cp	SVM	0.5814	0.4762	0.6818	0.5263
	RF	0.7674	0.7619	0.7727	0.7619
	KNN	0.7674	0.6667	0.8636	0.7368
Fl	SVM	0.7674	0.6667	0.8636	0.7368
	RF	0.8140	**0.8095**	0.8182	0.8095
	KNN	0.8140	0.6667	**0.9545**	0.7778
Fr	SVM	**0.8372**	0.7143	**0.9545**	**0.8108**
	RF	0.8140	**0.8095**	0.8182	0.8095
	KNN	0.6744	0.6190	0.7273	0.6500
O	SVM	0.7209	0.6667	0.7727	0.7000
	RF	0.7442	**0.8095**	0.6818	0.7556
	KNN	**0.8372**	0.7143	**0.9545**	**0.8108**
Pl	SVM	**0.8372**	0.7143	**0.9545**	**0.8108**
	RF	0.7674	0.7143	0.8182	0.7500
	KNN	0.5814	0.3810	0.7727	0.4706
Pr	SVM	0.5349	0.3810	0.6818	0.4444
	RF	0.7209	0.6667	0.7727	0.7000

The bold values indicate the optimal result under the same index, the same as follow.

4. DISCUSSION

The study showed that machine learning can effectively identify CI. In the time domain, the features of N100 had the best performance. In the aspect of spatiality, the features of the right frontal region and left parietal region had the best performance. Then, we discussed our results from these two dimensions, and particularly emphasized the influence of sensory contamination within the TEP. Finally, we discussed the limitations of our study.

Previous studies have used TMS to stimulate the motor cortex and have demonstrated that patients with AD have increased

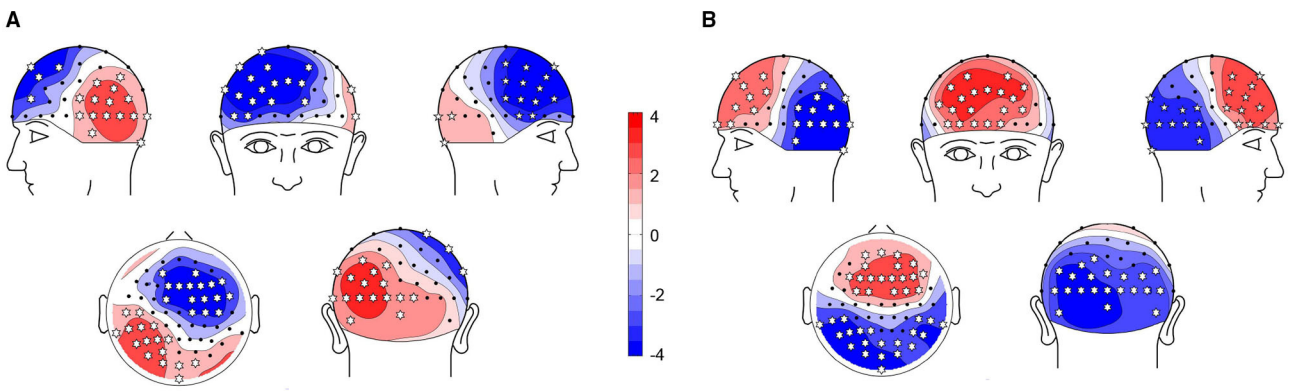


FIGURE 3 | Comparison of TEPs using cluster-based permutation tests. Red means TEP of cognitive impairment (CI) is higher than HC, blue means TEP of CI is lower than healthy controls (HC). The asterisk indicates that $p < 0.01$. **(A)** N100: CI vs. HC. **(B)** P200: CI vs. HC.

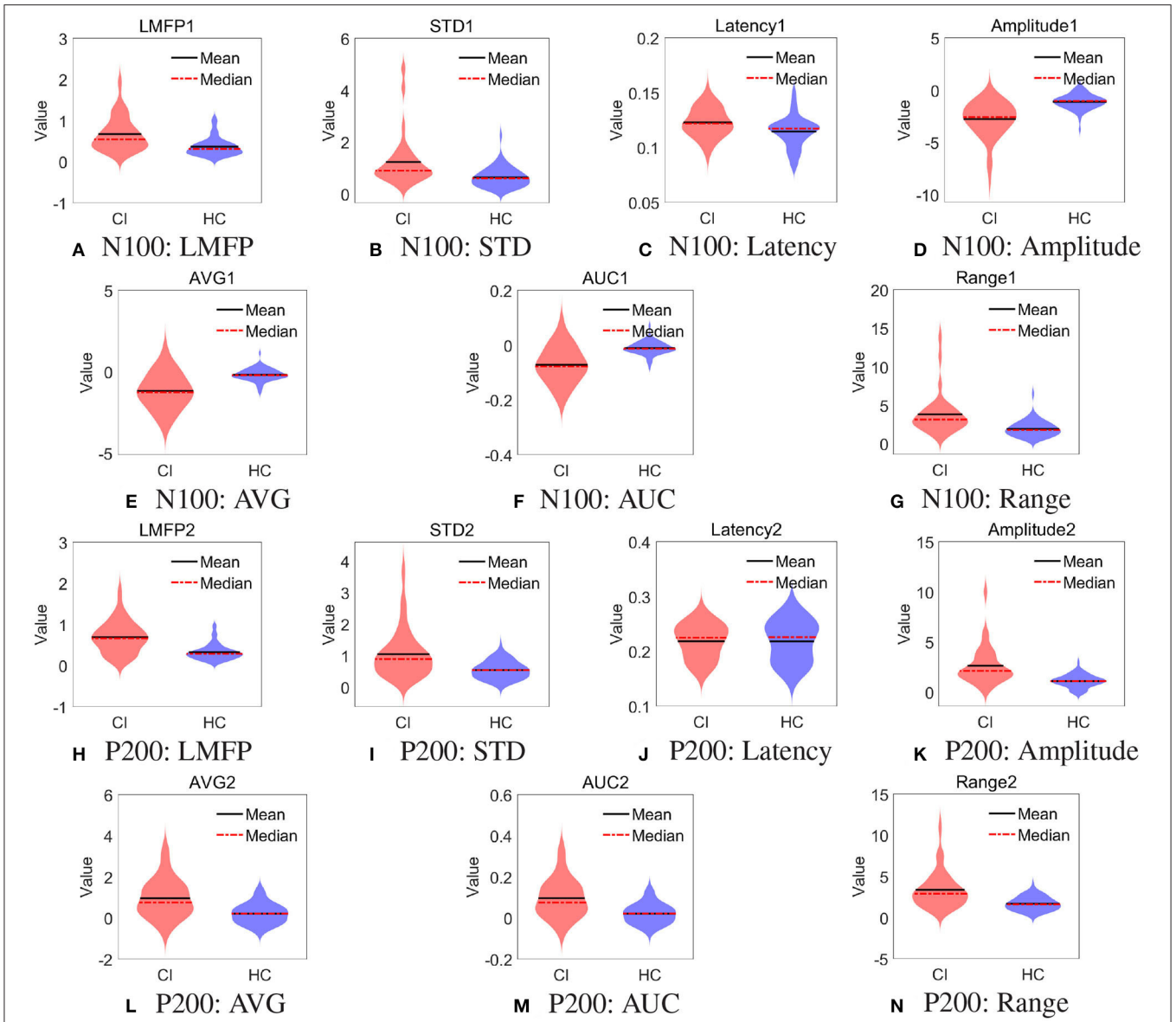


FIGURE 4 | Distribution of spatial features in right frontal region. **(A)** N100: LMFP, **(B)** N100: STD, **(C)** N100: Latency, **(D)** N100: Amplitude, **(E)** N100: AVG, **(F)** N100: AUC, **(G)** N100: Range, **(H)** P200: LMFP, **(I)** P200: STD, **(J)** P200: Latency, **(K)** P200: Amplitude, **(L)** P200: AVG, **(M)** P200: AUC, and **(N)** P200: Range.

TABLE 4 | Local features in right frontal region.

	N100			P200		
	CI	**HC**	***P*-value**	**CI**	**HC**	***P*-value**
LMFP	0.67 ± 0.40	0.37 ± 0.22	$p < 0.001^*$	0.70 ± 0.38	0.33 ± 0.19	$p < 0.001^*$
STD	1.23 ± 0.98	0.64 ± 0.38	$p < 0.001^*$	1.05 ± 0.76	0.54 ± 0.27	$p < 0.001^*$
Latency	0.12 ± 0.013	0.11 ± 0.014	0.001^*	0.22 ± 0.031	0.22 ± 0.036	0.961
Amplitude	−2.77 ± 1.68	−1.10 ± 0.72	$p < 0.001^*$	2.64 ± 1.92	1.11 ± 0.64	$p < 0.001^*$
AVG	−1.17 ± 1.10	−0.17 ± 0.39	$p < 0.001^*$	0.98 ± 0.93	0.22 ± 0.48	$p < 0.001^*$
AUC	−0.071 ± 0.067	−0.011 ± 0.023	$p < 0.001^*$	0.098 ± 0.094	0.022 ± 0.048	$p < 0.001^*$
Range	3.87 ± 2.71	1.97 ± 1.02	$p < 0.001^*$	3.38 ± 2.19	1.71 ± 0.81	$p < 0.001^*$

** means a significant difference with p = 0.01.*

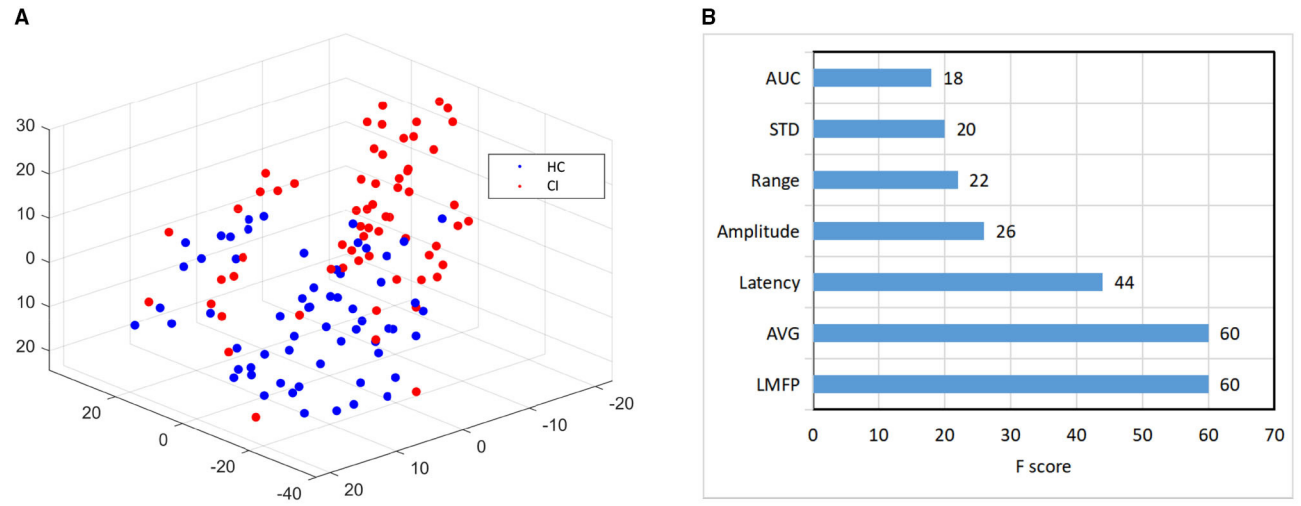


FIGURE 5 | Feature visualization and importance comparison. **(A)** The visualization features map based on t-Distributed Stochastic Neighbor Embedding (t-SNE) dimension-reduction. **(B)** The feature importance based on XGBoost.

cortical excitability, which is manifested by a decrease in RMT and an increase in MEP (Alagona et al., 2001; Di Lazzaro et al., 2004). There is also evidence of high excitability in patients with AD after TMS stimulating the left DLPFC, which is manifested as an increase in CEA, and this increase is negatively correlated with overall cognitive and executive functions (Joseph et al., 2021). In our study, the CI group had higher GMFP, suggesting higher cortical excitability which is consistent with previous studies. This hyper-excitability may reflect the plastic reorganization of the sensorimotor system and may be used as a compensatory mechanism to offset the loss of cortical volume and protect the motor function of the patient (Niskanen et al., 2011; Bagattini et al., 2019).

It was found in a previous study that the reliability of the potentials induced by TMS in the left DLPFC are highly consistent and the measurement error is small. The most reliable peaks are generally located at 100 and 200 ms (Kerwin et al., 2018). High reproducibility is necessary as a potential neurobiomarker. Although there was some prior knowledge that CI may affect P30 (Bagattini et al., 2019), considering the first peak may be affected by preprocessing (–5–15 ms removed and cubic interpolation), we still focused on the N100 and P200 two components to preserve the temporal features.

In the time domain, the features of N100 have the best performance according to the four indicators in the classifier. In the CI group, the change of N100 and P200 may be related to the alterations of GABAergic activity. Gamma-aminobutyric acid (GABA) is an inhibitory neurotransmitter. Its natural function is binding to GABA-A receptors and GABA-B receptors on the neurons to modulate and block impulses between nerve cells (Gou et al., 2013). There is some evidence that the amplitude of the TEPs component is related to GABAergic activity. GABA-A receptors agonists (alprazolam and diazepam) and GABA-B receptors agonists (baclofen) can modulate the amplitude of N100 or P200 (Premoli et al., 2014; Murphy et al., 2016). There are also some studies stated that the alterations of GABAergic circuits may contribute to CI by disrupting the overall network function (Li et al., 2016). This may be the neuropathological basis for the difference of N100 and P200 between the two groups, which was the precondition for classification.

In the aspect of spatiality, the features of the right frontal region and left parietal region had the best performance in the SVM classifier. From the statistical results, the right prefrontal region and the left parietal region both had significant differences in these two selected time windows, which may explain why the features of these two ROIs perform better in the classifiers. Interestingly, the significant difference between CI and HC is reflected in the anterior and posterior regions, and the trend is opposite over time. This seems to indicate the abnormal changes in the connectivity of the brains of patients with CI. Previous studies have shown that the damaged excitatory-inhibitory balance between anterior and posterior regions might represent a maladaptive pathogenic mechanism (Bagattini et al., 2019).

In addition, we want to emphasize the issue of sensory contamination within the TEP. There is no doubt that the auditory complex can overlap with N100 and P200 (Conde et al., 2019), but in fact we have used strict online and offline methods to avoid the impact of auditory and somatosensory evoked potential as much as possible. In the data collection process, we played white noise by earphones for the subjects. In addition, a foam layer was placed under the coil to inhibit bone conduction and weaken the scalp sensation caused by the vibration of the coil. During data preprocessing, we paid attention to the removal of auditory artifacts in the second ICA run which were characterized by a topography centering over Cz to Fz and a time course with bipolar peaks at approximately 100–200 ms (Rogasch et al., 2014). From another perspective, all subjects received

the same protocol and all the EEG were produced in the same environment and preprocessed by the same method. Therefore, even if the online and offline methods cannot completely remove the auditory and somatosensory evoked potentials, the remaining potentials have the same effect on both two groups, which does not contribute to the difference between groups, let alone the impact on classification. In summary, there is no reason to think that the difference in the selected time windows is related to sensory contamination.

In this study, we observed that all classification results had high specificity but unsatisfied sensitivity, meaning that CI subjects were not well distinguished from HC. It is due to the heterogeneity of cognitive related disease, time of illness, and disease progression may lead to more scattered features. There was a limitation that all patients with CI enrolled were highly suspected of AD with mild to moderate symptoms based on clinical diagnosis and MoCA, excluded vascular and other types of dementia, but not on biological markers of CSF or PET. The definition of AD is controversial throughout articles (Jellinger, 2020; Milà-Alomà et al., 2021). The focus of this study was the extraction method of TEPs' features and machine learning rather than strict AD diagnostic criteria. To be conservative and rigorous, although all subjects in the CI group were highly suspected of AD, we did not define them as AD but summarized it with CI in this study.

A further limitation is that the sample size in this study was small in the field of machine learning. Although we have used data augmentation and voting strategy to obtain a more robust model, more data will still be needed in subsequent studies to meet the real and complex clinical needs.

5. CONCLUSION

We found that the TEPs induced by TMS over the left DLPFC has significant differences between CI and HC. Machine learning based on the spatiotemporal features of TEPs is effective for the classification of CI and HC.

In the time domain, the features of N100 had the best performance in the SVM classifier, with an accuracy of 88.37%. In the aspect of spatiality, the features of the right frontal region and left parietal region had the best performance in the SVM classifier, with an accuracy of 83.72%. By using XGBoost to evaluate the importance of features, the LMFP, AVG, Latency, and Amplitude contributed the most in classification. It is suggested that clinicians should pay close attention to the important features above, which may be potential biomarkers for diagnosing CI.

In this study, the features selected were all simple and linear, the classification algorithms used were popular and sophisticated. Therefore, our research particularly emphasized the interpretability and clinical usability. These findings prove that machine learning based on spatiotemporal features of TEP has the potential to automatically clinical auxiliary diagnosis of CI.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Shenzhen People's Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JZ: writing—original draft preparation, TMS-EEG data artifacts removing, features extraction, and statistical analysis. HL: writing—original draft preparation, machine learning, and feature analysis. LZ: writing—reviewing, conceptualization, and methodology. HR: data collection. XSu and XL: TMS equipment operation and EEG acquisition. GD, XZ, XJ, JF, and XSh: investigation, writing—reviewing, and editing. TW and XH: writing—reviewing and editing, investigation, methodology, and supervision. YG: writing—reviewing and editing, clinical diagnosis, investigation, and methodology. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by Shenzhen Science and Technology Innovation Commission (KCXFZ20201221173400001 and KCXFZ20201221173411032), Natural Science Fund of Guangdong Province (2021A1515010983). Shenzhen Key Medical Discipline Construction Fund (No. SZXK005), and Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation (Climbing Program Special Funds, pdjh2021c0018).

REFERENCES

Ahmed, M. A., Darwish, E. S., Khedr, E. M., Ali, El Serogy, Y. M., and Ali, A. M. (2012). Effects of low versus high frequencies of repetitive transcranial magnetic stimulation on cognitive function and cortical excitability in alzheimer's dementia. *J. Neurol.* 259, 83–92. doi: 10.1007/s00415-011-6128-4

Alagona, G., Bella, R., Ferri, R., Carnemolla, A., Pappalardo, A., Costanzo, E., et al. (2001). Transcranial magnetic stimulation in alzheimer disease: motor cortex excitability and cognitive severity. *Neurosci. Lett.* 314, 57–60. doi: 10.1016/s0304-3940(01)02288-1

Bagattini, C., Mutanen, T. P., Fracassi, C., Manenti, R., Cotelli, M., Ilmoniemi, R. J., et al. (2019). Predicting alzheimer's disease severity by means of tms–eeg coregistration. *Neurobiol. Aging* 80, 38–45. doi: 10.1016/j.neurobiolaging.2019.04.008

Benussi, A., Grassi, M., Palluzzi, F., Cantoni, V., Cotelli, M. S., Premi, E., et al. (2021). Classification accuracy of tms for the diagnosis of mild cognitive impairment. *Brain Stimul.* 14, 241–249. doi: 10.1016/j.brs.2021.01.004

Buchhave, P., Minthon, L., Zetterberg, H., Wallin, Å. K., Blennow, K., Hansson, O., et al. (2012). Cerebrospinal fluid levels ofβ-amyloid 1-42, but not of tau, are fully changed already 5 to 10 years before the onset of alzheimer dementia. *Arch. Gen. Psychiatry* 69, 98–106. doi: 10.1001/archgenpsychiatry.2011.155

Cao, K.-X., Ma, M.-L., Wang, C.-Z., Iqbal, J., Si, J.-J., Xue, Y.-X., et al. (2021). Tms-eeg: an emerging tool to study the neurophysiologic biomarkers of psychiatric disorders. *Neuropharmacology* 197:108574. doi: 10.1016/j.neuropharm.2021.108574

Carlén, M. (2017). What constitutes the prefrontal cortex? *Science* 358, 478–482. doi: 10.1126/science.aan8868

Chang, C., and Lin, C. (2011). Libsvm: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–39. doi: 10.1145/1961189.1961199

Chen, T., and Guestrin, C. (2016). "Xgboost: a scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD '16, (Association for Computing Machinery)* (New York, NY), 785–794.

Ciesielska, N., Sokołowski, R., Mazur, E., Podhorecka, M., Polak-Szabela, A., Kdziora-Kornatowska, K., et al. (2016). Is the montreal cognitive assessment (moca) test better suited than the mini-mental state examination (mmse) in mild cognitive impairment (mci) detection among people aged over 60? meta-analysis. *Psychiatr Pol.* 50, 1039–1052. doi: 10.12740/PP/45368

Conde, V., Tomasevic, L., Akopian, I., Stanek, K., Saturnino, G. B., Thielscher, A., et al. (2019). The non-transcranial tms-evoked potential is an inherent source of ambiguity in tms-eeg studies. *Neuroimage* 185, 300–312. doi: 10.1016/j.neuroimage.2018.10.052

Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* 13, 21–27. doi: 10.1109/TIT.1967.1053964

Cullen, N. C., Leuzy, A., Palmqvist, S., Janelidze, S., Stomrud, E., Pesini, P., et al. (2021). Individualized prognosis of cognitive decline and dementia in mild cognitive impairment based on plasma biomarker combinations. *Nat. Aging* 1, 114–123. doi: 10.1038/s43587-020-00003-5

Delorme, A., and Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Di Lazzaro, V., Oliviero, A., Pilato, F., Saturno, E., Dileone, M., Marra, C., et al. (2004). Motor cortex hyperexcitability to transcranial magnetic stimulation in alzheimer's disease. *J. Neurol. Neurosurgery Psychiatry* 75, 555–559. doi: 10.1136/jnnp.2003.018127

Esser, S., Huber, R., Massimini, M., Peterson, M., Ferrarelli, F., Tononi, G., et al. (2006). A direct demonstration of cortical ltp in humans: a combined tms/eeg study. *Brain Res. Bull.* 69, 86–94. doi: 10.1016/j.brainresbull.2005.11.003

Frisoni, G. B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., et al. (2017). Strategic roadmap for an early diagnosis of alzheimer's disease based on biomarkers. *Lancet Neurol.* 16, 661–676. doi: 10.1016/S1474-4422(17)30159-X

Gou, Z., Wang, X., and Wang, W. (2013). Evolution of neurotransmitter gamma-aminobutyric acid, glutamate and their receptors. *Dongwuxue Yanjiu* 33, E75–E81. doi: 10.3724/SP.J.1141.2012.E05-06E75

Gwosdek, P., Grewenig, S., Bruhn, A., and Weickert, J. (2011). "Theoretical Foundations of Gaussian Convolution by Extended Box Filtering," in *Scale Space and Variational Methods in Computer Vision. SSVM 2011. Lecture Notes in Computer Science, vol 6667,* eds A. M. Bruckstein, B. M. ter Haar Romeny, A. M. Bronstein, M. M. Bronstein (Berlin: Springer), 447–458.

Hallett, M. (2007). Transcranial magnetic stimulation: a primer. *Neuron* 55, 187–199. doi: 10.1016/j.neuron.2007.06.026

Hansson, O., Seibyl, J., Stomrud, E., Zetterberg, H., Trojanowski, J. Q., Bittner, T., et al. (2018). Csf biomarkers of alzheimer's disease concord with amyloid-β pet and predict clinical progression: a study of fully automated immunoassays in biofinder and adni cohorts. *Alzheimers Dementia* 14, 1470–1481. doi: 10.1016/j.jalz.2018.01.010

Hjorth, B. (1970). Eeg analysis based on time domain properties. *Electroencephalogr. Clin. Neurophysiol.* 29, 306–310. doi: 10.1016/0013-4694(70)90143-4

Jellinger, K. (2020). Towards a biological definition of alzheimer disease. *Int. J. Neurol Neurother.* 7:95. doi: 10.23937/2378-3001/1410095

Jeong, J. (2004). Eeg dynamics in patients with alzheimer's disease. *Clin. Neurophysiol.* 115, 1490–1505. doi: 10.1016/j.clinph.2004.01.001

Joseph, S., Knezevic, D., Zomorrodi, R., Blumberger, D. M., Daskalakis, Z. J., Mulsant, B. H., et al. (2021). Dorsolateral prefrontal cortex excitability abnormalities in alzheimer's dementia: findings from transcranial magnetic stimulation and electroencephalography study. *Int. J. Psychophysiol.* 169, 55–62. doi: 10.1016/j.ijpsycho.2021.08.008

Kerwin, L. J., Keller, C. J., Wu, W., Narayan, M., and Etkin, A. (2018). Test-retest reliability of transcranial magnetic stimulation eeg evoked potentials. *Brain stimul.* 11, 536–544. doi: 10.1016/j.brs.2017.12.010

Kimiskidis, V. (2016). Transcranial magnetic stimulation (tms) coupled with electroencephalography (eeg): biomarker of the future. *Rev. Neurol.* 172, 123–126. doi: 10.1016/j.neurol.2015.11.004

Komssi, S., Aronen, H. J., Huttunen, J., Kesäniemi, M., Soinne, L., Nikouline, V. V., et al. (2002). Ipsi- and contralateral eeg reactions to transcranial magnetic stimulation. *Clin. Neurophysiol.* 113, 175–184. doi: 10.1016/s1388-2457(01)00721-0

Leuzy, A., Chiotis, K., Lemoine, L., Gillberg, P. G., Almkvist, O., Rodriguez-Vieitez, E., et al. (2019). Tau pet imaging in neurodegenerative tauopathies-still a challenge. *Mol. Psychiatry* 24, 1112–1134. doi: 10.1038/s41380-018-0342-8

Li, A.-S., Miao, C.-G., Han, Y., He, X., and Zhang, Y. (2018). Electrophysiological correlates of the effect of task difficulty on inhibition of return. *Front. Psychol.* 9:2403. doi: 10.3389/fpsyg.2018.02403

Li, Y., Sun, H., Chen, Z., Xu, H., Bu, G., and Zheng, H. (2016). Implications of gabaergic neurotransmission in alzheimer's disease. *Front. Aging Neurosci.* 8:31. doi: 10.3389/fnagi.2016.00031

Liaw, A., and Wiener, M. (2002). Classification and regression by randomforest. *R News* 23. Available online at: https://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of eeg-and meg-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024

Milà-Alomà, M., Salvadó, G., Shekari, M., Grau-Rivera, O., Sala-Vila, A., Sánchez-Benavides, G., et al. (2021). Comparative analysis of different definitions of amyloid-β positivity to detect early downstream pathophysiological alterations in preclinical alzheimer. *J. Prevent. Alzheimers Dis.* 8, 68–77. doi: 10.14283/jpad.2020.51

Montine, T. J., Bukhari, S. A., and White, L. R. (2021). Cognitive impairment in older adults and therapeutic strategies. *Pharmacol. Rev.* 73, 152–162. doi: 10.1124/pharmrev.120.000031

Murphy, S. C., Palmer, L. M., Nyffeler, T., Müri, R. M., and Larkum, M. E. (2016). Transcranial magnetic stimulation (tms) inhibits cortical dendrites. *Elife* 5:e13598. doi: 10.7554/eLife.13598

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x

Ng, T. K. S., Ho, C. S. H., Tam, W. W. S., Kua, E. H., and Ho, R. C.-M. (2019). Decreased serum brain-derived neurotrophic factor (bdnf) levels in patients with alzheimer's disease (ad): a systematic review and meta-analysis. *Int. J. Mol. Sci.* 20:257. doi: 10.3390/ijms20020257

Nikulin, V. V., Kičić, D., Kähkönen, S., and Ilmoniemi, R. J. (2003). Modulation of electroencephalographic responses to transcranial magnetic stimulation: evidence for changes in cortical excitability related to movement. *Eur. J. Neurosci.* 18, 1206–1212. doi: 10.1046/j.1460-9568.2003.02858.x

Niskanen, E., Könönen, M., Määttä, S., Hallikainen, M., Kivipelto, M., Casarotto, S., et al. (2011). New insights into alzheimer's disease progression: a combined tms and structural mri study. *PLoS ONE* 6:e26113. doi: 10.1371/journal.pone.0026113

Pellicciari, M. C., Brignani, D., and Miniussi, C. (2013). Excitability modulation of the motor system induced by transcranial direct current stimulation: a multimodal approach. *Neuroimage* 83, 569–580. doi: 10.1016/j.neuroimage.2013.06.076

Premoli, I., Castellanos, N., Rivolta, D., Belardinelli, P., Bajo, R., Zipser, C., et al. (2014). Tms-eeg signatures of gabaergic neurotransmission in the human cortex. *J. Neurosci.* 34, 5603–5612. doi: 10.1523/JNEUROSCI.5089-13.2014

Rabinovici, G. D., Gatsonis, C., Apgar, C., Chaudhary, K., Gareen, I., Hanna, L., et al. (2019). Association of amyloid positron emission tomography with subsequent change in clinical management among medicare beneficiaries with mild cognitive impairment or dementia. *JAMA* 321, 1286–1294. doi: 10.1001/jama.2019.2000

Rajji, T. K., Sun, Y., Zomorrodi-Moghaddam, R., Farzan, F., Blumberger, D. M., Mulsant, B. H., et al. (2013). Pas-induced potentiation of cortical-evoked activity in the dorsolateral prefrontal cortex. *Neuropsychopharmacology* 38, 2545–2552. doi: 10.1038/npp.2013.161

Rogasch, N. C., Daskalakis, Z. J., and Fitzgerald, P. B. (2015). Cortical inhibition of distinct mechanisms in the dorsolateral prefrontal cortex is related to working memory performance: A tms–eeg study. *Cortex* 64, 68–77. doi: 10.1016/j.cortex.2014.10.003

Rogasch, N. C., Sullivan, C., Thomson, R. H., Rose, N. S., Bailey, N. W., Fitzgerald, P. B., et al. (2017). Analysing concurrent transcranial magnetic stimulation and electroencephalographic data: a review and introduction to the open-source tesa software. *Neuroimage* 147, 934–951. doi: 10.1016/j.neuroimage.2016.10.031

Rogasch, N. C., Thomson, R. H., Farzan, F., Fitzgibbon, B. M., Bailey, N. W., Hernandez-Pavon, J. C., et al. (2014). Removing artefacts from tms-eeg recordings using independent component analysis: importance for assessing prefrontal and motor cortex network properties. *Neuroimage* 101, 425–39. doi: 10.1016/j.neuroimage.2014.07.037

Rosanova, M., Casarotto, S., Pigorini, A., Canali, P., Casali, A., and Massimini, M. (2012). Combining transcranial magnetic stimulation with electroencephalography to study human cortical excitability and effective connectivity. *Neuromethods* 67, 435–457. doi: 10.1007/7657_2011_15

Tremblay, S., Rogasch, N. C., Premoli, I., Blumberger, D. M., Casarotto, S., Chen, R., et al. (2019). Clinical utility and prospective of tms–eeg. *Clin. Neurophysiol.* 130, 802–844. doi: 10.1016/j.clinph.2019.01.001

van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-sne. *J. Mach. Learn. Res.* 9, 2579–2605.

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Trans. Neural Netw.* 10, 988–999. doi: 10.1109/72.788640

Ye, Q., and Bai, F. (2018). Contribution of diffusion, perfusion and functional mri to the disconnection hypothesis in subcortical vascular cognitive impairment. *Stroke Vascular Neurol.* 3, 131–139. doi: 10.1136/svn-2017-000080

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

OPINION
published: 11 January 2022
doi: 10.3389/fnagi.2021.827273

The Effective Cognitive Assessment and Training Methods for COVID-19 Patients With Cognitive Impairment

Dong Wen[1], Jian Xu[2], Zhonglin Wu[2], Yijun Liu[3], Yanhong Zhou[4], Jingjing Li[2], Shaochang Wang[2], Xianlin Dong[5], M. Iqbal Saripan[6] and Haiqing Song[7]**

[1] Brain Computer Intelligence and Intelligent Health Institution, Institute of Artificial Intelligence, University of Science and Technology Beijing, Beijing, China, [2] The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, School of Information Science and Engineering, Yanshan University, Qinhuangdao, China, [3] Department of Statistics, School of Science, Yanshan University, Qinhuangdao, China, [4] Department of Computer Science and Technology, School of Mathematics and Information Science and Technology, Hebei Normal University of Science and Technology, Qinhuangdao, China, [5] Department of Biomedical Engineering, Chengde Medical University, Chengde, China, [6] Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Malaysia, [7] Department of Neurology, Xuanwu Hospital of Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:
Peng Xu,
University of Electronic Science and Technology of China, China

Reviewed by:
Chao Chen,
Tianjin University of Technology, China
Yuliang Ma,
Hangzhou Dianzi University, China

***Correspondence:**
Yanhong Zhou
yhzhou168@163.com
Haiqing Song
songhq@xwhosp.org

Specialty section:
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

Received: *01 December 2021*
Accepted: *16 December 2021*
Published: *11 January 2022*

Citation:
Wen D, Xu J, Wu Z, Liu Y, Zhou Y, Li J, Wang S, Dong X, Saripan MI and Song H (2022) The Effective Cognitive Assessment and Training Methods for COVID-19 Patients With Cognitive Impairment. Front. Aging Neurosci. 13:827273. doi: 10.3389/fnagi.2021.827273

Keywords: cognitive assessment and training, COVID-19 patients, cognitive impairment, brain-computer interface, virtual reality

INTRODUCTION

Since late 2019, COVID-19 has been raging worldwide. Related studies have reported that many COVID-19 patients present cognitive sequelae (Ahmad and Rathore, 2020; Baschi et al., 2020; Gunasekaran et al., 2020; Heneka et al., 2020; Kieron et al., 2020; Koralnik and Tyler, 2020; Pinna et al., 2020; Vespignani et al., 2020; Woo et al., 2020; Taquet et al., 2021). However, the implementation of isolation measures greatly limits the traditional cognitive impairment assessment and treatment methods (Lara et al., 2020). Therefore, we need to explore better ways to assess and train cognitive impairment in patients with COVID-19.

The rapid development of a brain-computer interface (BCI), virtual reality (VR), and artificial intelligence has promoted the diagnosis and treatment of cognitive impairment in the direction of intellectual development. Studies have shown that BCI-VR technology can compensate for the limitations of BCI alone and provide new rehabilitation and assessment methods for patients with cognitive impairment, which has attracted increasing attention (Wen et al., 2018, 2020; Bauer and Andringa, 2020; Mancuso et al., 2020; Pinter et al., 2021).

This paper first discussed current cognitive impairment assessment and rehabilitation methods for patients with cognitive impairment. On this basis, we proposed a new comprehensive strategy for cognitive impairment assessment and rehabilitation based on BCI-VR. It will be a new approach for cognitive rehabilitation that fully integrates the accurate detection of BCI technology and the effective training of VR technology. Further, we discussed the advantages and challenges of BCI-VR in cognitive impairment assessment and rehabilitation training for COVID-19 patients in this current pandemic.

CURRENT COGNITIVE IMPAIRMENT OF COVID-19 PATIENTS

Since the start of the COVID-19 pandemic, a growing number of studies have reported neurological impairment in COVID-19 patients (Ahmad and Rathore, 2020; Gunasekaran et al., 2020; Heneka et al., 2020; Kieron et al., 2020; Koralnik and Tyler, 2020; Pinna et al., 2020; Vespignani et al., 2020; Taquet et al., 2021). We found that stroke and cognitive impairment are the most common

manifestations of neurological impairment in COVID-19 patients (Belani et al., 2020; Fara et al., 2020; Fatima et al., 2020; Haji Akhoundi et al., 2020; Heneka et al., 2020; Jain et al., 2020; Mahboob et al., 2020; Merkler et al., 2020; Rajdev et al., 2020; Sezgin et al., 2020; Taquet et al., 2021). Studies showed that the incidence of stroke in confirmed hospitalized COVID-19 patients ranged from 2.8 to 5.4%, the incidence of cognitive impairment was 26%, and patients with stroke often cause concurrent manifestations of cognitive impairment (Haji Akhoundi et al., 2020; Majidi et al., 2020; Oxley et al., 2020; Sun et al., 2020; Taquet et al., 2021). Other studies showed that patients with neurological impairment caused by COVID-19 tended to be younger (Ahmad and Rathore, 2020; Cavallieri et al., 2020; Oxley et al., 2020; Woo et al., 2020; Harrison et al., 2021).

In the current pandemic, both hospitals and patients are facing huge and severe challenges in the assessment and rehabilitation training of patients with cognitive impairment (Coetzer, 2020; Richardson et al., 2020). Studies showed, as the COVID-19 pandemic intensified, patients with cognitive impairment were limited in going out for rehabilitation training, which had many negative effects on the mental state of patients and his mental condition was deteriorating (Lara et al., 2020; Manca et al., 2020; Devita et al., 2021). Therefore, we need to explore better ways to assess and train cognitive impairment in patients with COVID-19.

CURRENT EVALUATION METHODS OF COGNITIVE IMPAIRMENT

Currently, common cognitive impairment methods include cognitive scale, neuroimaging technology and new wearable devices. In clinical practice, the most classic method of objective evaluation is the cognitive scale. The cognitive scale test achieves the evaluation effect by comparing the patient's test results with the scale indicators, include: Mini-Mental State Examination (MMSE) (Pangman et al., 2000), Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2019) and Activities of Daily Living (ADL) (Lopez Mongil, 2017), Auditory Verbal Learning Test (AVLT) (Stricker et al., 2021), Trail Marking Test (TMT) (Lunardini et al., 2020), etc. Recently, Burns et al. proposed a new hybrid scale—Free-COG, which could also be used to assess subjects' cognitive and executive functions (Burns et al., 2021). However, subjective factors of the testers reduced the accuracy of the results in the cognitive scale test.

The degree of cognitive impairment of patients is evaluated by observing the changes of brain structure through imaging (Knopman and Petersen, 2014). Among them, the commonly used neuroimaging techniques include: structural neuroimaging techniques (Zhang et al., 2019), functional neuroimaging techniques, positron emission tomography (PET), molecular imaging, and functional magnetic resonance imaging (fMRI) (Zhang et al., 2019; Xu et al., 2020). However, neuroimaging technology equipment is larger, and detection costs are higher, which significantly limits its application (Narayanan and Murray, 2016).

Relevant studies have introduced wearable devices into the assessment of cognitive impairment (Narayanan and Murray, 2016). Related study showed that wearable biosensor devices might be a viable tool to assess physiological changes in patients with AD, enabling remote and continuous monitoring of neurocognitive function in patients (Saif et al., 2019; Stavropoulos et al., 2020; Eggenberger et al., 2021). However, the evaluation indicators of the new wearable devices are uncertain, and there is no unified standard in use.

CURRENT REHABILITATION METHODS OF COGNITIVE IMPAIRMENT

Currently, commonly used cognitive rehabilitation methods include medication-assisted, cognitive rehabilitation training, and home rehabilitation.

Medication-assisted therapy can inhibit the induction of cognitive impairment or other diseases (Rejdak and Grieb, 2020; Zhaojun and Miaser, 2020). The implementation of medication-assisted therapy is costly, and it also only serves as an adjustment role in the rehabilitation of cognitive impairment and may be accompanied by other side effects (Jin-xuan et al., 2020).

General cognitive rehabilitation training usually refers to systematic and targeted training depending on the patient's cognitive function under face-to-face guidance by the rehabilitation therapist. Studies showed it can improve or maintain patients' cognitive abilities related to daily task performance, so as to prevent or delay cognitive decline (Irazoki et al., 2020). But it requires the participation of both the therapist and the patient. And there are many limitations in the rehabilitation plan, such as time, personnel, and cost.

During the COVID-19 pandemic, traditional rehabilitation training is limited, researchers have suggested remote home rehabilitation for patients with cognitive impairment (Chang and Boudier-Revéret, 2020) and adopting some remote home rehabilitation measures (Richardson et al., 2020). Through literature analysis and comparison, it is found that home rehabilitation provides great convenience for both the therapists and patients, which can meet the needs of patients with cognitive impairment (Natta et al., 2015; Gately et al., 2019). However, long-term home rehabilitation reduces contact between patients with cognitive impairment and the outside world and has a certain impact on the patient's psychological state.

BCI-VR STRATEGY FOR EVALUATING AND REHABILITATING COVID-19 PATIENTS WITH COGNITIVE IMPAIRMENT

It can be seen that there are many limitations in traditional cognitive impairment assessment and cognitive rehabilitation training during the COVID-19 pandemic. There is an urgent need for a novel and comprehensive strategy to overcome the shortcomings of traditional approaches. Thus, we propose a comprehensive rehabilitation strategy of BCI-VR, which combines the characteristics of accurate detection of BCI technology with the characteristics of effective training of

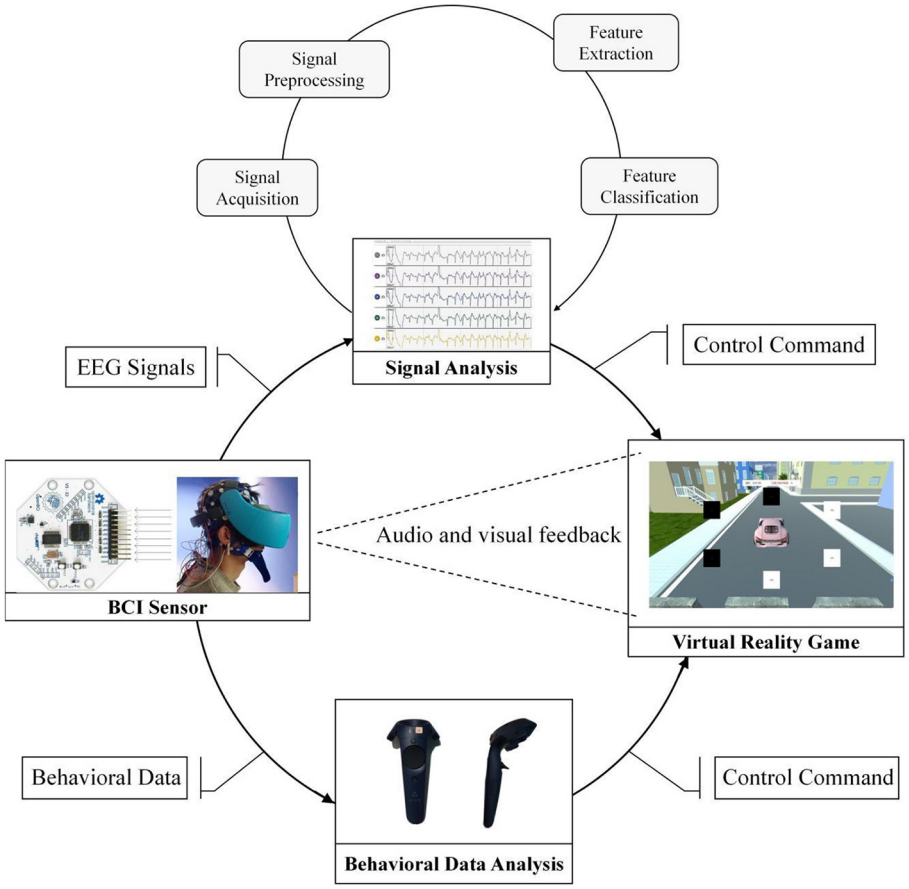


FIGURE 1 | Diagram of the rehabilitation strategy based on BCI-VR.

VR technology and provide one-stop service for cognitive impairment assessment and cognitive rehabilitation training for COVID-19 patients. This strategy is described in detail below.

In BCI technology, EEG signals can be used to objectively and accurately detect the brain specificity of patients with cognitive impairment, which could be performed in the community or even at home (PoŽar et al., 2020; San-Juan et al., 2020; Pinter et al., 2021). VR technology provides an immersive environment for patients with cognitive impairment, improves patient participation and training effect (Snider et al., 2013; Appel et al., 2020; Bauer and Andringa, 2020; Mancuso et al., 2020; Thielbar et al., 2020). This makes the new rehabilitation strategy of BCI-VR proposed in this research more feasible.

The BCI-VR rehabilitation strategy proposed in this study provides guarantee for cognitive impairment assessment and cognitive rehabilitation training during the COVID-19 pandemic, which can meet the cognitive rehabilitation needs of patients at home. Moreover, the application of VR technology would greatly alleviate the negative psychological state and mental state of patients with cognitive impairment caused by blocking (Gao et al., 2020). In the implementation process (as shown in **Figure 1**), BCI-VR requires computers, EEG acquisition instruments and VR wearable devices, which are relatively light and common compared with medical devices in hospitals.

The BCI-VR strategy is specifically divided into a behavioral data analysis module and an EEG analysis module. During cognitive rehabilitation training in a VR environment, the EEG signals of patients with cognitive impairment are collected synchronously for offline or real-time online analysis (Taquet et al., 2021). Rehabilitation training data of patients with cognitive impairment in the VR environment will be recorded in the behavioral data analysis module. With the advancement of the rehabilitation training process, behavior data analysis can be intuitive to see its effect. For example, after a month of spatial navigation ability training, the first day of the spatial navigation ability will compare the 30 days of spatial navigation ability to test the result of rehabilitation training (Vespignani et al., 2020).

The EEG signals of cognitive impairment patients during rehabilitation training will be recorded and processed in the EEG analysis module. Continuous rehabilitation training will gradually show the characteristics of brain regions that constantly change. In patients, these changes can play an evaluation role.

Compared with traditional methods, BCI-VR has the following advantages: It reduces the need for patients with cognitive impairment to go out, and they can receive effective

cognitive impairment assessment and rehabilitation at home; long-term home rehabilitation can relieve the mental state of patients with cognitive impairment, such as impatience and depression; it provides a low-cost cognitive rehabilitation strategy that uses relatively light and common equipment, which can be used in a wide range of applications.

DISCUSSION

More and more studies have been reported on the symptoms of cognitive impairment in COVID-19 patients (Baschi et al., 2020; Haji Akhoundi et al., 2020; Heneka et al., 2020; Jain et al., 2020). These patients need to undergo necessary cognitive impairment assessment and rehabilitation training after they are cured from COVID-19. Traditional cognitive impairment assessment and rehabilitation training have been greatly limited during the pandemic. BCI-VR provides a feasible method for patients in this situation. Recent studies have shown that the EEG signals of COVID-19 patients have certain characteristics (Sethi, 2020; Kubota et al., 2021). More and more researchers suggested that more attention should be paid to the EEG signals of patients during the epidemic (Haines et al., 2020). Perhaps BCI-VR may also monitor whether SARS-COV-2 virus reactivation occur while conducting cognitive rehabilitation training for COVID-19 patients after they are cured.

The application scenarios and implementation forms of the BCI-VR strategy are relatively flexible. It can be applied in rehabilitating various cognitive functions, such as memory, spatial cognition, or language perception, and multi-person interactive rehabilitation training, cross-scene interactive rehabilitation training, or cross-age rehabilitation training. Moreover, studies have shown that cognitive rehabilitation training with multi-person interaction in a VR environment has a better effect (Thielbar et al., 2020). Therefore, BCI-VR is better developed and applied in cognitive impairment rehabilitation.

BCI-VR in the assessment and rehabilitation of cognitive impairment is still in its early stages. In future research, BCI-VR will make great progress in integrating medical and industrial intelligence, which is not limited to cognitive impairment rehabilitation. However, the current optimized data fusion algorithm and feature extraction of high-dimensional data are still a bottleneck for BCI-VR development. In the following work, we will continue to solve the key BCI and VR technologies in monitoring, evaluating, and rehabilitating cognitive impairment.

CONCLUSION

Through literature analysis and summary, we will find that more and more patients, including young people with COVID-19, exhibit signs of cognitive impairment. We analyze some popular traditional cognitive impairment assessment and rehabilitation methods and summarize their limitations during the current pandemic. Moreover, the proposed new comprehensive rehabilitation BCI-VR strategy and the cognitive impairment assessment and rehabilitation process of BCI-VR are expounded. The advantages of BCI-VR in cognitive impairment assessment and rehabilitation are discussed, and the development trend of this technology in the future is evaluated. However, the optimized data fusion algorithm and feature extraction of high-dimensional data are still the bottlenecks of BCI-VR development. Nevertheless, we expect that BCI-VR will soon play an important role in many fields, such as medical rehabilitation, providing more service support for humans.

AUTHOR CONTRIBUTIONS

DW, YZ, and HS contributed to conception and design of the study. JX, ZW, and YL searched the database. DW, JX, and JL performed the analysis of literature. JX, DW, and ZW wrote the first draft of the manuscript. YL and SW wrote sections of the manuscript. XD and MS revised this paper and analyzed the literature. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported in part by National Natural Science Foundation of China (Nos. 61876165 and 61503326) and National Key Research and Development Program of China (No. 2021YFF1200603).

REFERENCES

Ahmad, I., and Rathore, F. A. (2020). Neurological manifestations and complications of COVID-19: a literature review. *J. Clin. Neurosci*. 77, 8–12. doi: 10.1016/j.jocn.2020.05.017

Appel, L., Appel, E., Bogler, O., Wiseman, M., Cohen, L., and Ein, N., et al. (2020). Older adults with cognitive and/or physical impairments can benefit from immersive virtual reality experiences: a feasibility study. *Front. Med*. 6:329. doi: 10.3389/fmed.2019.00329

Baschi, R., Luca, A., Nicoletti, A., Caccamo, M., Cicero, C. E., and D'Agate, C., et al. (2020). Changes in motor, cognitive, and behavioral symptoms in parkinson's disease and mild cognitive impairment during the COVID-19 lockdown. *Front. Psychiatry*. 11:590134. doi: 10.3389/fpsyt.2020.590134

Bauer, A. C. M., and Andringa, G. (2020). The potential of immersive virtual reality for cognitive training in elderly. *Gerontology*. 66, 614–623. doi: 10.1159/000509830

Belani, P., Schefflein, J., Kihira, S., Rigney, B., Delman, B. N., and Mahmoudi, K., et al. (2020). COVID-19 is an independent risk factor for acute ischemic stroke. *Am. J. Neuroradiol*. 41, 1361–1364. doi: 10.3174/ajnr.A6650

Burns, A., Harrison, J. R., Symonds, C., and Morris, J. (2021). A novel hybrid scale for the assessment of cognitive and executive function: the Free-Cog. *Int. J. Geriatr. Psych*. 36, 566–572. doi: 10.1002/gps.5454

Cavallieri, F., Marti, A., Fasano, A., Dalla Salda, A., Ghirarduzzi, A., and Moratti, C., et al. (2020). Prothrombotic state induced by COVID-19 infection as trigger for stroke in young patients: a dangerous association. *eNeurologicalSci*. 20:100247. doi: 10.1016/j.ensci.2020.100247

Chang, M. C., and Boudier-Revéret, M. (2020). Usefulness of telerehabilitation for stroke patients during the COVID-19 pandemic. *Am. J. Phys. Med. Rehab*. 99:582. doi: 10.1097/PHM.0000000000001468

Coetzer, R. (2020). First impressions of performing bedside cognitive assessment of COVID-19 inpatients. *J. Am. Geriatr. Soc*. 68, 1389–1390. doi: 10.1111/jgs.16561

Devita, M., Bordignon, A., Sergi, G., and Coin, A. (2021). The psychological and cognitive impact of Covid-19 on individuals with neurocognitive impairments: Research topics and remote intervention proposals. *Aging Clin. Exp. Res.* 33, 733–736. doi: 10.1007/s40520-020-01637-6

Eggenberger, P., Bürgisser, M., Rossi, R. M., and Annaheim, S. (2021). Body temperature is associated with cognitive performance in older adults with and without mild cognitive impairment: a cross-sectional analysis. *Front. Aging Neurosci.* 13:585904. doi: 10.3389/fnagi.2021.585904

Fara, M. G., Stein, L. K., Skliut, M., Morgello, S., Fifi, J. T., and Dhamoon, M. S. (2020). Macrothrombosis and stroke in patients with mild Covid-19 infection. *J. Thromb. Haemost.* 18, 2031–2033. doi: 10.1111/jth.14938

Fatima, N., Saqqur, M., Qamar, F., Shaukat, S., and Shuaib, A. (2020). Impact of COVID-19 on neurological manifestations: an overview of stroke presentation in pandemic. *Neurol. Sci.* 41, 2675–2679. doi: 10.1007/s10072-020-04637-6

Gao, Z., Lee, J. E., McDonough, D. J., and Albers, C. (2020). Virtual reality exercise as a coping strategy for health and wellness promotion in older adults during the COVID-19 pandemic. *J. Clin. Med.* 9:1986. doi: 10.3390/jcm9061986

Gately, M. E., Trudeau, S. A., and Moo, L. R. (2019). In-home video telehealth for dementia management: implications for rehabilitation. *Curr. Geriatr. Rep.* 8, 239–249. doi: 10.1007/s13670-019-00297-3

Gunasekaran, K., Amoah, K., Rajasurya, V., and Buscher, M. G. (2020). Stroke in a young COVID-19 patient. *QJM* 113, 573–574. doi: 10.1093/qjmed/hcaa177

Haines, S., Caccamo, A., Chan, F., Galaso, G., Catinchi, A., and Gupta, P. K. (2020). Practical considerations when performing neurodiagnostic studies on patients with COVID-19 and other highly virulent diseases. *Neurodiagn. J.* 60, 78–95. doi: 10.1080/21646821.2020.1756132

Haji Akhoundi, F., Sahraian, M. A., and Naser Moghadasi, A. (2020). Neuropsychiatric and cognitive effects of the COVID-19 outbreak on multiple sclerosis patients. *Multiple Scler. Relat. Disord.* 41:102164. doi: 10.1016/j.msard.2020.102164

Harrison, S. L., Fazio-Eynullayeva, E., Lane, D. A., Underhill, P., and Lip, G. Y. H. (2021). Higher mortality of ischaemic stroke patients hospitalized with COVID-19 compared to historical controls. *Cerebrovasc. Dis.* 50, 326–331. doi: 10.1159/000514137

Heneka, M. T., Golenbock, D., Latz, E., Morgan, D., and Brown, R. (2020). Immediate and long-term consequences of COVID-19 infections for the development of neurological disease. *Alzheimer's Res. Ther.* 12:69. doi: 10.1186/s13195-020-00640-3

Irazoki, E., Contreras-Somoza, L. M., Toribio-Guzmán, J. M., Jenaro-Río, C., van der Roest, H., and Franco-Martín, M. A. (2020). Technologies for cognitive training and cognitive rehabilitation for people with mild cognitive impairment and dementia. A systematic review. *Front. Psychol.* 11:648. doi: 10.3389/fpsyg.2020.00648

Jain, R., Young, M., Dogra, S., Kennedy, H., Nguyen, V., and Jones, S., et al. (2020). COVID-19 related neuroimaging findings: a signal of thromboembolic complications and a strong prognostic marker of poor patient outcome. *J. Neurol. Sci.* 414:116923. doi: 10.1016/j.jns.2020.116923

Jin-xuan, Z., Rui, S., Tian-hao, M., and Hong, W. (2020). Progress in rehabilitation treatment of mild cognitive impairment. *Chinese J. Alzheimer's Dis. Relat. Disord.* 3, 147–153. doi: 10.3969/j.issn.2096-5516.2020.02.011

Kieron, S., Laura, M., McColl, B. W., Elkind, M. S., Allan, S. M., Smith, C. J. (2020). Preceding infection and risk of stroke: an old concept revived by the COVID-19 pandemic. *Int. J. Stroke.* 15, 722–732. doi: 10.1177/1747493020943815

Knopman, D. S., and Petersen, R. C. (2014). Mild cognitive impairment and mild dementia: a clinical perspective. *Mayo Clin. Proc.* 89, 1452–1459. doi: 10.1016/j.mayocp.2014.06.019

Koralnik, I. J., and Tyler, K. L. (2020). COVID-19: a global threat to the nervous system. *Ann. Neurol.* 88, 1–11. doi: 10.1002/ana.25807

Kubota, T., Gajera, P. K., and Kuroda, N. (2021). Meta-analysis of EEG findings in patients with COVID-19. *Epilepsy Behav.* 115:107682. doi: 10.1016/j.yebeh.2020.107682

Lara, B., Carnes, A., Dakterzada, F., Benitez, I., and Pinol-Ripoll, G. (2020). Neuropsychiatric symptoms and quality of life in Spanish patients with Alzheimer's disease during the COVID-19 lockdown. *Eur. J. Neurol.* 27, 1744–1747. doi: 10.1111/ene.14339

Lopez Mongil, R. (2017). [Cognitive impairment/dementia: Assessment and follow-up. Functional aspects]. Deterioro cognitivo/demencia: Valoracion y seguimiento. Aspectos funcionales. *Revista espanola de geriatria y gerontologia.* 52, 28–33. doi: 10.1016/S0211-139X(18)30077-5

Lunardini, F., Luperto, M., Romeo, M., Basilico, N., Daniele, K., and Azzolino, D., et al. (2020). Supervised digital neuropsychological tests for cognitive decline in older adults: usability and clinical validity study. *JMIR mHealth uHealth* 8:e17963. doi: 10.2196/17963

Mahboob, S., Boppana, S. H., Rose, N. B., Beutler, B. D., and Tabaac, B. J. (2020). Large vessel stroke and COVID-19: case report and literature review. *eNeurologicalSci* 20:100250. doi: 10.1016/j.ensci.2020.100250

Majidi, S., Fifi, J. T., Ladner, T. R., Lara-Reyna, J., Yaeger, K. A., and Yim, B., et al. (2020). Emergent large vessel occlusion stroke during new york city's COVID-19 outbreak clinical characteristics and paraclinical findings. *Stroke.* 51, 2656–2663. doi: 10.1161/STROKEAHA.120.030397

Manca, R., De Marco, M., and Venneri, A. (2020). The impact of COVID-19 infection and enforced prolonged social isolation on neuropsychiatric symptoms in older adults with and without dementia: a review. *Front. Psychiatry.* 11:585540. doi: 10.3389/fpsyt.2020.585540

Mancuso, V., Stramba-Badiale, C., Cavedoni, S., Pedroli, E., Cipresso, P., and Riva, G. (2020). Virtual reality meets non-invasive brain stimulation: integrating two methods for cognitive rehabilitation of mild cognitive impairment. *Front. Neurol.* 11:566731. doi: 10.3389/fneur.2020.566731

Merkler, A. E., Parikh, N. S., Mir, S., Gupta, A., Kamel, H., and Lin, E., et al. (2020). Risk of ischemic stroke in patients with Coronavirus Disease 2019 (COVID-19) vs patients with influenza. *JAMA Neurol.* 77, 1366–1372. doi: 10.1001/jamaneurol.2020.2730

Narayanan, L., and Murray, A. D. (2016). What can imaging tell us about cognitive impairment and dementia? *World J. Radiol.* 8:240. doi: 10.4329/wjr.v8.i3.240

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., and Collin, I., et al. (2019). The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment (vol 53, pg 695, 2005). *J. Am. Geriatr. Soc.* 67:1991. doi: 10.1111/j.1532-5415.2005.53221.x

Natta, D. D. N., Alagnide, E., Kpadonou, G. T., Stoquart, G. G., Detrembleur, C., and Lejeune, T. M. (2015). Feasibility of a self-rehabilitation program for the upper limb for stroke patients in Benin. *Ann. Phys. Rehabil. Med.* 58, 322–325. doi: 10.1016/j.rehab.2015.08.003

Oxley, T. J., Mocco, J., Majidi, S., Kellner, C. P., Shoirah, H., and Singh, I. P., et al. (2020). Large-Vessel stroke as a presenting feature of covid-19 in the young. [Case Reports; Letter]. *New Engl. J. Med.* 382:e60. doi: 10.1056/NEJMc2009787

Pangman, V. C., Sloan, J., and Guse, L. (2000). An examination of psychometric properties of the mini-mental state examination and the standardized mini-mental state examination: Implications for clinical practice. *Appl. Nurs. Res.* 13, 209–213. doi: 10.1053/apnr.2000.9231

Pinna, P., Grewal, P., Hall, J. P., Tavarez, T., Dafer, R. M., and Garg, R., et al. (2020). Neurological manifestations and COVID-19: experiences from a tertiary care center at the Frontline. *J. Neurol. Sci.* 415:116969. doi: 10.1016/j.jns.2020.116969

Pinter, D., Kober, S. E., Fruhwirth, V., Berger, L., Damulina, A., and Khalil, M., et al. (2021). MRI correlates of cognitive improvement after home-based EEG neurofeedback training in patients with multiple sclerosis: a pilot study. *J. Neurol.* 268, 3808–3816. doi: 10.1007/s00415-021-10530-9

PoŽar, R., Giordani, B., and Kavcic, V. (2020). Effective differentiation of mild cognitive impairment by functional brain graph analysis and computerized testing. *PLoS ONE.* 15:e230099. doi: 10.1371/journal.pone.0230099

Rajdev, K., Lahan, S., Klein, K., Piquette, C. A., and Thi, M. (2020). Acute ischemic and hemorrhagic stroke in COVID-19: mounting evidence. *Cureus* 12:e10157. doi: 10.7759/cureus.10157

Rejdak, K., and Grieb, P. (2020). Adamantanes might be protective from COVID-19 in patients with neurological diseases: multiple sclerosis, parkinsonism and cognitive impairment. *Multiple Scler. Relat. Disord.* 42:102163. doi: 10.1016/j.msard.2020.102163

Richardson, S. J., Carroll, C. B., Close, J., Gordon, A. L., O'Brien, J., and Quinn, T. J., et al. (2020). Research with older people in a world with COVID-19: identification of current and future priorities, challenges and opportunities. *Age Ageing.* 49, 901–906. doi: 10.1093/ageing/afaa149

Saif, N., Yan, P., Niotis, K., Scheyer, O., Rahman, A., and Berkowitz, M., et al. (2019). Feasibility of using a wearable biosensor device in patients at risk for Alzheimer's disease dementia. *J. Prevent. Alzheimer's Dis.* 7, 104–111. doi: 10.14283/jpad.2019.39

San-Juan, D., Jiménez, C. R., Camilli, C. X., de la Cruz Reyes, L. A., Galindo, E. G. A., and Burbano, G. E. R., et al. (2020). Guidance for clinical neurophysiology examination throughout the COVID-19 pandemic. Latin American chapter of the IFCN task force –COVID-19. *Clin. Neurophysiol.* 131, 1589–1598. doi: 10.1016/j.clinph.2020.04.011

Sethi, N. K. (2020). EEG during the COVID-19 pandemic: what remains the same and what is different. *Clin. Neurophysiol.* 131:1462. doi: 10.1016/j.clinph.2020.04.007

Sezgin, M., Ekizoglu, E., Yesilot, N., and Coban, O. (2020). Stroke during covid-19 pandemic. *Arch. Neuropsychiatry*. 57, 83–84. doi: 10.29399/npa.27196

Snider, J., Plank, M., Lee, D., and Poizner, H. (2013). Simultaneous neural and movement recording in large-scale immersive virtual environments. *IEEE T. Biomed. Circ. S.* 7, 713–721. doi: 10.1109/TBCAS.2012.2236089

Stavropoulos, T. G., Papastergiou, A., Mpaltadoros, L., Nikolopoulos, S., and Kompatsiaris, I. (2020). IoT wearable sensors and devices in elderly care: a literature review. *Sensors Basel.* 20:2826. doi: 10.3390/s20102826

Stricker, N. H., Christianson, T. J., Lundt, E. S., Alden, E. C., Machulda, M. M., and Fields, J. A., et al. (2021). Mayo normative studies: regression-based normative data for the auditory verbal learning test for ages 30–91 years and the importance of adjusting for sex. *J. Int. Neuropsych. Soc.* 27, 211–226. doi: 10.1017/S1355617720000752

Sun, J., Lv, X., Gao, X., Chen, Z., Wei, D., and Ling, Y., et al. (2020). The association between serum uric acid level and the risk of cognitive impairment after ischemic stroke. *Neurosci. Lett.* 734:135098. doi: 10.1016/j.neulet.2020.135098

Taquet, M., Geddes, J. R., Husain, M., Luciano, S., and Harrison, P. J. (2021). 6-Month neurological and psychiatric outcomes in 236379 survivors of COVID-19: a retrospective cohort study using electronic health records. *Lancet Psychiatry* 8, 416–427. doi: 10.1016/S2215-0366(21)00084-5

Thielbar, K., Spencer, N., Tsoupikova, D., Ghassemi, M., and Kamper, D. (2020). Utilizing multi-user virtual reality to bring clinical therapy into stroke survivors' homes. *J. Hand Ther.* 33, 246–253. doi: 10.1016/j.jht.2020.01.006

Vespignani, H., Colas, D., Lavin, B. S., Soufflet, C., Maillard, L., and Pourcher, V., et al. (2020). Report on electroencephalographic findings in critically ill patients with COVID-19. *Ann. Neurol.* 88, 626–630. doi: 10.1002/ana.25814

Wen, D., Fan, Y., Hsu, S., Xu, J., Zhou, Y., and Tao, J., et al. (2020). Combining brain-computer interface and virtual reality for rehabilitation in neurological diseases: a narrative review. *Ann. Phys. Rehabil. Med.* 64:101404. doi: 10.1016/j.rehab.2020.03.015

Wen, D., Lan, X., Zhou, Y., Li, G., Hsu, S., and Jung, T. (2018). The study of evaluation and rehabilitation of patients with different cognitive impairment phases based on virtual reality and EEG. *Front. Aging Neurosci.* 10:88. doi: 10.3389/fnagi.2018.00088

Woo, M. S., Malsy, J., Pöttgen, J., Seddiq Zai, S., Ufer, F., and Hadjilaou, A., et al. (2020). Frequent neurocognitive deficits after recovery from mild COVID-19. *Brain Commun.* 2:fcaa205. doi: 10.1093/braincomms/fcaa205

Xu, M., Wang, Z., Zhang, H., Pantazis, D., Wang, H., and Li, Q. (2020). A new graph Gaussian embedding method for analyzing the effects of cognitive training. *PLoS Comput. Biol.* 16:e1008186. doi: 10.1371/journal.pcbi.1008186

Zhang, X., Su, J., Gao, C., Ni, W., Gao, X., and Li, Y., et al. (2019). Progression in vascular cognitive impairment: pathogenesis, neuroimaging evaluation, and treatment. *Cell Transplant.* 28, 18–25. doi: 10.1177/0963689718815820

Zhaojun, Y., Miaser, and Hongxia, C. (2020). Research progress on integrated chinese and western medicine treatment of mild cognitive impairment. *J. Pract. Tradit. Chin. Intern. Med.* 34, 67–71. doi: 10.13729/j.issn.1671-7813.Z20200342

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 21 March 2022
doi: 10.3389/fnagi.2022.826622

A Two-Stage Model for Predicting Mild Cognitive Impairment to Alzheimer's Disease Conversion

Peixin Lu[1], Lianting Hu[2,3], Ning Zhang[4], Huiying Liang[2,3], Tao Tian[5] and Long Lu[1]**

[1] School of Information Management, Wuhan University, Wuhan, China, [2] Medical Big Data Center, Guangdong Provincial People's Hospital, Guangzhou, China, [3] Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangzhou, China, [4] School of Business, Qingdao University, Qingdao, China, [5] The First Division of Psychiatry, Jingmen No. 2 People's Hospital, Jingmen, China

Early detection of Alzheimer's disease (AD), such as predicting development from mild cognitive impairment (MCI) to AD, is critical for slowing disease progression and increasing quality of life. Although deep learning is a promising technique for structural MRI-based diagnosis, the paucity of training samples limits its power, especially for three-dimensional (3D) models. To this end, we propose a two-stage model combining both transfer learning and contrastive learning that can achieve high accuracy of MRI-based early AD diagnosis even when the sample numbers are restricted. Specifically, a 3D CNN model was pretrained using publicly available medical image data to learn common medical features, and contrastive learning was further utilized to learn more specific features of MCI images. The two-stage model outperformed each benchmark method. Compared with the previous studies, we show that our model achieves superior performance in progressive MCI patients with an accuracy of 0.82 and AUC of 0.84. We further enhance the interpretability of the model by using 3D Grad-CAM, which highlights brain regions with high-predictive weights. Brain regions, including the hippocampus, temporal, and precuneus, are associated with the classification of MCI, which is supported by the various types of literature. Our model provides a novel model to avoid overfitting because of a lack of medical data and enable the early detection of AD.

Keywords: mild cognitive impairment, Alzheiemer's disease, contrastive learning, transfer leaning, MRI, deep learning

OPEN ACCESS

Edited by:
Peng Xu,
University of Electronic Science and Technology of China, China

Reviewed by:
Mingxia Liu,
University of North Carolina at Chapel Hill, United States
Penghui Song,
Capital Medical University, China

***Correspondence:**
Tao Tian
tiantao91@163.com
Long Lu
bioinfo@gmail.com

Specialty section:
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

Received: *01 December 2021*
Accepted: *17 February 2022*
Published: *21 March 2022*

Citation:
Lu P, Hu L, Zhang N, Liang H, Tian T and Lu L (2022) A Two-Stage Model for Predicting Mild Cognitive Impairment to Alzheimer's Disease Conversion. Front. Aging Neurosci. 14:826622. doi: 10.3389/fnagi.2022.826622

INTRODUCTION

Alzheimer's disease (AD), a severe neurodegenerative disease, is the most common type of dementia (Heun et al., 1997; Association, 2019). Nowadays, at least 50 million people worldwide suffer from AD or other types of dementia, and it is expected that this number will reach 131 million in 2050 (Livingston et al., 2017). This further increases the burden of the medical care system in aging societies. Mild cognitive impairment (MCI) is a stage between normal and AD, with 10–12% of people developing AD each year (Petersen, 2000). Based on the progression toward AD, it can be classified into two categories: progressive MCI (pMCI) and stable MCI (sMCI). Although there is no effective treatment for AD at present, its progression can be slowed by medication, memory training, exercise, and diet, which necessitates the early detection of potential patients (Roberson and Mucke, 2006). Neuroimaging techniques, which can detect disease-related neuropathological changes, are valuable tools for assessing and diagnosing AD (Johnson et al., 2012). MRI is one of

the most widely studied neuroimaging techniques because it is non-invasive, generally available, affordable, and capable of distinguishing between different soft tissues (Klöppel et al., 2008).

With the rapid development, deep learning has achieved remarkable progress in a variety of fields, especially in computer vision and medical imaging, where it outperforms traditional machine learning methods (Shen et al., 2017; Bernal et al., 2019; Abrol et al., 2021). Deep learning approaches perform feature selection during model training and loss function optimization without the need for domain experts' prior knowledge. As a result, individuals with no medical expertise can use them for research and applications, especially in the field of medical image analysis (Shen et al., 2017). Notably, Convolutional Neural Network (CNN) has achieved outstanding performance in the classification tasks of AD and normal control (NC) (Liu et al., 2019) and pMCI/sMCI (Choi and Jin, 2018; Spasov et al., 2019). In general, deep neural networks require large samples for model fitting, especially 3D deep neural network models with more parameters. However, as compared with existing million-sample natural image datasets, neuroimaging datasets have a relatively small sample size (Russakovsky et al., 2015), which can possibly be explained by the following factors. At first, collecting large training sets and labeling image data are costly and time consuming (Shen et al., 2017; Irvin et al., 2019). Furthermore, technical and privacy issues also constraints obstruct medical data collection (Irvin et al., 2019). Therefore, preventing model overfitting due to the scarcity of medical samples has become one of the hottest topics in deep learning of neuroimaging.

Transfer learning is a popular method for dealing with a small number of samples. It utilizes a pretrained model with supervised learning on a large labeled dataset (source domain, e.g., ImageNet) and then fine tunes it on the task of interest (target domain). Studies have shown that knowledge transferred from large-scale annotated natural images (ImageNet) to medical images can significantly improve the effectiveness of assisted diagnosis (Tajbakhsh et al., 2016; Raghu et al., 2019). However, standard medical images, such as MRI, CT, and positron emission tomography (PET), are in three dimensional (3D), preventing ImageNet-based pretrained models from being directly transferred to MRI. Converting 3D data into two-dimensional (2D) slices is a typical method, however, this ignores the rich 3D spatial anatomical information and inevitably affects the performance. To address this issue, several studies (Yang et al., 2017; Zeng and Zheng, 2018) have used pretrained 3D models based on natural video datasets (Tran et al., 2014; Carreira and Zisserman, 2017) to transfer to medical imaging tasks, but have not yet achieved the expected results because of the vast difference between these two domains.

Recently, contrastive learning, a self-supervised learning method, has recently been demonstrated to perform superiorly in various vision tasks (Wu et al., 2018; Zhuang C. et al., 2019; Chen X. et al., 2020). Momentum Contrast (MoCo) (He et al., 2020) is a state-of-the-art method in contrastive learning, which minimizes positive pairs variability while maximizing negative pairs variability. Based on existing research concerns, we proposed a two-stage model based on MoCo (He et al., 2020) to classify sMCI and pMCI. The main contributions of our study are as follows.

1) Systematic evaluation of 3D ResNet models with different structures and selection of the best model for sMCI and pMCI classification. Provides a reference for related studies.
2) A two-stage model is proposed to solve the problem of domain transfer between the source and target domains, which solves the problem of overfitting caused by small samples in sMCI and pMCI classification and improves the classification performance in AD diagnosis. To the best of our knowledge, we first introduce the MoCo in pMCI and sMCI classification.
3) Three-dimensional Gradient-weighted Class Activation Mapping (Grad-CAM), which is widely used for model interpretability, was introduced to get the heatmap that highlights the brain regions our model focuses on and increases the interpretability of the model.

MATERIALS AND METHODS

As indicated in **Figure 1**, our two-stage transfer learning model was divided into three main parts. In our framework, we did not directly transfer trained model on natural image sets or other medical image sets to our research such as previous studies, mainly for the following reasons: at first, the adoption of 3D CNN in this study to preserve more spatial information limits direct transfer learning from 2D natural images; second, the different components of medical image sets may harm the performance. For instance, the Med3D is composed of MRI and CT of brains, lungs, chests, and other organs (Chen et al., 2019), while our MCI data set only includes brain MRI data. The details of each step were described next.

Dataset and Data Preprocessing

Data used in our study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, which is the largest open-access AD database with wide popularity in AD-related research. It was launched in 2003 by the several public and private organizations to measure the progression of MCI and early AD by medical image (such as MRI, PET), genomics, biological markers, and neuropsychological assessments (Jack et al., 2008). More information can be found at http://adni.loni.usc.edu/.

As defined in this study, participants with MCI at baseline who developed or did not develop AD within 3 years were referred to as pMCI and sMCI, respectively. To prevent data leakage, only participants' baseline data were selected in this study. Finally, data from 577 MCI subjects (Means $\pm$ std age = 73.08 $\pm$ 7.25 years) were included in our study, and 297 of the MCI was classified as sMCI (51.5%) and the rest 280 were pMCI (48.5%). The demographics and the mini-mental state examination scores (MMSEs) information of subjects is summarized in **Table 1**. Differences in the median age and gender between groups were tested using ANOVA and Fisher's exact test, respectively.

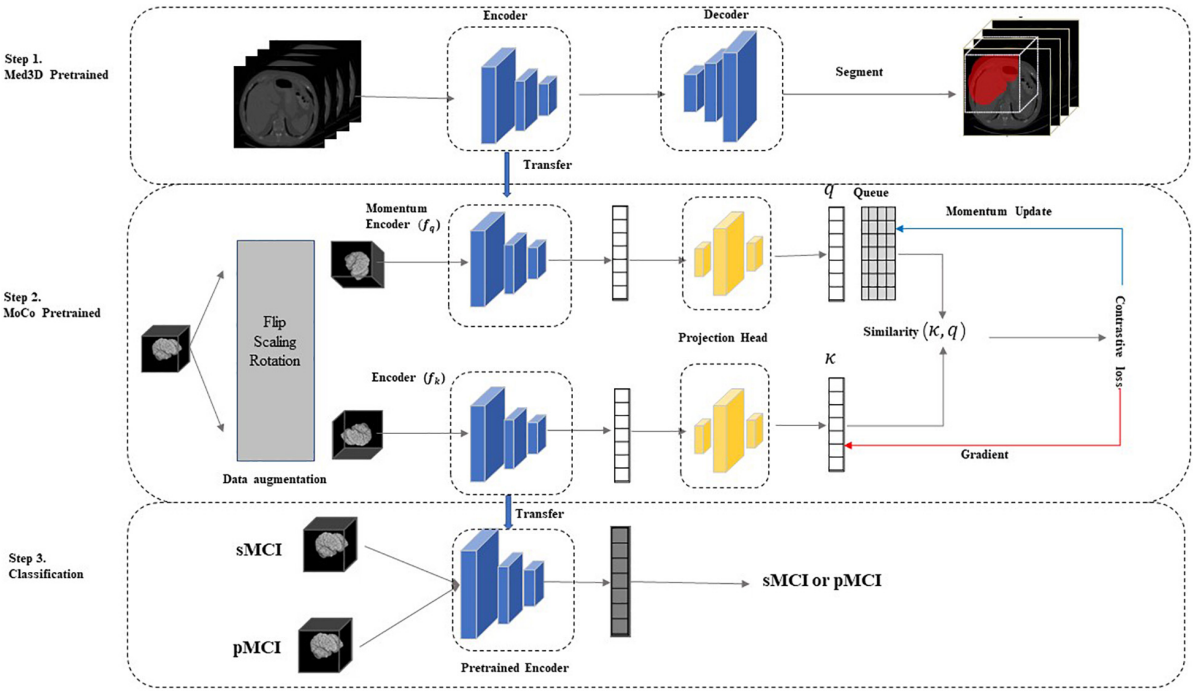


FIGURE 1 | The two-stage model. In step 1, we used the Med3D (Chen et al., 2019) pre-trained model to initialize network weights and learn the common medical image features. In step 2, we performed contrastive learning on unlabeled sMCI and pMCI samples using the improved MoCo to update the network weights, further increasing the correlation between the target and source domains and learning the specific medical features of the sMCI/pMCI classification task. Finally, the network was fine-tuned using labeled sMCI/pMCI samples to achieve sMCI and pMCI classification. MoCo, Momentum Contrast; pMCI, progressive mild cognitive impairment; sMCI, stable mild cognitive, impairment.

TABLE 1 | A summary of the demographic and clinical information of participants.

	Number	Age (years old)	Sex(M/F)	MMSE
sMCI	297	72.2 ± 7.4[55.0,88.4]	174/123	28.0 ± 1.7[23,30]
pMCI	280	73.9 ± 6.9[55.2,88.3]	172/108	26.8 ± 1.8[21,30]

Values are presented as Means ± Standard Deviation [Range]; M, Male; F, Female; MMSE, Mini-Mental State Examination.

These two interactions yielded no statistically significant results ($p > 0.05$).

All 1.5T and 3T structural MRI of the participant were downloaded. The detailed information of MRI, such as scanner protocols, can be found at http://adni.loni.usc.edu/methods/documents/mri-protocols/. Data are preprocessed using FSL[1] with three main steps: bias field correction using the N4 algorithm (Tustison et al., 2010); affine linear alignment of scans onto the MIN152 atlas; skull stripping of each image for 129 × 145 × 129 voxels. **Figure 2** shows the difference before and after preprocessing of the MRI from the same sample.

Network Weight Initialization Using Med3D

Many studies have demonstrated that using transfer learning parameter initialization can significantly improve the performance of models compared to training from scratch (Afzal et al., 2019; Mousavian et al., 2019; Naz et al., 2021). This study selected the Med3D network and its pretrained weights on eight segmented datasets (Chen et al., 2019).

The authors of Med3D integrated data from eight medical segmentation datasets to create the 3DSeg-8 dataset, which contains multiple modalities (MRI and CT), target organs (brain, heart, pancreas), and pathological conditions (CodaLab, 2021 Competition; Menze et al., 2015; Tobon-Gomez et al., 2015; Medical Segmentation Decathlon, 2021). Med3D uses a standard encode–decode partition structure, where the encoder uses the ResNet family. The main idea of ResNet is to introduce the residual block in the network, as illustrated in **Figure 3**, where F(x) is the residual mapping and X is the identity mapping, also called "shortcut." This helps train a deeper network to achieve higher accuracy without vanishing or explosion of the gradient. Notably, Med3D uses a parallel format for model training, which means the same encoder is used for eight datasets, and the decoder is composed of 8 branches accordingly in parallel. This allows the decoder to adapt to different organizational segmentation targets, while the encoder can learn universal features. **Figure 1**, Step1 depicts the Med3D structure. Med3D pretrained models can be used for classification, detection, and segmentation. We used the parameters of the models pretrained by the 3DSeg-8 dataset for the initialization of our network. Transfer learning strategy effects were evaluated in various

[1]https://fsl.fmrib.ox.ac.uk/

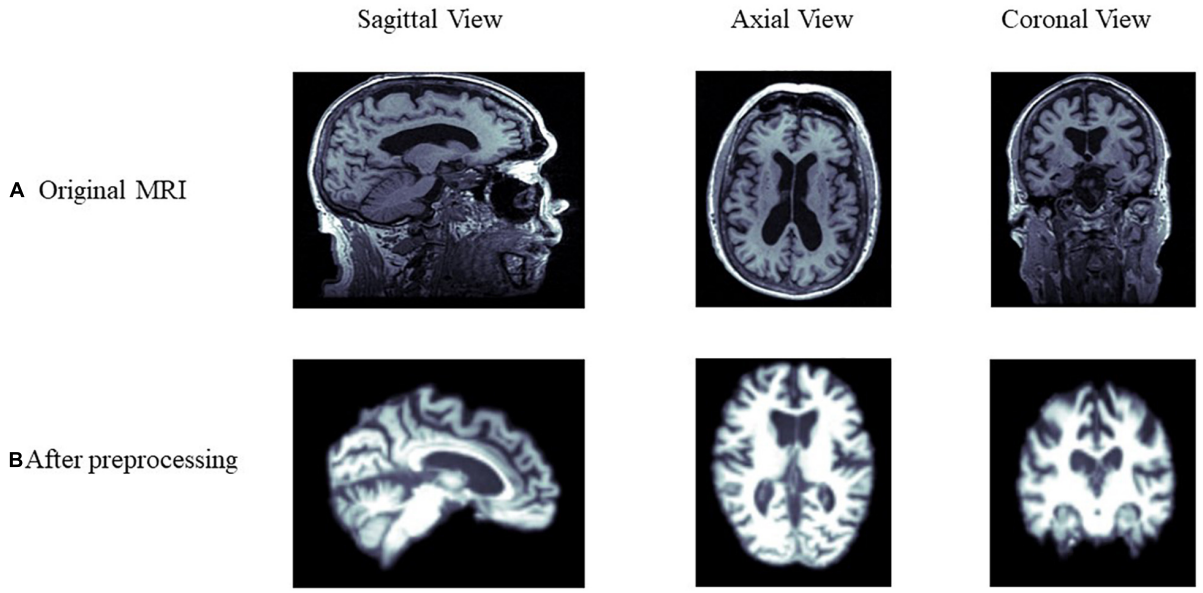


FIGURE 2 | Comparison of MRI data before and after preprocessing. **(A,B)** Show the differences before and after preprocessing from the sagittal, axial, and coronal view of the brain, respectively.

ResNet networks in Med3D, including, ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-152 (He et al., 2016).

Transfer Learning Using Contrastive Learning

We extracted general 3D medical image features by Med3D (Chen et al., 2019). However, while Med3D includes MRI and CT of brains, lungs, chests, and other organs, whereas our MCI data set only include brain MRI data, there are still domain transfer concerns between the dataset of Med3D and MRI of sMCI and pMCI (Chen et al., 2019), while our MCI data set is only composed of brain MRI data. Contrastive learning, a special unsupervised learning method with a supervisory function, was introduced in this study to further increase the correlation between the target and source domains. The labels of contrastive learning are generated by the data itself rather than by manual labeling (Liu et al., 2020). It uses unlabeled data to train models and learn embeddings of the data by maximizing the consistency between different augmented views of the same sample and minimizing the consistency between different samples through a contrast loss function (Tian et al., 2020).

Currently, there are various representative contrastive learning methods, such as MoCo (He et al., 2020), SimCLR (Chen T. et al., 2020), and PIRL (Misra and van der Maaten, 2020). Because of the sample size and computational resources constraints, we chose MoCo as our pretraining model in our study. Unlike the end-to-end gradient update of SimCLR, MoCo introduces a dynamic queueing dictionary, which is updated by adding new training batches to the queue while removing the oldest ones from the queue according to the first-in-first-out principle and keeping the length of the queueing dictionary unchanged. This approach allows MoCo to obtain good training results with small batch size.

Given and preprocessed sample x, contrastive learning obtains two augmented samples x_q and x_k by data augmentation of sample x. x_q and x_k are referred to as query data and key data, respectively. q and k are the latent representation of x_q and x_k using a query encoder $q=f_q(x_q;\theta_q)$ and key encoder $k=f_k(x_k;\theta_k)$ with weight θ_q and θ_k, respectively. If the query and the key belong to the same sample, it is marked as a positive pair k^+. Otherwise, it is a negative sample pair k^-. The auxiliary task in contrastive learning is: given a pair (x_q, x_k), determining whether it is a

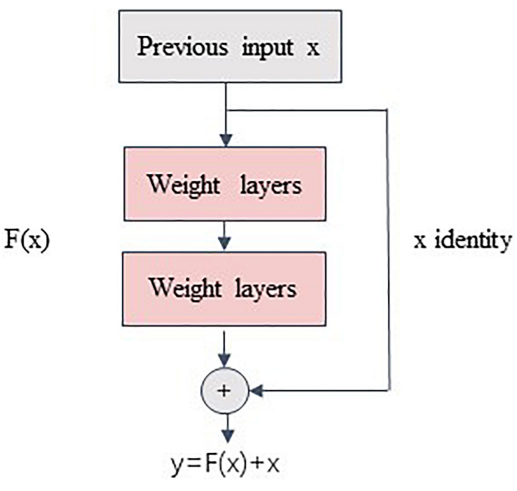


FIGURE 3 | Residual block. The ResNet's Core Modules.

positive or negative sample pair and making the positive samples closer together and the negative samples further apart.

Give a query q. MoCo applies a queue storing a set of keys k from different samples, including one $_{k^+}$ and several $_{k^-}$. The contrastive loss can be defined as:

$$\mathcal{L}_{q,k+,\{k^-\}} = -\log\frac{\exp\left(q\cdot k^+/\tau\right)}{\exp\left(q\cdot k^+/\tau\right)+\sum_{k^-}\exp\left(q\cdot k^-/\tau\right)} \quad (1)$$

Here, $_{\tau}$ is the temperature parameter. The model updates the encoder weights by minimizing the contrastive loss.

In MoCo, the key encoder is neither updated by back-propagation nor copied from the query encoder. Still, a running average of the key encoder is used, which is known as a momentum encoder. The updating of $_{\theta_q}$ and $_{\theta_k}$ can be formulated as:

$$\begin{aligned}\theta_q &\leftarrow \theta_q - \alpha\frac{\partial\mathcal{L}}{\partial\theta_q}\\ \theta_k &\leftarrow m\theta_k + (1-m)\theta_q\end{aligned} \quad (2)$$

Here, m $_{\in}$ [0, 1) is the momentum coefficient. As in Eq. (2), $_{\theta_k}$ is updated more smoothly than $_{\theta_q}$ which is updated by back-propagation.

It was shown that data augmentation methods such as Gaussian blur and color distortion for natural images might not be applicable in the medical image. For example, Gaussian blur on MRI can potentially change the label of the data. Therefore, we improved the data augmentation method in MoCo by using random rotation (± 10°), horizontal flip, and random scaling (± 10%) on 3D MRI. In detail, we rotated images at any angle between −10° and +10° along the three axes. Random scaling was also applied to randomly scale the image by ± 10%. If the image size is larger than the size of the original image, the same volume of the original image is extracted by cropping the center region of the image. If the volume is less than its original size, it is filled with 0 in the reduced region. **Figure 4** shows the schematic diagram of the three data augmentation methods.

We performed MoCo on the full unlabeled MRI data using the Med3D pretrained ResNet as the encoder. In addition, as it is shown that adding a projection head helps to learn feature representation better (Chen T. et al., 2020; Chen X. et al., 2020), we added a projection head, as shown in **Figure 1**. We used two MLP with 128-D hidden layers and a ReLU activation function as the projection head as:

$$\begin{aligned}z_q &= g(q) = W^2\sigma\left(W^1 q\right)\\ z_k &= g(k) = W^2\sigma\left(W^1 k\right)\end{aligned} \quad (3)$$

where $_{W^1}$ and $_{W^2}$ represent the hidden and output layer weights, and $_{\sigma}$ is the ReLU activation function.

Classification

The last step of our model is classification, where the labeled data were divided into training, validation, and test sets to fine-tune the pretrained encoders. We added a linear classifier to the frozen backbone model to complete the classification of sMCI and pMCI as proposed by Chen X. et al. (2020).

Model Evaluation

We first used different ResNet, including ResNet-10, ResNet-34, ResNet-50, ResNet-101 as our two-stage model backbone network, and selected the ResNet with the best performance (ResNet-50, see in results) as our backbone in the following comparative experiment.

Evaluation of Transferring Learning Strategies

We first conducted the following comparative experiments with different transfer learning strategies. All the models in different transfer learning strategies used the best ResNet selected by the first experiment.

λ Med3D, pretrained ResNet using Med3D and fine-tuned the model to complete the classification of sMCI and pMCI.

λ MoCo, random initialization of weights, followed with the modified MoCo in the method to train the ResNet without using sMCI/pMCI labels. Then fine-tune the model using labeled data and do the classification.

λ Only ResNet, random initialization of weights, and training ResNet from scratch.

λ Med3D+MoCo, our two-stage model.

Comparison With Previous Studies

To comprehensively understand the performance of our method, we reviewed the state-of-the-art literature, which utilized deep learning to predict the conversion from MCI to Alzheimer's using MRI. We selected studies that achieved criteria for a fair comparison, including (1) only used MRI. (2) published in the last 3 years. (3) the data were from ADNI.

We selected five evaluation metrics to evaluate our model accurately. (1) Accuracy (Acc) is used to measure the proportion of correctly classified samples. (2) Sensitivity (Sens), also known as the true positive rate, is the proportion of predicted positive results that are true positives. (3) Specificity (Spec) is the proportion of correctly identified negatives. (4) F1-score (F1) is the reconciled average of sensitivity and retrieval. (5) Area Under ROC Curve (AUC) represents how the false-positive rate increases with the true-positive rate and increases the area under the characteristic curve. The aforementioned evaluation metrics were calculated based on True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN), and these four indicators form a confusion matrix.

In our study, sMCI and pMCI were referred to as positive and negative examples, respectively. We can calculate accuracy, sensitivity, specificity, and F1 as follows:

$$Acc = \frac{TP+TN}{TP+TN+FN+FP} \quad (4)$$

$$SPE = \frac{TN}{TN+FP} \quad (5)$$

$$SEN = \frac{TP}{TP+FN} \quad (6)$$

$$\mathbb{F}1 = 2\times\frac{\mathrm{TP}}{(2\times\mathrm{TP}+\mathrm{FP}+\mathrm{FN})} \quad (7)$$

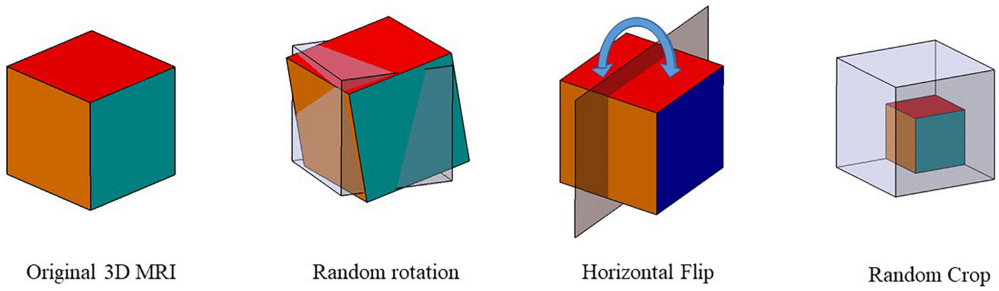


FIGURE 4 | Data augmentation schematic. Three data augmentation were used to augment the data size: random rotation, horizontal flip, and random scaling.

Finally, we used the non-parametric bootstrap to construct each evaluation metrics' CIs. A total of 10,000 bootstrap replicates were extracted from the test set, and the performance metrics of the model on each bootstrap replicate were calculated. This generated a distribution for each estimate and reported 95% bootstrap percentile intervals (2.5 and 97.5 percentile) (Efron and Tibshirani, 1993).

Experimental Settings

The model was implemented in PyTorch. We used Stochastic Gradient Descent (SGD) with a weight decay of 0.0001 and momentum of 0.99 as our optimizer. A minibatch size of 16 and a cosine annealing learning rate with an initialized value of 0.01 were used in training. Other hyperparameters are the same as default values in MoCo (He et al., 2020). All unlabeled data were used to train MoCo. We trained the classifier with optimized cross-entropy loss and a learning rate of 0.001 in 100 epochs. The dataset was randomly split into training and testing data with a ratio of 8:2. Optimal hyperparameters were selected using fivefold cross-validation on the training set, and the optimal model was chosen for model evaluation on the test set (**Figure 5**). All experiments were run on NVIDIA GTX 2080.

RESULTS

Results of Different ResNet Models Using a Two-Stage Model

In this part, we investigated the classification performance of different Med3D pretrained CNN backbones on the test set, including ResNet-18, ResNet-34, ResNet-50, and ResNet-101. As highlighted in **Table 2** and shown in **Figure 6A**, ResNet-50 achieved the best performance with an accuracy of 0.819 and an AUC of 0.835, indicating complex models with more parameters may not always work best. ResNet-50, the model with the best performance, was used in our following experiments. The confusion maps of different ResNet are shown in **Figure 6B**.

Results of Different Transfer Learning Strategies Using ResNet-50

Table 3 and **Figure 7** show the results of the comparison of accuracy, sensitivity, specificity, F1, and ROC for different transfer learning strategies (based on ResNet-50) mentioned in section "Experimental Settings," where the optimal results are highlighted. As **Table 3** and **Figure 7A** indicate, our method achieves better results compared to other methods in terms of accuracy (0.819), sensitivity (0.786), specificity (0.850), and F1 score (0.807). Similarly, **Figure 7A** shows ROC curves of different transfer learning strategies, where our method has the best AUC of 0.835 compared with other methods. The confusion maps of different transfer learning strategies are shown in **Figure 7B**.

Results of the Relevant Brain Region

In this study, the 3D Grad-CAM method was used to identify brain regions associated with sMCI or pMCI and improve the interpretability of the model. After weight back-propagation of trained models, we obtained average relevance heatmaps of each class in the test dataset. For comparison, we highlighted temporal superior, temporal middle, hippocampus, thalamus, precuneus, cingulate in the automated anatomical labeling (AAL2)[2] in **Figure 8** first row. **Figure 8** second and third rows show each class's last convolutional layer's attention heatmap. As shown in **Figure 8**, the hippocampus, temporal superior, temporal middle, thalamus, and cingulate are relevant for both sMCI and pMCI. But precuneus is recognized as a unique feature of pMCI.

Comparisons With Previous Studies

We further used four evaluation metrics to compare our results to previous state-of-the-art deep learning studies on sMCI/pMCI classification published in the last 3 years, including accuracy, sensitivity, specificity, and AUC. **Table 4** summarizes the results in relation to the literature, and the best results are indicated by the bold text. In the classification tasks of sMCI and pMCI using deep learning, our method achieves better or comparable classification results in terms of accuracy, specificity, and AUC.

DISCUSSION

This study proposed a two-stage method that combined contrastive learning and transfer learning for predicting conversion from MCI to AD based on MRI. Pretrained models from sizeable medical image datasets were used to initialize the

[2]http://www.gin.cnrs.fr/en/tools/aal-aal2/

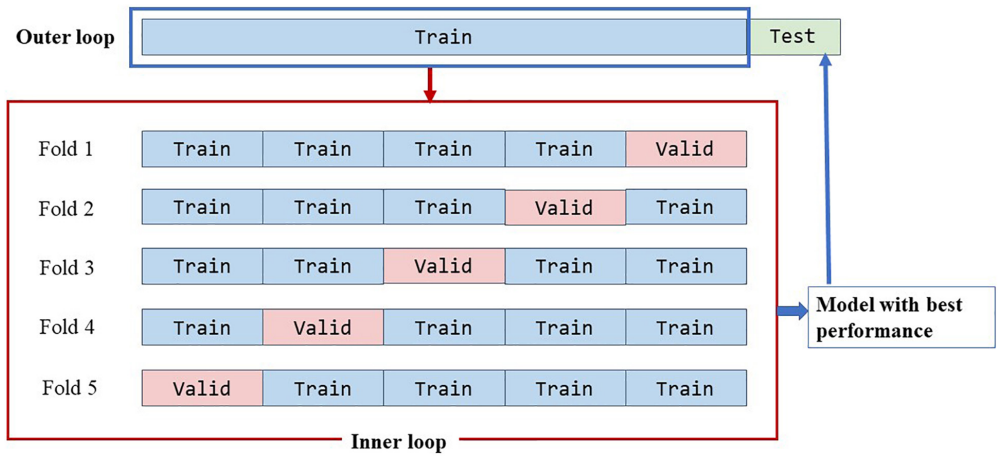


FIGURE 5 | Data spilt strategies.

TABLE 2 | Performance of different ResNet using two-stage model.

Encoder	Acc (95% CI)	Sens (95% CI)	Spec (95% CI)	F1 (95% CI)
ResNet-18	0.707 (0.674, 0.735)	0.679 (0.637, 0.720)	0.733 (0.698, 0.772)	0.691 (0.651, 0.723)
ResNet-34	0.716 (0.690, 0.741)	0.679 (0.644, 0.715)	0.750 (0.716, 0.790)	0.697 (0.670, 0.726)
ResNet-50	**0.819 (0.798, 0.841)**	**0.786 (0.754, 0.821)**	**0.850 (0.815, 0.877)**	**0.807 (0.783, 0.834)**
ResNet-101	0.759 (0.730, 0.783)	0.767 (0.731, 0.808)	0.750 (0.716, 0.785)	0.754 (0.724, 0.779)

The bold numbers denote the maximum value of each column.
Acc, Accuracy; CI, Confidence Interval; F1, F1-score; Sens, Sensitivity; Spec, Specificity.

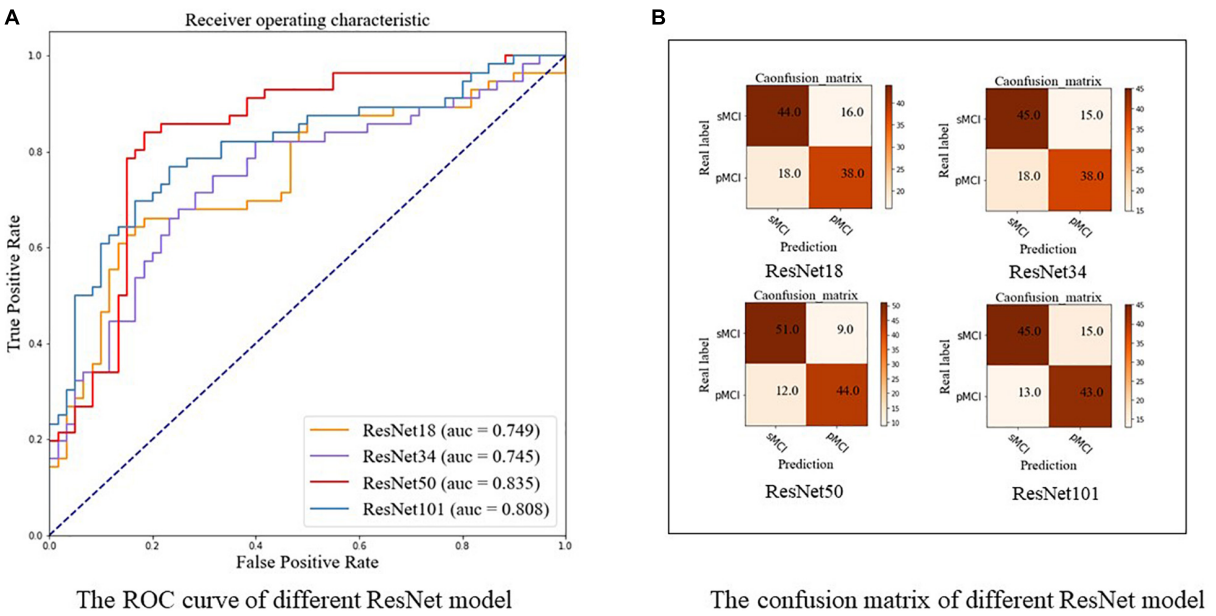


FIGURE 6 | The ROC curve and confusion matrix of different ResNet using our two-stage model. **(A)** The ROC curve. **(B)** The confusion matrix. pMCI, progressive mild cognitive impairment; sMCI, stable mild cognitive impairment.

model parameters and obtain general imaging features. Training on unlabeled target datasets using contrastive learning was used to get target imaging features. At last, the network was fine-tuned using the labeled target dataset to complete the classification. In addition, 3D Grad-CAM was used to highlight brain regions potentially associated with pMCI/sMCI classification. We demonstrated the validity of our model through multiple evaluation experiments. The experiments showed that our model

TABLE 3 | Performance of different transfer learning strategies using ResNet-50.

	Acc (95% CI)	Sens (95% CI)	Spec (95% CI)	F1 (95% CI)
Med3D	0.655 (0.630,0.682)	0.661 (0.623,0.702)	0.650 (0.609,0.685)	0.649 (0.619, 0.6768)
MoCo	0.733 (0.704, 0.756)	0.786 (0.747, 0.820)	0.683 (0.651, 0.725)	0.740 (0.712, 0.762)
Only ResNet-50	0.716 (0.694, 0.744)	0.750 (0.712, 0.789)	0.683 (0.656, 0.722)	0.718 (0.694, 0.744)
Med3D+MoCo	**0.819 (0.798, 0.841)**	**0.786 (0.754, 0.821)**	**0.850 (0.815, 0.877)**	**0.807 (0.783, 0.834)**

The bold numbers denote the maximum value of each column.
Acc, Accuracy; CI, Confidence Interval; F1, F1-score; MoCo, Momentum Contrast; Sens,Sensitivity; Spec, Specificity.

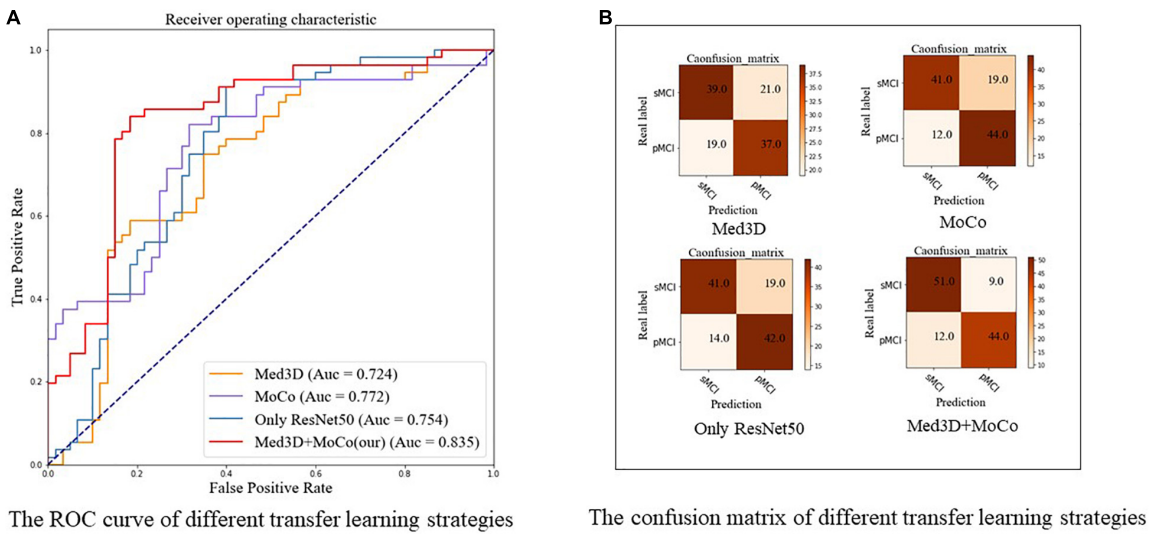


FIGURE 7 | The ROC curve and confusion matrix of different transfer learning strategies using ResNet-50. **(A)** The ROC curve. **(B)** The confusion matrix, respectively. All the models in this figure use ResNet-50. Med3D, pre-trained ResNet-50 using Med3D and fine-tune the model to complete the classification of sMCI and pMCI; MoCo, random initialization of weights, followed with the modified MoCo in the method to train the ResNet50 without using sMCI/pMCI labels. Then fine-tune the model using labeled data and complete the classification; Only ResNet, random initialization of weights and training ResNet from scratch; Med3D+MoCo, our two-stage transfer learning model; MoCo, Momentum Contrast; pMCI, progressive mild cognitive impairment; sMCI, stable mild cognitive impairment.

outperformed both transfer learning and contrastive learning individually and achieved better or comparable results than previous state-of-the-art studies.

Several factors improve the performance of our classification model. The first contribution is the proposal of a two-stage model. **Table 3** shows that the classification accuracy of ResNet-50 trained from scratch on sMCI and pMCI is 6.03% higher than that of ResNet-50 pre-trained in Med3D, implying that direct transfer learning for two data sets that are not highly correlated does not always achieve good results, and may even result in a negative transfer. The performance of transfer learning is affected by the various factors such as the size of pretrained samples, the relevance of the source and target domains. Thus, not all the transfer learning can improve the model's performance (Huh et al., 2016; Zhuang X. et al., 2019; Alzubaidi et al., 2020, 2021; Mustafa et al., 2021). For example, Alzubaidi et al. (2021) found that the model trained from scratch performed better than those pretrained by ImageNet using three different medical imaging datasets. This observation inspired us to develop a two-stage model. Our two-stage model is sample efficient when compared with existing transfer learning-based models for sMCI and pMCI classification (Oh et al., 2019; Gao et al., 2020; Bae et al., 2021). In each of these studies, additional AD and NC samples were collected for training. But our model does not require additional data collection, makes full use of each sample, and produces better or equivalent results. For example, Bae et al. (2021) developed a transfer learning model based on 3D ResNet29. In the source task, the model is pretrained using MRI scans of 2,084 normal samples and 1,406 AD samples. Then they used pMCI and sMCI samples to fine-tune the model to accomplish the target task of classifying pMCI and sMCI. In comparison to our results, they got the same accuracy but lower AUC.

To the best of our knowledge, we are the first to use the MoCo pretrained model for sMCI and pMCI classification. Compared with the models, only pretrained by Med3D and ResNet-50 trained from scratch, our method improved accuracy by 16.4 and 10.3%, respectively, further demonstrating the importance and efficiency of including contrastive learning into our method. Pretraining by contrastive learning allows the model to have a feature representation with better generalization at the same domain of the target task (Sun et al., 2019). Recent studies have shown that fine-tuning on a well-trained contrastive learning model can achieve comparable or even better results than fully supervised learning (Wu et al., 2018; Zhuang C. et al., 2019),

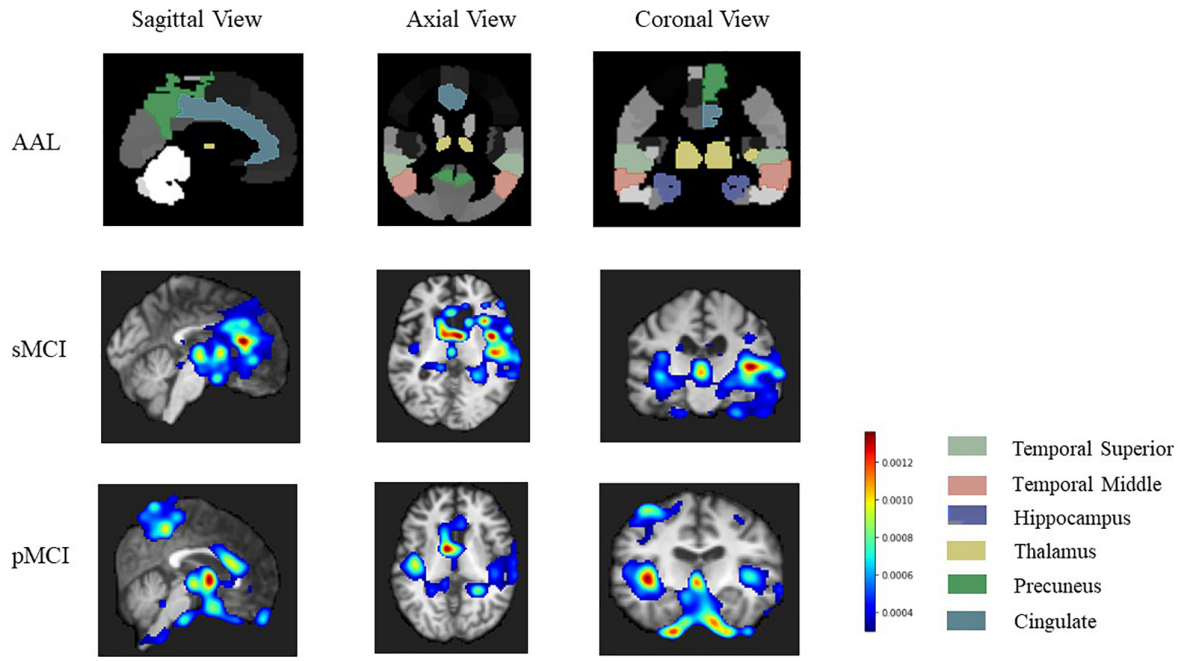


FIGURE 8 | The heatmap of related brain region our model focuses on. The first row is a golden standard of temporal superior, temporal middle, hippocampus, thalamus, precuneus, cingulate in automated anatomical labeling (AAL2, http://www.gin.cnrs.fr/en/tools/aal-aal2/). The second and third rows show brain regions that our model focuses on more on sMCI and pMCI, respectively. pMCI, progressive mild cognitive impairment; sMCI, stable mild cognitive impairment.

TABLE 4 | A summarized comparison of state-of-the-art research on MRI using deep learning in sMCI and pMCI classification.

Research	Conversion time	sMCI/pMCI number	Network	Acc	Sens	Spec	AUC
Liu M. et al. (2018)	36 Months	100/164	Landmark detection and 3D CNN	0.77	0.42	0.82	0.78
Liu J. et al. (2018)	18 Months	160/120	Whole brain hierarchical network	0.72	0.75	0.71	0.72
Lin et al. (2018)	36 Months	139/169	2D CNN	0.79	**0.86**	0.68	0.83
Shmulev and Belyaev (2018)	60 Months	532/327	3DResNet/VoxCNN	0.62	0.75	0.54	0.70
Shi et al. (2018)	18 Months	56/43	2D Deep polynomial network	0.79	0.68	0.87	0.80
Basaia et al. (2019)	36 months	253/510	3D CNN	0.75	0.75	0.75	NA
Li et al. (2019)	36 months	95/126	Self-weighting grading biomarker	0.69	0.82	0.51	0.70
Oh et al. (2019)	36 months	101/106	3D CNN + Transfer learning from CN/AD	0.73	0.77	0.71	0.79
Spasov et al. (2019)	36 months	228/181	3D CNN	0.72	0.63	0.81	0.79
Abrol et al. (2020)	24 Months	409/217	2D Multiscale Deep Neural Network	0.75	0.73	0.76	0.71
Gao et al. (2020)	36 Months	129/168	3D CNN + Transfer learning for AD age prediction	0.81	0.76	0.77	0.76
Pan et al. (2020)	18 Months	173/105	CNN and ensemble learning	0.62	NA	NA	0.59
Wen et al. (2020a)	36 Months	298/295	3D CNN	0.74	0.80	0.68	NA
Bae et al. (2021)	36 Months	222/228	3D ResNet29 + Transfer learning from CN /AD	**0.82**	0.72	NA	0.83
Guan et al. (2021)	36 Months	401/197	3D CNN	0.79	0.55	0.84	0.78
Zhang J. et al. (2021)	18 Months	251/162	3D DenseNet + Attention	0.79	0.75	0.82	**0.86**
Our	36 months	297/280	3D ResNet and transfer learning from self	**0.82**	0.79	**0.85**	0.84

The bold numbers denote the maximum value of each column.
Acc, Accuracy; AD, Alzheimer's disease; AUC, Area Under Curve; F1, F1-score; NC, Normal control; pMCI, progressive mild cognitive impairment; Sens, Sensitivity; Spec, Specificity; sMCI, stable mild cognitive impairment.

which is consistent with our findings. In addition, one of the critical factors limiting the performance of contrastive learning is the slow convergence rate (Chen T. et al., 2020; Chen X. et al., 2020; Tian et al., 2020). As shown in **Table 3**, compared with MoCo trained from scratch, our method improved accuracy by 6.89%, which indicates that transfer learning can accelerate the convergence of the MoCo model and improve the model performance. MoCo and transfer learning can reinforce and complement one other.

In addition, our model uses complete 3D MRI as model input. Unlike models using 2D slices, the 3D model makes full use of the spatial information of the brain to improve the accuracy of the

model (Wen et al., 2020b). Furthermore, some previous studies used feature engineering or cherry-picked regions of interest as input features (Gerardin et al., 2009; Ahmed et al., 2014; Basaia et al., 2019), which ignored the contributions of other features in the model, resulting in information loss in some cases. For example, Basaia et al. (2019) chose brain gray matter to train the model, neglecting the role of cerebrospinal fluid or white matter in early diagnosis of AD (Jack et al., 2010; Weiler et al., 2015). Our model differs from the previous studies by using an end-to-end model to learn from all possible features in medical images, which improves model performance.

In **Figure 8**, the hippocampus, temporal, and thalamus are highlighted in both sMCI and pMCI. Hippocampus and amygdala in the middle temporal lobe have been considered as crucial brain regions for the diagnosis of early AD (Visser et al., 2002; Braak and van Braak, 2004b; Burton et al., 2009; Costafreda et al., 2011). The hippocampus is essential for memory formation, and the recent studies have found that the hippocampus atrophy of pMCI is more pronounced than sMCI (Devanand et al., 2007; Risacher et al., 2009; Costafreda et al., 2011). Similarly, the amygdala, which is primarily responsible for emotion and expression, is intimately linked to emotional changes of AD, such as anxiety and irritability (Unger et al., 1991; Poulin et al., 2011). Thalamic damage is associated with decreased body movement and coordination, attention, and awareness in AD (Braak and van Braak, 2004a; de Jong et al., 2008; Cho et al., 2014; Aggleton et al., 2016). In addition to the hippocampus and temporal, which have been widely studied in AD, our heatmap also reveals that pMCI is also closely related to the precuneus with high-level memory and cognitive functions, which is in line with the previous studies (Whitwell et al., 2008; Bailly et al., 2015; Perez et al., 2015; Colangeli et al., 2016; Kato et al., 2016; Zhang H. et al., 2021). The results, as mentioned earlier, further indicate that some structural brain region abnormalities play an important role in predicting early AD. In summary, our discovery of important brain regions is supported by abundant literature, which helps construct a more comprehensive brain biomarker atlas to predict MCI progression.

CONCLUSION

In conclusion, our two-stage model increases both the accuracy of early AD detection as well as the transparency of the model. Notably, a comprehensive comparison of different 3D ResNet networks provides references for related research. Furthermore, the combination of transfer learning and contrastive learning solves the negative transfer problem and alleviates the model overfitting problem due to a lack of medical data. Notably, it also substantially improves the diagnostic performance of this tricky classification problem in neuroscience. Our model only uses low-invasive, low-cost, and widely available MRI data, which significantly expands the application scenarios of the model.

However, this study also has some limitations that merit additional exploration. First of all, we will explore more options for the model's various modules, such as different data augmentation methods and pretrained models on model effectiveness. When a larger dataset becomes available, we will also continue to validate our model. At final, it is worth noting that a direct comparison of different methods using the same evaluation metrics is straightforward but may not be the optimal solution. Factors such as sample size, dataset split strategy, sMCI, and pMCI definitions, and test data selection can have an impact on model outcomes. A more statistically robust comparison should be proposed in our future studies. Despite these limitations, our model provides a new solution to avoid overfitting because of the insufficient medical data and allows early identification of AD.

AUTHOR'S NOTE

The numerical calculations in this article have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

PL and LL designed the study and wrote the article. PL implemented the algorithm and preprocessed the data. LH, NZ, HL, and TT provided critical suggestions. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the National Natural Science Foundation of China (61772375, 61936013, and 71921002), the National Social Science Fund of China (18ZDA325), the National Key R&D Program of China (2019YFC0120003), the Natural Science Foundation of Hubei Province of China (2019CFA025), and the Independent Research Project of School of Information Management Wuhan University (413100032).

ACKNOWLEDGMENTS

We want to extend our sincere gratitude to Kang Jin for his editing and valuable suggestion on this work. We are also deeply indebted to Lei Li and Dingfeng Wu, whose recommendations make our work better.

REFERENCES

Abrol, A., Bhattarai, M., Fedorov, A., Du, Y., Plis, S., and Calhoun, V. (2020). Deep residual learning for neuroimaging: an application to predict progression to Alzheimer's disease. *J. Neurosci. Methods* 339:108701. doi: 10.1016/j.jneumeth.2020.108701

Abrol, A., Fu, Z., Salman, M. S., Silva, R. F., Du, Y., Plis, S., et al. (2021). Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning. *Nat. Commun.* 12:353. doi: 10.1038/s41467-020-20655-6

Afzal, S., Maqsood, M., Nazir, F., Khan, U., Aadil, F., Awan, K. M., et al. (2019). A Data Augmentation-Based Framework to Handle Class Imbalance Problem for Alzheimer's Stage Detection. *IEEE Access* 7, 115528–115539. doi: 10.1109/access.2019.2932786

Aggleton, J. P., Pralus, A., Nelson, A. J. D., and Hornberger, M. (2016). Thalamic pathology and memory loss in early Alzheimer's disease: moving the focus from the medial temporal lobe to Papez circuit. *Brain* 139, 1877–1890. doi: 10.1093/brain/aww083

Ahmed, O., Ben, Benois-Pineau, J., Allard, M., Amar, C., Ben, et al. (2014). Classification of Alzheimer_s disease subjects from MRI using hippocampal visual features. *Multimed. Tools Appl.* 74, 1249–1266. doi: 10.1007/s11042-014-2123-y

Alzubaidi, L., Duan, Y., Al-dujaili, A., Ibraheem, I. K., Alkenani, A. H., Santamaría, J., et al. (2021). Deepening into the suitability of using pre-trained models of ImageNet against a lightweight convolutional neural network in medical imaging: an experimental study. *PeerJ Comput. Sci.* 7:e715. doi: 10.7717/peerj-cs.715

Alzubaidi, L., Fadhel, M. A., Al-Shamma, O., Zhang, J., Santamaría, J., Duan, Y., et al. (2020). Towards a Better Understanding of Transfer Learning for Medical Imaging: a Case Study. *Appl. Sci.* 10:4523. doi: 10.1016/j.morpho.2019.09.001

Association, A. (2019). 2019 Alzheimer's disease facts and figures. *Alzheimer's & Dement.* 15, 321–387.

Bae, J., Stocks, J., Heywood, A., Jung, Y., Jenkins, L. M., Hill, V. B., et al. (2021). Transfer learning for predicting conversion from mild cognitive impairment to dementia of Alzheimer's type based on a three-dimensional convolutional neural network. *Neurobiol. Aging* 99, 53–64. doi: 10.1016/j.neurobiolaging.2020.12.005

Bailly, M., Destrieux, C., Hommet, C., Mondon, K., Cottier, J. P., Beaufils, É, et al. (2015). Precuneus and Cingulate Cortex Atrophy and Hypometabolism in Patients with Alzheimer's Disease and Mild Cognitive Impairment: MRI and 18F-FDG PET Quantitative Analysis Using FreeSurfer. *Biomed Res. Int* 2015:583931. doi: 10.1155/2015/583931

Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., et al. (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NEUROIMAGE-CLINICAL* 21:101645. doi: 10.1016/j.nicl.2018.101645

Bernal, J., Kushibar, K., Asfaw, D. S., Valverde, S., Oliver, A., Marti, R., et al. (2019). Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. *Artif. Intell. Med.* 95, 64–81. doi: 10.1016/j.artmed.2018.08.008

Braak, H., and van Braak, E. (2004b). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol.* 82, 239–259. doi: 10.1007/BF00308809

Braak, H., and van Braak, E. (2004a). Alzheimer's disease affects limbic nuclei of the thalamus. *Acta Neuropathol.* 81, 261–268. doi: 10.1007/BF00305867

Burton, E. J., Barber, R., Mukaetova-Ladinska, E. B., Robson, J. I., Perry, R. H., Jaros, E., et al. (2009). Medial temporal lobe atrophy on MRI differentiates Alzheimer's disease from dementia with Lewy bodies and vascular cognitive impairment: a prospective study with pathological verification of diagnosis. *Brain* 132, 195–203. doi: 10.1093/brain/awn298

Carreira, J., and Zisserman, A. (2017). *Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset. 2017 IEEE Conf. Comput. Vis. Pattern Recognit*. Piscataway: IEEE, 4724–4733.

Chen, S., Ma, K., and Zheng, Y. (2019). Med3D: transfer learning for 3D medical image analysis. *arXiv* [Preprint]. arXiv:1904.00625

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. E. (2020). *A Simple Framework for Contrastive Learning of Visual Representations. ArXiv abs/2002.0*.

Chen, X., Fan, H., Girshick, R. B., and He, K. (2020). Improved baselines with momentum contrastive learning. *arXiv* [Preprint]. arXiv:2003.04297

Cho, H., Kim, J., Kim, C., Ye, B. S., Kim, H. J., Yoon, C. W., et al. (2014). Shape changes of the basal ganglia and thalamus in Alzheimer's disease: a three-year longitudinal study. *J. Alzheimers. Dis.* 40, 285–295. doi: 10.3233/JAD-132072

Choi, H., and Jin, K. (2018). Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging. *Behav. Brain Res.* 344, 103–109. doi: 10.1016/j.bbr.2018.02.017

CodaLab. (2021). *Competition*. Available online at: https://competitions.codalab.org/competitions/17094#results (Accessed October 1, 2021).

Colangeli, S., Boccia, M., Verde, P. A., La, Guariglia, P., Bianchini, F., et al. (2016). Cognitive Reserve in Healthy Aging and Alzheimer's Disease. *Am. J. Alzheimer's Dis. Other Dementias* 31, 443–449.

Costafreda, S. G., Dinov, I. D., Tu, Z., Shi, Y., Liu, C.-Y., Kloszewska, I., et al. (2011). Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. *Neuroimage* 56, 212–219. doi: 10.1016/j.neuroimage.2011.01.050

de Jong, L. W., van der Hiele, K., Veer, I. M., Houwing, J. J., Westendorp, R. G. J., Bollen, E. L. E. M., et al. (2008). Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. *Brain* 131, 3277–3285. doi: 10.1093/brain/awn278

Devanand, D. P., Pradhaban, G., Liu, X., Khandji, A. G., de Santi, S., Segal, S., et al. (2007). Hippocampal and entorhinal atrophy in mild cognitive impairment. *Neurology* 68, 828–836. doi: 10.1212/01.wnl.0000256697.20968.d7

Efron, B., and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. New York: Chapman and Hall/CRC.

Gao, F., Yoon, H., Xu, Y., Goradia, D., Luo, J., Wu, T., et al. (2020). AD-NET: age-adjust neural network for improved MCI to AD conversion prediction. *Neuroimage* 27:102290. doi: 10.1016/j.nicl.2020.102290

Gerardin, E., Chetelat, G., Chupin, M., Cuingnet, R., Desgranges, B., Kim, H. S., et al. (2009). Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *Neuroimage* 47, 1476–1486. doi: 10.1016/j.neuroimage.2009.05.036

Guan, H., Wang, C., Cheng, J., Jing, J., and Liu, T. (2021). A parallel attention_augmented bilinear network for early magnetic resonance imaging-based diagnosis of Alzheimer's disease. *Hum. Brain Mapp.* 43, 760–772. doi: 10.1002/hbm.25685

He, K., Fan, H., Wu, Y., Xie, S., and Girshick, R. B. (2020). "Momentum contrast for unsupervised visual representation learning," in *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, 9726–9735.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Piscataway, NJ: IEEE), 770–778.

Heun, R., Mazanek, M., Atzor, K.-R., Tintera, J., Gawehn, J., Burkart, M., et al. (1997). Amygdala-Hippocampal Atrophy and Memory Performance in Dementia of Alzheimer Type. *Dement. Geriatr. Cogn. Disord.* 8, 329–336. doi: 10.1159/000106651

Huh, M., Agrawal, P., and Efros, A. A. (2016). What makes ImageNet good for transfer learning? *arXiv* [Preprint]. arXiv:1608.08614

Irvin, J. A., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., et al. (2019). *CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison*. California: AAAI Press.

Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., et al. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J. Magn. Reson. IMAGING* 27, 685–691. doi: 10.1002/jmri.21049

Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., et al. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol.* 9, 119–128. doi: 10.1016/S1474-4422(09)70299-6

Johnson, K. A., Fox, N. C., Sperling, R. A., and Klunk, W. E. (2012). Brain imaging in Alzheimer disease. *Cold Spring Harb. Perspect. Med.* 2:a006213.

Kato, T., Inui, Y., Nakamura, A., and Ito, K. (2016). Brain fluorodeoxyglucose (FDG) PET in dementia. *Ageing Res. Rev.* 30, 73–84. doi: 10.1016/j.arr.2016.02.003

Klöppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., et al. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain* 131), 681–689. doi: 10.1093/brain/awm319

Li, Y., Fang, Y., Zhang, H., and Hu, B. (2019). Self-Weighting Grading Biomarker Based on Graph-Guided Information Propagation for the Prediction of Mild Cognitive Impairment Conversion. *IEEE Access* 7, 116632–116642.

Lin, W. M., Tong, T., Gao, Q. Q., Guo, D., Du, X. F., Yang, Y. G., et al. (2018). Convolutional Neural Networks-Based MRI Image Analysis for the Alzheimer's Disease Prediction From Mild Cognitive Impairment. *Front. Neurosci* 12:777. doi: 10.3389/fnins.2018.00777

Liu, J., Li, M., Lan, W., Wu, F.-X., Pan, Y., and Wang, J. (2018). Classification of Alzheimer's Disease Using Whole Brain Hierarchical Network. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 15, 624–632. doi: 10.1109/TCBB.2016.2635144

Liu, M., Zhang, J., Adeli, E., and Shen, D. (2018). Landmark-based deep multi-instance learning for brain disease diagnosis. *Med. Image Anal* 43, 157–168. doi: 10.1016/j.media.2017.10.005

Liu, X., Zhang, F., Hou, Z., Wang, Z., Mian, L., Zhang, J., et al. (2020). Self-supervised learning: generative or contrastive. *arXiv* [Preprint]. arXiv:2006.08218

Liu, Y., Li, Z., Ge, Q., Lin, N., and Xiong, M. (2019). Deep Feature Selection and Causal Analysis of Alzheimer's Disease. *Front. Neurosci.* 13:1198. doi: 10.3389/fnins.2019.01198

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., et al. (2017). Dementia prevention, intervention, and care. *Lancet* 390, 2673–2734.

Medical Segmentation Decathlon. (2021). *Medical Segmenation Decathlon*. Available online at: http://medicaldecathlon.com/index.html (Accessed October 1, 2021)

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J. S., et al. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans. Med. Imaging* 34, 1993–2024.

Misra, I., and van der Maaten, L. (2020). "Self-supervised learning of pretext-invariant representations," in *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, 6706–6716. doi: 10.1109/CVPR42600.2020.00674

Mousavian, M., Chen, J., and Greening, S. G. (2019). "Depression detection using feature extraction and deep learning from sMRI images," in *Proceedings of the 2019 18th IEEE International Conference On Machine Learning and Applications (ICMLA)*, Boca Raton, FL, 1731–1736.

Mustafa, B., Loh, A., von Freyberg, J., MacWilliams, P., Wilson, M., McKinney, S. M., et al. (2021). Supervised transfer learning at scale for medical imaging. *arXiv* [Preprint]. arXiv:2101.05913

Naz, S., Ashraf, A., and Zaib, A. (2021). Transfer learning using freeze features for Alzheimer neurological disorder detection using ADNI dataset. *Multimed. Syst.* 28, 85–94.

Oh, K., Chung, Y.-C., Kim, K. W., Kim, W.-S., and Oh, I.-S. (2019). Classification and Visualization of Alzheimer's Disease using Volumetric Convolutional Neural Network and Transfer Learning. *Sci. Rep.* 9, 18150.

Pan, D., Zeng, A., Jia, L., Huang, Y., Frizzell, T., and Song, X. (2020). Early Detection of Alzheimer's Disease Using Magnetic Resonance Imaging: a Novel Approach Combining Convolutional Neural Networks and Ensemble Learning. *Front. Neurosci.* 14:259. doi: 10.3389/fnins.2020.00259

Perez, S. E., He, B., Nadeem, M., Wuu, J., Scheff, S. W., Abrahamson, E. E., et al. (2015). Resilience of Precuneus Neurotrophic Signaling Pathways Despite Amyloid Pathology in Prodromal Alzheimer's Disease. *Biol. Psychiatr.* 77, 693–703. doi: 10.1016/j.biopsych.2013.12.016

Petersen, R. C. (2000). Mild cognitive impairment: transition between aging and Alzheimer's disease. *Neurologia* 15, 93–101.

Poulin, S., Dautoff, R., Morris, J. C., Barrett, L. F., and Dickerson, B. C. (2011). Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. *Psychiatry Res. Neuroimaging* 194, 7–13. doi: 10.1016/j.pscychresns.2011.06.014

Raghu, M., Zhang, C., Kleinberg, J., and Bengio, S. (2019). "Transfusion: understanding transfer learning for medical imaging," in *Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, Vancouver, BC.

Risacher, S. L., Saykin, A. J., West, J. D., Shen, L., Firpi, H. A., and McDonald, B. C. (2009). Baseline MRI Predictors of Conversion from MCI to Probable AD in the ADNI Cohort. *Curr. Alzheimer Res.* 6, 347–361. doi: 10.2174/156720509788929273

Roberson, E. D., and Mucke, L. (2006). 100 Years and Counting: prospects for Defeating Alzheimer's Disease. *Science (80-.).* 314, 781–784. doi: 10.1126/science.1132813

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* 115, 211–252. doi: 10.1007/s11263-015-0816-y

Shen, D. G., Wu, G. R., and Suk, H. I. (2017). "Deep Learning in Medical Image Analysis,". 19, 221–248. doi: 10.1146/annurev-bioeng-071516-044442

Shi, J., Zheng, X., Li, Y., Zhang, Q., and Ying, S. H. (2018). Multimodal Neuroimaging Feature Learning With Multimodal Stacked Deep Polynomial Networks for Diagnosis of Alzheimer's Disease. *IEEE J. Biomed. Heal. INFORMATICS* 22, 173–183. doi: 10.1109/JBHI.2017.2655720

Shmulev, Y., and Belyaev, M. (2018). "Predicting conversion of mild cognitive impairments to Alzheimer's disease and exploring impact of neuroimaging," in *Graphs in Biomedical Image Analysis and Integrating Medical Imaging and Non-Imaging Modalities. GRAIL 2018, Beyond MIC 2018. Lecture Notes in Computer Science*, Vol. 11044, eds D. Stoyanov, et al. (Cham: Springer). doi: 10.1007/978-3-030-00689-1_9

Spasov, S. E., Passamonti, L., Duggento, A., Liu, P., and Toschi, N. (2019). A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. *Neuroimage* 189, 276–287. doi: 10.1016/j.neuroimage.2019.01.031

Sun, Y., Tzeng, E., Darrell, T., and Efros, A. A. (2019). Unsupervised domain adaptation through self-supervision. *arXiv* [Preprint]. arXiv:1909.11825

Tajbakhsh, N., Shin, J. Y., Gurudu, S., Hurst, R. T., Kendall, C. B., Gotway, M., et al. (2016). Convolutional Neural Networks for Medical Image Analysis: full Training or Fine Tuning? *IEEE Trans. Med. Imaging* 35, 1299–1312. doi: 10.1109/TMI.2016.2535302

Tian, Y., Sun, C., Poole, B., Krishnan, D., Schmid, C., and Isola, P. (2020). What makes for good views for contrastive learning. *arXiv* [Preprint]. arXiv:2005.10243

Tobon-Gomez, C., Geers, A. J., Peters, J., Weese, J., Pinto, K., Karim, R., et al. (2015). Benchmark for Algorithms Segmenting the Left Atrium From 3D CT and MRI Datasets. *IEEE Trans. Med. Imaging* 34, 1460–1473. doi: 10.1109/TMI.2015.2398818

Tran, D., Bourdev, L. D., Fergus, R., Torresani, L., and Paluri, M. (2014). *C3D: Generic Features for Video Analysis. ArXiv abs/1412.0*.

Tustison, N., Avants, B., Cook, P., Zheng, Y., Egan, A., Yushkevich, P., et al. (2010). N4ITK: improved N3 Bias Correction. *IEEE Trans. Med. Imaging* 29, 1310–1320. doi: 10.1109/TMI.2010.2046908

Unger, J. W., Lapham, L. W., McNeill, T. H., Eskin, T. A., and Hamill, R. W. (1991). The amygdala in Alzheimer's disease: neuropathology and Alz 50 Immunoreactivity. *Neurobiol. Aging* 12, 389–399. doi: 10.1016/0197-4580(91)90063-p

Visser, P. J., Verhey, F. R. J., Hofman, P. A. M., Scheltens, P., and Jolles, J. (2002). Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. *J. Neurol. Neurosurg. Psychiatr.* 72, 491–497. doi: 10.1136/jnnp.72.4.491

Weiler, M., Agosta, F., Canu, E., Copetti, M., Magnani, G., Marcone, A., et al. (2015). Following the Spreading of Brain Structural Changes in Alzheimer's Disease: a Longitudinal. *Multimodal MRI Study. J. Alzheimers. Dis.* 47, 995–1007. doi: 10.3233/JAD-150196

Wen, J., Thibeau-Sutre, E., Samper-González, J., Routier, A., Bottani, S., Durrleman, S., et al. (2020b). Convolutional Neural Networks for Classification of Alzheimer's Disease: overview and Reproducible Evaluation. *Med. Image Anal.* 63:101694. doi: 10.1016/j.media.2020.101694

Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-González, J., Routier, A., Bottani, S., et al. (2020a). Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation. *Med. Image Anal.* 63:101694. doi: 10.1016/J.MEDIA.2020.101694

Whitwell, J. L., Shiung, M. M., Przybelski, S. A., Weigand, S. D., Knopman, D. S., Boeve, B. F., et al. (2008). MRI patterns of atrophy associated with progression to AD in amnestic mild cognitive impairment. *Neurology* 70, 512–520. doi: 10.1212/01.wnl.0000280575.77437.a2

Wu, Z., Xiong, Y., Yu, S. X., and Lin, D. (2018). "Unsupervised feature learning via non-parametric instance discrimination," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, 3733–3742. doi: 10.1109/CVPR.2018.00393

Yang, X., Bian, C., Yu, L., Ni, D., and Heng, P. (2017). "Hybrid Loss Guided Convolutional Networks for Whole Heart Parsing," in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*, eds M. Pop, et al. (Cham: Springer).

Zeng, G., and Zheng, G. (2018). Deep Learning-Based Automatic Segmentation of the Proximal Femur from MR Images. *Adv. Exp. Med. Biol.* 1093, 73–79. doi: 10.1007/978-981-13-1396-7_6

Zhang, H., Wang, Y., Lyu, D., Li, Y., Li, W., Wang, Q., et al. (2021). Cerebral blood flow in mild cognitive impairment and Alzheimer's disease: a systematic review and meta-analysis. *Ageing Res. Rev* 71:101450. doi: 10.1016/j.arr.2021.101450

Zhang, J., Zheng, B., Gao, A., Feng, X., Liang, D.-P., and Long, X. (2021). A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification. *Magn. Reson. Imaging* 78, 119–126. doi: 10.1016/j.mri.2021.02.001

Zhuang, C., Zhai, A., and Yamins, D. (2019). "Local aggregation for unsupervised learning of visual embeddings," in *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 6001–6011. doi: 10.1109/ICCV.2019.00610

Zhuang, X., Li, Y., Hu, Y., Ma, K., Yang, Y., and Zheng, Y. (2019). "Self-supervised Feature Learning for 3D Medical Images by Playing a Rubik's Cube," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. MICCAI 2019*, eds D. Shen, et al. (Cham: Springer). doi: 10.1016/j.media.2020.101746

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

METHODS
published: 13 April 2022
doi: 10.3389/fnagi.2022.865558

Sleep EEG-Based Approach to Detect Mild Cognitive Impairment

Duyan Geng[1,2], Chao Wang[2], Zhigang Fu[3], Yi Zhang[2], Kai Yang[2] and Hongxia An[2]*

[1] State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin, China, [2] Key Laboratory of Electromagnetic Field and Electrical Apparatus Reliability of Hebei Province, Hebei University of Technology, Tianjin, China, [3] Physical Examination Center, The 983 Hospital of Joint Logistics Support Force of the Chinese People's Liberation Army, Tianjin, China

Mild Cognitive Impairment (MCI) is an early stage of dementia, which may lead to Alzheimer's disease (AD) in older adults. Therefore, early detection of MCI and implementation of treatment and intervention can effectively slow down or even inhibit the progression of the disease, thus minimizing the risk of AD. Currently, we know that published work relies on an analysis of awake EEG recordings. However, recent studies have suggested that changes in the structure of sleep may lead to cognitive decline. In this work, we propose a sleep EEG-based method for MCI detection, extracting specific features of sleep to characterize neuroregulatory deficit emergent with MCI. This study analyzed the EEGs of 40 subjects (20 MCI, 20 HC) with the developed algorithm. We extracted sleep slow waves and spindles features, combined with spectral and complexity features from sleep EEG, and used the SVM classifier and GRU network to identify MCI. In addition, the classification results of different feature sets (including with sleep features from sleep EEG and without sleep features from awake EEG) and different classification methods were evaluated. Finally, the MCI classification accuracy of the GRU network based on features extracted from sleep EEG was the highest, reaching 93.46%. Experimental results show that compared with the awake EEG, sleep EEG can provide more useful information to distinguish between MCI and HC. This method can not only improve the classification performance but also facilitate the early intervention of AD.

Keywords: mild cognitive impairment, sleep EEG, sleep slow waves, sleep spindles, machine learning

OPEN ACCESS

Edited by:
Peng Xu,
University of Electronic Science and Technology of China, China

Reviewed by:
Carmen Jiménez-Mesa,
University of Granada, Spain
Carlos Tomaz,
Universidade Ceuma, Brazil

****Correspondence:***
Duyan Geng
dygeng@hebut.edu.cn

Specialty section:
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

Received: *30 January 2022*
Accepted: *07 March 2022*
Published: *13 April 2022*

Citation:
Geng D, Wang C, Fu Z, Zhang Y, Yang K and An H (2022) Sleep EEG-Based Approach to Detect Mild Cognitive Impairment. Front. Aging Neurosci. 14:865558. doi: 10.3389/fnagi.2022.865558

INTRODUCTION

As the aging of the population becomes increasingly serious, Alzheimer's disease has become a major challenge to human health and a serious social problem. Alzheimer's disease (AD) is the most common type of dementia, accounting for roughly 70% of all dementias worldwide. It is an irreversible neurodegenerative disease marked by cognitive, behavioral, and intellectual impairments (Prince, 2015). Mild cognitive impairment (MCI) is a pre-dementia condition in which daily functioning is usually maintained despite objectively measured cognitive impairment in one or more cognitive domains. As MCI is the primary stage of cognitive impairment, about 10–15% of MCI patients will progress to AD every year on average, and about 2/3 of AD patients

are developed from MCI (Tsai et al., 2016). As a result, early detection of MCI is critical for early intervention in the preclinical stage of AD and has attracted much attention from researchers in recent decades (Jiang et al., 2020).

According to recent studies, patients with MCI may return to normal over time, therefore early detection and diagnosis of MCI are critical (Amezquita-Sanchez et al., 2019). Early detection of cognitive decline can lead to appropriate interventions before further cognitive impairment occurs, thus delaying or even preventing the progression of dementia as much as possible. It is estimated that the average annual cost per patient for mild dementia is $15,889, for moderate dementia is $26,859, and for severe dementia is $36,180. Prevention of the disease is therefore important for better health care, as well as for national financial interests, and for controlling the progression of cognitive impairment.

Finding biomarkers with low cost, high specificity, and sensitivity has been the focus of MCI research. Magnetic resonance imaging, such as functional magnetic resonance imaging (fMRI) (Ni et al., 2017), magnetic resonance spectral imaging (MRS) (Gao and Barker, 2014), diffusion-weighted imaging (DWI) (Ge, 2017), diffusion tensor imaging (DTI) (Ahmed et al., 2017), positron emission tomography, such as fluorodeoxyglucose positron emission tomography (FDG-PET) (Karow et al., 2010), etc., cerebrospinal fluid markers (Handels et al., 2017), such as Aβ40, Aβ42, total tau protein (t-tau), etc., are currently the main methods for early diagnosis of MCI. However, these methods are expensive, and the equipment is large, with high expertise requirements. As a result, researchers are looking for non-invasive, quick, low-cost, and dependable approaches for disease detection (Alberdi et al., 2016).

Biomarkers based on EEG have emerged as a viable tool in the research of AD. In terms of EEG acquisition methods, most of the published work relies on the analysis of closed resting state EEG (rsEEG) recordings. Waninger et al. (2016) used fast Fourier transform to calculate Power Spectral Density (PSD) to study the difference between MCI and HC and found that there were significant differences in theta and alpha frequency bands, which were classified by linear discriminant analysis. The final classification result was 85.11%. Rodrigues et al. (2021) conducted statistical analysis of lacstral distances between EEG subbands and found a metric that could identify AD at all stages and characterize AD activity in each electrode, achieving a classification accuracy of 98.06% with an artificial neural network. Cassani and Falk (2020) proposed a new feature characterized by a two-dimensional modulation spectrum domain based on rsEEG signals, collected EEG signals of 20 channels, and obtained classification accuracy of 88.1% by SVM classification of MCI and HC. Other papers extract evoked potentials by giving specific stimuli to the nervous system to detect and classify MCI and other disorders that affect cognitive states. For instance, Khatun et al. (2019) proposed an MCI detection method based on single-channel EEG, which stimulated auditory speech signals, extracted features from event-related potential (ERP), and obtained an accuracy of 87.9% by SVM classification. Although the literature has reported levels of accuracy above 80%, it has been difficult to evaluate studies and determine the most advanced approaches due to variances in experimental settings, data collection methods, and database sizes.

Changes in sleep electrophysiology may be linked to the cognitive condition of AD and MCI patients, according to recent research (Gorgoni et al., 2020). Local sleep EEG oscillations have a critical function in learning and plasticity mechanisms, it's worth highlighting. Several electrophysiological aspects of NREM (such as slow waves, sleep spindles, and hippocampal ripples) and REM sleep (such as θ activity) are particularly active in memory consolidation (Klinzing et al., 2019). Sleep EEG can identify the sleep changes associated with AD and MCI pathology, and is low-cost and portable, so it can be utilized to make quick and precise diagnoses (D'Atri et al., 2021). This is the main motivation of this study. Various studies have employed multi-channel EEG data to characterize MCI or AD using EEG signals. Although there are several EEG-based studies in the literature, no one has attempted to detect and classify MCI using two-channel sleep EEG signals, as far as we are aware. Compared with the multi-channel, the number of two-channel leads is less, and the measurement method is simple. Moreover, compared with the EEG signal during waking, the EEG signal collected during sleep is stable and easy to be disturbed, which is more conducive to the study of neurodegenerative diseases.

This study proposes a new method to distinguish the EEG signals of MCI and health control (HC). We use from the C3 and C4 (central electrode) dual channel sleep EEG data with labels, based on the sleep slow waves, spindles, power spectral density and complexity, the use of machine learning and deep learning methods classifying MCI, and the classification results were compared with the awake EEG classification results without sleep features. The paper is organized as follows: section "Materials" describes the data set used for this work, and the section "Methodology" introduces the signal processing methods, including the detection of sleep slow waves and spindles, as well as the method of feature extraction and classification. Section "Result and Discussion" is the experimental results and the discussion of the paper. Finally, the conclusions are contained in section "Conclusion."

MATERIALS

Data and Materials

All data were obtained from NSRR (Zhang et al., 2018). To balance the data, 40 subjects (20 patients with MCI and 20 healthy subjects as controls, all women) with polysomnography (PSG) signals were randomly selected in the SOF study (Spira et al., 2008). All data were approved by the local institutional review board of the institution, and each participant provided written, informed consent before participation. The data included functional tests, cognitive exams, use of medication, health habits, and much more. All subjects underwent the Mini-mental State Examination (MMSE) to determine the severity of their impairment or dementia. According to the classification method of dementia severity, the MMSE score between 21 and 26 was considered MCI, and the MMSE score greater than 26 was

normal. The comprehensive demographic information of the subjects in this study is shown in **Table 1**.

MCI exclusion criteria are as follows:

(i) a history of depression (mild to moderate or major depression) or a history of adolescent paroxysmal mental illness; (ii) a history of major stroke or neurological symptoms; (iii) Other mental disorders, frontotemporal dementia, Lewy body dementia, vascular dementia, epilepsy, alcohol dependence; (iv) The use of psychoactive drugs, which modulate EEG markers; And (v) current or previously uncontrolled or complex systemic diseases (including diabetes), or traumatic brain injury (Moretti et al., 2013).

SOF data includes EDF and annotation files that include manually graded sleep stages in 30-s epochs, as well as manual annotations for arousal, limb movement, and signal artifacts. All experiments employed the American Academy of Sleep Medicine (AASM) staging, with NREM3 and NREM4 compressed to the N3 stage, and electrode labels were taken from the International 10-20 system.

EEG Data and Preprocessing

EEG channels are selected from the multi-channel PSG signals, and the signals are down-sampled to 100 Hz to speed up the calculation time. A text file of the sleep stage vector (also known as hypnogram) is then loaded at a sampling frequency of 1/30 and sampled upward to match the sampling frequency and length of the EEG signal.

Sleep changes are a core component of MCI and AD patients. The decrease in slow-wave activity (SWA) during NREM is due in part to amyloid disease and leads to cognitive decline in older adults (Rosinvil et al., 2020). Furthermore, studies have revealed that patients with MCI often have a substantial decrease in spindles, which is associated with cognitive decline in dementia patients (Gorgoni et al., 2016). Because sleep quality degradation is one of the primary symptoms of MCI, a variation in EEG activity during NREM sleep could be a possible biomarker (Romanella et al., 2021). Spindles are particularly noticeable during N2 sleep and are a distinguishing feature of this stage, while sleep slow waves are present during both N2 and N3. Therefore, for the main analysis, a hypnogram was used to extract the EEG signals of the N2 and N3 stages and the EEG signals of the waking stage for comparison, so as to verify the importance of sleep features.

First, the EEG signals were bandpass filtered between 0.1 and 30 Hz, and the artifact signals were eliminated using independent component analysis (ICA). The signals of each channel are divided into 5-min non-overlapping segments, and the abnormal segment is removed using the standard deviation-based rejection method. This method is speedier and is based solely on the standard deviation distribution of each segment. First, the standard deviation of each segment and channel is determined, and the resulting array of standard deviations is log-transformed and z-scored. Any epochs that have one or more channels that surpass the threshold will be labeled as an artifact. Because this method is more sensitive to the effects of noise, any segment with overlapping wake, motion, or signal artifact annotations have been removed before using this method. Accordingly, the segment of each subject was evaluated which accounts for a total of 2063 segments for NREM sleep and a total of 768 segments for wakefulness are analyzed.

TABLE 1 | Corresponding statistical information of subjects.

	Subjects	Age (year)	MMSE
HC	20 (female)	82.95 ± 2.71	29.6 ± 0.73
MCI	20 (female)	84.05 ± 3.51	24.4 ± 0.97

METHODOLOGY

Figure 1 describes a proposed algorithm for classifying EEG segments from MCI and HC. As shown in the figure, the algorithm consists of three steps. The first step is to calculate the power spectral density of each frequency band. For the sleep EEG signals, we mainly focus on sleep slow waves and spindles during NREM sleep, and use the YASA algorithm to detect sleep slow waves and spindles. Secondly, extract the features of sleep slow waves and spindles and calculate the spectral and complexity features. Thirdly, train the classifier with the extracted features and evaluate the test results. In addition, to confirm the validity of the sleep features reported in this work in MCI classification, we extracted the spectral and complexity features from the awake EEG for comparison. Two classification methods were used to verify the classification effect of EEG signals during wakefulness and sleep, and the test results were evaluated and compared.

Power Spectral Density

PSD is a frequency-dependent measure of the mean power distribution. Because EEG slowing is the main linear indicator of cognitive decline, spectral analysis is a critical parameter for measuring neurocognitive impairment (Sharma et al., 2016). Welch's periodogram is the most generally used method for calculating the estimated value of PSD (Welch, 1967), It involves averaging sequential Fourier transforms of small windows of the signal, with or without overlapping.

The PSD was calculated every 5 min by the Welch method, using 5 s of hamming windows, with 50% overlap, and median-averaging to limit the influence of artifacts. A commonly used method of determining the window width is to adopt a window long enough to contain at least two full minimum frequency periods of interest. The lowest frequency of interest here is 0.5 Hz, so we can choose a window that is greater than or equal to 4 s. Here we chose a 5 s-long window. The following is the formula for calculating power spectral density using the Welch method:

$$P(\omega) = \frac{1}{MUL}\sum_{i=1}^{L}\left|\sum_{n=0}^{M-1} x^{i}(n)d_2(n)e^{-j\omega n}\right|^2$$

Where M is the window length, $U = \frac{1}{M}\sum_{n=0}^{M-1} {d_2}^2(n)$, $L = \frac{N-M/2}{M/2}$, the $x^i(n)$ is the signal for each window, and the $d_2(n)$ is the function of Hamming window.

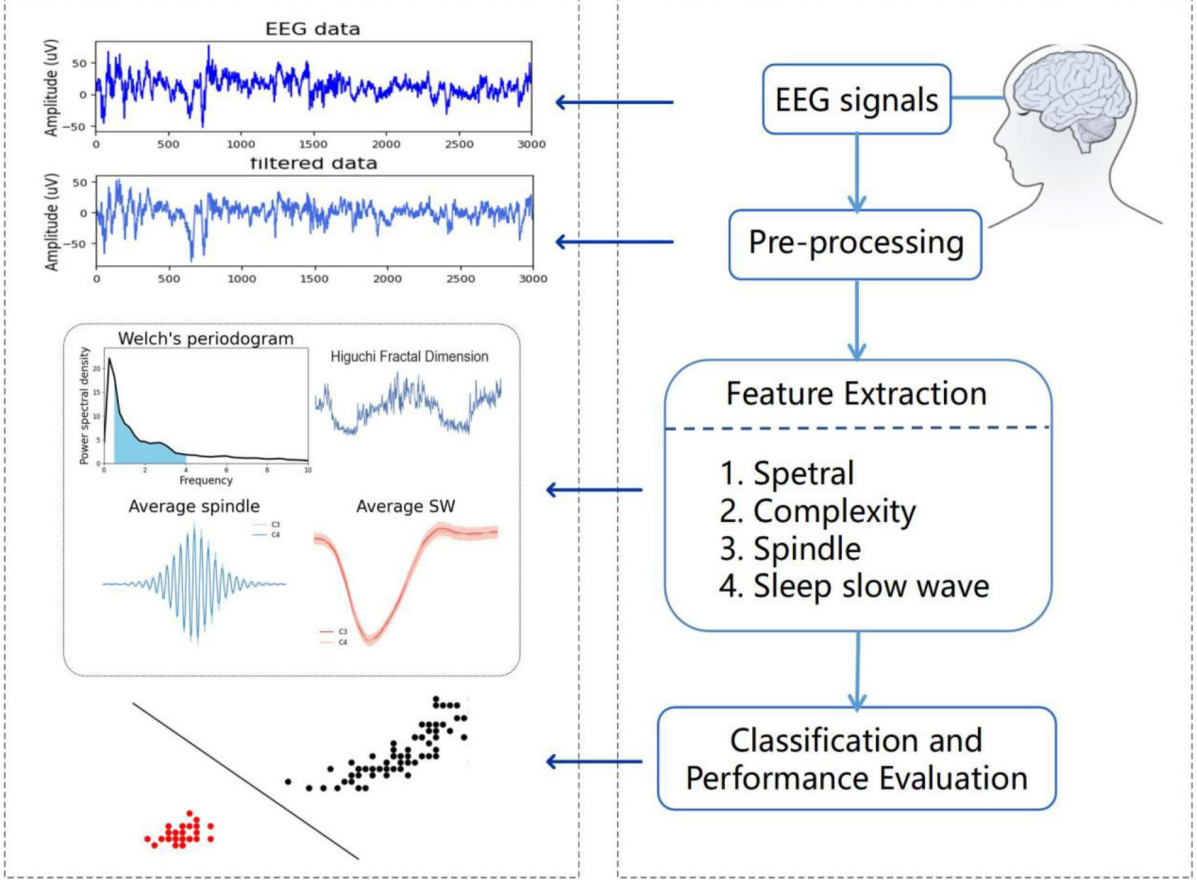


FIGURE 1 | Schematic diagram of proposed algorithm.

Spindle

Sleep spindles are a characteristic of N2 sleep, consisting of a succession of separate waves with frequencies ranging from 11 to 16 Hz (most typically 12–14 Hz), duration 0.5 s, usually using the maximum amplitude of the center deviation (Peyrache and Seibt, 2020). YASA (Yet Another Spindle Algorithm) is an open-source Python package for sleep analysis (Vallat, 2019). Spindles are detected using the YASA algorithm. The main idea of the algorithm is to calculate different thresholds from broadband filtering signals (1–30 Hz, EEG_{bf}) and sigma filtering signals (11–16 Hz, EEG_{σ}). **Figure 2A** shows sleep spindles as detected by a 30-s EEG segment. The algorithm consists of three steps.

Step I: The FIR filter was used to bandpass the EEG segments at 1–30 and 12–15 Hz.

Step II: Three thresholds are calculated.

Threshold 1: the relative power of the sigma band is the power of the sigma band relative to the total power of broadband (1–30 Hz). Calculated using the short-time Fourier transform (STFT), the continuous period is 2 s and the overlap is 200 ms. The first threshold is exceeded whenever the segment's relative power is greater than or equal to 0.2.

Threshold 2: movement correlation. Pearson correlation coefficient was obtained by moving the sliding window of 300 ms and the step of 100 ms. The correlation value $r \geq 0.65$ will exceed the second threshold.

Threshold 3: The moving root mean square is defined by calculating the moving root mean square (RMS) of EEG_{σ}. The window width is 300 ms and the step of 100 ms. The third threshold is exceeded whenever the RMS value of the segment $RMS \geq RMS_{thresh}$. Where,

$$\boldsymbol{RMS_{thresh} = RMS_{mean} + 1.5{*}RMS_{std}}$$

Step III: Decision making. Each EEG segment detected above three thresholds is considered a potential sleep spindle. The soft threshold is calculated by smoothing the decision vector of the 100 ms window. The true start and end times of the spindles are then found in the decision vector by finding the parameters that two of the three critical values are exceeded. Finally, spindles that are close to each other (less than 500 ms) are merged, and spindles that are too short or too long are removed.

Sleep Slow Wave

Sleep slow waves, defined as those with slow frequencies (<2 Hz) and high amplitudes (> 75 mV), have been linked to a drop in steady-state sleep pressure and the protective impact of arousal (Westerberg et al., 2012). **Figure 2B** shows slow sleep waves

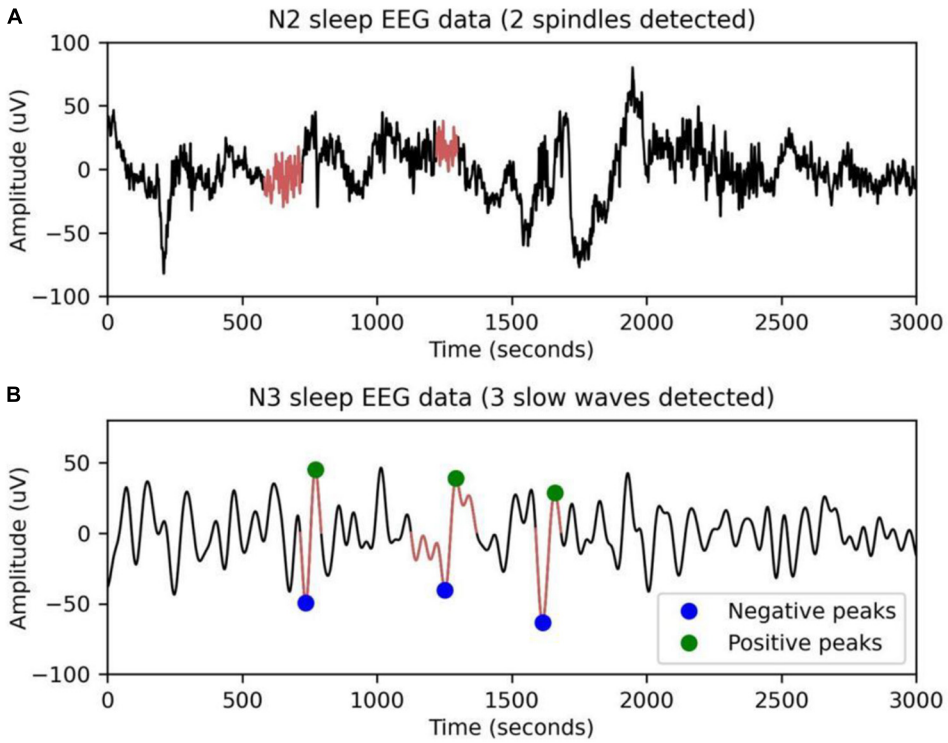


FIGURE 2 | EEG waveforms detected during 30-s NREM sleep **(A)** spindles and **(B)** slow sleep waves.

detected in 30-s EEG segments by the YASA algorithm. The algorithm consists of four steps.

Step I: Bandpass filtering is carried out between 0.3 and 2 Hz using an FIR filter with a transition band of 0.2 Hz.

Step II: Detection of all the negative peaks with amplitudes between −40 and −300 μV and all the positive peaks with amplitudes between 10 and 150 μV in the filtered signal.

Step III: For each negative peak (= slow-wave trough), find the closest positive peak and calculate several metrics, including peak-to-peak (PTP) amplitude, duration of the negative and positive phases, slope, etc.

Step IV: Apply a set of logical thresholds to determine true slow waves. PTP amplitudes need to be between 75 and 400 μV. The positive and negative phase duration is calculated by the zero-crossing value to ensure that the positive phase duration is 0.1–1 s and the negative phase duration is 0.3–1.5 s. The slope between the trough and the midline is greater than 0.

Feature Extraction

We classify EEG data using four major types of features in this paper: spectral, complexity, spindles, and sleep slow waves. **Table 2** shows the whole list of computed features.

For each EEG segment, absolute band power, total signal power, and relative band power are calculated using PSD. The term "relative band power" refers to the normalization of each frequency band's power in relation to the total power of 0.5–40 Hz (Kim and Kim, 2018). A total of 13 features are spectral features. Nonlinear approaches using fractal dimension or entropy methods may facilitate the identification of MCI. We use entropy and fractal dimension methods to calculate the features of EEG segments as complexity features (Ma Y. et al., 2018). And the properties of spindles and sleep slow waves are calculated after they are successfully detected.

Spectral, complexity, spindles, and sleep slow waves features were extracted from sleep EEG. In order to test the validity of sleep features, spectral and complexity features were extracted from EEG signals during wakefulness as controls. All the features were standardized. Then, these features are then utilized to train the classifier to automatically identify between MCI and HC EEG data.

Classification

The classifier receives the extracted features from the EEG data as input. The classifier determines which category the new observation belongs to. In this study, we use support vector machine (SVM) and gated recurrent unit (GRU) to classify the data separately.

We used two classifiers to predict the true category of subjects. The predicted outcome variable is binary (0 for healthy controls, 1 for patients with MCI), and the predicted scores range from 0 to 1. The predicted scores of the subjects were utilized as MCI scores. All data sets were randomly separated into two divisions, 80% of the data were used for training and 20% for testing. The MCI classification model was fitted using the training data. We fitted the MCI classification model on the whole training data using the optimal hyperparameter configuration to determine the

TABLE 2 | List of computed features for each 5-min segment.

Feature Group	Features
Spectral(absolute power and relative power)	Delta
	Theta
	Alpha
	Sigma
	Beta
	Gamma
	Total power
Complexity	Permutation entropy
	Singular value decomposition
	Sample entropy
	Detrended fluctuation analysis
	Petrosian FD algorithms
	Katz fractal dimension
	Higuchi's fractal dimension
	Lempel-Ziv complexity
Spindle	Density
	Duration
	Amplitude
	RMS
	Abspower
	Relpower
	Frequency
	Oscillations
Sleep slow wave	Density
	Duration
	ValNegPeak
	ValPosPeak
	PTP
	Slope
	Frequency

performance of the MCI classification model in the training data. These optimal hyperparameters were determined using 10-fold cross-validation. The generated model was directly applied to the test data, and the MCI score was used to determine whether the subject was HC or MCI. We use two classifiers to perform MCI detection respectively, in order to analyze detection results according to the structure of different types of classifiers.

Support Vector Machine

The SVM classifier is a supervised learning approach for separating two classes by finding the best separation hyperplane in the feature space (Safi and Safi, 2021). For N training samples $\{(x_i, y_i), i = 1, \cdots, N\}$, where x_i is the i th input vector and y_i is the known target, SVM training is the same as figuring out how to solve the following optimization problem:

$$\min_{w,b,\xi} J(w, \xi) = \frac{1}{2} w^T w + c \sum_{i=1}^{N} \xi_i$$

subject to:

$$y_i \left[w^T \varphi(x_i) + b \right] \geq 1 - \xi_i, \xi_i \geq 0$$

Where ξ_i is slack variable, indicating the tolerance of misclassification. C is a punishment parameter that is used to penalize mistakes during training, b is a bias term, w is the weight applied for input data x_i. The kernel function $\varphi(x)$ is a nonlinear transformation function that maps the input vectors into a high-dimensional feature space (Madusanka et al., 2019).

Gate Recurrent Unit

GRU is a variation structure of the Recurrent Neural Network (RNN). RNN will remember past information and apply it to the current output computation. Furthermore, RNN suffers from the problem of vanishing and exploding gradients (Chung et al., 2014), which causes the model to learn and train slowly. These concerns are addressed by taking into account its versions, such as GRU, which works on gated mechanisms. A GRU has two gates, the update gate controls how much prior state information is brought into the present state, while the reset gate controls how much previous state information is ignored. The following expressions show how a GRU calculates the result.

$$r = \sigma(wv_r i_t + xrhs_{t-1})$$
$$u = \sigma(wvui_t + xrhs_{t-1})$$
$$h_t = tanh(wvi_t + x(r \odot hs_{t-1}, x_t))$$
$$hs_t = [(1-u)\, hs_{t-1}] + uh_t$$

where, r is a reset gate and u is the update gate, σ is the sigmoid function, h_t for the hidden state, element multiplication denoted by $\odot$.

Bidirectional GRU (Bi-GRU) can not only take advantage of past information, but also capture subsequent information (Ma C. et al., 2018).

$$h_t = \left(\overrightarrow{h_t} || \overleftarrow{h_t} \right)$$

where, h_t for output states, $\overleftarrow{h_t}$ a backward and $\overrightarrow{h_t}$ forward states in the opposite direction.

Performance Evaluation

To decide which classifier method is the most successful, the sensitivity, specificity, F1 score, and accuracy of each one should be calculated. The confusion matrix gives an exact idea of the number of correctly classified and unclassified samples. The parameters are calculated by the following equations:

$$Sensitivity = \frac{TP}{TP + FN} * 100\%$$
$$Specificity = \frac{TN}{TP + FN} * 100\%$$
$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} * 100\%$$
$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

RESULTS AND DISCUSSION

According to the steps of this method, EEG segments of 40 different subjects with HC and MCI were analyzed. Then, in the analysis step, the features of the signals during wakefulness and NREM sleep are extracted, respectively. For the training of the SVM classifier and the GRU network, all datasets were randomly separated into two divisions of the training set and testing set for evaluating the accuracy of the classifiers, with 80 percent of the data treated as training data and 20% of the data considered as testing data. The test accuracy, sensitivity, specificity, and F1 score were finally acquired with 10-fold cross-validation was used to find the best hyperparameters values.

Results

To construct SVM classifiers, the "fitcsvm" module of MATLAB is used. Bayesian hyperparameter optimization is used to find the hyperparameter that minimizes the cross-validation loss to optimize the classifier. In this study, the cubic polynomial kernel function was selected, gamma = 2.15/C = 1 for the features of NREM sleep and gamma = 13.2/C = 2 for the features of wakefulness was used to obtain the best results. GRU classifier construction, training, and testing were carried out with the assistance of the PyTorch library. The default GRU sequential class was updated to add different layers based on the models provided. For two GRU networks, we used batch size 32, Adam optimizer, and binary cross-entropy loss, and varied dropout, and the number of hidden layers using random search (Yasir et al., 2021). For the features of NREM sleep, the best results for GRU were achieved with the dropout parameter value 0.5, the output layer has 2 nodes, the input layer has 36 nodes, and finally three hidden layers have 89 nodes. For the features of wakefulness, the best results for GRU were achieved with the dropout parameter value 0.5, the output layer has 2 nodes, the input layer has 21 nodes, and finally three hidden layers have 59 nodes. We trained 200 epochs for GRU and tested on the epoch that had the best cross-validation accuracy.

The mean of each metric was used to objectively evaluate classification performance, and the classification results comparison between EEG during NREM sleep and wakefulness are shown in **Table 3**.

As shown in **Table 3**, for different feature sets, the classification accuracy of the sleep EEG feature set is higher than that of the awake EEG feature set, and the sensitivity, specificity, and F1 score are also higher than that of the awake EEG. The results of the two classifiers showed that sleep EEG had better performance in MCI and HC classification than awake EEG, and the common features of sleep and awake EEG combined with sleep features could significantly improve the classification accuracy. For different classifiers, the accuracy of the GRU network is 5 and 3% higher than that of the SVM classifier for the same features, indicating that the classification performance of GRU is better than that of the SVM classifier. These results suggest that sleep features can reflect cognitive performance in patients with MCI and emphasize that altered sleep is a component of mild cognitive impairment.

The classification performance was also confirmed by ROC curve analysis, especially by calculating the area under the ROC curve (AUC). As shown in **Figure 3**, when awake and sleep EEG features are used as inputs to the classifiers, ROC curves and corresponding AUC values for MCI and HC classification are calculated. The ROC curve of sleep EEG was closer to the upper left corner, and the AUC value of GRU was higher than that of awake EEG, indicating that the GRU network with sleep EEG feature input achieved the best performance (AUC value was as high as 0.981).

The results showed that the GRU classifier with the sleep features had the best effect, with an accuracy of 93.46%, sensitivity of 93.33%, specificity of 93.60%, F1 value of 93.56%, and AUC value of 0.98. In conclusion, the spectral and complexity analysis of sleep EEG, combined with the features of sleep slow waves and spindles, and the classification of sleep EEG by GRU are effective for the early detection of MCI.

Discussion

AD is an irreversible neurodegenerative disease, so early screening and diagnosis of MCI is particularly important, which is the key to effective early intervention of AD and delay the progression of dementia. MCI is currently diagnosed by specialists through extensive testing, including neurophysiological assessment, blood analysis, cerebrospinal fluid analysis, and imaging techniques. However, the evaluation of these medical records is not only costly and complex to implement, but also requires experienced physicians. Therefore, the automated decision method which only needs one physiological parameter can not only objectively evaluate the patients, but also ensure high diagnostic accuracy. In addition, it will be economical, portable, and more suitable for the elderly population.

Several innovative research in recent years have only focused on awake EEG to detect MCI, and they collected about 20 channels of EEG and extracted features for classification. The DWT decomposition method was utilized to analyze 20 channels of EEG signals, and the Hjorth parameter and KNN classifier were incorporated to achieve 97.64% accuracy (Safi and Safi, 2021). Although this method has high classification accuracy, the acquisition channels of the EEG signal are numerous, and the measurement method is complex. Khatun et al. (2019) proposed an MCI detection method based on single-channel EEG, which stimulated auditory speech signals, extracted features from event-related potential (ERP), and obtained an accuracy of 87.9% by SVM classification. This method uses a single-channel signal, but it requires the collection of evoked potentials, which is complicated and difficult for patients to cooperate with. Cejnek et al. (2021) proposed a new approach for detecting MCI based on EEG recordings. The highest accuracy was 91.62 percent among the results of cross-validation classification between HC and MCI patients by each channel. All of the work above on EEG acquisition methods relied on the analysis of EEG recordings of waking states. It is worth noting that in this study, we proposed using the features extracted from sleep EEG to detect MCI and achieved an accuracy of 93.46%. Studies have shown that changes in sleep electrophysiology may be linked to the

TABLE 3 | Classification performance comparison between EEG during NREM sleep and wakefulness for distinguishing MCI from HC.

	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 score (%)
W	Cubic SVM	85.47	88.57	82.14	84.35
	Bi-GRU	90.26	90.00	90.54	90.57
NREM	Cubic SVM	90.51	92.02	88.89	89.87
	Bi-GRU	93.46	93.33	93.60	93.56

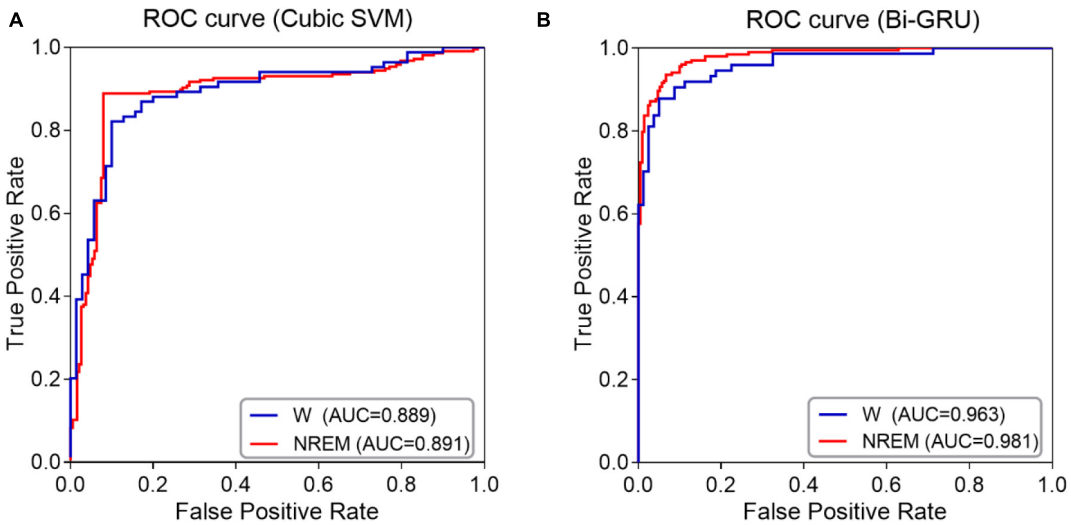


FIGURE 3 | ROC curves and AUC values of NREM sleep EEG and awake EEG **(A)** SVM classifier **(B)** GRU network.

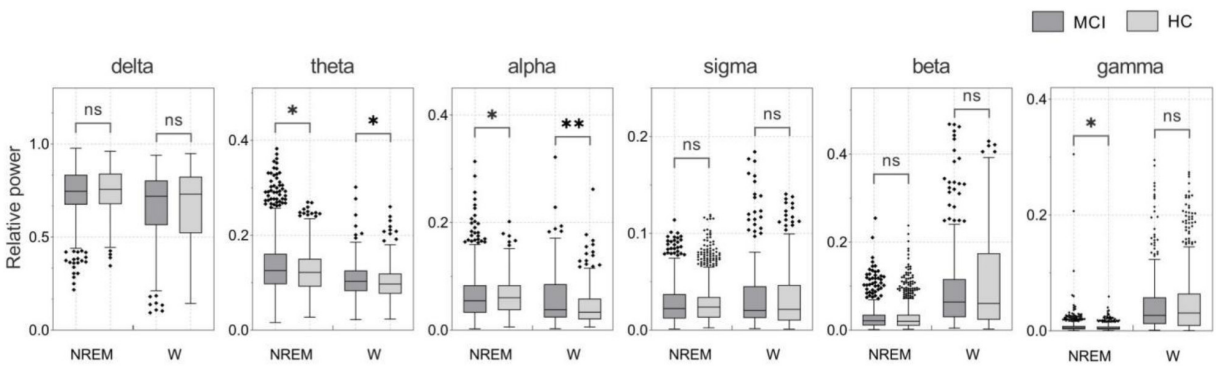


FIGURE 4 | Relative power of NREM sleep and awake EEG in MCI and HC subjects (*$p < 0.05$; **$p < 0.01$).

cognitive condition of AD and MCI patients. We used spindles and sleep slow waves, along with other common features, to improve classification accuracy by the GRU network. Although the accuracy is lower than that of multi-channel EEG signal classification, it is superior to other single-channel results and is not limited by the experimental site. Therefore, it is especially suitable for head-mounted wearable devices.

In this study, we proposed a diagnostic method for patients with MCI based on sleep EEG signals. We have shown that incorporating features of sleep slow waves and spindles as new features of the MCI detection significantly improves the accuracy of the MCI detection over traditional features during wakefulness. The proposed method is 5 and 3% better than traditional features in the classification results of the SVM classifiers and GRU network, respectively. In this study, the detection of MCI using the GRU network with the features of NREM sleep achieved the highest accuracy, reaching 93.46%. The following three aspects are considered superior to other methods. Firstly, the features of spindles and slow sleep waves during NREM sleep are an effective supplement to traditional features, and the data during sleep contains more information than wakefulness. As shown in **Figure 4**, SPSS software was used for statistical analysis. Non-normally distributed variables were compared by the Mann-Whitney U test. *P*-values less than 0.05

were considered statistically significant. During NREM sleep, the MCI group showed significant differences in theta, alpha, and gamma bands compared with the HC group. During wakefulness, a similar phenomenon was observed, but only the theta and alpha bands were significantly different (both $p < 0.05$). This suggests that significant changes in power spectral density can be detected during both NREM sleep and awake sleep, and that sleep EEG does not reduce the difference between MCI and HC. Therefore, related features of sleep may be an important biomarker of MCI. Secondly, this algorithm adopted two-channel EEG, which is simple compared with the multi-channel acquisition, and does not require additional nursing staff to take care of patients when we collect the EEG signals during sleep. Finally, the GRU network significantly improved the classification performance in MCI recognition.

While the proposed approach is encouraging, there are some limitations that should be addressed. In this study, only MCI and HC were categorized, requiring further consideration of other stages of cognitive impairment. In addition, we were only able to analyze sleep EEG signals from 40 different subjects in this study, so this research should be regarded as preliminary, and future studies should include larger datasets to validate the suggested method's stability and generalizability.

CONCLUSION

EEG signals are non-stationary, nonlinear, and noisy, so it is a challenging problem to distinguish between MCI and HC based on EEG signals. In this study, MCI was detected and classified using two-channel sleep EEG signals. Based on traditional features, the features of sleep slow waves and spindles were extracted. Unlike the existing features, the proposed features are not restricted by the use of traditional EEG bands. In particular, sleep features combined with traditional features outperformed traditional features on classification tasks, proved to be more accurate in predicting MCI and performed better with sleep EEG signals than with wakeful signals. Although studies have shown that changes in EEG activity during NREM sleep are associated with MCI, no studies have used it for the recognition of MCI. The high classification accuracy obtained in this paper once again proves that sleep slow waves and spindles can be used as early biomarkers for the development of AD (Romanella et al., 2020). Early diagnosis would also provide patients access to available treatment, while possibly initiating an earlier treatment. In addition, this study is based on sleep EEG, which is non-invasive, portable, and low-cost, and therefore has high value as a diagnostic tool.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://sleepdata.org/datasets/sof; Study of Osteoporotic Fractures.

ETHICS STATEMENT

The study was conducted using an open source database. The procedures for information collection in the database were in accordance with the ethical standards of the institutional and national Non-invasive Clinical Research Medical Ethics Review Board. Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

DG and CW were responsible for the study concept and design. DG and ZF contributed to the acquisition of data and funding. CW determined software and analytical methods as well as manuscript writing. YZ assisted with data analysis and interpretation of findings. KY and HA provided critical revision of the manuscript for important intellectual content. All authors critically reviewed the content and approved the final version for publication.

ACKNOWLEDGMENTS

We thank the National Sleep Research Resource and the Study of Osteoporotic Fractures database. The Study of Osteoporotic Fractures (SOF) was supported by National Institutes of Health grants (AG021918, AG026720, AG05394, AG05407, AG08415, AR35582, AR35583, AR35584, RO1 AG005407, R01 AG027576-22, 2 R01 AG005394-22A1, 2 RO1 AG027574-22A1, HL40489, and T32 AG000212-14). The National Sleep Research Resource was supported by the National Heart, Lung, and Blood Institute (R24 HL114473 and 75N92019R002).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.865558/full#supplementary-material

REFERENCES

Ahmed, O. B., Benois-Pineau, J., Allard, M., Catheline, G., and Amar, C. B. (2017). Recognition of alzheimer's disease and mild cognitive impairment with multimodal image-derived biomarkers and multiple kernel learning. *Neurocomputing* 220, 98–110. doi: 10.1016/j.neucom.2016.08.041

Alberdi, A., Aztiria, A., and Basarab, A. (2016). On the early diagnosis of alzheimer's disease from multimodal signals: a survey. *Artif. Intell. Med.* 71, 1–29. doi: 10.1016/j.artmed.2016.06.003

Amezquita-Sanchez, J. P., Mammone, N., Morabito, F. C., Marino, S., and Adeli, H. (2019). A novel methodology for automated differential diagnosis of mild

cognitive impairment and the Alzheimer's disease using EEG signals. *J Neurosci. Methods* 322, 88–95. doi: 10.1016/j.jneumeth.2019.04.013

Cassani, R., and Falk, T. H. (2020). Alzheimer's disease diagnosis and severity level detection based on electroencephalography modulation spectral "Patch" Features. *IEEE J. Biomed. Health Inform.* 24, 1982–1993. doi: 10.1109/JBHI.2019.2953475

Cejnek, M., Vysata, O., Valis, M., and Bukovsky, I. (2021). Novelty detection-based approach for alzheimer's disease and mild cognitive impairment diagnosis from EEG. *Med. Biol. Eng. comp.* 59, 2287–2296. doi: 10.1007/s11517-021-02427-6

Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *Comp. Sci.* doi: 10.48550/arXiv.1412.3555

D'Atri, A., Scarpelli, S., Gorgoni, M., Truglia, I., Lauri, G., Cordone, S., et al. (2021). EEG alterations during wake and sleep in mild cognitive impairment and alzheimer's disease. *iScience* 24:102386. doi: 10.1016/j.isci.2021.102386

Gao, F., and Barker, P. B. (2014). Various MRS application tools for alzheimer disease and mild cognitive impairment. *AJNR Am. J. Neuroradiol.* 35, S4–S11. doi: 10.3174/ajnr.A3944

Ge, J. W. (2017). Research progress of functional magnetic resonance imaging in mild cognitive impairment. *Med. J. Qilu* 32, 616–619. doi: 10.13362/j.qlyx.201705034

Gorgoni, M., D'Atri, A., Scarpelli, S., Reda, F., and De Gennaro, L. (2020). Sleep electroencephalography and brain maturation: developmental trajectories and the relation with cognitive functioning. *Sleep Med.* 66, 33–50. doi: 10.1016/j.sleep.2019.06.025

Gorgoni, M., Lauri, G., Truglia, I., Cordone, S., Sarasso, S., Scarpelli, S., et al. (2016). Parietal fast sleep spindle density decrease in alzheimer's disease and amnesic mild cognitive impairment. *Neural Plast.* 2016:8376108. doi: 10.1155/2016/8376108

Handels, R. L. H., Vos, S. J. B., Kramberger, M. G., Jelic, V., Blennow, K., van Buchem, M., et al. (2017). Predicting progression to dementia in persons with mild cognitive impairment using cerebrospinal fluid markers. *Alzheimer's Dementia* 13, 903–912. doi: 10.1016/j.jalz.2016.12.015

Jiang, J., Kang, L., Jianjun, H., and Zhang, T. (2020). Deep learning based mild cognitive impairment diagnosis using structure MR images. *Neurosci. Lett.* 730:134971. doi: 10.1016/j.neulet.2020.134971

Karow, D. S., McEvoy, L. K., Fennema-Notestine, C., Hagler, D. J. Jr., Jennings, R. G., Brewer, J. B., et al. (2010). Relative capability of MR imaging and FDG PET to depict changes associated with prodromal and early alzheimer disease. *Radiology* 256, 932–942. doi: 10.1148/radiol.10091402

Khatun, S., Morshed, B. I., and Bidelman, G. M. (2019). A single-channel EEG-based approach to detect mild cognitive impairment via speech-evoked brain responses. *IEEE Trans. Neural Syst. Rehabil. Eng.* 27, 1063–1070. doi: 10.1109/TNSRE.2019.2911970

Kim, D., and Kim, K. (2018). Detection of early stage alzheimer's disease using eeg relative power with deep neural network. *Annu Int Conf. IEEE Eng. Med. Biol. Scoc.* 2018, 352–355. doi: 10.1109/embc.2018.8512231

Klinzing, J. G., Niethard, N., and Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature Neurosci.* 22, 1598–1610. doi: 10.1038/s41593-019-0467-3

Ma, C., Yang, C., Yang, F., Zhuang, Y., Zhang, Z., Jia, H., et al. (2018). "Trajectory factory: tracklet cleaving and re-connection by deep siamese bi-gru for multiple object tracking"," in *Proceeding of the: 2018 IEEE International Conference on Multimedia and Expo (ICME)* (San Diego, CA: IEEE), 1–6.

Ma, Y., Shi, W., Peng, C. K., and Yang, A. C. (2018). Nonlinear dynamical analysis of sleep electroencephalography using fractal and entropy approaches. *Sleep Med. Rev.* 37, 85–93. doi: 10.1016/j.smrv.2017.01.003

Madusanka, N., Choi, H. K., So, J. H., and Choi, B. K. (2019). Alzheimer's disease classification based on multi-feature fusion. *Curr. Med. Imaging Rev.* 15, 161–169. doi: 10.2174/1573405614666181012102626

Moretti, D. V., Paternico, D., Binetti, G., Zanetti, O., and Frisoni, G. B. (2013). EEG upper/low alpha frequency power ratio relates to temporo-parietal brain atrophy and memory performances in mild cognitive impairment. *Front. Aging Neurosci.* 5:63. doi: 10.3389/fnagi.2013.00063

Ni, H., Qin, J., Zhou, L., Zhao, Z., Wang, J., and Hou, F. (2017). Network analysis in detection of early-stage mild cognitive impairment. *Phy. A: Stat. Mech. Appl.* 478, 113–119. doi: 10.1016/j.physa.2017.02.044

Peyrache, A., and Seibt, J. (2020). A mechanism for learning with sleep spindles. *Philos. Trans. R. Soc. Lond B Biol. Sci.* 375:20190230. doi: 10.1098/rstb.2019.0230

Prince, M. (2015). *World Alzheimer Report 2015: the Global Impact of Dementia*. London: Alzheimer's Disease International.

Rodrigues, P. M., Bispo, B. C., Garrett, C., Alves, D., Teixeira, J. P., and Freitas, D. (2021). Lacsogram: a new EEG tool to diagnose alzheimer's disease. *IEEE J. Biomed. Health Inform.* 25:3384–3395. doi: 10.1109/jbhi.2021.3069789

Romanella, S. M., Roe, D., Paciorek, R., Cappon, D., Ruffini, G., Menardi, A., et al. (2020). Sleep, noninvasive brain stimulation, and the aging brain: challenges and opportunities. *Ageing Res. Rev.* 61:101067. doi: 10.1016/j.arr.2020.101067

Romanella, S. M., Roe, D., Tatti, E., Cappon, D., Paciorek, R., Testani, E., et al. (2021). The sleep side of aging and alzheimer's disease. *Sleep Med.* 77, 209–225. doi: 10.1016/j.sleep.2020.05.029

Rosinvil, T., Bouvier, J., Dubé, J., Lafrenière, A., Bouchard, M., Cronier, J., et al. (2020). Are age and sex effects on sleep slow waves only a matter of EEG amplitude? *Sleep* 44:zsaa186. doi: 10.1093/sleep/zsaa186

Safi, M. S., and Safi, S. M. M. (2021). Early detection of alzheimer's disease from EEG signals using hjorth parameters. *Biomed. Signal Process. Control* 65:102338. doi: 10.1016/j.bspc.2020.102338

Sharma, N., Kolekar, M., and Chandra, S. (2016). The role of EEG signal processing in detection of neurocognitive disorders. *Int. J. Behav. Healthcare Res.* 6, 15–27. doi: 10.1504/IJBHR.2016.10002016

Spira, A. P., Blackwell, T., Stone, K. L., Redline, S., Cauley, J. A., Ancoli-Israel, S., et al. (2008). Sleep-disordered breathing and cognition in older women. *J. Am. Geriatr. Soc.* 56, 45–50. doi: 10.1111/j.1532-5415.2007.01506.x

Tsai, C. L., Pai, M. C., Ukropec, J., and Ukropcová, B. (2016). The Role of physical fitness in the neurocognitive performance of task switching in older persons with mild cognitive impairment. *J. Alzheimer's Dis.* 53, 143–159. doi: 10.3233/JAD-151093

Vallat, R. (2019). YASA (yet another spindle algorithm): a fast and open-source sleep spindles and slow-waves detection toolbox. *Sleep Med.* 64:S396. doi: 10.1016/j.sleep.2019.11.1104

Waninger, S., Ahlfors, S., Stikic, M., Korszen, S., Berka, C., Salat, D. H., et al. (2016). P3-163: Identification of neurophysiological biomarkers of MCI using resting state EEG. *Alzheimer's Dementia* 12, 882–882. doi: 10.1016/j.jalz.2016.06.1823

Welch, P. (1967). The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans.Audio Electroacoustics* 15, 70–73. doi: 10.1109/TAU.1967.1161901

Westerberg, C. E., Mander, B. A., Florczak, S. M., Weintraub, S., Mesulam, M. M., Zee, P. C., et al. (2012). Concurrent impairments in sleep and memory in amnestic mild cognitive impairment. *J. Int. Neuropsychol. Soc.* 18, 490–500. doi: 10.1017/S135561771200001X

Yasir, M., Chen, L., Khatoon, A., Malik, M. A., and Abid, F. (2021). Mixed script identification using automated DNN hyperparameter optimization. *Comput. Intell. Neurosci.* 2021, 8415333–8415333. doi: 10.1155/2021/8415333

Zhang, G. Q., Cui, L., Mueller, R., Tao, S., Kim, M., Rueschman, M., et al. (2018). The national sleep research resource: towards a sleep data commons. *J. Am. Med. Inform. Assoc.* 25, 1351–1358. doi: 10.1093/jamia/ocy064

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 26 April 2022
doi: 10.3389/fnagi.2022.871706

Alzheimer's Disease Diagnosis With Brain Structural MRI Using Multiview-Slice Attention and 3D Convolution Neural Network

Lin Chen[1]*, *Hezhe Qiao*[1,2] *and Fan Zhu*[1]

[1] Chongqing Key Laboratory of Big Data and Intelligent Computing, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, China, [2] University of Chinese Academy of Sciences, Beijing, China

Numerous artificial intelligence (AI) based approaches have been proposed for automatic Alzheimer's disease (AD) prediction with brain structural magnetic resonance imaging (sMRI). Previous studies extract features from the whole brain or individual slices separately, ignoring the properties of multi-view slices and feature complementarity. For this reason, we present a novel AD diagnosis model based on the multiview-slice attention and 3D convolution neural network (3D-CNN). Specifically, we begin by extracting the local slice-level characteristic in various dimensions using multiple sub-networks. Then we proposed a slice-level attention mechanism to emphasize specific 2D-slices to exclude the redundancy features. After that, a 3D-CNN was employed to capture the global subject-level structural changes. Finally, all these 2D and 3D features were fused to obtain more discriminative representations. We conduct the experiments on 1,451 subjects from ADNI-1 and ADNI-2 datasets. Experimental results showed the superiority of our model over the state-of-the-art approaches regarding dementia classification. Specifically, our model achieves accuracy values of 91.1 and 80.1% on ADNI-1 for AD diagnosis and mild cognitive impairment (MCI) convention prediction, respectively.

Keywords: Alzheimer's disease (AD), disease prognosis, multi-view-slice attention, 3D convolution neural network, brain sMRI image

OPEN ACCESS

Edited by:
Peng Xu,
University of Electronic Science and Technology of China, China

Reviewed by:
Xianmin Wang,
Guangzhou University, China
Dianlong You,
Yanshan University, China

***Correspondence:**
Lin Chen
chenlin@cigit.ac.cn

Specialty section:
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

Received: *08 February 2022*
Accepted: *17 March 2022*
Published: *26 April 2022*

Citation:
Chen L, Qiao H and Zhu F (2022) Alzheimer's Disease Diagnosis With Brain Structural MRI Using Multiview-Slice Attention and 3D Convolution Neural Network. Front. Aging Neurosci. 14:871706. doi: 10.3389/fnagi.2022.871706

1. INTRODUCTION

Alzheimer's disease (AD) is the most common cause of dementia that causes progressive and permanent memory loss and brain damage. It is critical to initiate treatment for slowing down AD development in early AD. As a non-contact diagnostic method, structural magnetic resonance imaging (sMRI) is regarded as a typical imaging biomarker in quantifying the stage of neurodegeneration (Kincses et al., 2015; Bayram et al., 2018; Shi et al., 2018). Based on the examination of the brain's sMRI images, numerous artificial intelligence (AI) technologies, including conventional voxel-based machine learning methods and deep-learning-based approaches, have been performed for assisting the cognitive diagnosis (Martí-Juan et al., 2020; Tanveer et al., 2020; Wu et al., 2021a,b).

In the early attempts, traditional statistical methods based on voxel-based morphology (VBM) were introduced to measure the brain's morphologic changes. VBM-based studies determine the

intrinsic characteristics of specific biomarkers, such as the hippocampus volumes (Fuse et al., 2018), cortex sickness (Luk et al., 2018), subcortical volumes (Vu et al., 2018), and frequency features with non-subsampled contourlets (Feng et al., 2021), to calculate the regional, anatomical volume of the brain. However, most VBM-based approaches relying on domain knowledge and expert's experience need a complex handcrafted feature extraction procedure, which is independent of the subsequent classifiers, resulting in potential diagnostic performance degradation.

With the advancement of deep learning, especially the successful applications of convolution neural networks (CNN), in recent years, a growing body of research employed deep learning to analyze the MR images by training an end-to-end model without handcrafted features (Zhang et al., 2020; AbdulAzeem et al., 2021; Qiao et al., 2021). Since the 3D volumetric nature of sMRI, 3D-CNN could be directly applied to capture the structural changes of the whole brain at the subject-level (Jin et al., 2019). However, there is much useless information in the complete MRI with millions of voxels. Furthermore, it is hard to fully train the CNNs with only a few labeled MRI data available at the subject level. Many deep-learning-based methods turn to exact pre-determination of regions-of-interest (ROI) for training the models with 3D-Patch or 2D-slice (Ebrahimighahnavieh et al., 2020). Liu et al. (2020b) extract multi-scale image patches based on the pre-determined anatomical landmarks from sMRI for training an end-to-end CNN. Lian et al. (2020a,b) trained multiple classifiers with multilevel discriminative sMRI features from the whole sMRI with a hybrid network to capture local-to-global structural information. Compared with the modeling in the subject level, the patches or slices carry more local features but lose some global information. In addition, some studies try to exclude irrelevant regions by emphasizing specific brain tissues with the help of segmentation technology. Cui and Liu (2019) and Poloni and Ferrari (2022) focus on the specific biomarker from specific regions, such as the hippocampus, to capture the structural changes in 3D MR images for AD and mild cognitive impairment (MCI) classification. Chen and Xia (2021) design a sparse regression module to identify the critical cortical regions, such as the amygdala, posterior temporal lobe, and propose a deep feature extraction module to integrate the features landmarked regions for the diagnosis process. However, such methods need extra tissue segmentation operations, which inevitably increase the complexity of the diagnostic model.

Although the existing models have achieved outstanding results so far, it is still a challenging work for AD diagnosis due to a large number of volumes in 3D MR images and a subtle difference between abnormalities and normality brains, i.e., it is vital to extract subtle changes in disease progression from MRI sequence data with a high denominational. Previous studies focus on extracting features from the whole brain or individual slices separately, ignoring the feature complementarity from different views. As illustrated in **Figure 1**, each slice of the brain sMRI in different views contains a certain amount of local information that could also be valuable for dementia diagnosis. Considering both global structure changes of whole brain and fine-grained local distinctions of slices could be both crucial, this study proposes a novel fusion model for AD classification, named multiView-slice attention and 3D convolution neural network (MSA3D), which organically integrates multiple slices features and 3D structural information.

The main contributions of this study are three-fold:

(1) We proposed an MSA3D model to combine the 2D multi-view-slice levels and global 3D subject-level features for fully mining the subtle changes in different views and dimensions.
(2) We propose a slice-level attention module to help the CNN focus on specific slices to obtain more discriminative features representations from abundant vowels.
(3) We perform two classification tasks, i.e., AD diagnosis and MCI conversion prediction, on two ADNI datasets. Our model achieves superior diagnostic results compared with other tested models, demonstrating our model's efficacy in aiding dementia prediction.

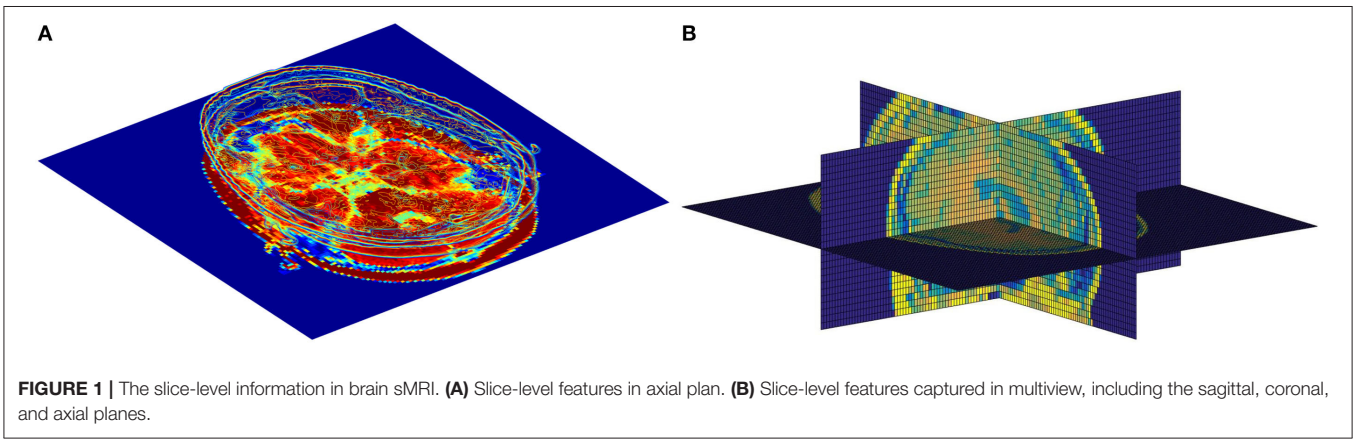


FIGURE 1 | The slice-level information in brain sMRI. **(A)** Slice-level features in axial plan. **(B)** Slice-level features captured in multiview, including the sagittal, coronal, and axial planes.

2. MATERIALS AND DATA PREPROCESSING

2.1. Studied Subjects

Following the previous studies (Liu et al., 2019; Lian et al., 2020b), we employed two public sMRI data sets, i.e., ADNI-1 and ADNI-2, for empirical study. Both of them can be found on the Alzheimer's Disease Neuroimaging Initiative (ADNI) website (Jack et al., 2008). This study employed the ADNI data only for model validation but did not involve any patient interaction or data acquisition. More detailed data acquisition protocols are available at http://adni.loni.usc.edu/. We collected a total of 1,451 subjects from the ADNI database with baseline T1 weighted (T1W) brain MRI scans, which are divided into four categories:

- Cognitively Normal (CN): Subjects diagnosed with CN at baseline and showed no cognitive decline.
- Stable MCI (sMCI): Subjects diagnosed with MCI remain stable and have not converted to AD at all time-points (0–90 months).
- Progressive MCI (pMCI): Subjects are diagnosed with MCI who would gradually progress to AD within 0–36 months.
- Alzheimer's disease: Subjects diagnosed as AD at baseline and whose conditions would not change during the follow-up period.

To avoid data leakage problems mentioned in Wen et al. (2020), we also remove the subjects exited in both ADNI-1 and ADNI-2. More specifically, the ADNI-1 dataset is formed of 808 subjects with 1.5 T T1W sMR brain images, including 183 AD, 229 CN, 167 pMCI, and 229 sMCI. The ADNI-2 dataset has 643 3T T1W sMR brain images, including 143 AD, 184 CN, 75 pMCI, and 241 sMCI. **Table 1** summarizes the detailed clinical information of the studied subjects, including age, sex, and the scores of the mini-mental state examination (MMSE). In our experiments, these two independent datasets will be employed as the training dataset and testing dataset, repetitively, to perform cross-validation. More specifically, we first trained the model on the ADNI-1 and evaluated it on ADNI-2. Subsequently, we reversed the experimentation and used the ADNI-2 for model learning, and then the trained model was assessed on ADNI-1. Note that we employed the ADNI data only for empirical analysis but this study did not employ any patient interaction or data acquisition.

TABLE 1 | Detailed clinical information of the studied subjects in ADNI-1 and ADNI-2 (± means the SD).

Dataset	Label	Total number	Age (Years)	Sex (M/F)	MMSE
ADNI-1	NC	229	76.2 ± 5.1	119/110	29.2±1.0
	sMCI	229	74.8 ± 7.6	153/76	27.2 ± 1.7
	pMCI	167	74.9 ± 7.2	102/65	26.9 ± 1.7
	AD	183	75.6 ± 7.6	96/87	23.1 ± 2.5
ADNI-2	NC	184	77.3 ± 6.7	87/97	28.8 ± 1.7
	sMCI	241	71.3 ± 7.5	134/107	28.3 ± 1.5
	pMCI	75	71.9 ± 7.2	40/35	27.0 ± 1.6
	AD	143	75.6 ± 7.8	85/58	21.9 ± 3.8

2.2. Data Preprocessing

The standard preprocessing pipeline was performed on all the T1W brain MRIs as follows: First, all MRIs were performed in an axial orientation parallel to the line through anterior commissure (AC)-posterior commissure (PC) correction. Then the invalid volumes of the sMRI, i.e., the blank regions, were removed, leaving only the brain tissues. Subsequently, the intensity of brain images was corrected and normalized with the N3 algorithm after the skull dissection (Wang et al., 2011). Finally, all the aligned images are resized into the same spatial resolution for facilitating the CNN training. The model's inputs are fixed to $91 \times 101 \times 91$(i.e., $2mm \times 2mm \times 2mm$ cubic size) in our experiment, following the previous study (Jin et al., 2020).

3. METHODOLOGY

The overall architecture of our model is presented in **Figure 2**, which is composed of five main parts: the MRI sequences input, multi-view-slice sub-network (MVSSN), slices attention module (SAM), subject-level 3D-CNN (S3D-CNN), and a softmax classifier with full connection layer. The following sections provide more details for each module.

3.1. Multi-View-Slice 2D Sub-Networks

In this subsection, we introduce the MVSSN module for extracting multiview 2D-slice level features. As shown in **Figure 3**, the inputs of MVSSN are consist of the MR slices in three views, i.e., the sagittal, coronal, and axial imaging planes. Since discriminative features may exist in different slices, we employ a 2D-CNN to extract the multiview slice features from each slice. Let's denote the x, y, and z as the MRI planes of sagittal, coronal, and axial, respectively, particularly, $\mathbf{S}_x = [s_x^1, s_x^2, ..., s_x^{M_x}]$ denotes the slice cluster in the x plane, where M_x is the total slice number of the cluster $\mathbf{S}_x$. After using the multiple 2D-CNNs on each slice to generate the feature maps in different views separately, the input $I \in R^{D \times H \times W}$ can be transformed as the feature maps F_x, F_y, F_z in three dimensions. For example, each feature map F_x^i in sagittal view is calculated by Equation (1):

$$F_x^i = f_x^i(s_x^i, w_x^i) \tag{1}$$

where f_x^i is a independent 2D-based CNN, w_x^i is the weight of CNN f_x^i, and $i \in [1, M_x]$ means the ith slice in the x-direction. Each f_x^i contains three CNN blocks, each with a conventional layer, a barch normalize (BN) layer, a rectified linear unit (RELU) operator, and a maxpooling layer. Detailed parameters of our 2D-based CNN are listed in **Table 2**.

After the Global-Avg-Pooling (GAP) operation, the feature map F_x^i can be pooled as a vector denoted as I_x^i. In the end, all the feature maps in x view can be cascaded as $\mathbf{I}^x = [I_1^x, I_2^x, ..., I_{M_x}^x]$. The same conventional operation can be applied on y and z views to generate the corresponding feature map clusters.

3.2. Slices Attention Module

Each vector in $\mathbf{I}^k$ can be regarded as a class-specific response after extracting the multiple slices-level features using the MVSSN. Considering that the volumetric MRI data contains

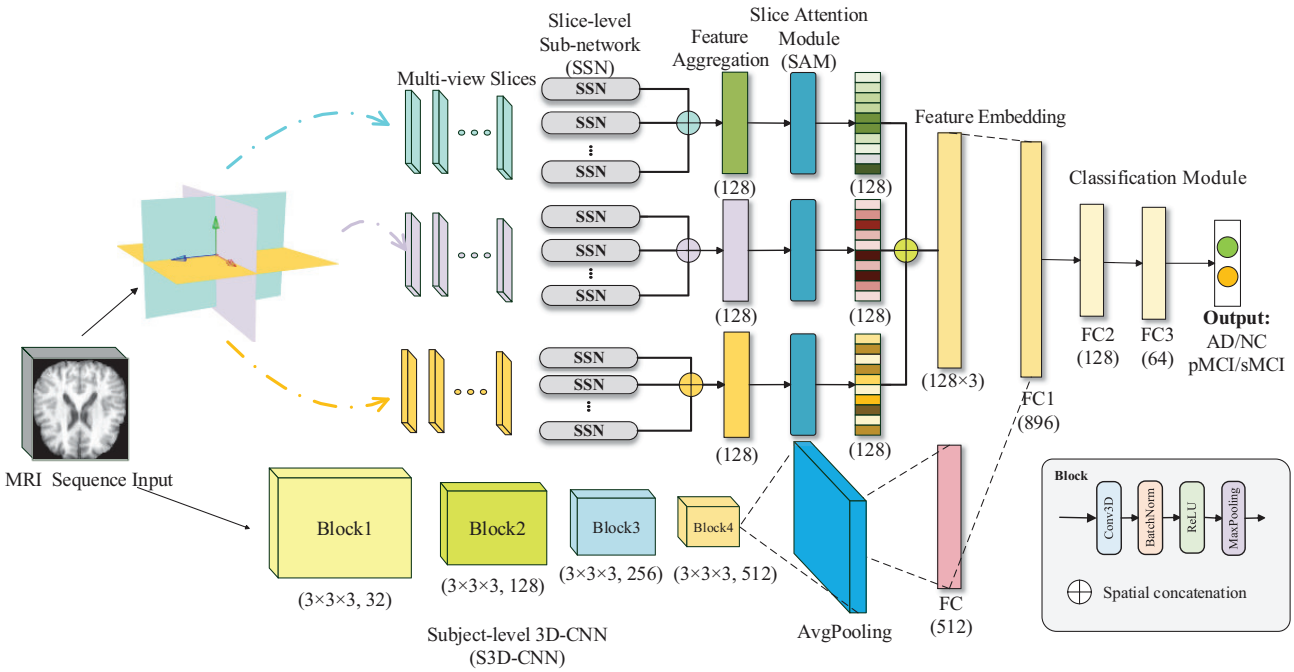


FIGURE 2 | Illustration of the proposed multiview-slice attention and 3D convolution neural network (MSA3D) model.

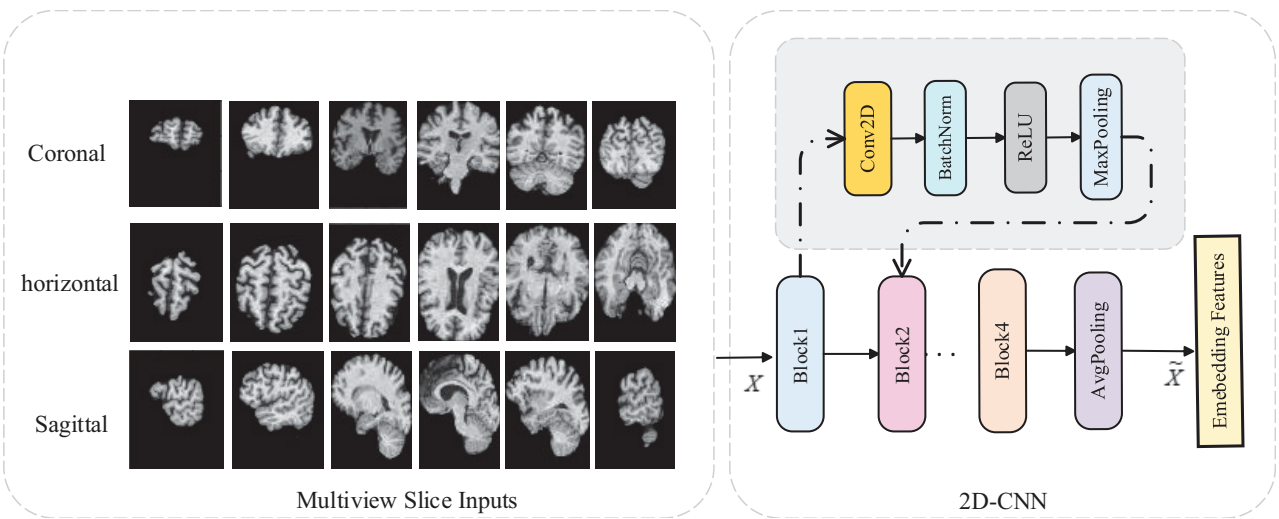


FIGURE 3 | 2D-CNN slice sub-network for multi-view slice features extraction.

different slices, many of them may not contain the most representative information relevant to dementia (Lian et al., 2021). To address this issue, we proposed a SAM to help the CNN focus on the specific features by exploiting the interdependencies among slices.

As shown in **Figure 2**, given a set of features embedding of the j^{th} direction, denoted as $\mathbf{I}^k \in {}^{M_k \times C}$, where C = 8 is the feature channels of each slice, and $k \in \{x, y, z\}$ means the MR plane. By employing an attention mechanism, we can obtain the slice attention $\mathbf{A}^k \in {}^{M_k \times M_k}$, which can build the dynamic correlations between the target diagnosis label and slice-level features with the following equation:

$$a_{ij}^k = \frac{\exp(I_i^k \cdot I_j^k)}{\sum_{i=1}^{M_k} \exp(I_i^k \cdot I_j^k)} \tag{2}$$

TABLE 2 | Detailed parameters of our 2D-CNN slice sub-network.

Layer	Kernel	Stride	Activation	Output channels
Conv2D	3 × 3	2	BachNorm+Relu	8
MaxPooling2D	2 × 2			8
Conv2d	3 × 3	2	BachNorm+Relu	32
MaxPooling2D	2 × 2			32
Conv2D	3 × 3	2	BachNorm+Relu	64
MaxPooling2D	2 × 2			64
Global-Avg-Pooling2D	1 × 1	1		128
Full connected				128
Full connected				8

TABLE 3 | Detailed parameters of our 3D-CNN subject sub-network.

Layer name	Kernel	Stride	Activation	Output channels
Conv3D	3× 3× 3	1	BachNorm3d+Relu	32
MaxPooling3D	3× 3× 3	2		32
Conv3D	3× 3× 3	1	BachNorm3d+Relu	128
MaxPooling3D	3× 3× 3	2		128
Conv3D	3× 3× 3	1	BachNorm3d+Relu	256
MaxPooling3D	3× 3× 3	2		256
Conv3D	2× 2× 2	2	BachNorm3d+Relu	512
MaxPooling3D	5× 5× 5	2		512
Globel-Avg-Pooling3D	1× 1× 1			512

where $a_{ij}^k \in \mathbf{A}^k$ is the score that semantically represents the impact of i^{th} slice feature on the j^{th} slice in the k^{th} direction. The final output of the weighted slice features $\widetilde{\mathbf{I}}^k \in {}^{M_k \times C}$ can be calculated by:

$$\widetilde{I}_j^k = \beta \sum_{i=1}^{M_k} (a_{ij}^k I_i^k) + I_j^k \tag{3}$$

where β is a learnable parameter that will gradually increase from 0, note that the final output feature maps are the sum of all the weighted features of the slices in one direction so that the SAM can adaptively emphasize the most relevant slices to produce a better AD inference.

After the SAM module, we fuse all the slice features in three directions using concatenation operation to form the final slice-level features $\mathbf{F_s} = [\widetilde{\mathbf{I}}^x, \widetilde{\mathbf{I}}^y, \widetilde{\mathbf{I}}^z]$, where $\mathbf{F_s}$ represents the cascaded weighted features which can capture the multiple views of local changes of the brain in three directions in 2D MRI images.

3.3. Subject-Level 3D Neural Network

The brain MRI data can be regarded as 3D data with an input size of $H \times W \times D$, where H and W denote the height and width of the MRI, repetitively, and D is the image sequence. In order to explore the global structure changes of the brain, all of the convolution operations and pooling layers are reformed from 2D to 3D. The 3D CNN operator is given in Equation (4):

$$u_j^l(x,y,z) = \sum_{\delta_x}\sum_{\delta_y}\sum_{\delta_z} F_i^{l-1}(x+\delta_x, y+\delta_y, z+\delta_z) \times W_{ij}^l(\delta_x, \delta_y, \delta_z) \tag{4}$$

where (x, y, z) refers to the 3D coordinates in sMRI data, F_i^{l-1} is the ith feature map of the $l-1$ layer. $W_{ij}^l(\delta_x, \delta_y, \delta_z)$ is a 3D convolution kernel slides in 3 dimensions, thus the new jth feature map $u_j^l(x, y, z)$ of the l layer can be generated after 3D convolution across the F_i^{l-1} from the $l-1$ layer. Similar to the 2D-CNN, our 3D-CNN includes four network blocks, and each block has a 3D-CNN layer, 3D BN layer, ReLu activation, and 3D max-pooling layer. Finally, the 3D convolutional feature maps are pooled into one 1D vector using a 3D-GAP layer with a kernel size of 1 × 1 × 1. The produced vector represents the global subject-level features. Detailed parameters of our 3D-CNN subject-level subnetwork are shown in **Table 3**.

3.4. Fully Connected Layer and Loss for Classification

To exploit both the slice-level and subject-level features generated by 2D and 3D-CNNs, a fully connected (FC) layer is employed to concatenate all the 2D and 3D features maps, followed by a final FC layer and a softmax classifier, which outputs the prediction probability of the diagnostic labels. The cross-entropy (CE) is widely adopted as the training loss function for image classification (Liu et al., 2021), which is given as follows:

$$L = -\frac{1}{C}\sum_{c=1}^{C}\frac{1}{N}\sum_{X_i \in \mathbf{X}} \mathrm{I}\left\{Y_i^c = c\right\} \log(\mathrm{P}(Y_i^c = c | X_i; \mathbf{W})) \tag{5}$$

where $\mathrm{I}\{\cdot\} = 1$ if $\{\cdot\}$ is true, otherwise $\mathrm{I}\{\cdot\} = 0$. N is the total number of test subjects and X_i means the ith sample with the corresponding label Y_i in the training datasets $\mathbf{X}$, and $i \in [1, N]$. $\mathrm{P}(Y_i^c = c | X_i; \mathbf{W})$ measures the probability of the input sample X_i that is correctly classified as the Y_i^c by the trained network with weights $\mathbf{W}$.

3.5. Complexity Analysis

We further analyze our proposed model's complexity by reporting the two branches of subnetworks, respectively. For the aspect of the global subject-level 3D-CNN model, the computational complexity of 3D-CNN layer is $O(D_x D_y D_z K_{global}^3)$, where K_{global} is 3D-CNN kernel size, while D_x, D_y, D_z is the feature map dimensions of the layer. For the aspect of the slice-level 2D-CNN model, since the 2D feature maps are fused in three dimensions, the time complexity of the 2D-CNN layer is $O(M_z D_x D_y K_{slice}^2 + M_x D_y D_z K_{slice}^2 + M_y D_x D_z K_{slice}^2)$, where M_x, M_y, M_z denotes the total number of slices in three MR planes, receptively, and K_{slice} is the 2D-CNN kernel size.

4. EXPERIMENTAL RESULTS

4.1. Competing Methods

We first compare our proposed MSA3D method with multiple deep-learning-based diagnosis approaches that we reproduced

TABLE 4 | Classification results of AD vs. CN and MCI convention on ADNI-2.

Method	AD vs. CN					pMCI vs. sMCI				
	ACC	SEN	SPE	AUC	F1	ACC	SEN	SPE	AUC	F1
Voxel+SVM	0.759	0.677	0.810	0.729	0.705	0.736	0.107	0.769	0.609	0.162
3D-CNN	0.872	0.874	0.839	0.933	0.856	0.769	0.427	0.831	0.721	0.467
Multi-Slice	0.838	0.755	0.826	0.894	0.813	0.728	0.267	0.792	0.620	0.317
Multi-Patch	0.841	0.790	0.844	0.924	0.803	0.722	0.373	0.821	0.698	0.438
MSA3D	**0.911**	**0.888**	**0.914**	**0.950**	**0.898**	**0.801**	**0.520**	**0.856**	**0.789**	**0.553**

All the models are trained on ADNI-1. The best results are highlighted in bold.

and evaluated on the same training and testing datasets including (1) a statistical method based on VBM with SVM [denoted as VBM+SVM, proposed by Ashburner and Friston (2000)], (2) a method using 3D-CNN features [denoted as 3D-CNN, proposed by Wen et al. (2020)], (3) a method using multi-slice 2D features, i.e., the features extracted from all the slices in three directions (denoted as Multi-Slice), and (4) a method using 3D-CNN with 3D patch-level features (denoted as Multi-Patch).

(1) Voxel+SVM: As a conventional statistical-based model, Voxel+SVM performed sMRI analyses at the voxel level (Ashburner and Friston, 2000). Using a non-linear image registration approach, we first normalized all MRIs with the automated anatomical atlas (AAL) template. Then, we segmented the gray matter (GM) from sMRI data. In the end, we mapped the density of GM tissue into one vector and used the support vector machine (SVM) as the classifier for AD diagnosis.
(2) 3D convolution neural network: As an important part of MSA3D, 3D-CNN can extract global subject-level changes of sMRI for dementia diagnosis (Wen et al., 2020). Thus, it can be regarded as the baseline model in our study. In this model, we only give the 3D MRI data as the input for training the 3D-CNN.
(3) Multi-Slice: As another essential component of MSA3D, the multi-slice model focus on the local slice-level features, which consist of all the features extracted by using the 2D-CNN with the 2D slices in sagittal, coronal, and axial MR planes.
(4) Multi-Patch: In this method, multiple 3D-patches are partitioned from the whole brain according to the landmarks defined in Zhang et al. (2016) and Liu et al. (2020b) to extract region-scale features (ROI), and then we train a 3D-CNN as the feature extractor for each patch. In the end, all the ROI-based features were cascaded to obtain the final embedded feature for the entire sMRI.

4.2. Experimental Setting

All the tested models are implemented with Python on Pytorch using one NVIDIA GTX1080TI-11G GPU. During the training stage, the batch size is set to the same value of 12 for all models for a fair comparison. Stochastic gradient descent (SGD) with an initial learning rate of 0.01 and a weighted delay of 0.02 is adopted as the optimization approach, along with an early stopping mechanism for avoiding over-fitting. The following five criteria are calculated to investigate the performance of the tested models, including accuracy (ACC), specificity (SPE), sensitivity (SEN), the area under the ROC curve (AUC), and F1-values (F1).

4.3. Results on ADNI-2

We first present the comparison results of two classification tasks (i.e., AD vs. NC and pMCI vs. sMCI) on ADNI-2 in **Table 4** and **Figure 4**, with the tested methods trained on the ADNI-1. As we can inform from **Table 4**, Multi-Patch shows a better performance than Multi-Slice on AD prediction, especially on the challenging pMCI vs. sMCI. The results indicate that local discriminative features are important for MCI prediction, and only the 2D-slice level features may not be a good option for CNNs. In addition, 3D-CNN achieved the second-best results on both AD and MCI prediction tasks. We can also find that all the deep-learning-based models perform better than the conventional Voxel+svm method. The main reason is that the deep-learning-based technique can achieve a better feature extraction with an end-to-end framework. In general, our model consistently yields better performance than the tested methods, e.g., in the case of MSA3D vs. 3D-CNN baseline, our model resulted in 7 and 5.6% improvements in terms of ACC and AUC for classifying AD/NC, and 7.3% and 16.9% improvements in terms of ACC and AUC for determining pMCI/sMCI. This result shows that after fusion of the 2D and 3D information through two branches of CNNs, our model can capture more discriminative changes in both multiview 2D-slices and 3D whole-brain volumes in the progress of AD and MCI conversion. So that our model generates significant improvements in terms of all the metrics compared to other methods in comparison.

4.4. Results on ADNI-1

In order to further investigate the effectiveness of the test models, we also perform a cross-valuation on ADNI datasets, i.e., we trained the models on ADNI-2 and tested them on ADNI-1. It needs to be pointed out that because of the lack of sufficient pMCI samples in ADNI-2 (75 in ADNI-2 vs. 167 in ADNI-1), we only conduct the experiments of AD diagnosis on ADNI-1. The comparison results are summarized in **Table 5** and **Figure 5**, from which we can observe similar results compared to the models tested on the ADNI-2. Our model still produces the best values in terms of all the metrics compared with the other methods.

Meanwhile, we can find a significant performance drop for all models when trained on ADNI-2, which leads to a relatively small improvement of AUC achieved by our model compared with the 3D-CNN. The main reason for this is that ADNI-1 and

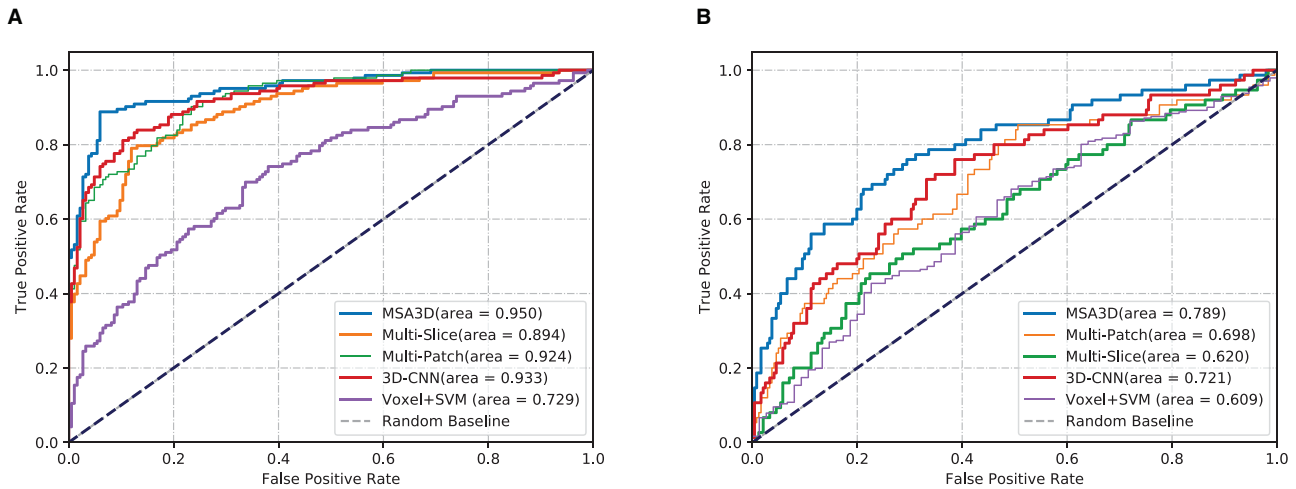


FIGURE 4 | Comparisons results in terms of ROC curves. The models are trained on ADNI-1 and tested on ADNI-2. **(A)** AD vs. NC. **(B)** pMCI vs. sMCI.

ADNI-2 were collected using 1.5 and 3.0 Tesla MRI scanners, respectively. The strength of a 3.0 T magnet is two times that of a 1.5 T magnet, which could cause the overestimation of brain parenchymal volume at 1.5 T (Chu et al., 2016). The variable image quality between different scanners directly impacts the models for diagnosis. However, our model still outperforms the 3D-CNN baseline by 5.3% of the F1 value in this scenario. All of these findings suggest the proposed model's efficacy and reliability.

4.5. Comparison With Other Methods in Literature

In this section, we give a brief description of our MSA3D method with the previous study reported in the literature for AD diagnosis using the ADNI database. The state-of-the-art comparison studies contain:

(1) The conventional statistical-based methods include: SVM trained with Voxel-based features (VBF; Salvatore et al., 2015); landmark-based morphometric features extracted from a local patch (LBM; Zhang et al., 2016); SVM trained with landmark-based features (SVM-landmark; Zhang et al., 2017).

(2) The deep-learning-based methods include: 3D-CNN based on the whole brain sMRI data (whole-3DCNN; Korolev et al., 2017); Multi-layer perception + recurrent neural network using the longitudinal sMRI features (MLP-RNN; Cui et al., 2018); 3D-CNN based on the multiple-modality inputs including sMRI, PET, and MD-DTI data (multi-3DCNN; Khvostikov et al., 2018); 3D-DenseNet based on the 3D-patches features extraction from the hippocampal areas (3D-DenseNet; Liu et al., 2020a); hierarchical fully convolutional network based on 3D-patch and regions features extracted with prior landmarks (wH-FCN; Lian et al., 2020b).

TABLE 5 | Classification results of AD vs. CN on ADNI-1.

Methods	ACC	SEN	SPE	AUC	F1
Voxel+SVM	0.754	0.728	0.781	0.774	0.741
3D-CNN	0.833	0.738	0.813	0.905	0.796
Multi-slice	0.774	0.776	0.812	0.832	0.753
Multi-patch	0.808	0.710	0.793	0.890	0.767
MSA3D	**0.864**	**0.858**	**0.884**	**0.912**	**0.849**

All the models are trained on ADNI-2. The best results are highlighted in bold.

As shown in **Table 6**, We can draw the following conclusions: (1) deep-learning-based methods, especially the CNN-based models, perform much better than most of the conventional statistical methods in terms of ACC. The main reason is that CNN has more feature representation power than handcrafted features. (2) The local features, including ROI-based, landmark-based, and hippocampal segmentation, are also essential to improve the performance of dementia prediction, which indicates that the local changes in whole-brain images provide some valuable clues for AD diagnosis. However, most of these models need predefined landmarks or segmentation regions, which could be hard to obtain potentially informative ROIs due to the local differences between subjects. (3) Different from existing deep-learning-based models (Korolev et al., 2017; Khvostikov et al., 2018; Lian et al., 2020b; Liu et al., 2020a), our proposed model can extract more discriminative features from both local 2D-slice level and 3D-subject level sMRI data using 2D-slice attention network and 3D-CNN, it generates the best ACC, SEN values on AD vs. CN task, and the best SPE and AUC values for predicting pMCI vs. sMCI.

It is noteworthy that our model does not need any predefined landmarks or extra location modules (e.g., hippocampus segmentation), but it achieved better or at least comparative diagnostic results than that of existing deep-learning-based AD

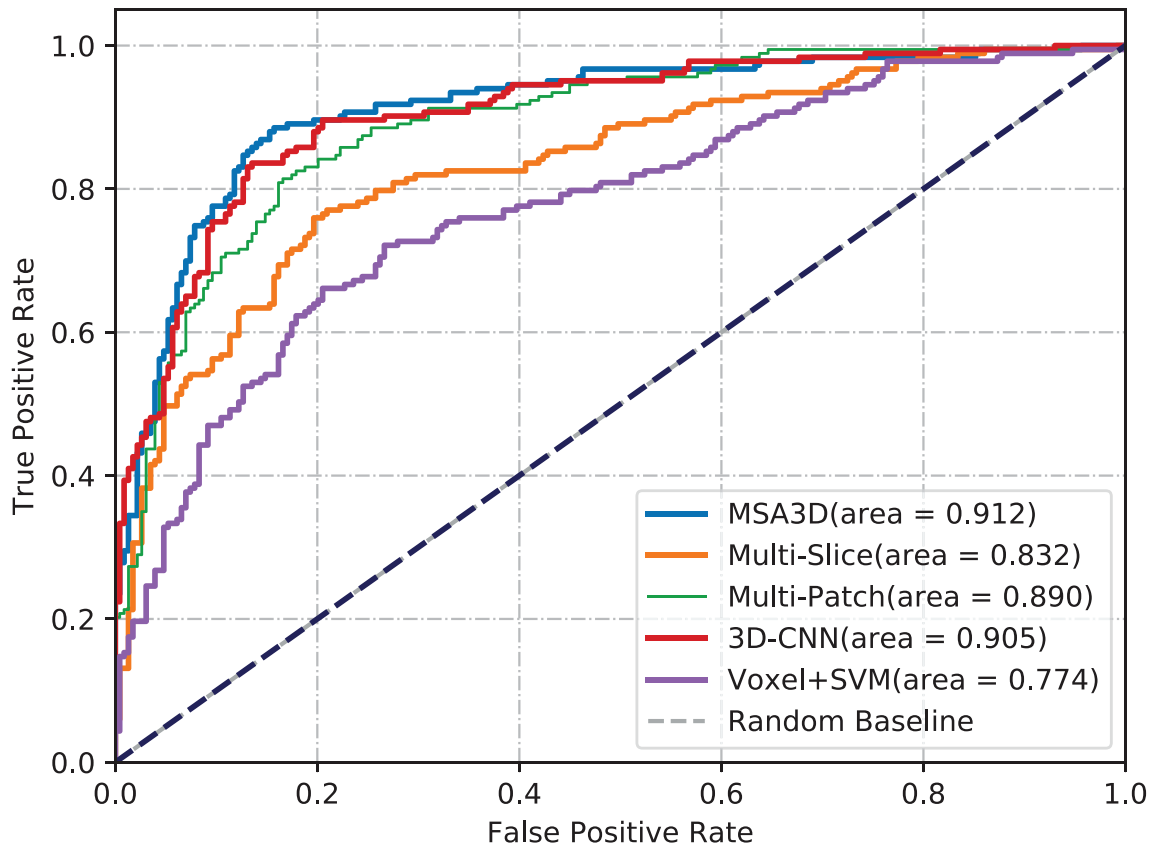


FIGURE 5 | Comparisons of ROC curves. The models are trained on ADNI-2 and tested on ADNI-1.

TABLE 6 | The performance comparison of our model with other state-of-the-art studies report in the literature using the ADNI database for prediction of AD vs. CN and pMCI vs. sMCI.

Method	Test subjects	AD vs. CN				pMCI vs. sMCI			
		ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
VBF	137AD+76sMCI+ 134pMCI+162CN	0.760	–	–	–	0.660			
SVM-Landmark	154 AD+346 MCI +207 CN	0.822	0.774	0.861	0.881	–	–	–	–
LBM	385AD+465sMCI+ 205pMCI+429CN	0.822	0.774	0.861	0.881	0.686	0.395	0.732	0.636
MLP-RNN	198AD+229CN	0.897	0.868	0.925	0.921				
Whole-3DCNN	50AD+77sMCI+ 43pMCI+61CN	0.800	–	–	0.870	0.520	–	–	0.520
Multi-3DCNN	48AD+58CN	0.850	0.880	0.900	–	–	–	–	–
3D-DenseNet	97AD+233MCI +119CN	0.889	0.866	0.808	0.925				
wH-FCN	385AD+465sMCI +205pMCI+429CN	0.903	0.824	0.965	**0.951**	**0.809**	0.526	**0.854**	0.781
Our model	326AD++470sMCI +242pMCI+413CN	**0.911**	**0.888**	0.914	**0.950**	0.801	0.520	**0.856**	**0.789**

The best results are highlighted in bold.

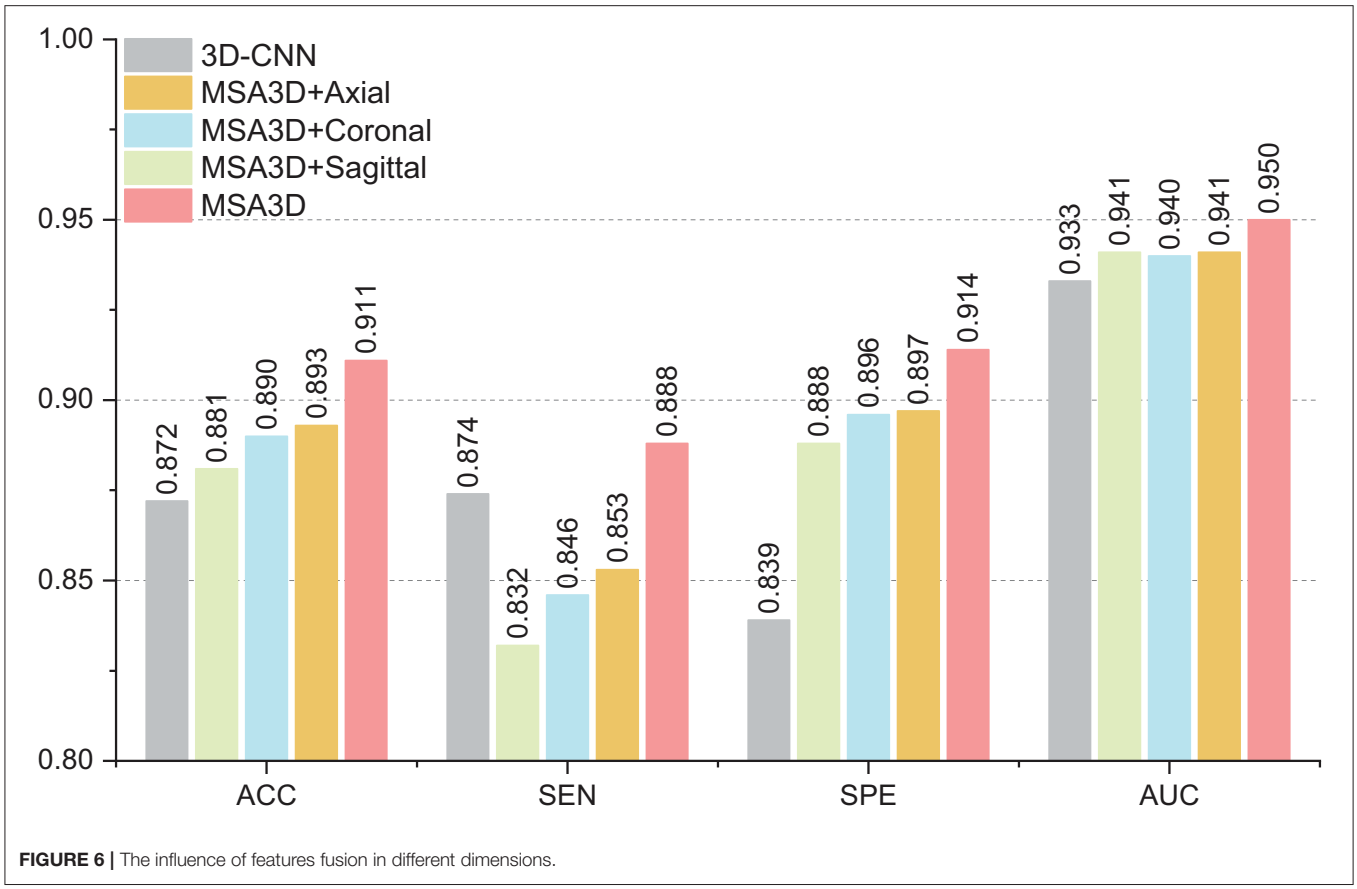


FIGURE 6 | The influence of features fusion in different dimensions.

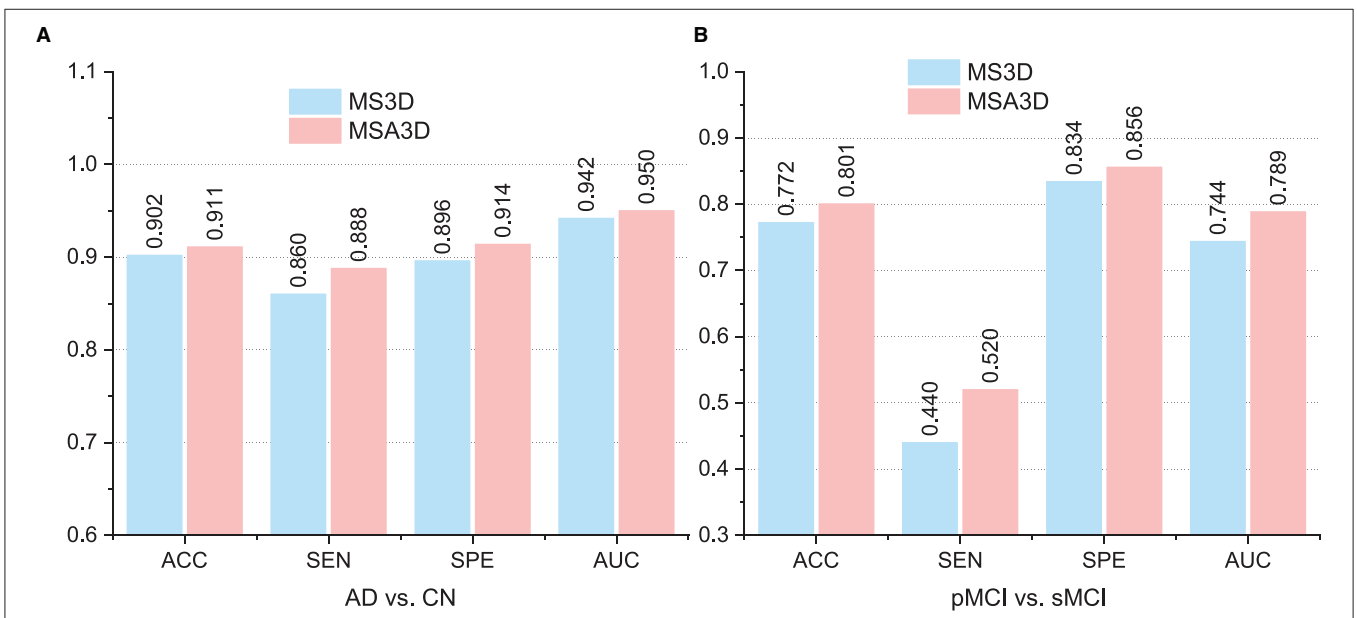


FIGURE 7 | Comparison between multi-view slice fusion without SAM (i.e., MS3D) and multi-view slice fusion with SAM (i.e., MSA3D). **(A,B)** Show the classification results for AD vs. CN and pMCI vs. sMCI, respectively.

diagnosis methods. For example, compared with the second-best wH-FCN model, which extracts features from multiple 3D-patches with hierarchical landmarks proposals, MAS3D generates better results in terms of ACC, SEN, and yields almost the same AUC values on AD vs. CN task. For the aspect of the pMCI vs. sMCI task, our model performs slightly worse than the wH-FCN in terms of ACC and SEN. The possible reason is that wH-FCN adopts more prior knowledge to improve the model's recognition capability, i.e., wH-FCN constrains the distances between landmarks and initializes the network parameters of the MCI prediction model from the task of AD classification.

5. DISCUSSION

5.1. Influence of Features in Different Dimensions

In this section, we investigate the effects of models using multiple slice-level features in different views for AD classification. As shown in **Figure 6**, compared with the model combined with features in the axial plane generates much better results than that of the sagittal and the coronal planes in terms of ACC and SEN. Moreover, after combining the features in three dimensions, our proposed MAS3D outperforms all the tested models, especially yielding significantly better SEN values than the tested methods. This result demonstrates that our 2D- and 3D-features fusion strategy can organically integrate the multi-view-slices features in all directions.

5.2. Influence of Slice Attention Module

As introduced in Section 3.2, the SAM was employed in our MSA3D model to assist the slice-level feature extraction by exploiting the relationships among the slices, i.e., to filter out uninformative slices efficiently. In this subsection, we conducted an ablation experiment for comparison, in which the SAM is removed from our MSA3D, defined as MS3D, to investigate the effectiveness of the proposed SAM, and all the models are trained using ADNI-1 and obtained the test results on ADNI-2.

The comparison results are illustrated in **Figure 7**, from which we can inform that: (1) the two variants of our methods (i.e., MS3D and MSA3D) consistently perform better than the baseline model (i.e., 3D-CNN), which means the fusion of 2D -slice level and 3D subject features provides richer feature representation power for AD diagnosis. (2) the SAM further improved the performance of slice level feature extraction, especially on the challenging MCI prediction task, e.g., The proposed MSA3D generally had better classification performances than MS3D (the ACC and SEN is 0.772 vs. 0.801 and 0.440 vs. 0.520, respectively). This indicates that the proposed SAM can help the neural network focus on specific slices and learn more discriminative 2D-slice level features from abundant slices.

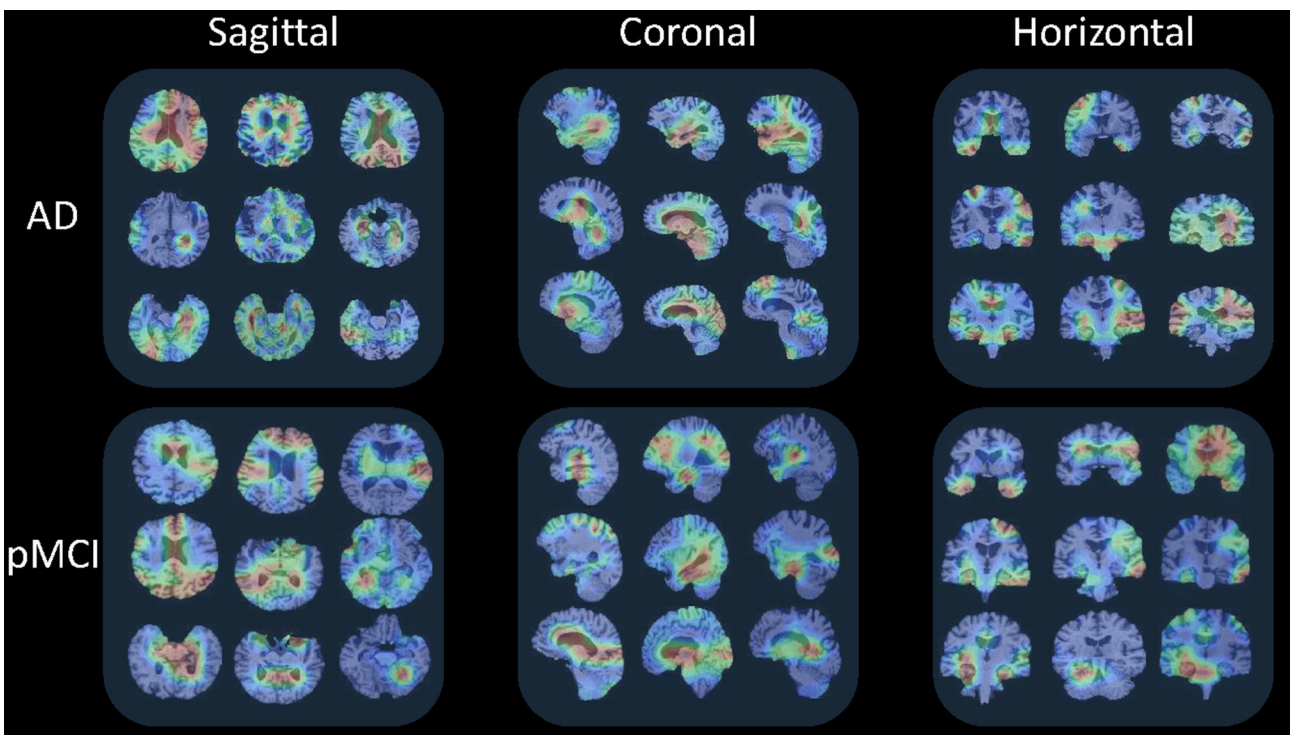


FIGURE 8 | Attention maps of our MAS3D method for predicting multiple subjects selected from the ADNI database with different stages of dementia (i.e., AD and pMCI), respectively. Each subject's attention map is displayed in three MR planes (i.e., sagittal, coronal, and horizontal), where red and blue colors denote high and low discriminative features in sMRI, respectively.

5.3. Visualization of Slices Features

This section visualizes the attention maps produced by our MAS3D method using the Grad-cam (Selvaraju et al., 2020) technology for predicting the subjects with AD and pMCI. The first, second, and third columns of **Figure 8** show the different 2D-slices of sMRI in different views, including sagittal, coronal, and horizontal, respectively, where the corresponding model is trained on ADNI-1, and three AD and three pMCI subjects are randomly selected from ADNI-2 for testing.

From **Figure 8**, we can infer that our model can identify discriminative atrophy areas for different subjects with different stages of dementia, especially for the regions that affect human memory and decision making in the brain. For example, our model emphasizes the atrophy of the frontoparietal cortex, ventricle regions, and hippocampus in the brain. It needs to be pointed out that these highlighted brain regions located by our model in AD diagnosis are consistent with previous clinical research (Chan et al., 2002; Zhang et al., 2021), which have reported the potential sensitive markers for neurodegeneration. All of these results suggest our proposed model can more precisely learn more discriminative features from the brain sMRI for precise dementia diagnosis.

5.4. Limitation and Future Study

While the experimental results suggested our proposed model performed well in automatic dementia detection, its performance and generalization might be potentially enhanced in the future by addressing the limitations listed below.

First, we take advantage of both 2D-slice and 3D-subject features in an integrated MSA3D model. However, the numerous 2D slices observably increased the computational complexity. Since not all the slices help determine the prediction, we could reduce the complexity by using an online feature selection module (Wu D. et al., 2021) to select the 2D slices dynamically. Second, the difference distributions between ADNI-1 and ADNI-2 were not taken into account, i.e., 1.5 T scanners and 3 T scanners for ADNI-1 and ADNI-2, repetitively, which might have a detrimental impact on the model's performance, i.e., the model trained on ADNI-2 and assessed on ADNI-1 performed worse than that trained on ADNI-1 and evaluated on ADNI-2. We could potentially introduce the domain adaption technique into our model to reduce the domain gap between different ADNI datasets. Finally, To further verify the generalization capacity of the proposed model, we will investigate more deep-learning-based methods and test our model on other AD datasets for more AD-related prediction tasks, such as dementia status estimation.

6. CONCLUSION

This study explores a 2D-slice-level and 3D subject-level fusion model for AI-based AD diagnosis using brain sMRI. In addition, a slice attention module is proposed to select the most discriminative slice-level features adaptively from the brain sMRI data. The effectiveness of our model is validated on ADNI-1 and ADNI-2, repetitively, for dementia classification. Specifically, our model achieves 91.1 and 80.1% ACC values on ADNI-1 in AD diagnosis and MCI convention precondition, respectively.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

Ethical review and approval were not required for the study on human participants because all the data in this study were downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.

AUTHOR CONTRIBUTIONS

LC and HQ implemented and optimized the methods and wrote the manuscript. LC and FZ designed the experiment and algorithm. All authors contributed to the article and approved the submitted version.

FUNDING

Publication costs are funded by the National Nature Science Foundation of China under grants (61902370 and 61802360) and are also by the key cooperation project of the Chongqing Municipal Education Commission (HZ2021008 and HZ2021017).

ACKNOWLEDGMENTS

Data used in the preparation of this article were obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) database. The investigators within the ADNI contributed to the design and implementation of ADNI and provided data but did not participate in the analysis or writing of this article. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

REFERENCES

AbdulAzeem, Y. M., Bahgat, W. M., and Badawy, M. M. (2021). A CNN based framework for classification of Alzheimer's disease. *Neural Comput. Appl.* 33, 10415–10428. doi: 10.1007/s00521-021-05799-w

Ashburner, J., and Friston, K. J. (2000). Voxel-based morphometry–the methods. *Neuroimage* 11, 805–821. doi: 10.1006/nimg.2000.0582

Bayram, E., Caldwell, J. Z., and Banks, S. J. (2018). Current understanding of magnetic resonance imaging biomarkers and memory in Alzheimer's disease. *Alzheimer's Dement. Transl. Res. Clin. Intervent.* 4, 395–413. doi: 10.1016/j.trci.2018.04.007

Chan, D., Fox, N., and Rossor, M. (2002). Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology* 58, 838–838. doi: 10.1212/WNL.58.5.838

Chen, Y., and Xia, Y. (2021). Iterative sparse and deep learning for accurate diagnosis of Alzheimer's disease. *Pattern Recogn.* 116:107944. doi: 10.1016/j.patcog.2021.107944

Chu, R., Tauhid, S., Glanz, B. I., Healy, B. C., Kim, G., Oommen, V. V., et al. (2016). Whole brain volume measured from 1.5 t versus 3T MRI in healthy subjects and patients with multiple sclerosis. *J. Neuroimaging* 26, 62–67. doi: 10.1111/jon.12271

Cui, R., and Liu, M. (2019). Hippocampus analysis by combination of 3-d densenet and shapes for Alzheimer's disease diagnosis. *IEEE J. Biomed. Health Informatics* 23, 2099–2107. doi: 10.1109/JBHI.2018.2882392

Cui, R., Liu, M., and Li, G. (2018). "Longitudinal analysis for Alzheimer's disease diagnosis using RNN," in *15th IEEE International Symposium on Biomedical Imaging, ISBI 2018* (Washington, DC), 1398–1401. doi: 10.1109/ISBI.2018.8363833

Ebrahimighahnavieh, M. A., Luo, S., and Chiong, R. (2020). Deep learning to detect Alzheimer's disease from neuroimaging: a systematic literature review. *Comput. Methods Programs Biomed.* 187:105242. doi: 10.1016/j.cmpb.2019.105242

Feng, J., Zhang, S., Chen, L., and Xia, J. (2021). Alzheimer's disease classification using features extracted from nonsubsampled contourlet subband-based individual networks. *Neurocomputing* 421, 260–272. doi: 10.1016/j.neucom.2020.09.012

Fuse, H., Oishi, K., Maikusa, N., and Fukami, T. (2018). "Detection of Alzheimer's disease with shape analysis of MRI images," in *2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS)* (Toyama), 1031–1034. doi: 10.1109/SCIS-ISIS.2018.00171

Jack, C. R. Jr, Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., et al. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27, 685–691. doi: 10.1002/jmri.21049

Jin, D., Xu, J., Zhao, K., Hu, F., Yang, Z., Liu, B., et al. (2019). "Attention-based 3D convolutional network for Alzheimer's disease diagnosis and biomarkers exploration," in *16th IEEE International Symposium on Biomedical Imaging, ISBI 2019* (Venice), 1047–1051. doi: 10.1109/ISBI.2019.8759455

Jin, D., Zhou, B., Han, Y., Ren, J., Han, T., Liu, B., et al. (2020). Generalizable, reproducible, and neuroscientifically interpretable imaging biomarkers for Alzheimer's disease. *Adv. Sci.* 7:2000675. doi: 10.1002/advs.202000675

Khvostikov, A. V., Aderghal, K., Benois-Pineau, J., Krylov, A. S., and Catheline, G. (2018). 3D CNN-based classification using sMRI and MD-DTI images for Alzheimer disease studies. *CoRR, abs/1801.05968*. doi: 10.48550/arXiv.1801.05968

Kincses, Z. T., Kiraly, A., Veréb, D., and Vécsei, L. (2015). Structural magnetic resonance imaging markers of Alzheimer's disease and its retranslation to rodent models. *J. Alzheimer's Dis.* 47, 277–290. doi: 10.3233/JAD-143195

Korolev, S., Safiullin, A., Belyaev, M., and Dodonova, Y. (2017). "Residual and plain convolutional neural networks for 3d brain MRI classification," in *14th IEEE International Symposium on Biomedical Imaging, ISBI 2017* (Melbourne, QLD), 835–838. doi: 10.1109/ISBI.2017.7950647

Lian, C., Liu, M., Pan, Y., and Shen, D. (2020a). Attention-guided hybrid network for dementia diagnosis with structural MR images. *IEEE Trans. Cybern.* 1–12. doi: 10.1109/TCYB.2020.3005859

Lian, C., Liu, M., Wang, L., and Shen, D. (2021). Multi-task weakly-supervised attention network for dementia status estimation with structural MRI. *IEEE Trans. Neural Netw. Learn. Syst.* 1–13. doi: 10.1109/TNNLS.2021.3055772

Lian, C., Liu, M., Zhang, J., and Shen, D. (2020b). Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. *IEEE Trans. Pattern Anal. Mach. Intell.* 42, 880–893. doi: 10.1109/TPAMI.2018.2889096

Liu, M., Chen, L., Du, X., Jin, L., and Shang, M. (2021). Activated gradients for deep neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* 1–13. doi: 10.1109/TNNLS.2021.3106044

Liu, M., Li, F., Yan, H., Wang, K., Ma, Y., Shen, L., et al. (2020a). A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. *NeuroImage* 208:116459. doi: 10.1016/j.neuroimage.2019.116459

Liu, M., Zhang, J., Adeli, E., and Shen, D. (2019). Joint classification and regression *via* deep multi-task multi-channel learning for Alzheimer's disease diagnosis. *IEEE Trans. Biomed. Eng.* 66, 1195–1206. doi: 10.1109/TBME.2018.2869989

Liu, M., Zhang, J., Lian, C., and Shen, D. (2020b). Weakly supervised deep learning for brain disease prognosis using MRI and incomplete clinical scores. *IEEE Trans. Cybern.* 50, 3381–3392. doi: 10.1109/TCYB.2019.2904186

Luk, C. C., Ishaque, A., Khan, M., Ta, D., Chenji, S., Yang, Y.-H., et al. (2018). Alzheimer's disease: 3-dimensional MRI texture for prediction of conversion from mild cognitive impairment. *Alzheimer's Dement.* 10, 755–763. doi: 10.1016/j.dadm.2018.09.002

Martí-Juan, G., Sanroma-Guell, G., and Piella, G. (2020). A survey on machine and statistical learning for longitudinal analysis of neuroimaging data in Alzheimer's disease. *Comput. Methods Prog. Biomed.* 189:105348. doi: 10.1016/j.cmpb.2020.105348

Poloni, K. M., and Ferrari, R. J. (2022). Automated detection, selection and classification of hippocampal landmark points for the diagnosis of Alzheimer's disease. *Comput. Methods Prog. Biomed.* 214:106581. doi: 10.1016/j.cmpb.2021.106581

Qiao, H., Chen, L., Ye, Z., and Zhu, F. (2021). Early Alzheimer's disease diagnosis with the contrastive loss using paired structural MRIs. *Comput. Methods Prog. Biomed.* 208:106282. doi: 10.1016/j.cmpb.2021.106282

Salvatore, C., Cerasa, A., Battista, P., Gilardi, M. C., Quattrone, A., and Castiglioni, I. (2015). Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. *Front. Neurosci.* 9:307. doi: 10.3389/fnins.2015.00307

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2020). Grad-CAM: visual explanations from deep networks *via* gradient-based localization. *Int. J. Comput. Vis.* 128, 336–359. doi: 10.1007/s11263-019-01228-7

Shi, J., Zheng, X., Li, Y., Zhang, Q., and Ying, S. (2018). Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. *IEEE J. Biomed. Health Informatics* 22, 173–183. doi: 10.1109/JBHI.2017.2655720

Tanveer, M., Richhariya, B., Khan, R., Rashid, A., Khanna, P., Prasad, M., and Lin, C. (2020). Machine learning techniques for the diagnosis of Alzheimer's disease: a review. *ACM Trans. Multim. Comput. Commun. Appl.* 16, 1–35. doi: 10.1145/3344998

Vu, T. D., Ho, N., Yang, H., Kim, J., and Song, H. (2018). Non-white matter tissue extraction and deep convolutional neural network for Alzheimer's disease detection. *Soft Comput.* 22, 6825–6833. doi: 10.1007/s00500-018-3421-5

Wang, Y., Nie, J., Yap, P., Shi, F., Guo, L., and Shen, D. (2011). "Robust deformable-surface-based skull-stripping for large-scale studies," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011 - 14th International Conference*, Vol. 6893 of Lecture Notes in Computer Science, eds G. Fichtinger, A. L. Martel, and T. M. Peters (Toronto, ON: Springer), 635–642. doi: 10.1007/978-3-642-23626-6_78

Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-González, J., Routier, A., Bottani, S., et al. (2020). Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation. *Med. Image Anal.* 63:101694. doi: 10.1016/j.media.2020.101694

Wu, D., He, Y., Luo, X., and Zhou, M. (2021). A latent factor analysis-based approach to online sparse streaming feature selection. *IEEE Trans. Syst. Man Cybern.* 1–15. doi: 10.1109/TSMC.2021.3096065

Wu, E. Q., Hu, D., Deng, P.-Y., Tang, Z., Cao, Y., Zhang, W.-M., et al. (2021a). Nonparametric bayesian prior inducing deep network for automatic detection of cognitive status. *IEEE Trans. Cybern.* 51, 5483–5496. doi: 10.1109/TCYB.2020.2977267

Wu, E. Q., Lin, C.-T., Zhu, L.-M., Tang, Z. R., Jie, Y.-W., and Zhou, G.-R. (2021b). Fatigue detection of pilots' brain through brains cognitive map and multilayer latent incremental learning model. *IEEE Trans. Cybern.* 1–13. doi: 10.1109/TCYB.2021.3068300

Zhang, J., Gao, Y., Gao, Y., Munsell, B. C., and Shen, D. (2016). Detecting anatomical landmarks for fast alzheimer's disease diagnosis. *IEEE Trans. Med. Imaging* 35, 2524–2533. doi: 10.1109/TMI.2016.2582386

Zhang, J., Liu, M., An, L., Gao, Y., and Shen, D. (2017). Alzheimer's disease diagnosis using landmark-based features from longitudinal

structural MR images. *IEEE J. Biomed. Health Inform.* 21, 1607–1616. doi: 10.1109/JBHI.2017.2704614

Zhang, T., Li, C., Li, P., Peng, Y., Kang, X., Jiang, C., et al. (2020). Separated channel attention convolutional neural network (SC-CNN-attention) to identify ADHD in multi-site RS-fMRI dataset. *Entropy* 22:893. doi: 10.3390/e22080893

Zhang, Y., Jiang, X., Qiao, L., and Liu, M. (2021). Modularity-guided functional brain network analysis for early-stage dementia identification. *Front. Neurosci.* 15:720909. doi: 10.3389/fnins.2021.720909

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 24 May 2022
doi: 10.3389/fnagi.2022.911513

Time-Varying Effective Connectivity for Describing the Dynamic Brain Networks of Post-stroke Rehabilitation

Fangzhou Xu[1], Yuandong Wang[1,2], Han Li[1,2], Xin Yu[1,2], Chongfeng Wang[1,2], Ming Liu[1,2], Lin Jiang[3,4], Chao Feng[1], Jianfei Li[1], Dezheng Wang[5], Zhiguo Yan[2]*, Yang Zhang[5]* and Jiancai Leng[1]**

[1] International School for Optoelectronic Engineering, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China, [2] School of Electrical Engineering and Automation, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China, [3] The Clinical Hospital of Chengdu Brain Science Institute, Ministry of Education Key Lab for Neuroinformation, University of Electronic Science and Technology of China, Chengdu, China, [4] School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China, [5] The Department of Physical Medicine and Rehabilitation, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China

OPEN ACCESS

Edited by:
Zehong Jimmy Cao,
University of South Australia, Australia

Reviewed by:
Dong Wen,
University of Science and Technology Beijing, China
Minpeng Xu,
Tianjin University, China

***Correspondence:**
Fangzhou Xu
xfz@qlu.edu.cn
Zhiguo Yan
yanzg500@sina.com
Yang Zhang
zhangyang982003@163.com
Jiancai Leng
jiancaileng@qlu.edu.cn

Specialty section:
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

Received: 02 April 2022
Accepted: 19 April 2022
Published: 24 May 2022

Citation:
Xu F, Wang Y, Li H, Yu X, Wang C, Liu M, Jiang L, Feng C, Li J, Wang D, Yan Z, Zhang Y and Leng J (2022) Time-Varying Effective Connectivity for Describing the Dynamic Brain Networks of Post-stroke Rehabilitation. Front. Aging Neurosci. 14:911513. doi: 10.3389/fnagi.2022.911513

Hemiplegia is a common motor dysfunction caused by a stroke. However, the dynamic network mechanism of brain processing information in post-stroke hemiplegic patients has not been revealed when performing motor imagery (MI) tasks. We acquire electroencephalography (EEG) data from healthy subjects and post-stroke hemiplegic patients and use the Fugl-Meyer assessment (FMA) to assess the degree of motor function damage in stroke patients. Time-varying MI networks are constructed using the adaptive directed transfer function (ADTF) method to explore the dynamic network mechanism of MI in post-stroke hemiplegic patients. Finally, correlation analysis has been conducted to study potential relationships between global efficiency and FMA scores. The performance of our proposed method has shown that the brain network pattern of stroke patients does not significantly change from laterality to bilateral symmetry when performing MI recognition. The main change is that the contralateral motor areas of the brain damage and the effective connection between the frontal lobe and the non-motor areas are enhanced, to compensate for motor dysfunction in stroke patients. We also find that there is a correlation between FMA scores and global efficiency. These findings help us better understand the dynamic brain network of patients with post-stroke when processing MI information. The network properties may provide a reliable biomarker for the objective evaluation of the functional rehabilitation diagnosis of stroke patients.

Keywords: stroke, motor imagery, time-varying network, graph theory, Fugl-Meyer assessment

INTRODUCTION

Stroke, also known as cerebrovascular accident, is a disease of the blood vessels supplying the brain are damaged. It can lead to avascular necrosis or hemorrhage of our brain tissue. Stroke has high morbidity, disability, and mortality rates, 40% of stroke survivors still suffer from various disabilities, and the incidence of stroke increases disproportionately with age. Moreover, aging is a stroke risk factor (Egorova et al., 2019).

Two types of Motor Imagery (MI) can be divided: Kinesthetic Motor Imagery (KMI) and Visual Motor Imagery (VMI). KMI is defined as the thought process of imagining a given movement without any motor output. VMI mainly relies on the visualization of the execution of that movement (Rimbert et al., 2019). MI is regarded as a mental process involving a variety of advanced cognitive functions (Li et al., 2019). The MI-based brain-computer interface (BCI) has been widely used in motor function rehabilitation, motor skill learning, and other fields (Long et al., 2011; Mane et al., 2020; Xu et al., 2021b). Patients with motor cortex damage can get better functional recovery by MI therapy (Xu et al., 2021a). Researchers aim to obtain good performance from MI recognition (Xu et al., 2016, 2021c; Wang et al., 2020). Electroencephalography (EEG), as a method of recording brain activity using electrophysiological indicators, has the characteristics of high time resolution, low cost, and easy operation (Zhang et al., 2015; Xu et al., 2020a). EEG is the most commonly used brain signal for BCI (Xu et al., 2021d). In recent years, EEG-based BCI systems have developed rapidly, and the number of commands that BCI can process has increased from the initial 30 to more than 100 (Xu et al., 2020b). Recently, the measurement precision of BCI first reached the level of the sub-microvolt in amplitude (Xu et al., 2018), which significantly broadened the category of BCI. EEG is also the commonly used brain signal for clinical rehabilitation. Ding et al. (2022) have used transcranial magnetic stimulation and electroencephalography (TMS-EEG) to directly measure cortical responses in aging stroke patients after intermittent theta-burst stimulation (iTBS) and found that iTBS can normalize natural frequency in aging stroke patients, which can be utilized in stroke rehabilitation.

The human brain is a complex network consisting of a large number of interconnected cortical regions. Recently, the brain network method has attracted much attention and has been widely used in decoding related cognitive functions. The main methods of brain networks are effective and functional connectivity. Functional connectivity is an undirected network that represents the coordination mechanism between different neurons (Reid et al., 2019). Effective connectivity is a directed network defined as the direct or indirect influence from one brain function area to another brain function area (Park et al., 2018).

Based on EEG analysis, directed networks have directional information compared to undirected networks. The directed networks can more accurately assess the information flow between brain nodes and better understand the brain's information processing mechanism when performing MI recognition. Directed analysis methods such as granger causality analysis (GCA), partial directed coherence (PDC), and directed transfer function (DTF) have significant advantages in capturing directional coupling between different brain regions (Jastreboff, 1990; Maudoux et al., 2012). Based on the DTF method, Vecchio and Babiloni (2011) have found that the directionality of frontal-parietal EEG synchronization in Alzheimer's Disease (AD) and Amnestic Mild Cognitive Impairment (MCI) is abnormal.

EEG has millisecond-level time resolution, which leads to different network structures corresponding to different stages of the brain processing information. Therefore, the study of time-varying networks helps us to explore the dynamic process of brain information processing in MI recognition and to capture the time-varying connections of cognitive processes. Including time-varying granger causality analysis (tv-GCA), time-varying partial directed coherence (tv-PDC), and adaptive directed transfer function (ADTF) can get different network connection structures in different cognitive procedures (Li et al., 2015; Manomaisaowapak et al., 2015). Li et al. (2016) have used an adaptive directed transfer function to construct a time-varying network of P300 and found that different stages of P300 induce different brain network structures. Based on the ADTF method, Si et al. (2019) have studied the role of the frontal cortex in the decision-making stage and the different network structures in different decision-making stages.

Fugl-Meyer assessment (FMA) is an authoritative method to assess the motor function of stroke patients. It can provide a visual representation of motor function after stroke, and can play an important role in the baseline assessment, as well as monitor and quantify longitudinal changes in motor function (Riahi et al., 2020). FMA is a reliable and effective method for measuring motor dysfunction, a higher score corresponds to better motor function (Saes et al., 2019). All patients have been completed the FMA to ensure the consistency of the FMA scores and the EEG recording. Saes et al. (2021) have used the resting state EEG parameters of stroke patients to predict FMA scores, and they have proved that resting-state EEG parameters can be used as a biomarker for predicting stroke recovery. A challenge associated with this assessment is the availability of trained doctors to conduct the evaluation. The study of biomarkers can estimate that FMA may help to solve the problem.

The network mechanism of stroke patients based on the ADTF method has been studied. The dynamic reorganization and compensation of the brain network have been revealed. The correlation between network properties and FMA scores has been analyzed. Our proposed method provides a new neuroregulatory index for diagnosis and treatment of post-stroke patients.

MATERIALS AND METHODS

Participants

After receiving a detailed explanation of the purpose and potential risks of the experiment, all subjects have provided written informed consent. The study protocols have been approved by the medical ethics committee of Qilu Hospital, Cheeloo College of Medicine, Shandong University. The study is carried out in accordance with relevant guidelines and regulations. Twenty-one right-handed subjects have been recruited in our current study, consisting of seven male patients with left hemiplegic stroke (marked as LS, age 49 $\pm$ 12 years); five male patients with right hemiplegic stroke (marked as RS, age 54 $\pm$ 8 years); nine male health control (marked as HC, age 45 $\pm$ 12 years). All subjects have normal hearing and vision, and no psychiatric drugs are taken for healthy subjects.

Experimental Procedures

The experiment is conducted in a separate relatively shielded room. The room is lighted with soft luminance. In addition, during the acquisition of EEG signals, the indoor temperature

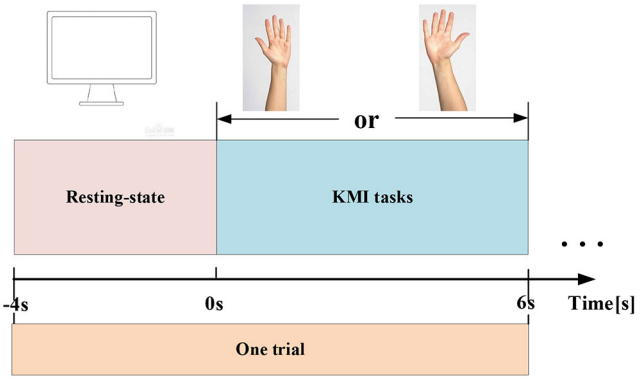


FIGURE 1 | EEG experimental paradigm. One KMI trial includes a 4-s resting state (represented by a blank screen), and a 6-s KMI task (represented by the left or right arrow on the screen).

is maintained at ∼21°C by the air conditioner, and the doors and windows are tightly closed to avoid the influence of noise. Throughout the experiment, all subjects are asked to stay relaxed to avoid real hand movements affecting the validity of the data. The experiment adopts the KMI paradigm. Each subject has performed 70 independent experiments, including 30 MIs for each of the left and right hands, 10 actual exercises, and EEG data have been acquired from 64 electrodes. Each KMI trial has a total of 10 s. The first 4 s are resting, and a blank screen appears to remind the subjects to prepare, and the next 6 s are the task state. When the KMI recognition starts, a left or right arrow appears on the screen to remind the subjects to imagine the left-hand or right-hand lifting action. The left-hand or right-hand KMI trials are randomly presented to the subjects. The experimental paradigm is shown in **Figure 1**.

Signal Recording

A BrainAmp 67-node amplifier from Brain Products (Australia) has been used to record EEG. All 64 Ag/AgCl electrodes are placed according to the 10–20 international system. The REF electrode between the CZ electrode and the CPZ electrode is used as a reference. In all experiments, the sampling rate is 1000 Hz.

Data Analysis

In this study, the preprocessing procedure and analysis procedure are shown in **Figure 2**, the time-varying network analysis has been performed and the correlation between the global efficiency (GE) and the FMA score has been calculated.

Preprocessing

The purpose of preprocessing is to acquire clean EEG data for subsequent analysis. The detailed procedures include 8–30 Hz band-pass filtering, performing reference electrode standardization technique (REST) processing on the filtering data (Yao, 2001; Dong et al., 2017), segmenting data with a time window of [−4 s, 6 s] (0 s corresponds to the stimulus onsets), and removing bad trials [±70 μV as the threshold for ocular artifacts (Li et al., 2019, 2021)]. Then, the data has been down-sampled to 100 Hz. To reduce the influence of the volume conduction between network nodes, 21 electrodes (i.e., Fp1, Fpz, Fp2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1, O2, and Oz) of the 64 electrodes have been selected to construct the brain functional network.

Time-Varying Network Pattern Analysis

For each subject, the preprocessed EEG is used to further construct time-varying KMI networks based on ADTF (Li et al., 2016). Then, the left-hand and right-hand time-varying KMI networks corresponding to each trial are averaged for each subject. Therefore, a time-varying network of two-classes KMI tasks is generated. The detailed description of ADTF in our study is as follows:

Time-Varying Multivariable Adaptive Autoregressive Model

For the time series of each subject's trial, the following formula can be used to construct a corresponding time-varying multivariable adaptive autoregressive (tv-MVAAR) model to describe the dataset:

$$X(t) = \sum_{i=1}^{p} A(i,t)\, X(t-i) + E(t)$$

where $X(t)$ is the data vector of each trial at time t, $A(i,t)$ denotes the model coefficient matrix estimated by the Kalman filter algorithm (Arnold et al., 1998; Pagnotta and Plomp, 2018), $E(t)$ denotes the multivariate independent white noise, p is the optimal model order automatically determined by the Akaike information criterion (AIC) within the range of 2–20.

$$AIC(p) = \ln\left[\det(\varepsilon)\right] + 2\beta^2 p/\alpha$$

where β is the number of nodes, p is the order of the best model of tv-MVAAR, α is the number of sampling points in the time of [−4 s, 6 s] (0 s corresponds to the stimulus onsets), and ε is the corresponding covariance matrix.

Adaptive Directed Transfer Function

The time-varying model coefficient matrix $A(i,t)$ can be transformed in the frequency domain to obtain the transfer matrix $H(f,t)$ of the time-varying model, which can be further derived $H_{ij}(f,t)$ is the directional information flow from the jth node to the ith node at time t. Then, the time-frequency representations of $X(t)$ and $A(i,t)$ are described as follows:

$$A(f,t)\, X(f,t) = E(f,t)$$
$$X(f,t) = A^{-1}(f,t) E(f,t) = H(f,t)\, E(f,t)$$

where $A(f,t) = \sum_{k=0}^{p} A_k(t)\, e^{-j2\pi f \Delta t k}$ with the $A(t)$ denotes the matrix of model coefficients, $X(f,t)$ and $E(f,t)$ are the representations of $X(t)$ and $E(t)$ in the frequency domain, respectively.

Under the premise of a given frequency f and corresponding time point t, the ADTF value describes the directional causal interaction from the jth node to the ith node is normalized and defined as:

$$r_{ij}^2(f,t) = \frac{\left|H_{ij}(f,t)\right|^2}{\sum_{m=1}^{n}\left|H_{im}(f,t)\right|^2}$$

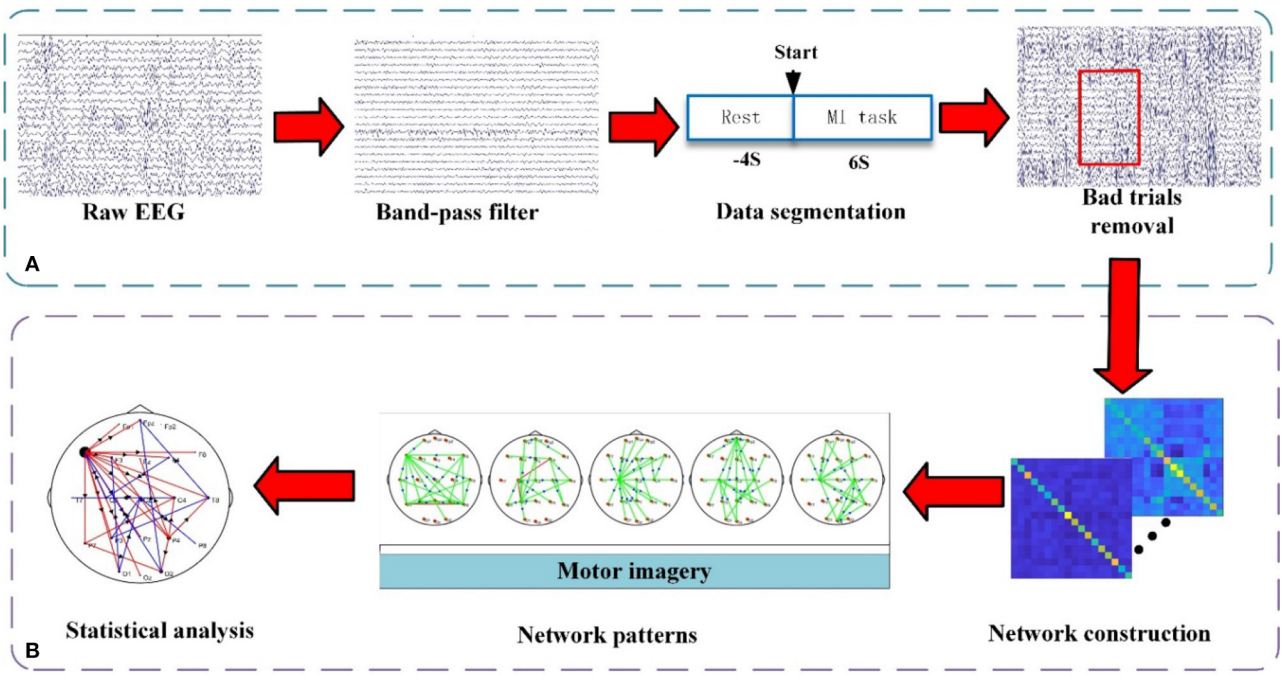


FIGURE 2 | The framework of EEG processing procedure. **(A)** Preprocessing. **(B)** Time-varying network pattern analysis. Between the two electrodes, the connecting edge represents the coupling relationship and the arrow represents the flow direction.

Finally, the ADTF values on the frequency band of interest containing MI-related rhythms at 8–30 Hz are averaged to evaluate the directional information flow of two different nodes (Burianová et al., 2013; Zhang et al., 2018):

$$\Theta_{ij}^{2}(t) = \frac{\sum_{k=f_1}^{f_2} r_{ij}^{2}(k,t)}{f_2 - f_1}$$

For each subject, all trials connectivity networks are averaged across all of these artifact-free trials and then induce the final time-varying network model. When exploring the group-wise networks' differences, the time-varying networks of the LS, RS, and HC have been binarily thresholded into the time-varying binary networks with a connectivity cost of 5% to illustrate the time-varying network architectures. The networks have been also statistically compared by using the non-parametric Wilcoxon rank-sum test. Some previous studies have shown that the difference between Θ_{ij} and Θ_{ji} determines the direction of information flow in time-varying networks (Babiloni et al., 2009; Vecchio and Babiloni, 2011). As ADTF captures the dynamic networks for each time point, and nearby time points have shown highly similar network patterns. In our study, we describe the KMI time-varying networks with a time interval of 1.5 s and reveal the dynamic KMI network mechanism by evaluating the time-varying networks corresponding to different KMI stages.

Time-Varying Network Properties

According to the obtained adjacency matrix, Brain Connectivity Toolbox (BCT, http://www.nitrc.org/projects/bct/) has been employed to calculate the GE of all subjects at each time point (Zhang et al., 2020), the time-varying KMI networks are analyzed through graph theory. The GE describes the ability of the brain network to process information. The GE calculation formula is as follows:

$$GE = \frac{1}{n}\sum_{i\in N}\frac{\sum_{j\in N, j\neq i}\left(d_{\overrightarrow{ij}}\right)^{-1}}{n-1}$$

Here, n represents the node number, $d_{\overrightarrow{ij}}$ represents the shortest characteristic path length, and N denotes the set of current network nodes.

Correlation Analysis Between Time-Varying Network and FMA

According to the FMA scores, 12 stroke patients have been divided into three classes: severe (FMA: 0–20), moderate (FMA: 20–40), and mild (FMA: 40–60). The 12 patients are ranked from lowest to highest score. The high scores correspond to better motor function, and the low scores correspond to poor motor function. Pearson correlation analysis has been used to explore the potential relationship between each patient's GE and FMA scores to reveal whether the network properties can be used as potential biomarkers to indicate the degree of motor function rehabilitation.

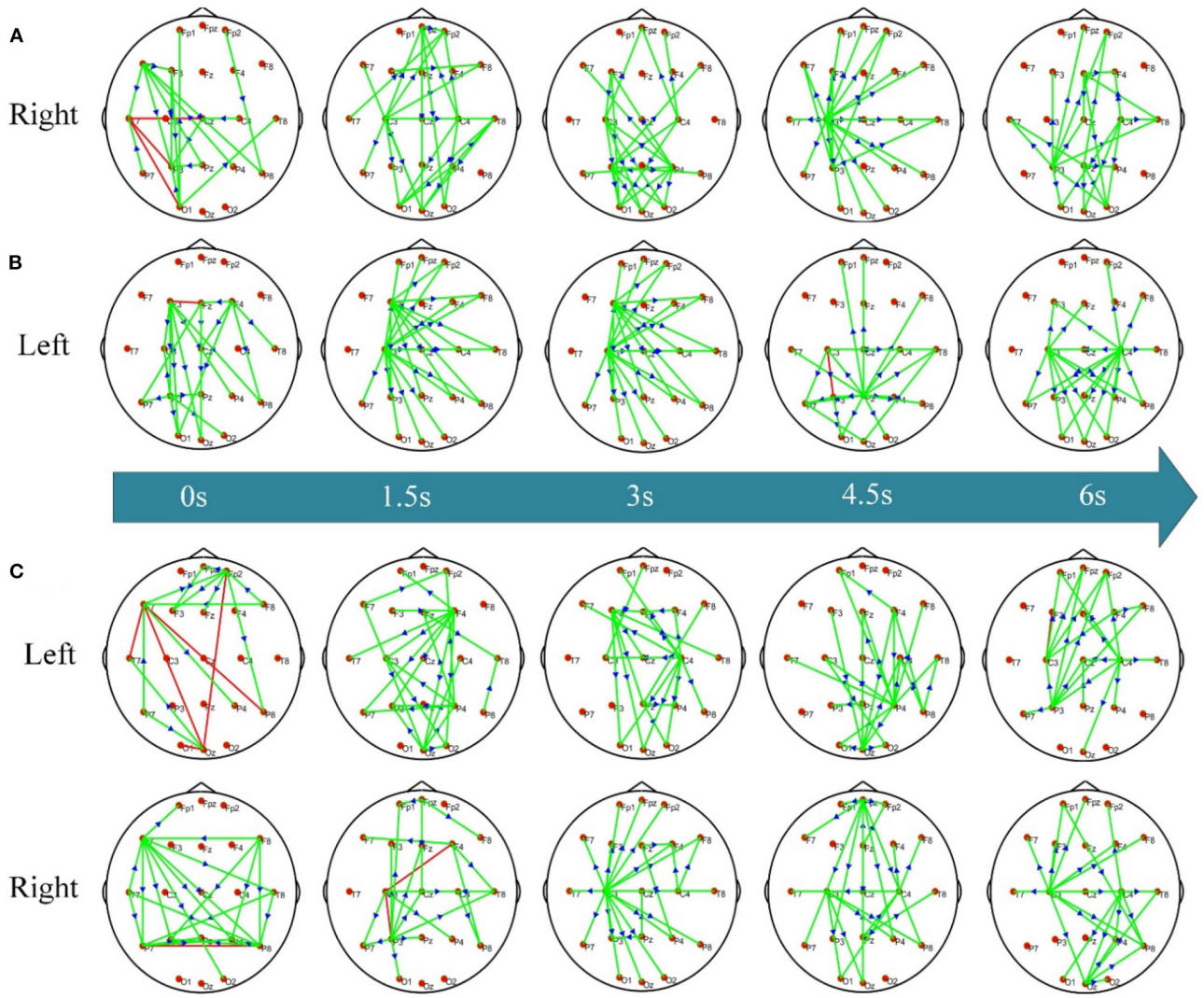


FIGURE 3 | The dynamic KMI network patterns of RS/LS/HC. **(A)** Time-varying network pattern in the right hand of the RS group; **(B)** time-varying network pattern of the left hand in LS group; **(C)** time-varying network pattern of left/right hand in HC group. The connecting edge in the figure represents the coupling relationship between the two electrodes, the red edge represents the two-way connection between the two nodes, the green edge represents the one-way connection between the nodes, and the arrow represents the flow direction between them.

RESULTS

Dynamic KMI Network Patterns

To investigate the dynamic network difference between post-stroke hemiplegic patients and healthy subjects when performing KMI recognition, the ADTF function has been used to calculate the time-varying network matrix of LS, RS, and HC groups in the 8–30 Hz frequency band of interest, and take the sparsity of 5% (i.e., the connection edge with the strongest weight remaining 5%) to display the transient topology. When performing the right-hand KMI tasks, the crucial hubs for the RS subjects (**Figure 3A**) are located at the contralateral P4 and ipsilateral P3. The motor areas of the stroked hemisphere (i.e., right hemisphere) for the LS subjects (**Figure 3B**) have been shown the weaker connectivity when executing the left-hand KMI tasks, but the contralateral F3 and C3 electrodes (i.e., at the left hemisphere) extend to the occipital lobe have been shown the stronger connectivity. However, the electrodes C3 or C4 for the HC subjects (**Figure 3C**) have served as the important hub to control the KMI recognition, and then have transferred to the joint control from bilateral C3 and C4 electrodes.

Dynamic Network Differences

To further explore the differential dynamic network patterns of the time-varying networks between post-stroke hemiplegic patients and healthy subjects, **Figure 4** shows the corresponding statistical network topology diagrams at different time points. Compared to HC subjects (**Figure 4A**), stronger information flow in the LS group has transferred from the occipital lobe

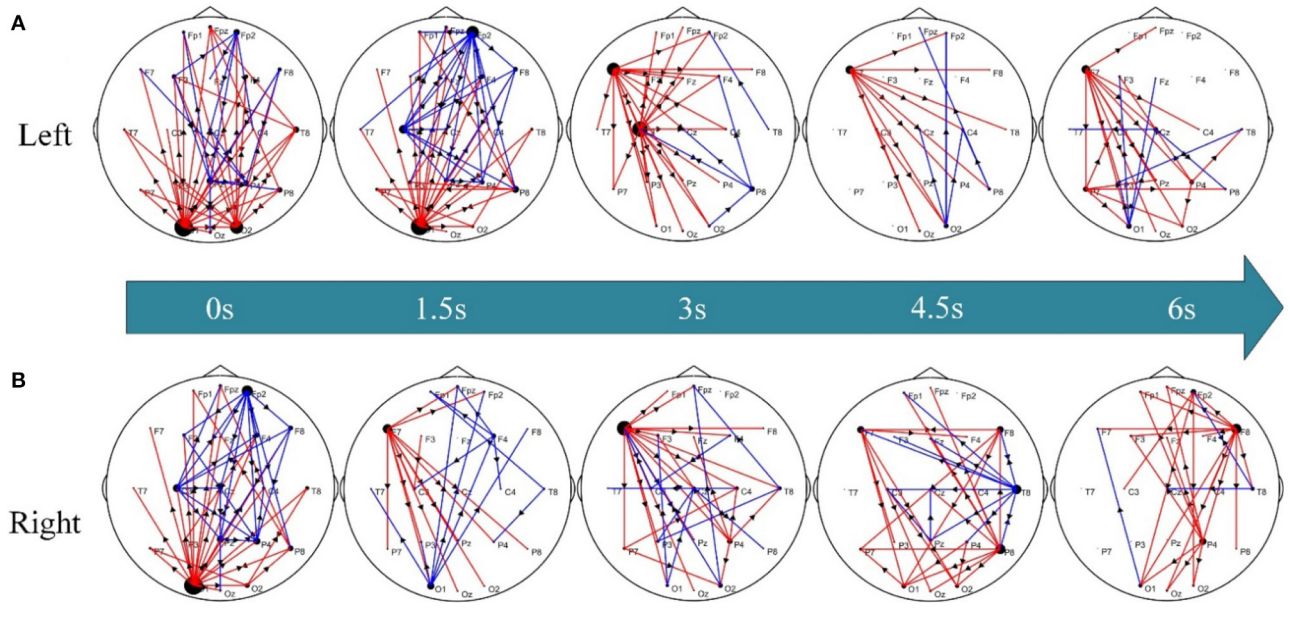


FIGURE 4 | Differential time-varying network topologies between the pairwise groups. **(A)** LS vs. HC groups and **(B)** RS vs. HC groups. Here, the red edge represents the connection edge where LS/RS is stronger than HC, the blue edge denotes the connection edge where HC is stronger than LS/RS, and the arrow indicates the information flow between nodes.

(e.g., O1, O2) to the left frontal lobe (e.g., F7); however, these phenomena in the RS group (**Figure 4B**) have occurred from occipital lobe (e.g., O1) and left frontal lobe (e.g., F7) to the right frontal lobe (e.g., F8).

Dynamic of the Time-Varying GE

To further explore the connection pattern of the time-varying network, the values of GE at each time point are averaged for the subjects in the three groups of LS, RS, and HC. **Figure 5A** shows the GE increase along with the progress of the KMI recognition. When performing right-hand KMI tasks, the GE of the HC group is greater than that of the RS group ($p < 0.05$), as shown in **Figure 5B**.

Correlation of GE and FMA Scores

Clinically, the higher the FMA scores are responding to the less severe the damage of motor function. **Figure 6A** shows the average GE of 12 stroke patients, and the x-axis represents the 12 stroke patients who have been ranked in ascending order of FMA scores. The scatter plot of GE and FMA of 12 stroke patients and the positive correlation ($r = 0.61$, $p = 0.035$) are shown in **Figure 6B**.

DISCUSSION

Stroke causes damage to the motor functional areas of the brain, which in turn leads to motor dysfunction. Compared with healthy subjects, the functional connections between different brain regions of stroke patients are more complicated in performing KMI. Moreover, the brain processes information very efficiently, which leads to different network structures corresponding to different cognitive stages. To evaluate the network reorganization and compensation of brain function after stroke, the ADTF has been employed to better explore the dynamic network mechanism of post-stroke hemiplegic patients and healthy subjects during the execution of KMI.

Time-varying network topology diagrams under different conditions are calculated to study the interaction patterns between different brain regions of post-stroke hemiplegic patients and healthy subjects. **Figure 3** shows the dynamic network patterns of the RS, LS, and HC groups when performing KMI recognition. When the patients with left brain damage perform right-hand KMI tasks, the connection between the motor areas on the stroked left hemisphere and other functional brain areas is enhanced, and the hub node has transferred from node C3 to node C4, as shown in **Figure 3A**. The enhancement of the bilateral occipital lobe (i.e., P3 and P4) connection is enhanced during the later stage of KMI. These phenomena might further indicate that the contralateral brain areas of the stroked hemisphere have functional compensation, and the ipsilateral non-motor areas that are responsible for the high-level cognition also have functional compensation, such as motor planning and attention (Li et al., 2021). When the patients with right brain damage imagine the left-hand movement, stronger functional connectivity has existed between the frontal and parietal-occipital lobe, while seldom connectivity of the stroked right hemisphere has been observed, as shown in **Figure 3B**. The phenomena may account for the deficits in performing left-hand KMI tasks and left-hand wrist extension of the LS patients. The frontal and parietal lobe is responsible for the advanced regulation of limb movement. Right brain damage causes human motor

dysfunction, the left brain areas increase the response to provide compensation for motor function (Li et al., 2020). Thereafter, the bilateral motor areas C3 and C4 are more connected. Because the brain of stroke patients is damaged, the response pattern and functional connection of the brain are different from healthy subjects. The location and severity of brain damage affect the degree of brain function network remodeling (Arun et al., 2020). For healthy subjects, when performing left-hand and right-hand KMI tasks, the brain function networks appear a network pattern from the opposite side to the bilateral connection, as shown in **Figure 3C**. When performing the left-hand KMI tasks, the network connection of the right motor areas is enhanced, and then the network connection gradually appears in the bilateral motor areas. During the right-hand KMI procedure, the left motor areas have presented significantly stronger connectivity, which has switched to a bilateral connectivity architecture. During KMI procedure, the brain functional areas involved in healthy subjects include the main motor areas, medial frontal gyrus, parietal lobe, and primary motor cortex (Zhou et al., 2022). The KMI procedure of health subjects mainly responds to the contralateral brain areas (Sharma and Baron, 2013).

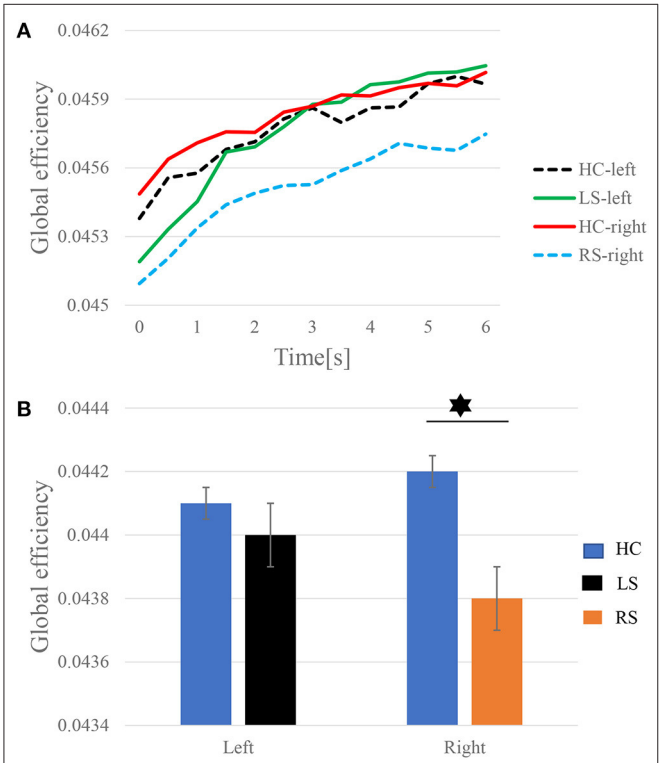


FIGURE 5 | The time-varying GE of left-hand and right-hand KMI recognition. **(A)** Dynamics of the GE. **(B)** Statistics of the average GE. The asterisk represents significant differences in GE between the two groups ($p < 0.05$).

The functional compensation and plasticity of the brain after stroke are related to the functional connection difference between stroke and healthy subjects, and are related to the response between different brain regions (Bundy and Nudo, 2019). The study further explores the abnormal networks connection status of stroke patients. Under the premise of the LS group and HC group, the connecting edge of LS is significantly stronger than HC from the occipital lobe to the left frontal lobe, as shown in **Figure 4A**. At the beginning of the KMI recognition, when the subjects see the prompt instruction, the LS is relative to the HC, the connection of the occipital lobe is stronger at this time. The occipital lobe is the center of the visual cortex (Chu et al., 2021). The damage to the brain motor function areas of stroke patients leads to paying more attention to prompt instructions. Therefore, the patients' attention to action prompt instructions is also a good compensation effect for the motor dysfunction (Rowe et al., 2002). In addition, the stronger connectivity of the occipital lobe plays an important role in improving the performance of SSVEP-based BCI systems (Gao et al., 2018; Sun et al., 2020). During KMI recognition, the functional connections of the LS brain are enhanced from the left frontal lobe (i.e., F7 node) to the bilateral

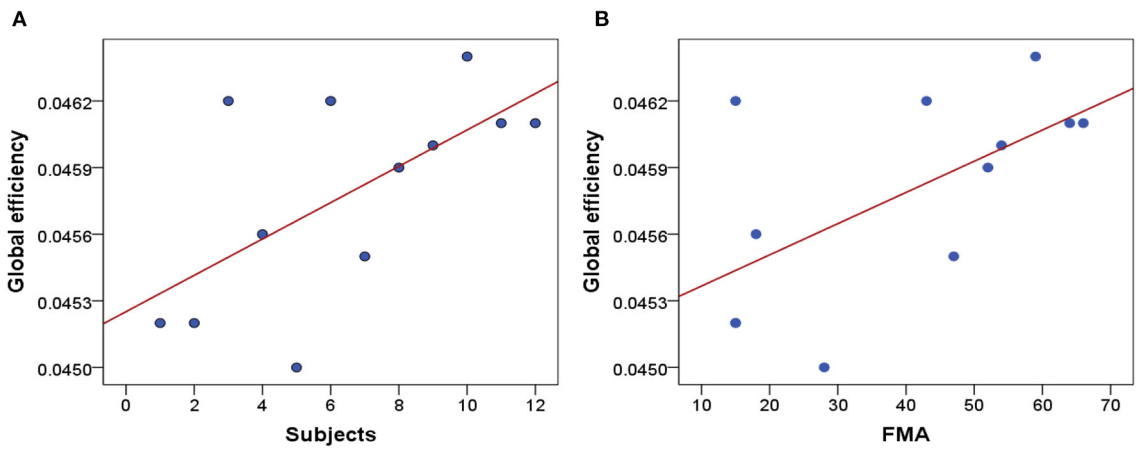


FIGURE 6 | Correlation analysis. **(A)** The average GE of 12 stroke patients. **(B)** Correlation between FMA scores and GE of 12 stroke patients.

parieto-occipital lobes. The connection of the frontal and parietal brain areas plays an important role in motor planning, decision-making, etc. (Bowling et al., 2020). The frontal lobe is related to the movement of the limbs. Right brain stroked in LS leads to increased connections between the left frontal lobe and other brain regions. The phenomena show that the left brain areas participate in motor planning and regulation as compensation when performing KMI recognition. As shown in **Figure 4B**, the significantly stronger connection edges of RS have transferred from the occipital lobe and the left frontal lobe to the right frontal lobe compared to HC. Stroke results in the dysfunction of the patients' motor network, more non-injured brain areas and non-motor areas of the damaged brain areas can participate in the completion of KMI recognition (Li et al., 2020). The connection of the brain network of the frontal lobe and the occipital lobe is abnormal, the functional compensation of the brain to the damaged motor network is indicated.

Based on the time-varying networks of the three groups of subjects, the time series of the dynamic GE of the different groups in the KMI stage are shown in **Figure 5**. The GE is the average efficiency of related brain networks and is usually used to estimate the potential of information transfer among brain regions. As illustrated in **Figure 5A**, the time-varying network efficiency of the HC, LS, and RS groups has increased along with the execution. When subjects are asked to perform KMI recognition, more advanced cognitive functions in the brain are gradually recruited, so network efficiency gradually increases (Zhang et al., 2016). Throughout the KMI recognition stage, the gradual increase in network efficiency can guarantee the completion of KMI recognition. Sabaté et al. (2004) have found that after left hemisphere stroke, a person's limb movement speed is significantly slowed down, and after right hemisphere stroke, the brain activity during KMI is stronger than that in the left hemisphere stroke. The information transfer rate between brain regions in the RS group is lower than the LS group when performing KMI recognition. LS has stronger brain compensatory and remodeling capabilities. After a stroke, plastic changes occur between different brain areas, the interaction between brain areas is enhanced to compensate for the damaged brain areas. The patients need to activate other brain areas as compensation to complete the KMI recognition. And indeed, when performing the right-hand KMI tasks, the average GE of the HC group is significantly larger than that of the RS group, as shown in **Figure 5B**.

To further investigate whether the GE is correlated with the FMA scores, we have performed one correlation analysis. As shown in **Figure 6B**, there is a positive correlation between GE and FMA. The higher the FMA scores, the higher the corresponding global efficiency. It proves that the GE can reflect the severity of clinical motor function damage. We can conclude that GE may be used as a potential biomarker to reflect the severity of motor function damage and objectively evaluate the efficacy of neuromodulation therapy. And it can also be used as a feedback indicator to guide the development of more effective KMI rehabilitation therapies in the future.

Our current study also has some limitations. The number of patients is scarce, and the subjects between males and females are unbalanced. To promote clinical treatment and effective intervention for stroke, more subjects will be recruited, meanwhile, the balanced male and female subjects will be considered for analysis.

CONCLUSIONS

In our study, we have constructed the time-varying KMI networks between post-stroke hemiplegic patients and healthy subjects based on the ADTF method. In post-stroke hemiplegic patients, the connection between the damaged brain areas and other motor areas is weaker when performing KMI recognition. The effective connection between the non-damaged brain areas and other motor areas is stronger. The connection between the frontal-parietal lobe and the occipital lobe is enhanced to provide compensation for motor dysfunction in stroke patients, and FMA scores are closely correlated with GE. These findings allow us to better understand the mechanism of movement disorders in patients with post-stroke hemiplegic. It also shows that the brain network may provide a more reliable quantitative analysis method for the clinical diagnosis and treatment of stroke.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of Qilu Hospital, Cheeloo College of Medicine, Shandong University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XY, CW, ML, CF, and JLi have designed the experiment. DW and YZ have conducted the experiments. FX, YW, HL, and LJ have conducted a brain function analysis. FX, YW, HL, ZY, and JLe have performed correlation and statistical analysis. ZY, YZ, and JLe have reviewed the manuscript. FX is a major contributor to the writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The project was supported in part by the Introduce Innovative Teams of 2021 New High School 20 Items Project under Grant No. 2021GXRC071, in part by the Program for Youth Innovative Research Team in the University of Shandong Province in China under Grant No. 2019KJN010, in part by the Natural Science Foundation of China under Grant Nos. 82172535, 61877062,

61977043, 62122059, and 61976152, in part by the Clinical Research Cross-Project of Shandong University under Grant No. 2020SDUCRCB004, in part by the Natural Science Foundation of Shandong Province of China under Grant Nos. ZR2019MA037, ZR2019PF002, and ZR202102200383, in part by the Research Leader Program of Jinan Science and Technology Bureau under Grant No. 2019GXRC061, in part by the Graduate Education and Teaching Reform Research Project of Qilu University of Technology in 2019 under Grant No. YJG19007.

ACKNOWLEDGMENTS

We would like to thank all our authors for their interests and time investment.

REFERENCES

Arnold, M., Milner, X. H. R., Witte, H., Bauer, R., and Braun, C. (1998). Adaptive AR modeling of nonstationary time series by means of Kalman filtering. *IEEE Trans. Biomed. Eng.* 45, 553–562. doi: 10.1109/10.668741

Arun, K. M., Smitha, K. A., Sylaja, P. N., and Kesavadas, C. (2020). Identifying resting-state functional connectivity changes in the motor cortex using fNIRS during recovery from stroke. *Brain Topogr.* 33, 710–719. doi: 10.1007/s10548-020-00785-2

Babiloni, C., Ferri, R., Binetti, G., Vecchio, F., Frisoni, G. B., Lanuzza, B., et al. (2009). Directionality of EEG synchronization in Alzheimer's disease subjects. *Neurobiol. Aging* 30, 93–102. doi: 10.1016/j.neurobiolaging.2007.05.007

Bowling, J. T., Friston, K. J., and Hopfinger, J. B. (2020). Top-down versus bottom-up attention differentially modulate frontal–parietal connectivity. *Hum. Brain Mapp.* 41, 928–942. doi: 10.1002/hbm.24850

Bundy, D. T., and Nudo, R. J. (2019). Preclinical studies of neuroplasticity following experimental brain injury: an update. *Stroke* 50, 2626–2633. doi: 10.1161/STROKEAHA.119.023550

Burianová, H., Marstaller, L., Sowman, P., Tesan, G., Rich, A. N., Williams, M., et al. (2013). Multimodal functional imaging of motor imagery using a novel paradigm. *Neuroimage* 71, 50–58. doi: 10.1016/j.neuroimage.2013.01.001

Chu, C., Zhang, Z., Wang, J., Liu, S., Wang, F., Sun, Y., et al. (2021). Deep learning reveals personalized spatial spectral abnormalities of high delta and low alpha bands in EEG of patients with early Parkinson's disease. *J. Neural Eng.* 18, 066036. doi: 10.1088/1741-2552/ac40a0

Ding, Q., Chen, S., Chen, J., Zhang, S., Peng, Y., Chen, Y., et al. (2022). Intermittent theta burst stimulation increases natural oscillatory frequency in ipsilesional motor cortex post-stroke: a transcranial magnetic stimulation and electroencephalography study. *Front. Aging Neurosci.* 14, 818340. doi: 10.3389/fnagi.2022.818340

Dong, L., Li, F., Liu, Q., Wen, X., Lai, Y., Xu, P., et al. (2017). MATLAB toolboxes for reference electrode standardization technique (REST) of scalp EEG. *Front. Neurosci.* 11, 601. doi: 10.3389/fnins.2017.00601

Egorova, N., Liem, F., Hachinski, V., and Brodtmann, A. (2019). Predicted brain age after stroke. *Front. Aging Neurosci.* 11, 348. doi: 10.3389/fnagi.2019.00348

Gao, Z., Zhang, K., Dang, W., Yang, Y., Wang, Z., Duan, H., et al. (2018). An adaptive optimal-kernel time-frequency representation-based complex network method for characterizing fatigued behavior using the SSVEP-based BCI system. *Knowl. Based Syst.* 152, 163–171. doi: 10.1016/j.knosys.2018.04.013

Jastreboff, P. J. (1990). Phantom auditory perception (tinnitus): mechanisms of generation and perception. *Neurosci. Res.* 8, 221–254. doi: 10.1016/0168-0102(90)90031-9

Li, F., Chen, B., Li, H., Zhang, T., Wang, F., Jiang, Y., et al. (2016). The time-varying networks in P300: a task-evoked EEG study. *IEEE Trans. Neural Syst. Rehabil. Eng.* 24, 725–733. doi: 10.1109/TNSRE.2016.2523678

Li, F., Jiang, L., Zhang, Y., Huang, D., Wei, X., Jiang, Y., et al. (2021). The time-varying networks of the wrist extension in post-stroke hemiplegic patients. *Cogn. Neurodyn.* doi: 10.1007/s11571-021-09738-2. [Epub ahead of print].

Li, F., Peng, W., Jiang, Y., Song, L., Liao, Y., Yi, C., et al. (2019). The dynamic brain networks of motor imagery: time-varying causality analysis of scalp EEG. *Int. J. Neural Syst.* 29, 1850016. doi: 10.1142/S0129065718500168

Li, H., Huang, G., Lin, Q., Zhao, J., Fu, Q., Li, L., et al. (2020). EEG changes in time and time-frequency domain during movement preparation and execution in stroke patients. *Front. Neurosci.* 14, 827. doi: 10.3389/fnins.2020.00827

Li, Y., Pan, J., Long, J., Yu, T., Wang, F., Yu, Z., et al. (2015). Multimodal BCIs: target detection, multidimensional control, and awareness evaluation in patients with disorder of consciousness. *Proc. IEEE* 104, 332–352. doi: 10.1109/JPROC.2015.2469106

Long, J., Li, Y., Yu, T., and Gu, Z. (2011). Target selection with hybrid feature for BCI-based 2-D cursor control. *IEEE Trans. Biomed. Eng.* 59, 132–140. doi: 10.1109/TBME.2011.2167718

Mane, R., Chouhan, T., and Guan, C. (2020). BCI for stroke rehabilitation: motor and beyond. *J. Neural Eng.* 17, 041001. doi: 10.1088/1741-2552/aba162

Manomaisaowapak, P., Nartkulpat, A., and Songsiri, J. (2015). Granger causality inference in EEG source connectivity analysis: a state-space approach. *IEEE Trans. Neural Netw. Learn. Syst.* (2021) 1. doi: 10.1101/2020.10.07.329276. [Epub ahead of print].

Maudoux, A., Lefebvre, P., Cabay, J. E., Demertzi, A., Vanhaudenhuyse, A., Laureys, S., et al. (2012). Auditory resting-state network connectivity in tinnitus: a functional MRI study. *PLoS ONE* 7, e36222. doi: 10.1371/journal.pone.0036222

Pagnotta, M. F., and Plomp, G. (2018). Time-varying MVAR algorithms for directed connectivity analysis: critical comparison in simulations and benchmark EEG data. *PLoS ONE* 13, e0198846. doi: 10.1371/journal.pone.0198846

Park, H. J., Friston, K. J., Pae, C., Park, B., and Razi, A. (2018). Dynamic effective connectivity in resting state fMRI. *NeuroImage* 180, 594–608. doi: 10.1016/j.neuroimage.2017.11.033

Reid, A. T., Headley, D. B., Mill, R. D., Sanchez-Romero, R., Uddin, L. Q., Marinazzo, D., et al. (2019). Advancing functional connectivity research from association to causation. *Nat. Neurosci.* 22, 1751–1760. doi: 10.1038/s41593-019-0510-4

Riahi, N., Vakorin, V. A., and Menon, C. (2020). Estimating Fugl-Meyer upper extremity motor score from functional-connectivity measures. *IEEE Trans. Neural Syst. Rehabil. Eng.* 28, 860–868. doi: 10.1109/TNSRE.2020.2978381

Rimbert, S., Gayraud, N., Bougrain, L., Clerc, M., and Fleck, S. (2019). Can a subjective questionnaire be used as brain-computer interface performance predictor? *Front. Hum. Neurosci.* 12, 529. doi: 10.3389/fnhum.2018.00529

Rowe, J., Friston, K., Frackowiak, R., and Passingham, R. (2002). Attention to action: specific modulation of corticocortical interactions in humans. *Neuroimage* 17, 988–998. doi: 10.1006/nimg.2002.1156

Sabaté, M., González, B., and Rodriguez, M. (2004). Brain lateralization of motor imagery: motor planning asymmetry as a cause of movement lateralization. *Neuropsychologia* 42, 1041–1049. doi: 10.1016/j.neuropsychologia.2003.12.015

Saes, M., Meskers, C. G., Daffertshofer, A., van Wegen, E. E., and Kwakkel, G. (2021). Are early measured resting-state EEG parameters predictive for upper limb motor impairment six months poststroke? *Clin. Neurophysiol.* 132, 56–62. doi: 10.1016/j.clinph.2020.09.031

Saes, M., Meskers, C. G. M., Daffertshofer, A., de Munck, J. C., Kwakkel, G., and van Wegen, E. E. H. (2019). How does upper extremity Fugl-Meyer motor score relate to resting-state EEG in chronic stroke? A power spectral density analysis. *Clin. Neurophysiol.* 130, 856–862. doi: 10.1016/j.clinph.2019.01.007

Sharma, N., and Baron, J. C. (2013). Does motor imagery share neural networks with executed movement: a multivariate fMRI analysis. *Front. Hum. Neurosci.* 7, 564. doi: 10.3389/fnhum.2013.00564

Si, Y., Wu, X., Li, F., Zhang, L., Duan, K., Li, P., et al. (2019). Different decision-making responses occupy different brain networks for information processing: a study based on EEG and TMS. *Cereb. Cortex.* 29, 4119–4129. doi: 10.1093/cercor/bhy294

Sun, J., He, J., and Gao, X. (2020). Neurofeedback training of the control network improves children's performance with an SSVEP-based BCI. *Neuroscience*. 478, 24–38. doi: 10.1016/j.neuroscience.2021.08.010

Vecchio, F., and Babiloni, C. (2011). Direction of information flow in Alzheimer's disease and MCI patients. *Int. J. Alzheimers Dis.* 2011:214580. doi: 10.4061/2011/214580

Wang, K., Xu, M., Wang, Y., Zhang, S., Chen, L., and Ming, D. (2020). Enhance decoding of pre-movement EEG patterns for brain–computer interfaces. *J. Neural Eng.* 17, 016033. doi: 10.1088/1741-2552/ab598f

Xu, F., Miao, Y., Sun, Y., Guo, D., Xu, J., and Wang, Y. (2021a). A transfer learning framework based on motor imagery rehabilitation for stroke. *Sci. Rep.* 11, 1–9. doi: 10.1038/s41598-021-99114-1

Xu, F., Rong, F., Leng, J., Sun, T., Zhang, Y., Siddharth, S., et al. (2021b). Classification of left-versus right-hand motor imagery in stroke patients using supplementary data generated by CycleGAN. *IEEE Trans. Neural Syst. Rehabil. Eng.* 29, 2417–2424. doi: 10.1109/TNSRE.2021.3123969

Xu, F., Rong, F., Miao, Y., Sun, Y., Dong, G., and Li, H. (2021c). Representation learning for motor imagery recognition with deep neural network. *Electronics* 10, 112. doi: 10.3390/electronics10020112

Xu, F., Zhou, W., Zhen, Y., Yuan, Q., and Wu, Q. (2016). Using fractal and local binary pattern features for classification of ECOG motor imagery tasks obtained from the right brain hemisphere. *Int. J. Neural Syst.* 26, 1650022. doi: 10.1142/S0129065716500222

Xu, L., Xu, M., Ke, Y., An, X., Liu, S., and Ming, D. (2020a). Cross-dataset variability problem in EEG decoding with deep learning. *Front. Hum. Neurosci.* 14, 103. doi: 10.3389/fnhum.2020.00103

Xu, M., Han, J., Wang, Y., Jung, T. P., and Ming, D. (2020b). Implementing over 100 command codes for a high-speed hybrid brain-computer interface using concurrent P300 and SSVEP features. *IEEE. Trans. Biomed. Eng.* 67, 3073–3082. doi: 10.1109/TBME.2020.2975614

Xu, M., He, F., Jung, T. P., Gu, X., and Ming, D. (2021d). Current challenges for the practical application of electroencephalography-based brain–computer interfaces. *Engineering* 7, 1710–1712 doi: 10.1016/j.eng.2021.09.011

Xu, M., Xiao, X., Wang, Y., Qi, H., Jung, T. P., and Ming, D. (2018). A brain–computer interface based on miniature-event-related potentials induced by very small lateral visual stimuli. *IEEE. Trans. Biomed. Eng.* 65, 1166–1175. doi: 10.1109/TBME.2018.2799661

Yao, D. (2001). A method to standardize a reference of scalp EEG recordings to a point at infinity. *Physiol. Meas.* 22, 693. doi: 10.1088/0967-3334/22/4/305

Zhang, R., Yao, D., Valdes-Sosa, P. A., Li, F., Li, P., and Zhang, T. (2015). Efficient resting-state EEG network facilitates motor imagery performance. *J. Neural Eng.* 12, 066024. doi: 10.1088/1741-2560/12/6/066024

Zhang, T., Li, M., Zhang, L., Biswal, B., Yao, D., and Xu, P. (2018). The time-varying network patterns in motor imagery revealed by adaptive directed transfer function analysis for fMR. *IEEE Access* 6, 60339–60352. doi: 10.1109/ACCESS.2018.2875492

Zhang, T., Liu, T., Li, F., Li, M., Liu, D., Zhang, R., et al. (2016). Structural and functional correlates of motor imagery BCI performance: insights from the patterns of fronto-parietal attention network. *Neuroimage* 134, 475–485. doi: 10.1016/j.neuroimage.2016.04.030

Zhang, X., Jiang, Y., Zhang, S., Li, F., Pei, C., He, G., et al. (2020). Correlation analysis of EEG brain network with modulated acoustic stimulation for chronic tinnitus patients. *IEEE Trans. Neural Syst. Rehabil. Eng.* 29, 156–162. doi: 10.1109/TNSRE.2020.3039555

Zhou, L., Zhu, Q., Wu, B., Qin, B., Hu, H., and Qian, Z. (2022). A comparison of directed functional connectivity among fist-related brain activities during movement imagery, movement execution, and movement observation. *Brain Res.* 1777, 147769. doi: 10.1016/j.brainres.2021.147769

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 07 June 2022
doi: 10.3389/fnagi.2022.872530

Investigation of Underlying Association Between Whole Brain Regions and Alzheimer's Disease: A Research Based on an Artificial Intelligence Model

*Shui Liu, Chen Jie, Weimin Zheng, Jingjing Cui and Zhiqun Wang**

Department of Radiology, Aerospace Center Hospital, Beijing, China

Alzheimer's disease (AD) is the most common form of dementia, causing progressive cognitive decline. Radiomic features obtained from structural magnetic resonance imaging (sMRI) have shown a great potential in predicting this disease. However, radiomic features based on the whole brain segmented regions have not been explored yet. In our study, we collected sMRI data that include 80 patients with AD and 80 healthy controls (HCs). For each patient, the T1 weighted image (T1WI) images were segmented into 106 subregions, and radiomic features were extracted from each subregion. Then, we analyzed the radiomic features of specific brain subregions that were most related to AD. Based on the selective radiomic features from specific brain subregions, we built an integrated model using the best machine learning algorithms, and the diagnostic accuracy was evaluated. The subregions most relevant to AD included the hippocampus, the inferior parietal lobe, the precuneus, and the lateral occipital gyrus. These subregions exhibited several important radiomic features that include shape, gray level size zone matrix (GLSZM), and gray level dependence matrix (GLDM), among others. Based on the comparison among different algorithms, we constructed the best model using the Logistic regression (LR) algorithm, which reached an accuracy of 0.962. Conclusively, we constructed an excellent model based on radiomic features from several specific AD-related subregions, which could give a potential biomarker for predicting AD.

Keywords: Alzheimer's disease, magnetic resonance imaging, radiomics, machine learning, structural MRI (sMRI)

OPEN ACCESS

Edited by:
Zehong Jimmy Cao,
University of South Australia, Australia

Reviewed by:
Xuetao Mu,
Chinese People's Armed Police General Hospital, China
Jiliang Fang,
China Academy of Chinese Medical Sciences, China

***Correspondence:**
Zhiqun Wang
wangzhiqun@126.com

Specialty section:
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

Received: *09 February 2022*
Accepted: *29 April 2022*
Published: *07 June 2022*

Citation:
Liu S, Jie C, Zheng W, Cui J and Wang Z (2022) Investigation of Underlying Association Between Whole Brain Regions and Alzheimer's Disease: A Research Based on an Artificial Intelligence Model. Front. Aging Neurosci. 14:872530. doi: 10.3389/fnagi.2022.872530

INTRODUCTION

Alzheimer's disease (AD) is the most common form of dementia, characterized by episodic memory decline. Its incidence is rising as the population ages and there are approximately 50 million people suffering from AD worldwide at present, which imposes a heavy burden on the society (Nichols et al., 2019; Breijyeh and Karaman, 2020). Due to the lack of sensitive diagnoses and effective treatments, it is of great theoretical significance and of potential clinical value to establish reliable radiologic biomarkers for early detection of AD by using new technologies, which can improve the prognosis of the disease.

Neuroimaging studies in AD have revealed the relationship between AD and structural atrophy in the temporal lobe, the entorhinal cortex, the hippocampus, and the limbic system, which reflects

the different stages of the disease and predicts the progress of AD (Sørensen et al., 2016; Wolk et al., 2017; Li et al., 2018). Previous studies used the analysis based on manually labeled regions of interest (ROI) to explore the subtle structural atrophy leading to AD (Last et al., 2020). In recent years, machine learning provides an automated and objective classification framework that includes feature extraction, algorithm selection, predictive model building, model validation, and so on. In the AD classification field, machine learning has attracted increasing attention by using the multimodal quantify patterns of atrophy together with different algorithms in recent years (Tang et al., 2015). For the radiomic feature extraction, brain atrophy was most often quantified *via* volume, texture, and geometric shape measures from structural magnetic resonance imaging (sMRI), which has achieved promising results.

For instance, one study combined hippocampal volume information and machine learning, suggesting that the volumetric reduction in the hippocampus was an important indicator of early AD (Uysal and Ozturk, 2020). Texture analysis had been successfully employed to search for imaging biomarkers for AD (De Oliveira et al., 2011; Anandh et al., 2015; Chincarini et al., 2016). Several studies had confirmed hippocampal texture abnormalities in AD and early stages of AD, indicating that texture might predict early cognitive impairment (Sørensen et al., 2016). Recently, Feng et al. employed the radiomic features of hippocampal subregions using a support vector machine (SVM) model to distinguish AD from healthy control (HC) and showed good performance (Feng et al., 2018). Besides the hippocampus, one study analyzed the texture features of the corpus callosum and the thalamus, suggesting that these regions could be used for the early diagnosis of AD (De Oliveira et al., 2011). Another study used the random forest (RF) classifier to identify the subcortical regions and found that the radiomic features from the hippocampus and amygdala regions have the greatest discriminative ability, which could differentiate AD from HC with the best performance (Chaddad et al., 2018). For geometric shape measures, a recent study used the geometric shape of the corpus callosum and multilayer perceptron (MLP) classifier to differentiate AD from HC, in which the classifier showed a high accuracy value (Dadsena et al., 2019).

As listed above, the research of AD prediction based on structural radiologic features and machine learning has made promising progress. However, most studies have explored a single structure or local features. Few studies have focused on the analysis of the whole brain subregions, which might be restricted by the manual annotation method. However, it is extremely important to analyze the whole brain radiomic features using a machine learning method, because it can facilitate an objective and comprehensive evaluation of brain atrophy patterns, which may provide more effective and sensitive markers for the early diagnosis of AD.

Since different brain subregions can be affected by AD in a distinct manner, it is very essential to investigate the radiomic features of whole brain structures. In this study, by using a machine learning method, we first explored radiomic features of whole brain in different subregions between AD and HC and identified key subregions, which showed significant differences between the two groups. Second, we constructed classification models based on the radiomic features of selected subregions and different algorithms. Finally, by calculating the classification accuracy and evaluating the model performances, we identified the best model to predict AD. Based on the pathology of AD and previous studies, we hypothesized that this classification was driven by a distributed atrophy pattern of several subregions and mainly includes the hippocampus and other limbic systems, which might be affected early in the disease course. We expected that the model based on radiomic features of specific subregions can be applied as a valuable radiologic biomarker for the early diagnosis of AD.

MATERIALS AND METHODS

Patient Information

In total, 160 right-handed subjects had participated in the study, i.e., 80 patients with AD and 80 healthy controls (HCs). This study was carried out in accordance with the recommendations of the Medical Research Ethics Committee of Aerospace Center Hospital. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Medical Research Ethics Committee of Aerospace Center Hospital. The AD subjects were recruited randomly from patients who had consulted the memory clinic at Aerospace Center Hospital for memory complaints. The HCs were recruited from the local community by recruitment advertisements. All the participants were required to complete the regular form, which includes age, gender, education, clinical history, family genetic history, previous examination results, and other clinical information.

All participants underwent a complete physical examination, neurological examination, and neuropsychological assessment. The neuropsychological examinations included the Mini-Mental State Examination (MMSE), the Clinical Dementia Rating (CDR) score, and other examinations. The patients with AD fulfilled the new research criteria for possible or probable AD (Dubois et al., 2007, 2010).

The HC fulfilled the following criteria: (a) no abnormal findings in routine brain Magnetic Resonance Imaging (MRI); (b) no findings of stroke, depression, or epilepsy, and other neurological or psychiatric disorders; (c) no visual loss or hearing loss and other neurological deficiencies; (d) no complaints about cognitive and memory; and (e) CDR score of 0.

The excluded criteria were as follows: participants with contraindications for MRI were excluded. For example, the subjects who have a cardiac defibrillator, a pacemaker, vascular clips, or a mechanical heart valve cannot take part in the examination; in addition, subjects with neurological or psychiatric diseases or with a history of cerebrovascular attacks or other degenerative disorders were excluded.

Structural Magnetic Resonance Imaging Data Acquisition

Magnetic Resonance Imaging examinations were performed at the department of radiology using a 3.0T Siemens Skyra MR

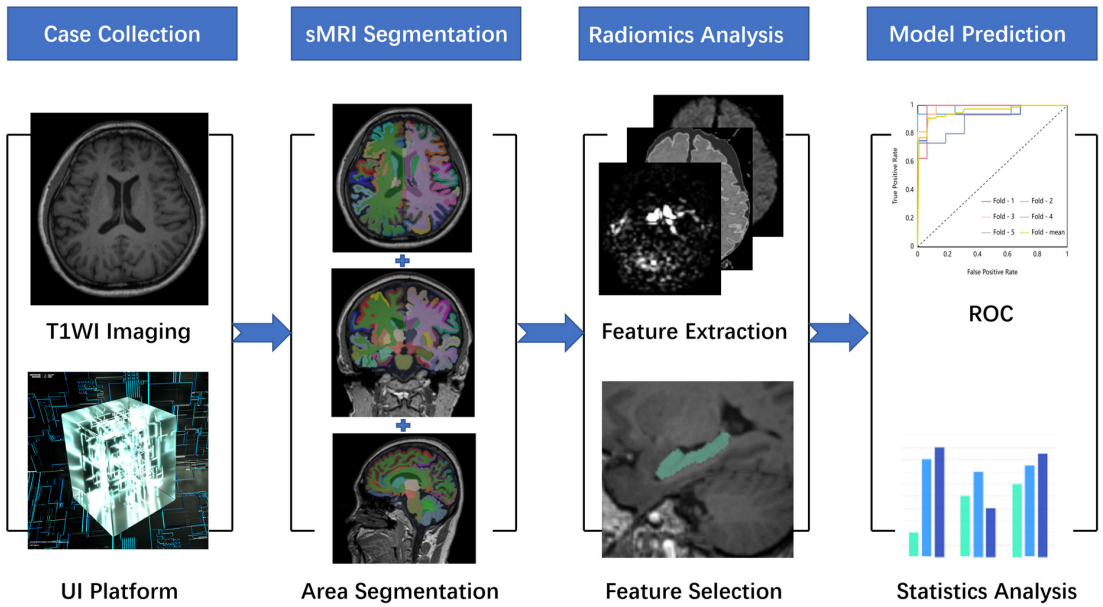


FIGURE 1 | A research map. This study is mainly divided into four parts. The first part is data collection and preprocessing, the second part is whole-brain structure segmentation, the third part is radiomic analysis, and the fourth part is model construction and evaluation.

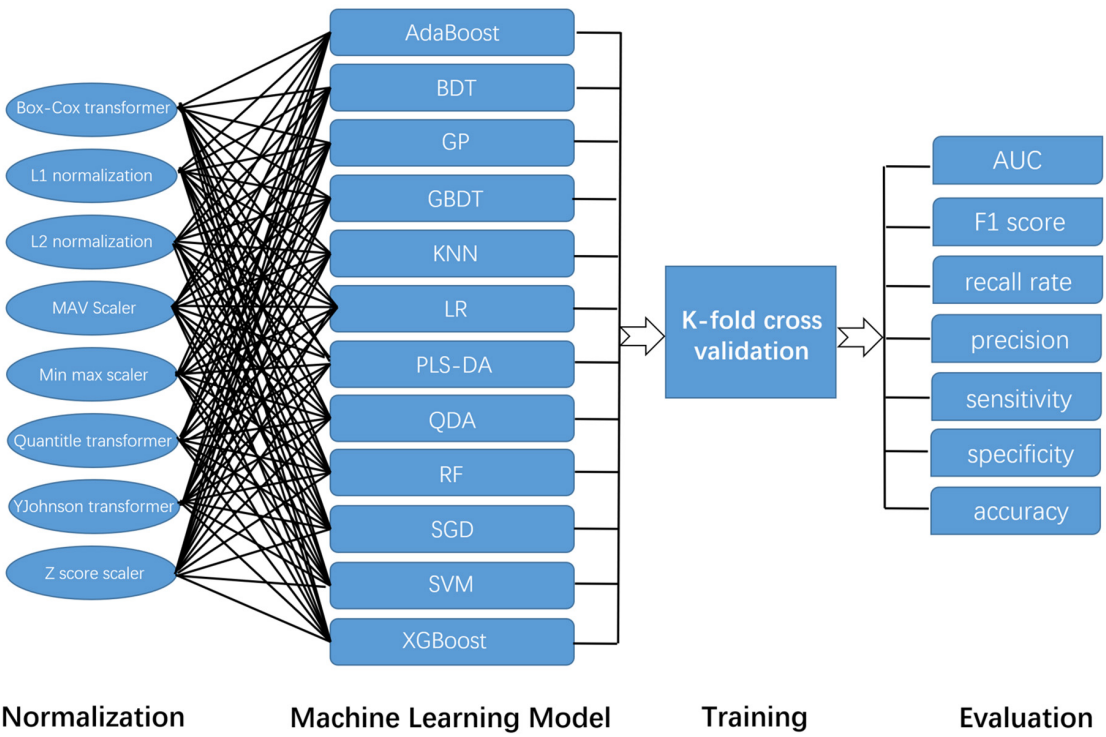


FIGURE 2 | The construction and evaluation of the machine learning model in this study. Because this study adopts a variety of normalization methods and machine learning models, their combination is shown in this figure.

System (Siemens, Germany) with a 20-channel head coil. Sagittal T1-weighted structural images were acquired for each subject using a magnetization-prepared rapid gradient echo (MPRAGE) sequence. Three-dimensional (3D) MPRAGE sagittal images were obtained with following parameters: Time of Repetition (TR)/Time of Echo (TE)/Time of Inversion (TI)/Flip Angle

(FA) = 1900 ms/2.2 ms/900 ms/9°, image matrix = 256 × 256, slice number = 176, and thickness = 1 mm. The obtained 3D images had a resolution of 1 mm × 1 mm × 1 mm.

Segmentation and Evaluation of Subregions

The whole brain subregions of each patient were extracted automatically by the given deep learning model first. The model was trained by the United Imaging platform.[1] The training process and the reference method of the model were similar to the previously published research (Desikan et al., 2006). The results of automatic segmentation included 22 temporal lobe structures, 20 frontal lobe structures, 12 parietal lobe structures, 8 occipital lobe structures, 8 cingulate gyrus structures, 2 insular structures, 12 subcortical gray matter structures, cerebral white matter structures, ventricles, the cerebellum, and other structures, with a total of 106 subregions. In particular, left and right structures were identified as different individuals. Once the automatic segmentation was done by the deep learning model, the result would be evaluated by two senior radiologists with more than 5 years of experience in radiologic diagnosis. In the segmentation results of 106 subregions of 160 patients in this study, two senior radiologists had no different opinions on the accuracy of the segmentation results.

Radiomics Feature Extraction and Model Construction

All preprocessing steps were performed using The United Imaging platform. Briefly, the radiomic features were extracted at first. Second, the features that were most consistent across different radiomics were selected to ensure robustness. Then, the dimension of extracted features was reduced using the Select K Best (K-Best) algorithm and traditional least absolute shrinkage and the selection operator (LASSO) algorithm, in which the two algorithms were used in series. Finally, the relevant parameters of these selected features would be used to build a machine learning model in order to successfully predict AD and HCs. These selected radiomic features of the most relevant subregions would be used in training sets and test sets in the form of 10-fold cross verification. The overall process is shown in **Figure 1**.

Model Validation

Before these data were used for model training, we used a variety of data normalization methods, such as Box-Cox transformer, L1 normalization, L2 normalization, max Absolute value scaler, min-max scaler, quantile transformer, YeoJohnson transformer, and Z score scaler. In this way, we tried to ensure the accuracy of the results. Then, we used a variety of common algorithms to predict AD and HCs, such as Adaptive Boosting (AdaBoost), Bagging Decision Tree (BDT), Gaussian Process (GP), Gradient Boosting Decision Tree (GBDT), K-Nearest Neighbor (KNN) algorithm, Logistic regression (LR), Partial Least Squares Discriminant Analysis (PLSDA), Quadratic Discriminant Analysis (QDA), RF, Stochastic Gradient Descent (SGD), SVM, and Extreme Gradient Boosting (XGBoost). The specific arrangement and combination ways are shown in **Figure 2**. The area under the curve (AUC) value, F1 score, recall rate, precision, sensitivity, specificity, and accuracy of each combination were evaluated separately.

Statistical Analysis

Statistical analyses were performed using SPSS software 22.0 (IBM, Armonk, NY, United States). For numerical data in AD and HC groups, a Wilcoxon test was used to evaluate the differences between AD and HC groups. For categorical data, such as gender, a Fisher's exact test was used to evaluate differences between AD and HC groups. Statistical significance was considered as $p < 0.05$.

RESULTS

Basic Characteristics of the Patients

In total, 80 patients with AD and 80 HCs with high-resolution sMRI data were collected retrospectively, adjusted for age, sex, MMSE, and CDR, among others. The detail is shown in **Table 1**. There were no significant differences in age and sex between AD and HCs. There were significant differences in MMSE between the two groups.

Automatic Segmentation Results of Whole Brain Subregions

As the result, concrete 106 subregions included temporal lobe structures (the hippocampus, the para hippocampal gyrus, the amygdala, the entorhinal gyrus, the fusiform, the temporal pole, the superior temporal gyrus, the middle temporal gyrus, the inferior temporal gyrus, and the transverse temporal gyrus), frontal lobe structures (the precentral cortex, the superior frontal gyrus, the frontal middle rostral, the frontal middle caudal, the frontal pole, the lateral orbitofrontal lobe, the medial orbitofrontal lobe, the pars opercularis, the pars orbitalis, and the pars triangularis), the parietal lobe structures (the postcentral cortex, the paracentral cortex, the superior parietal lobule, vinferior parietal lobule, the precuneus, and the supramarginal gyrus), the occipital lobe structures (the cuneus gyrus, the lingual gyrus, the pericalcarine gyrus, and the lateral occipital gyrus), the cingulate gyrus (the anterior cingulate gyrus, the middle cingulate gyrus, the posterior cingulate gyrus, and the cingulate gyrus of isthmus), the insular lobe structures, the subcortical gray matter structures (the caudate, the putamen, the pallidum, the thalamus, the nucleus accumbens, and the claustrum), the cerebral white matter, ventricles (lateral ventricle, 3rd ventricle, 4th ventricle, and cerebrospinal fluid), the cerebellum (the cerebellum cortex and the cerebellum white matter), and other structures (the choroid plexus, the inferior horn of lateral ventricle, the

[1] http://urp.united-imaging.com:8080/

TABLE 1 | Clinical characteristics of AD patients and HC.

	AD(*N* = 80)	HC(*N* = 80)	*p*-value
Age, median(min–max)	65(46–88)	64.5(48–83)	*9.54×10^{-1}
Sex, male/female	42/38	40/40	**9.87×10^{-1}
MMSE, median(min-max)	15(0–25)	28(12–30)	*7.26×10^{-5}
CDR	0.5–3	0	/

**Wilcoxon rank test; **Fisher exact test.*
AD, Alzheimer's disease; CDR, Clinical Dementia Rating; HC, healthy control; MMSE, Mini-Mental State Examination.

brainstem, the optic chiasm, and the corpus callosum). The specific segmentation results of 106 brain areas are shown in **Figure 3**.

Identification of Subregions Related to Alzheimer's Disease

Each subregion was extracted from 104 radiomic features and for a person, there would be, in total, 11,024 features with all 106 brain areas, which would be discussed in detail in the next section. We used two methods of dimensionality reduction in series, the Select K Best and LASSO. In the first step, we applied Select K Best to the 11,024 different features of subregions and screened the 3,660 specific features of subregions that may be related to AD (shown in **Figure 4**, only the top 10 features). Then, we applied LASSO to further screen the remained 3,660 specific radiomic features of subregions. After this, only 5 different radiomic features of 4 specific regions remained (shown in **Figure 4**). In the whole process of radiomic feature analysis, the rad scores of the training set and test set are shown in **Figure 5**.

As the above figures showed, there were several subregions related to AD. The most related subregions were the hippocampus, the inferior parietal lobule, the precuneus, and the lateral occipital gyrus.

Identification of Important Radiomic Features

As mentioned above, some specific subregions played an important role in the prediction of AD. At the same time, different radiomic features of specific subregions also played different roles in the prediction of AD. In the dimension reduction process of Select K Best, we selected the relevant features of top 100 for further analysis. The low-order radiomic features we used include the following categories: first order statistics, shape-based features, gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), GLSZM, neighboring gray tone difference matrix (NGTDM), and GLDM. The result is shown in **Figure 6**. Furthermore, after LASSO screening analysis, we noticed that only 5 remained were the most important features in predicting AD and NCs, such as first order statistics, GLSZM (i.e., two different subtypes), shape, and GLDM, which are mentioned in **Figure 5**.

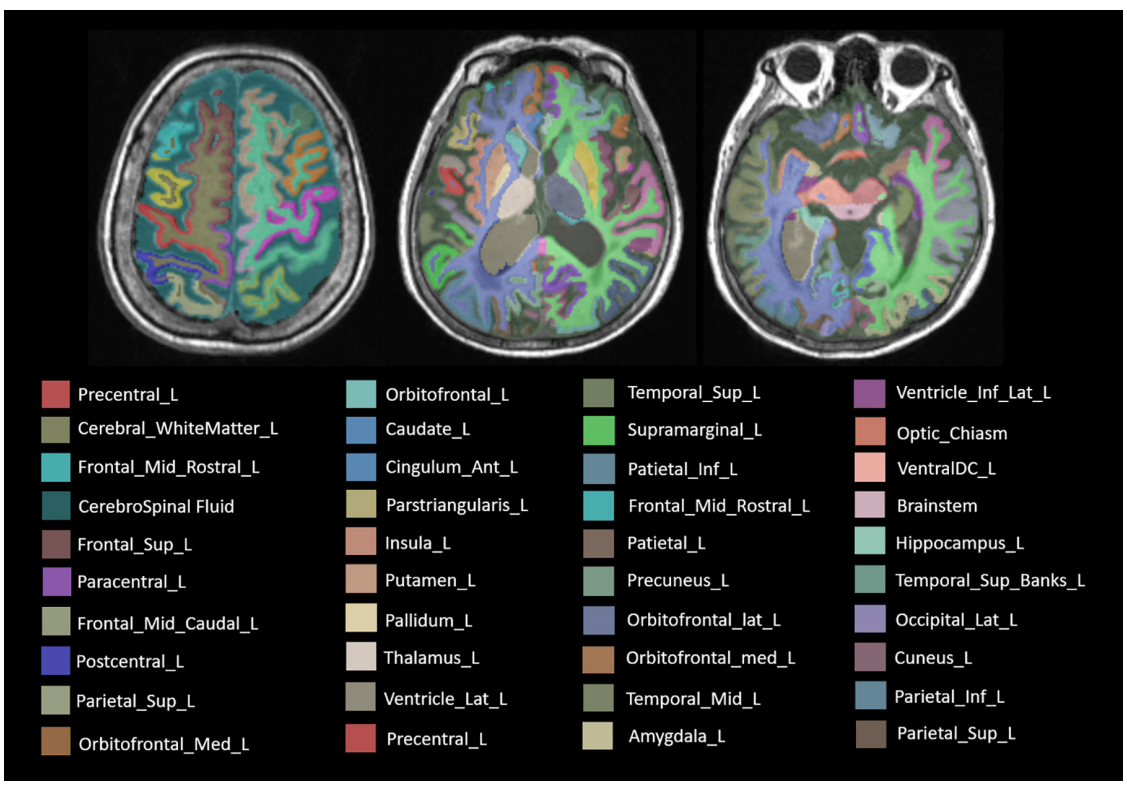


FIGURE 3 | The label of main brain regions in structural MRI (sMRI) after automatic segmentation. Specially, the symmetrical structure is divided into left and right and has different labels. Only the structure on the left is marked above.

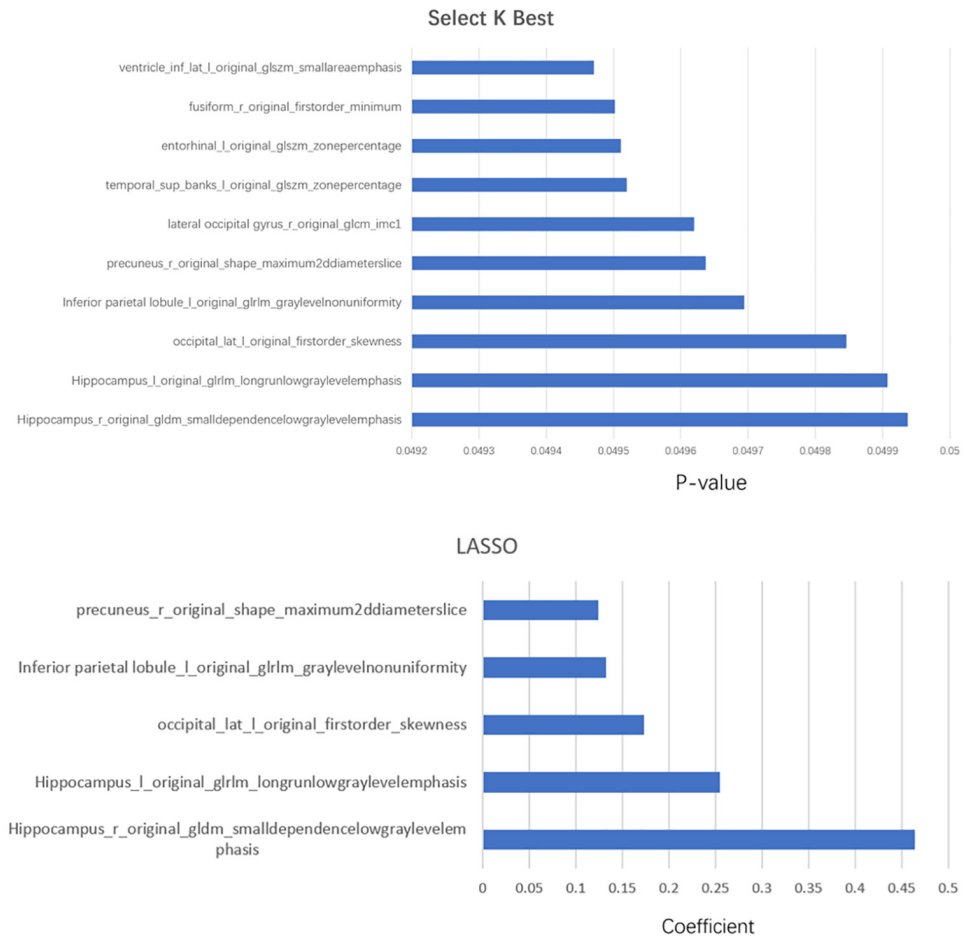


FIGURE 4 | The combination of statistically significant brain regions and radiomic features in the two-step dimensionality reduction process. In the first step, Select K Best, more than 10 radiomic features, are screened out; only the top 10 are listed in this figure.

The Evaluation and Comparison of Different Algorithms

We compared the performances of different algorithms for predicting AD. **Table 2** summarizes the AUC, F1 score, recall rate, precision, sensitivity, specificity, and accuracy of the train set and test set using different algorithms. Each normalization method was listed at the model's best performance. Especially, since we used the method of K-fold cross-validation, we got the above indicators for each training set and test set. Here, we use k-mean to reflect the average level of these models.

As listed above, the best of all is an LR model with the Box-Cox transformer, which has an accuracy of 0.962 in the test set, followed by KNN of 0.950 and SVM of 0.950. The receiver operator curves (ROCs) of each model are listed in **Figure 7**.

DISCUSSION

For classifying AD and HC subjects, clinical evaluation (i.e., CDR and MMSE score) or imaging volume features were commonly used, which were not very accurate (Hu et al., 2016; Deters et al., 2017). Nomograms based on gene expression signatures, cerebral spinal fluid (CSF), and pathological features are not yet ready to be used in daily practice. Radiomic features extracted from MR scans provide a noninvasive means to predict AD (Wen et al., 2020; Yun et al., 2020; Feng et al., 2021).

In this study, we segmented the whole brain subregions of the enrolled cases and extracted the radiomic features of each segmented subregion. Then, we comprehensively analyzed all the radiomic features of the whole brain subregions and identified the subregions and radiomic features most related to AD. Using the relevant areas and radiomic parameters obtained, we built a variety of machine learning models, such as LR, SVM, and RF. Then, we evaluated the diagnostic efficiency of each model. Finally, we found the best model for predicting AD. In this study, the subregions most relevant to AD included the hippocampus, the inferior parietal lobe, the precuneus, and the lateral occipital gyrus. The radiomic features extracted from these subregions had the greatest differences between subjects with AD and HCs.

Among the four most relevant subregions found in our study, the relationship between the hippocampus and AD had

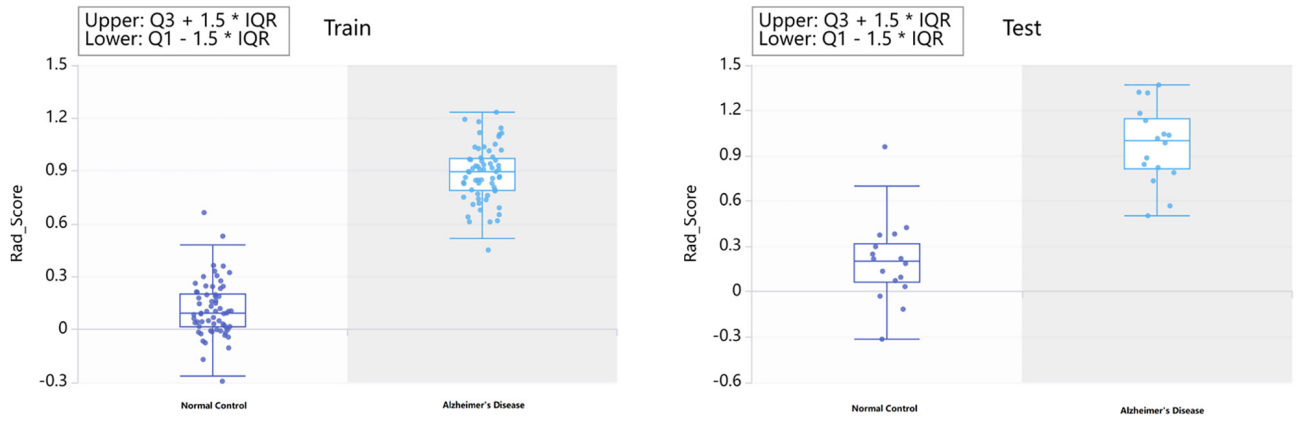


FIGURE 5 | The rad score in radiomics feature selection.

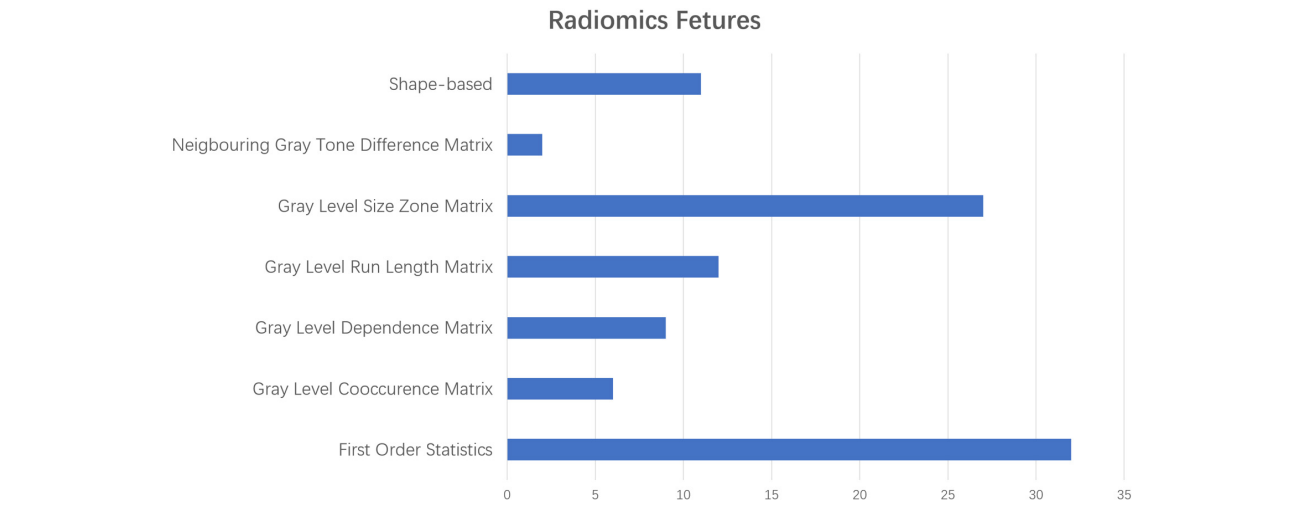


FIGURE 6 | The different proportions of radiomic characteristics in the process of predicting ad with radiomic characteristics in the low-order dimension.

been confirmed by many studies. The reduction of hippocampal volume had been well studied in individuals with AD (Stricker et al., 2012). Hippocampal atrophy was one of the core markers of AD in the revised national Alzheimer's association diagnostic criteria (Albert et al., 2011; Catani et al., 2013). In addition to volume reduction, abnormal metabolic levels, interruption of brain activity, and microstructure characteristics in the hippocampus had also been well reported (Catani et al., 2013). The discriminative ability of radiomic features from hippocampal regions, found by our analysis, was consistent with recent studies (Sørensen et al., 2016, 2017). For example, the hippocampal texture was shown to be a strong biomarker for differentiating HCs from patients with AD or mild cognitive impairment (MCI) (Feng et al., 2018). Pathologically, a previous study has confirmed that hippocampal shape alterations were associated with regional Aβ load in normal elderly individuals (Schroeder et al., 2017). We could speculate that the radiomic feature changes of the hippocampus might result from AD pathological changes, such as Aβ deposition, which might be taken as a potential biomarker to differentiate between patients with AD and HCs.

The inferior parietal lobule had begun to show promise as an important locus in AD in recent years (Greene et al., 2010). For example, several studies had shown AD-related alterations of the inferior parietal lobule, such as gray matter atrophy (Desikan et al., 2009; Jacobs et al., 2011), metabolic dysfunction (Walhovd et al., 2010), disrupted spontaneous brain functional activity and connectivity (Wang et al., 2015), and pathological changes (Nelson et al., 2009). These findings had important implications for the underlying neurobiology of AD. Compared to previous studies, our analysis investigated the link between AD and radiomic features of inferior parietal lobule regions, which added a piece of new evidence for the mechanism of AD. In addition, our study might contribute to the early detection of AD to some extent.

As for the precuneus, its atrophy played a special role in early-onset AD (Karas et al., 2007). Many neuroimaging studies

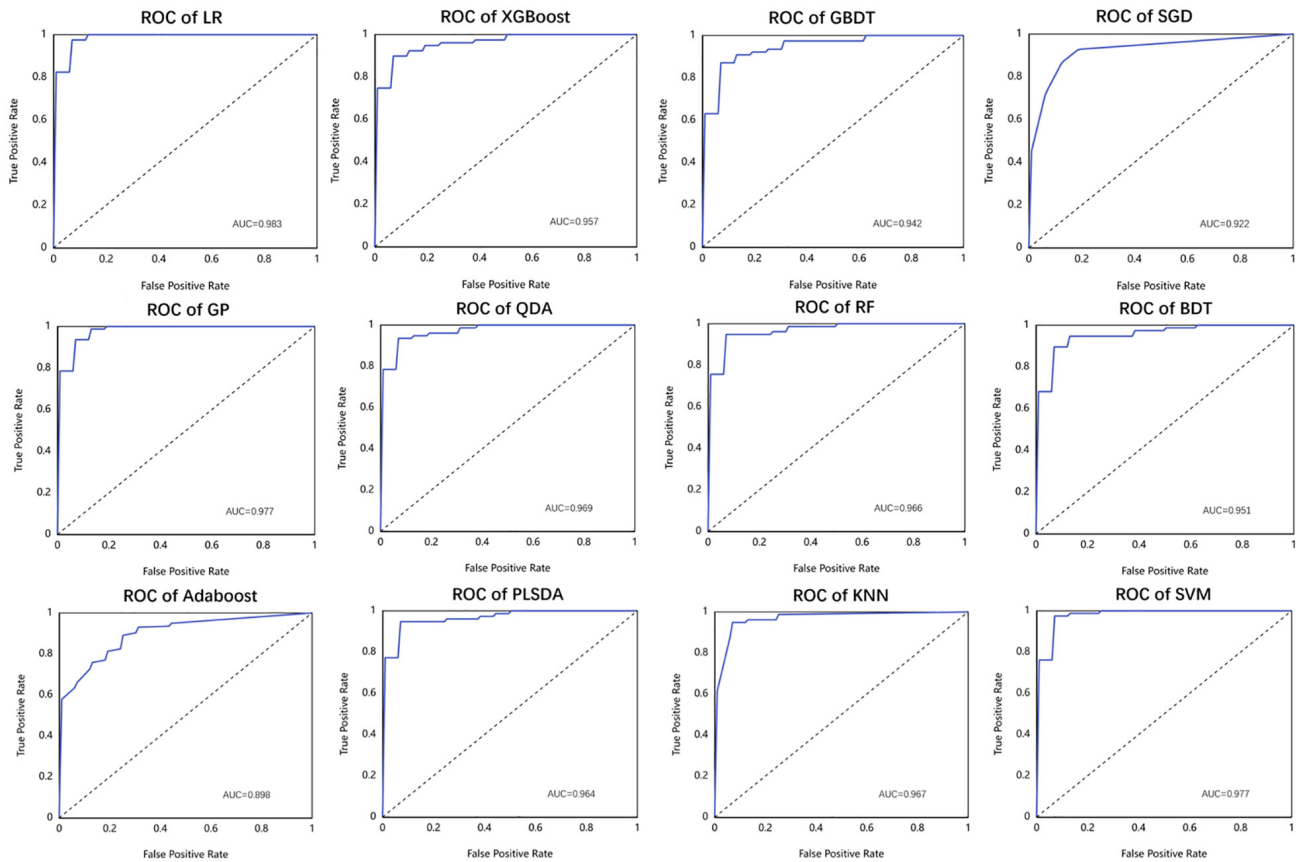


FIGURE 7 | The performance of different machine learning models used in this study in the form of receiver operator curve (ROC).

had demonstrated the structural and functional abnormalities of precuneus regions in AD, such as cortical thinning (Dickerson and Sperling, 2009), amyloid deposition (Buckner et al., 2009), decreased intrinsic brain activity (He et al., 2007; Wang et al., 2011), and disrupted functional connectivity of the region (Greicius et al., 2004). A previous fMRI study demonstrated reduced precuneus deactivation during object naming in patients with mild cognitive impairment, AD, and frontotemporal lobar degeneration (Frings et al., 2010). At the molecular level, precuneus amyloid burden was also associated with reduced cholinergic activity in AD (Ikonomovic et al., 2011), which might contribute to the cognitive decline. Our result on the link between AD and radiomic features in precuneus regions was consistent with the previous study, indicating the crucial role of precuneus in the early diagnosis of AD.

Finally, our study found that there is a specific relationship between the lateral occipital gyrus and AD. The machine learning model, which contained the lateral occipital gyrus, had higher diagnostic efficiency than the one without the lateral occipital gyrus constructed by radiomic features. The occipital gyrus is located in the primary visual cortex and plays a critical role in visual cognition. By using the fMRI method (Sala-Llonch et al., 2015), it was reported that the occipital gyrus presented higher activity during the task of visuo-perceptual working memory. Using diffusion tensor imaging (DTI) and tractography, a previous study demonstrated that the structural disconnection in the ventral occipital temporal cortex contributed to the deficit in facial recognition (Thomas et al., 2009). Visual cognition deficits were consistently reported to accompany the development of AD (Cronin-Golomb, 1995; Bokde et al., 2006). In our study, the potential relationship between the lateral occipital gyrus and AD put forward a possible new direction for the study of AD. The model constructed by integrating the most relevant structural areas provides a new idea for the prediction of AD.

However, there are still several issues that need further consideration in our study. First, in the current study, we mainly focused on structural analysis. Further studies that simultaneously combine the sMRI and other data, such as fMRI and clinical laboratory examination, might obtain a powerful and high-quality biomarker for clinical application. We plan to analyze the relationship between the radiomic features and the cognitive performances in the future to achieve early diagnosis and monitor the progress of the disease. Second, recent studies had paid more attention to individuals at high risk for AD, such as amnestic mild cognitive impairments, and ApoE-4 allele carriers. Exploring these populations would provide valuable biomarkers for the early diagnosis of AD. Finally, a longitudinal study with

TABLE 2 | Performance of different machine learning algorithms.

Model	Normalization	auc_ train	auc_ test	f1score_ train	f1score_ test	recall_ train	recall_test	precision_ train	precision_ test	sensitivity_ train	sensitivity_ test	specificity_ train	specificity_ test	accuracy_ train	accuracy_ test
AdaBoost	Quantitle transformer	0.995	0.898	0.997	0.851	0.997	0.846	0.997	0.864	0.997	0.846	0.997	0.862	0.997	0.855
BDT	BoxCox transformer	0.995	0.951	0.995	0.909	0.990	0.910	1.000	0.913	0.990	0.910	1.000	0.912	0.995	0.911
GP	Min-max scaler	0.995	0.977	0.991	0.936	0.997	0.937	0.985	0.938	0.997	0.937	0.984	0.938	0.991	0.937
GBDT	BoxCox transformer	0.995	0.942	1.000	0.876	1.000	0.872	1.000	0.883	1.000	0.872	1.000	0.888	1.000	0.880
KNN	Z-score scaler	0.995	0.967	1.000	0.949	1.000	0.949	1.000	0.951	1.000	0.949	1.000	0.950	1.000	0.950
LR	BoxCox transformer	0.995	0.983	0.998	0.962	1.000	0.962	0.997	0.962	1.000	0.962	0.997	0.962	0.998	0.962
PLS-DA	Quantitle transformer	0.988	0.964	0.744	0.743	1.000	1.000	0.593	0.593	1.000	1.000	0.322	0.312	0.659	0.653
QDA	BoxCox transformer	0.974	0.969	0.934	0.935	0.924	0.923	0.945	0.949	0.924	0.923	0.947	0.950	0.936	0.937
RF	BoxCox transformer	0.989	0.966	0.958	0.936	0.949	0.936	0.968	0.938	0.949	0.936	0.969	0.938	0.959	0.937
SGD	Max-abs scaler	0.992	0.922	0.994	0.923	1.000	0.923	0.988	0.924	1.000	0.923	0.988	0.925	0.994	0.924
SVM	Z-score scaler	0.995	0.977	1.000	0.950	1.000	0.950	1.000	0.951	1.000	0.950	1.000	0.950	1.000	0.950
XGBoost	L1 normalization	0.995	0.957	1.000	0.903	1.000	0.898	1.000	0.911	1.000	0.898	1.000	0.912	1.000	0.906

AdaBoost, Adaptive Boosting; BDT, Bagging Decision Tree; GP, Gaussian Process; GBDT, Gradient Boosting Decision Tree; KNN, K-Nearest Neighbor algorithm; LR, Logistic regression; PLSDA, Partial Least Squares Discriminant Analysis; QDA, Quadratic Discriminant Analysis; RF, random forest; SGD, Stochastic Gradient Descent; SVM, support vector machine; XGBoost, Extreme Gradient Boosting.

a large multicenter sample size is needed to confirm the stability and reliability of the model.

CONCLUSION

In conclusion, our study specifically focused on the potential relationship between AD and the whole brain subregions based on sMRI. The machine learning model constructed with the radiomic features of the hippocampus, the inferior parietal lobe, the precuneus, and the lateral occipital gyrus could be used as a potential sMRI marker for predicting AD and had outstanding performance.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Review Committee of Aerospace Center Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

SL: data acquisition, analysis, and interpretation for the research. CJ and WZ: image quality evaluation and evaluation of automatic segmentation results of whole brain areas. JC: extraction of radiomic features, construction of machine learning model, and parameter adjustment. ZW: drafting the work, providing final approval of the version to be published, and providing agreement to be accountable for all aspects of the work. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Scientific Foundation of China (Grant No. 81873892), Beijing Natural Scientific Foundation of China (No. 7222320), the Capital Health Development Scientific Research special Project (No. 2022-2-6081), and the Scientific Research Fund of Aerospace Center Hospital (No. YN201901).

ACKNOWLEDGMENTS

The authors thank the study investigators and patients who participated in the original trial. Especially, we appreciate the assistance and suggestion from JingJing Cui of United Imaging Intelligence (Beijing).

REFERENCES

Albert, M. S., Dekosky, S. T., and Dickson, D. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging and Alzheimer's Association workgroup. *Alzheimers Dement. J. Alzheimers Assoc.* 7, 1–10.

Anandh, K., Sujatha, C., and Ramakrishnan, S. (2015). Segmentation and analysis of corpus callosum in Alzheimer MR images using total variation based diffusion filter and level set method. *Biomed. Sci. Instr.* 51, 355–361.

Bokde, A., Lopez-Bayo, P., Meindl, T., Pechler, S., Born, C., Faltraco, F., et al. (2006). Functional connectivity of the fusiform gyrus during a face-matching task in subjects with mild cognitive impairment. *Brain* 129, 1113–1124. doi: 10.1093/brain/awl051

Breijyeh, Z., and Karaman, R. (2020). Comprehensive review on Alzheimer's disease: causes and treatment. *Molecules* 25:5789. doi: 10.3390/molecules25245789

Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., et al. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J. Neurosci.* 29, 1860–1873. doi: 10.1523/JNEUROSCI.5062-08.2009

Catani, M., Acqua, F. D., and Schotten, M. (2013). A revised limbic system model for memory, emotion and behavior. *Neurosci. Biobehav. Rev.* 37, 1724–1737.

Chaddad, A., Desrosiers, C., and Niazi, T. (2018). Deep radiomic analysis of MRI related to Alzheimer's disease. *IEEE Access* 6, 58213–58221.

Chincarini, A., Sensi, F., Rei, L., Gemme, G., Squarcia, S., Longo, R., et al. (2016). Integrating longitudinal information in hippocampal volume measurements for the early detection of Alzheimer's disease. *NeuroImage* 125, 834–847. doi: 10.1016/j.neuroimage.2015.10.065

Cronin-Golomb, A. (1995). Vision in Alzheimer's disease. *Gerontologist* 35, 370–376.

Dadsena, R., Rohini, P., and Ramakrishnan, S. (2019). Proposal of a machine learning approach to differentiate mild and Alzheimer's condition in MR images using shape changes in corpus callosum. *Stud Health Technol. Inform.* 258, 243–244.

De Oliveira, M., Balthazar, M., D'abreu, A., Yasuda, C., Damasceno, B., Cendes, F., et al. (2011). MR imaging texture analysis of the corpus callosum and thalamus in amnestic mild cognitive impairment and mild Alzheimer disease. *Am. J. Neuroradiol.* 32, 60–66. doi: 10.3174/ajnr.A2232

Desikan, R. S., Cabral, H. J., Fischl, B., Guttmann, C. R., Blacker, D., Hyman, B. T., et al. (2009). Temporoparietal MR imaging measures of atrophy in subjects with mild cognitive impairment that predict subsequent diagnosis of Alzheimer disease. *Am. J. Neuroradiol.* 30, 532–538. doi: 10.3174/ajnr.A1397

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021

Deters, K. D., Risacher, S. L., Kim, S., Nho, K., West, J. D., Blennow, K., et al. (2017). Plasma tau association with brain atrophy in mild cognitive impairment and Alzheimer's disease. *J. Alzheimers Dis.* 58, 1245–1254. doi: 10.3233/JAD-161114

Dickerson, B. C., and Sperling, R. A. (2009). Large-scale functional brain network abnormalities in Alzheimer's disease: insights from functional neuroimaging. *Behav. Neurol. 21* 1, 63–75. doi: 10.3233/BEN-2009-0227

Dubois, B., Feldman, H. H., Jacova, C., Cummings, J. L., DeKosky, S. T., Barberger-Gateau, P., et al. (2010). Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol.* 9, 1118–1127. doi: 10.1016/S1474-4422(10)70223-4

Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., et al. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS–ADRDA criteria. *Lancet Neurol.* 6, 734–746. doi: 10.1016/S1474-4422(07)70178-3

Feng, F., Wang, P., Zhao, K., Zhou, B., Yao, H., Meng, Q., et al. (2018). Radiomic features of hippocampal subregions in Alzheimer's disease and amnestic mild cognitive impairment. *Front. Aging Neurosci.* 10:290. doi: 10.3389/fnagi.2018.00290

Feng, J., Zhang, S., and Chen, L. (2021). *Extracting ROI-Based Contourlet Subband Energy Feature from the sMRI Image for Alzheimer's Disease Classification*. Available online at: https://ieeexplore.ieee.org/document/9320583 (accessed February 9, 2022).

Frings, L., Dressel, K., Abel, S., Saur, D., Kümmerer, D., Mader, I., et al. (2010). Reduced precuneus deactivation during object naming in patients with mild cognitive impairment, Alzheimer's disease, and frontotemporal lobar degeneration. *Dement. Geriatr. Cogn. Disord.* 30, 334–343. doi: 10.1159/000320991

Greene, S. J., Killiany, R. J., and Alzheimer's Disease Neuroimaging Initiative (2010). Subregions of the inferior parietal lobule are affected in the progression to Alzheimer's disease. *Neurobiol. Aging* 31, 1304–1311. doi: 10.1016/j.neurobiolaging.2010.04.026

Greicius, M. D., Srivastava, G., Reiss, A. L., and Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4637–4642. doi: 10.1073/pnas.0308627101

He, Y., Wang, L., Zang, Y., Tian, L., Zhang, X., Li, K., et al. (2007). Regional coherence changes in the early stages of Alzheimer's disease: a combined structural and resting-state functional MRI study. *Neuroimage* 35, 488–500. doi: 10.1016/j.neuroimage.2006.11.042

Hu, Y.-B., Li, C.-B., Song, N., Zou, Y., Chen, S.-D., Ren, R.-J., et al. (2016). Diagnostic value of microRNA for Alzheimer's disease: a systematic review and meta-analysis. *Front. Aging Neurosci.s* 8:13.

Ikonomovic, M., Klunk, W., Abrahamson, E., Wuu, J., Mathis, C., Scheff, S., et al. (2011). Precuneus amyloid burden is associated with reduced cholinergic activity in Alzheimer disease. *Neurology* 77, 39–47. doi: 10.1212/WNL.0b013e3182231419

Jacobs, H. I., Van Boxtel, M. P., Uylings, H. B., Gronenschild, E. H., Verhey, F. R., and Jolles, J. (2011). Atrophy of the parietal lobe in preclinical dementia. *Brain Cogn.* 75, 154–163. doi: 10.1016/j.bandc.2010.11.003

Karas, G., Scheltens, P., Rombouts, S., Van Schijndel, R., Klein, M., Jones, B., et al. (2007). Precuneus atrophy in early-onset Alzheimer's disease: a morphometric structural MRI study. *Neuroradiology* 49, 967–976. doi: 10.1007/s00234-007-0269-2

Last, B. S., Rizvi, B., and Brickman, A. M. (2020). "Structural MRI in Alzheimer's disease," in *Vascular Disease, Alzheimer's Disease, and Mild Cognitive Impairment: Advancing an Integrated Approach* (Oxford: Oxford University Press), 208–204.

Li, Y., Jiang, J., Shen, T., Wu, P., and Zuo, C. (2018). "Radiomics features as predictors to distinguish fast and slow progression of Mild Cognitive Impairment to Alzheimer's disease," in *Proceeding of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (Piscataway, NJ: IEEE), 127–130. doi: 10.1109/EMBC.2018.8512273

Nelson, P. T., Abner, E. L., Scheff, S. W., Schmitt, F. A., Kryscio, R. J., Jicha, G. A., et al. (2009). Alzheimer's-type neuropathology in the precuneus is not increased relative to other areas of neocortex across a range of cognitive impairment. *Neurosci. Lett.* 450, 336–339. doi: 10.1016/j.neulet.2008.11.006

Nichols, E., Szoeke, C. E., Vollset, S. E., Abbasi, N., Abd-Allah, F., Abdela, J., et al. (2019). Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* 18, 88–106. doi: 10.1016/S1474-4422(18)30403-4

Sala-Llonch, R., Palacios, E. M., Junqué, C., Bargalló, N., and Vendrell, P. (2015). Functional networks and structural connectivity of visuospatial and visuoperceptual working memory. *Front. Hum. Neurosci.* 9:340. doi: 10.3389/fnhum.2015.00340

Schroeder, C., Park, M. T. M., Germann, J., Chakravarty, M. M., Michels, L., Kollias, S., et al. (2017). Hippocampal shape alterations are associated with regional Aβ load in cognitively normal elderly individuals. *Eur. J. Neurosci.* 45, 1241–1251. doi: 10.1111/ejn.13408

Sørensen, L., Igel, C., Liv Hansen, N., Osler, M., Lauritzen, M., Rostrup, E., et al. (2016). Early detection of Alzheimer's disease using M RI hippocampal texture. *Hum. Brain Mapp.* 37, 1148–1161. doi: 10.1002/hbm.23091

Sørensen, L., Igel, C., Pai, A., Balas, I., Anker, C., Lillholm, M., et al. (2017). Differential diagnosis of mild cognitive impairment and Alzheimer's disease using structural MRI cortical thickness, hippocampal shape, hippocampal texture, and volumetry. *NeuroImage Clin.* 13, 470–482. doi: 10.1016/j.nicl.2016.11.025

Stricker, N. H., Dodge, H. H., Dowling, N. M., Han, S. D., Erosheva, E. A., Jagust, W. J., et al. (2012). CSF biomarker associations with change in hippocampal

volume and precuneus thickness: implications for the Alzheimer's pathological cascade. *Brain Imaging Behav.* 6, 599–609. doi: 10.1007/s11682-012-9171-6

Tang, X., Holland, D., Dale, A. M., Younes, L., Miller, M. I, and Alzheimer's Disease Neuroimaging Initiative (2015). Baseline shape diffeomorphometry patterns of subcortical and ventricular structures in predicting conversion of mild cognitive impairment to Alzheimer's disease. *J. Alzheimers Dis.* 44, 599–611.

Thomas, C., Avidan, G., Humphreys, K., Jung, K.-J., Gao, F., and Behrmann, M. (2009). Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nat. Neurosci.* 12, 29–31. doi: 10.1038/nn.2224

Uysal, G., and Ozturk, M. (2020). Hippocampal atrophy based Alzheimer's disease diagnosis via machine learning methods. *J. Neurosci. Methods* 337:108669. doi: 10.1016/j.jneumeth.2020.108669

Walhovd, K., Fjell, A., Brewer, J., McEvoy, L., Fennema-Notestine, C., Hagler, D., et al. (2010). Combining MR imaging, positron-emission tomography, and CSF biomarkers in the diagnosis and prognosis of Alzheimer disease. *Am. J. Neuroradiol.* 31, 347–354. doi: 10.3174/ajnr.A1809

Wang, Z., Xia, M., Dai, Z., Liang, X., Song, H., He, Y., et al. (2015). Differentially disrupted functional connectivity of the subregions of the inferior parietal lobule in Alzheimer's disease. *Brain Struct. Funct.* 220, 745–762. doi: 10.1007/s00429-013-0681-9

Wang, Z., Yan, C., Zhao, C., Qi, Z., Zhou, W., Lu, J., et al. (2011). Spatial patterns of intrinsic brain activity in mild cognitive impairment and Alzheimer's disease: a resting-state functional MRI study. *Hum. Brain Mapp.* 32, 1720–1740. doi: 10.1002/hbm.21140

Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-González, J., Routier, A., Bottani, S., et al. (2020). Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation. *Med. Image Anal.* 63:101694. doi: 10.1016/j.media.2020.101694

Wolk, D. A., Das, S. R., Mueller, S. G., Weiner, M. W., Yushkevich, P. A., and Alzheimer's Disease Neuroimaging Initiative (2017). Medial temporal lobe subregional morphometry using high resolution MRI in Alzheimer's disease. *Neurobiol. Aging* 49, 204–213. doi: 10.1016/j.neurobiolaging.2016.09.011

Yun, J. Y., Kim, K. T., and Choi, J. Y. (2020). Using 3D deep convolutional neural network with MRI biomarker patch Images for Alzheimer's disease diagnosis. *J. Korea Multimedia Soc.* 23, 940–952.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 10 June 2022
doi: 10.3389/fnhum.2022.875851

A Regional Smoothing Block Sparse Bayesian Learning Method With Temporal Correlation for Channel Selection in P300 Speller

Xueqing Zhao[1], Jing Jin[1,2], Ren Xu[3], Shurui Li[1], Hao Sun[1], Xingyu Wang[1] and Andrzej Cichocki[4,5,6]*

[1] The Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai, China, [2] Shenzhen Research Institute of East China University of Technology, Shenzhen, China, [3] g.tec medical engineering GmbH, Graz, Austria, [4] Skolkovo Institute of Science and Technology, Moscow, Russia, [5] Systems Research Institute of Polish Academy of Science, Warsaw, Poland, [6] Department of Informatics, Nicolaus Copernicus University, Toruń, Poland

The P300-based brain–computer interfaces (BCIs) enable participants to communicate by decoding the electroencephalography (EEG) signal. Different regions of the brain correspond to various mental activities. Therefore, removing weak task-relevant and noisy channels through channel selection is necessary when decoding a specific type of activity from EEG. It can improve the recognition accuracy and reduce the training time of the subsequent models. This study proposes a novel block sparse Bayesian-based channel selection method for the P300 speller. In this method, we introduce block sparse Bayesian learning (BSBL) into the channel selection of P300 BCI for the first time and propose a regional smoothing BSBL (RSBSBL) by combining the spatial distribution properties of EEG. The RSBSBL can determine the number of channels adaptively. To ensure practicality, we design an automatic selection iteration strategy model to reduce the time cost caused by the inverse operation of the large-size matrix. We verified the proposed method on two public P300 datasets and on our collected datasets. The experimental results show that the proposed method can remove the inferior channels and work with the classifier to obtain high-classification accuracy. Hence, RSBSBL has tremendous potential for channel selection in P300 tasks.

Keywords: channel selection, sparse bayesian learning, temporal correlation, brain-computer interface, EEG, P300

OPEN ACCESS

Edited by:
Björn H. Schott,
Leibniz Institute for Neurobiology (LG), Germany

Reviewed by:
Christoph Reichert,
Leibniz Institute for Neurobiology (LG), Germany
Jianjun Meng,
Shanghai Jiao Tong University, China

***Correspondence:**
Jing Jin
jinjingat@gmail.com

Specialty section:
This article was submitted to Brain Health and Clinical Neuroscience, a section of the journal Frontiers in Human Neuroscience

Received: *14 February 2022*
Accepted: *18 May 2022*
Published: *10 June 2022*

Citation:
Zhao X, Jin J, Xu R, Li S, Sun H, Wang X and Cichocki A (2022) A Regional Smoothing Block Sparse Bayesian Learning Method With Temporal Correlation for Channel Selection in P300 Speller. Front. Hum. Neurosci. 16:875851. doi: 10.3389/fnhum.2022.875851

INTRODUCTION

Brain–computer interface (BCI) is a direct interactive pathway designed to establish a non-muscle connection between the human brain and the computer (Wolpaw et al., 2002; Jin et al., 2015). It provides a new way to communicate with the outside, for example, daily communication (Sorbello et al., 2017; He et al., 2019) and wheelchair control (Kim et al., 2016; Deng et al., 2019). In addition, BCIs can also be used to aid in the diagnosis of disorders of consciousness (Maestú et al., 2019; Ando et al., 2021). BCIs can be divided into invasive and non-invasive ones. Electroencephalography (EEG) is a non-invasive technique that records brain signals through electrodes placed on the scalp. Generally, users' brain signals are recorded, amplified, and pre-processed with an EEG recorder, and then the signals are converted to commands *via* classifiers (Bashashati et al., 2007). Currently, BCIs

based on the Event-Related Potential (ERP) (Hoffmann et al., 2008a; Lopez-Calderon and Luck, 2014), Steady-State Visual Evoked Potential (SSVEP) (Nakanishi et al., 2017), and Motor Imagery (MI) (Padfield et al., 2019) are the three main research directions. The oddball paradigm is a typical paradigm of P300, where standard and deviant stimuli are included. These two kinds of stimuli appear randomly with large and small probabilities, and deviant stimuli are the targets in small probability events that correspond to the spelling character (Donchin et al., 2000). The spelling paradigms and algorithms based on P300 have been widely developed in recent years (Cecotti and Graser, 2010; Townsend et al., 2010; Hammer et al., 2018; Arvaneh et al., 2019; Jin et al., 2019; Huang et al., 2022). This study is focused on the P300 BCI system.

To provide a complete coverage of regions related to EEG activity, a large number of electrodes are used for EEG acquisition. An electrode is regarded as a channel. However, a realistic EEG system typically uses the data of a small number of channels during computation to minimize the preparation time and cost (Cecotti et al., 2011). Channel selection helps to exclude the weak task-relevant and noisy channels, thus improving the classification accuracy and reducing the classifier training time. Inter-participant differences and equipment differences can make the best subset of channels in the same paradigm different. The flexibility of selecting a subset of empirical channels in the complex BCI data is insufficient, and the data-based channel selection method is more conducive to giving the optimal channel selection. Therefore, the method of automatically determining a subset of channels has better application prospects than selecting a fixed subset.

Different evaluation approaches, such as filter, wrapper, embedded, hybrid, and human-based techniques have been widely used to select features and the subset of channels in the P300 speller (Alotaiby et al., 2015). Filters like Fisher Score (Lal et al., 2004) are usually independent of the classifier and select channels based on the relevance. A CCA spatial filter also proved to be effective in event-related signal processing (Reichert et al., 2017). On the other hand, wrappers select the channel set according to the algorithm effect and search for channels through continuous heuristic methods. Support Vector Machine based recursive channel elimination (SVM-RCE) can be considered a typical example of a wrapper (Rakotomamonjy and Guigue, 2008). The hybrid approach is a combination of filter and wrapper and uses the wrapper to obtain a subset of the available channels after handling the filter (Liu and Yu, 2005). The human-based approaches are the methods in which the experienced experts select channels by analyzing certain technical indicators (Tekgul et al., 2005). In addition, some channel selection algorithms are based on evolutionary algorithms, such as Particle Swarm Optimization (PSO), which also belong to wrappers (Martinez-Cagigal and Hornero, 2017; Arican and Polat, 2019). For embedded methods, the selection is usually implicit and integrated with the learner training process. By giving sparse weight to features or channels, sparse methods can obtain a classifier that needs fewer selected features or channels. The Least Absolute Shrinkage and Selectionator operator (LASSO), a linear regressor with L_1 regularization, can be regarded as an embedded method (Tibshirani, 1996). In EEG research, LASSO has also become a commonly used feature selection algorithm and extended to channel selection (Tomioka and Müller, 2010). Yuan extended the LASSO method to groups in 2006, giving birth to the group LASSO (GLASSO), which allows us to group all variables and then penalize the L_2 parametrization of each group in the objective function, thus achieving the effect of eliminating a whole group of coefficients to zero at the same time (Yuan and Lin, 2006). The Bayesian framework-based feature selection and classification methods are widely used in EEG. Studies have shown the outstanding performance of Bayesian linear discriminant analysis (BLDA) in EEG decoding (Hoffmann et al., 2008a; Lei et al., 2009; Manyakov et al., 2011). Tipping et al. proposed a sparse Bayesian learning (SBL) method under the Bayesian framework to solve the regression problem (Tipping, 2001). SBL can complete the feature selection of P300 through sparsity (Hoffmann et al., 2008b) and has been used for channel selection (Wu et al., 2014; Zhang et al., 2017; Dey et al., 2020). EEG is a common response of regional neurons (Hassan and Wendling, 2018). However, the channel optimization approach described above does not consider the spatial structure between the channels of EEG signals. In addition, a few existing algorithms consider the temporal correlation in a single channel, which means the amplitude correlation between time points within each channel.

This paper proposes a regional smoothing SBL (RSBSBL) method for channel selection of the P300 signal. Block sparse Bayesian learning (BSBL) was first proposed for sparse signal recovery (Zhang and Rao, 2011). It is the first time that the BSBL is applied to EEG channel selection. The P300 features are usually filtered and down-sampled in the temporal series, and features from the same channel are correlated. In this method, we combine BSBL with the spatial distribution properties of EEG to propose an RSBSBL. To ensure practicality, we design an automatic selection iteration strategy model to reduce the time cost caused by the inverse operation of large-size matrices.

For verification, RSBSBL was compared with some other methods with similar principles on the three BCI datasets. We used BLDA as a unified classifier for a fair comparison. The effectiveness of the proposed method was verified by the effectiveness of channel subsets and the character recognition performance.

We organize the rest of the paper as follows. Section "Materials and Methods" describes the principle and calculation process of the proposed algorithm. Section "Materials and Experiments" describes the dataset used and the data processing framework. Section "Results" shows the experimental results. Section "Discussion" further discusses the effectiveness of the selected channel subsets, character recognition performance, effectiveness of regional smoothing, time cost, and future work. Finally, section "Conclusion" gives the conclusion.

METHODS

Here, we show the principle and solution process of RSBSBL and give its flow of selecting channels. The input features of

one channel are regarded as a block. Based on the BSBL, we considered the spatial distribution of EEG and divided different regions according to the location of the electrodes. The automatic selection mode of the iterative strategy is used to ensure practicality.

Regional Smoothing Sparse Bayesian Learning

The EEG signals collected by the device are generally two-dimensional data after pre-processing. N_c is denoted as the number of channels and N_t as temporal points. Input data $\mathbf{X}$ contains N samples $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \ldots\ldots \mathbf{x}_N \in R^D$, where $D = N_t N_c$ represents the number of features in each sample. Then, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots \mathbf{x}_N]^{\mathrm{T}} \in R^{N \times D}$ and $\mathbf{y} = \left[y_1, y_2, y_3, \ldots, y_N\right]^{\mathrm{T}} \in R^N$ represent the corresponding labels, where $y_i \in \{1, -1\}$ is the class label. Its mathematical model can be expressed linearly as follows:

$$\mathbf{y} = \mathbf{X}\mathrm{w} + \varepsilon \tag{1}$$

where $\mathbf{w} = [w_1, w_2, w_3 \ldots\ldots w_D]^{\mathrm{T}}$ is a learnable weight vector, ε is noise, and $\mathbf{X}$ can be replaced by $\Phi(\mathbf{X})$ expressed in the form of a kernel function. Assume $\varepsilon \sim \mathcal{N}\left(0, \sigma^2 \mathbf{I}_N\right)$, then $\mathbf{y} \sim \mathcal{N}\left(\mathbf{X}\mathrm{w}, \sigma^2 \mathbf{I}_N\right)$ and its probabilistic framework is.

$$p\left(\mathbf{y}|\mathbf{w}, \sigma^2\right) = (2\pi\sigma^2)^{-\frac{N}{2}} exp\left(-\frac{1}{2\sigma^2}||\mathbf{y}-\mathbf{X}\mathrm{w}||_2^2\right) \tag{2}$$

The RSBSBL adds the symmetric positive definite matrix in the variance term of the distribution that $\mathbf{w}$ obeys. The input data of one channel are regarded as a block. So, for the mathematical model (1), assume that $\mathbf{w}_b\left(\forall b\right)$ is mutually independent and Gaussian distributed.

$$p\left(\mathbf{w}_b|\gamma_b, \mathbf{B}_b, \forall b\right) \sim \mathcal{N}\left(0, \gamma_b \mathbf{B}_b\right), b \in 1, \ldots, N_b \tag{3}$$

where $\mathbf{w}_b$ containing several w_i is bth block of $\mathbf{w}$, γ_b is a non-negative scalar that controls the variance of $\mathbf{w}_b$, $\mathbf{B}_b$ is a positive definite matrix reflecting the intra-block correlation, and N_b is the number of blocks. Since the features of a channel are considered to be a block, $N_b = N_c$.

In our case of EEG signal, b is the index of channels. In a channel of EEG signal with corresponding weight $\mathbf{w}_b$, it is assumed that all its feature weights share the same γ_b to control the variance of their distribution, and $\mathbf{B}_b$ controls the intra-block correlation.

In this case, we express the prior of $\mathbf{w}$ as $p(\mathbf{w}|\gamma, \mathbf{B}) \sim \mathcal{N}(0, \Sigma_0)$, where Σ_0 is

$$\Sigma_0 = \begin{bmatrix} \gamma_1 \boldsymbol{B}_1 & & \\ & \ddots & \\ & & \gamma_{N_b} \boldsymbol{B}_{N_b} \end{bmatrix} \tag{4}$$

the posterior probability has been calculated by the Bayesian rule,

$$p\left(\mathbf{w}|\mathbf{y}, \sigma^2, \gamma, \mathbf{B}\right) = \frac{p\left(\mathbf{y}|\mathbf{w}, \sigma^2\right) p(\mathbf{w}|\gamma, \mathbf{B})}{p\left(\mathbf{y}|\sigma^2, \gamma, \mathbf{B}\right)} \tag{5}$$

and the corresponding variance and mean of the posterior probability density $p\left(\mathbf{w}|\mathbf{y}, \sigma^2, \gamma, \mathbf{B}\right) \sim \mathcal{N}(\mu_{\mathbf{w}}, \Sigma_{\mathbf{w}})$ can be described as

$$\Sigma_{\mathbf{w}} = \left(\sigma^{-2}\mathbf{X}^{\mathrm{T}}\mathbf{X} + \Sigma_0^{-1}\right)^{-1} \tag{6}$$

$$\mu_{\mathbf{w}} = \sigma^{-2}\Sigma_{\mathbf{w}}\mathbf{X}^{\mathrm{T}}\mathbf{y} \tag{7}$$

When $N \geq D$, the Eqs (6) and (7) are suitable because the maximum size of the inverse matrix is D in this case. Now, we give the iterative ways when $N < D$. According to the matrix inversion formula and the matrix identity.

$$(\mathbf{E} + \mathbf{FGH})^{-1} = \mathbf{E}^{-1} - \mathbf{E}^{-1}\mathbf{FG}\left(\mathbf{I} + \mathbf{HE}^{-1}\mathbf{FG}\right)^{-1}\mathbf{HE}^{-1} \tag{8}$$

$$(\mathbf{I} + \mathbf{EF})^{-1}\mathbf{E} = \mathbf{E}(\mathbf{I} + \mathbf{FE})^{-1} \tag{9}$$

we replace the Eqs (6) and (7) with the following equations:

$$\Sigma_{\mathbf{w}} = \Sigma_0 - \Sigma_0\mathbf{X}^{\mathrm{T}}\left(\sigma^2\mathbf{I} + \mathbf{X}\Sigma_0\mathbf{X}^{\mathrm{T}}\right)^{-1}\mathbf{X}\Sigma_0 \tag{10}$$

$$\mu_{\mathbf{w}} = \Sigma_0\mathbf{X}^{\mathrm{T}}\left(\sigma^2\mathbf{I} + \mathbf{X}\Sigma_0\mathbf{X}^{\mathrm{T}}\right)^{-1}\mathbf{y} \tag{11}$$

To find the iterative equation of the parameters $\Theta = \left\{\gamma, \mathbf{B}, \sigma^2\right\}$, the expectation–maximization (EM) is used to maximize $\log p\left(\mathbf{y}|\Theta\right)$. The Q function is.

$$\begin{aligned} Q(\Theta) &= E_{\mathbf{w}|\mathbf{y},\Theta_{\mathrm{old}}}\left[\log p\left(\mathbf{y}, \mathbf{w}|\Theta\right)\right] \\ &= E_{\mathbf{w}|\mathbf{y},\Theta_{\mathrm{old}}}\left[\log p\left(\mathbf{y}|\mathbf{w}, \sigma^2\right)\right] \\ &\quad + \mathrm{E}_{\mathbf{w}|\mathbf{y},\Theta_{\mathrm{old}}}\left[\log p\left(\mathbf{w}|\gamma, \mathbf{B}\right)\right] \end{aligned} \tag{12}$$

The first term of the Q function is related to σ^2 and the second term is related to γ and $\mathbf{B}$. Then, we can get the parameters iteratively by maximizing the $\mathbf{Q}$ function.

$$\sigma^2 = \frac{||\mathbf{y}-\mathbf{X}\mu_{\mathbf{w}}||_2^2 + \sigma_{\mathrm{old}}^2\left[D - \mathrm{Tr}\left(\Sigma_{\mathbf{w}}\Sigma_0^{-1}\right)\right]}{N} \tag{13}$$

$$\gamma_b = \frac{\mathrm{Tr}\left[\mathbf{B}_b^{-1}\left(\Sigma_{\mathbf{w}}^b + \mu_{\mathbf{w}}^b\left(\mu_{\mathbf{w}}^b\right)^{\mathrm{T}}\right)\right]}{d_b}, \forall b \tag{14}$$

$$\mathbf{B}_{re} = \frac{1}{g_{re}}\sum_{b \in G_{re}} \frac{\Sigma_{\mathbf{w}}^b + \mu_{\mathbf{w}}^b\left(\mu_{\mathbf{w}}^b\right)^{\mathrm{T}}}{\gamma_b}, \forall re \tag{15}$$

where $^*_{\mathrm{old}}$ represents the hyperparameter in the previous iteration, and the superscript b of $\mu_{\mathbf{w}}^b$ and $\Sigma_{\mathbf{w}}^b$ indicates the bth block in $\mu_{\mathbf{w}}$ and $\Sigma_{\mathbf{w}}$ with the size of $d_b \times 1$ and $d_b \times d_b$ (d_b is the number of elements in $\mathrm{w_b}$).

The potential similarity exists in the adjacent electrode signals considering the volume conduction effects in the brain (Hassan and Wendling, 2018). We assign the same $\mathbf{B}_{re}$ for channels with close locations for regional smoothing, and the region G_{re} contains g_{re} channels. As shown in **Figure 1**, all the channels are divided into 13 regions by position, and each region contains at least three channels. $\mathbf{B}_{re}$ is the average of blocks in region re ($re \in [1, 13]$).

We use a first-order Auto-Regressive (AR) process to model the intra-block correlation. Many applications have used the AR process to express it (Zhang and Rao, 2011; Zhang et al., 2013; Yin et al., 2020). Thus, to find a symmetric positive definite matrix to

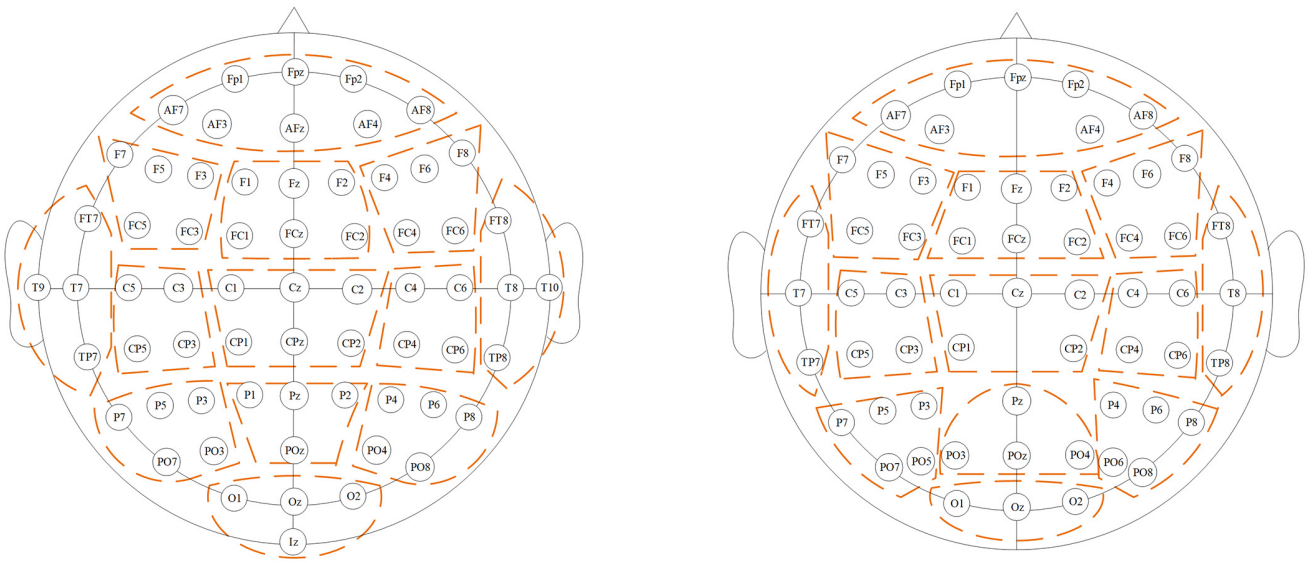


FIGURE 1 | Region division. Channels belonging to a region are circled with the dotted line. The left subfigure shows the division for DS1 and DS2, while the right subfigure shows the division for DS3.

approximate **B**, it can be constrained to the following form of the Toeplitz matrix.

$$\mathbf{B}_{re} \triangleq \text{Toeplitz}\left(\left[1,\ r,\ldots,r^{d_b-1}\right]\right) = \begin{bmatrix} 1 & \cdots & r^{d_b-1} \\ \vdots & \ddots & \vdots \\ r^{d_b-1} & \cdots & 1 \end{bmatrix} \quad (16)$$

Empirically calculate $r = \frac{m_1}{m_0}$, where m_0 is the average of the main diagonal of $\mathbf{B}_{re}$ and m_1 is the average of the main sub-diagonal.

Channel Selection Based on Regional Smoothing BSBL

Regarding the feature extracted from the same channel as a block, we perform RSBSBL to get the weight vector of features and design a channel selection based on the weight vector as **Algorithm 1**.

As shown in **Algorithm 1**, the parameters are initialized, and the shear threshold τ is set. Then, from Line 3 to Line 12, the algorithm iteratively solves BSBL and prunes the γ. Line 4 to Line 8 decide the calculation of $\Sigma_{\mathbf{w}}$, $\mu_{\mathbf{w}}$, so that the large time cost caused by finding the inverse matrix of a large-size matrix can be alleviated. The parameters are updated on Line 9 and Line 10. **Figure 2** illustrates the relationship between the parameters in a single iteration, where the parameters calculated simultaneously have the same color. The solid line indicates the passing relationship between the parameters of this iteration, and the dashed line indicates the passing relationship between the parameters of this iteration and the next iteration. After the parameters are calculated, in order to achieve the sparse block effect, make γ_b to 0 when γ_b is less than the threshold τ. Then, it comes into the next iteration until the convergence criterion is satisfied. Line 13 automatically selects the channels with γ_b greater than the shear threshold τ. Finally, the algorithm returns the selected channel and the corresponding weight vector.

The off-diagonal matrix **B** makes the weights **w** in the same block relevant in distribution. It means that the correlation of the features from the same channel can be reflected during the process. Moreover, the components of the temporal correlation of different channels in close locations are the same because the $\mathbf{B}_{re}$ of channels in the same region are shared. The sparsity of weights will form the units of channels. The

Algorithm 1: Regional Smoothing Sparse Bayesian Learning (RSBSBL).

Input: features $\mathbf{X}_{N \times (N_c N_t)}$ and labels $\mathbf{Y}_{N \times 1}$, where N denotes the number of samples, N_c represents the number of channels, and N_t is the number of features (sampling points) in one channel.

Output: sparse weights **w** and selected channels C_s.

1: Choose an initial setting for σ^2, γ , **B**. The block size is N_t.
2: Set a shear threshold τ to obtain the sparsity weights.
3: **While** the convergence criterion is not satisfied, **do**
4: **If** $N \geq N_c \times N_t$, **then**
5: Calculate $\Sigma_{\mathbf{w}}$, $\mu_{\mathbf{w}}$, according to (6)(7).
6: **Else**
7: Calculate $\Sigma_{\mathbf{w}}$, $\mu_{\mathbf{w}}$, according to (10)(11).
8: **End if**
9: Update σ^2, γ , **B** according to (13)(14)(15)and (16).
10: **If** $\gamma_b < \tau$, **then** $\gamma_b = 0$, $\gamma_b \in \gamma$.
11: $\sigma^2_{\text{old}} = \sigma^2$, $\mathbf{B}_{\text{old}} = \mathbf{B}$, $\gamma_{\text{old}} = \gamma$.
12: **End while**
13: $C_s = \{b | \gamma_b > \tau, b \in 1, 2, \ldots, N_c, \gamma_b \in \gamma\}$.
14: **Return w** $= \{\mu^b_{\mathbf{w}} | b \in C_s\}$ and C_s.

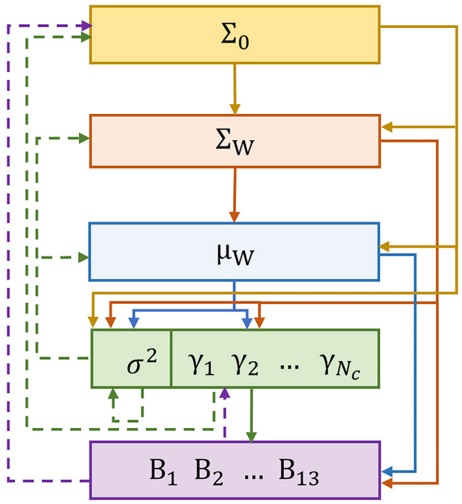


FIGURE 2 | Parameter relationship graphical model in a single iteration. Parameters of the same color can be iterated simultaneously.

TABLE 1 | The stimulus numbers for each participant of DS1, DS2, and DS3.

Dataset	Stimulus category	Training dataset size	Test dataset size
DS1	Target	1260	930
P1.1	Non-target	6300	4650
DS2	Target	2550	3000
P2.1/P2.2	Non-target	12750	15000
DS3	Target	144	144
P3.1-P3.12	Non-target	720	720

Pi.j represents the jth participant in the ith dataset.

features from one channel share the same weight distribution whose variance is controlled by γ. For practicality, up to five channels are removed in a single iteration when making a channel selection.

MATERIALS AND EXPERIMENTS

Data Descriptions

Three datasets were used in this study to validate the proposed method. DS1 is BCI Competition II dataset IIb (one participant) (Blankertz et al., 2004) and DS2 is BCI Competition III dataset II (two participants) (Blankertz et al., 2006). DS3 is the EEG signal collected in our lab (12 participants). The stimulus numbers for each participant of the above three datasets are shown in **Table 1**.

DS1 and DS2 provided by the BCI Competition are public datasets and follow the same experimental paradigm of Farwell and Donchin, as shown in **Figure 3**. In a six-by-six character matrix containing 26 characters and 10 numbers, participants were asked to focus on a specified character in each trial (a trial is a set of stimuli that can support the output of a recognized character). They could do this by mentally counting the target stimuli' number of flashes (intensifications). The paradigm continuously intensified and randomly scanned all rows and columns of the matrix at a rate of 5.7 Hz. Each row and column in the matrix was randomly intensified for 100 ms and was left blank for 75 ms. DS1 contained 42 training characters and 31 testing characters. The training set of DS2 contained 85 characters, and the testing set contained 100 characters. A trial for each character had 15 epochs to apply reliable spelling, and each epoch was comprised of 12 intensifications. Both datasets were collected using a 64-channel cap, filtered by 0.1–60 Hz, and digitized at a sampling rate of 240 Hz. DS1 and DS2 can be downloaded from the websites: http://www.bbci.de/competition/ii/ and http://www.bbci.de/competition/iii/.

DS3 was collected in our lab. Its paradigm was similar to the BCI Competition. It contained 26 characters and 10 numbers. DS3 consisted of 12 participants who were graduate students between the ages of 20 and 26 years, with normal or corrected-to-normal vision. The experiments used a 64-channel wireless EEG acquisition system (Neuracle, NeuSen W series, 59 EEG, 4EOG, 1ECG) to acquire data at the sampling rate of 1,000 Hz. In the paradigm, each row and column in the six-by-six matrix was randomly intensified for 80 ms and kept extinguished for 80 ms. A trial for each target character included four epochs, and each epoch had 12 intensifications. Participants were required to spell 36 characters. We randomly selected 18 characters as the training dataset and the rest as the test dataset.

The Framework of Data Processing

Considering that some channels contain less task-relevant information but more noise, it is vital to use a reasonable method to select the most effective channels. This study compares the proposed RSBSBL with two empirical channel sets (Set 1 and Set 2) (Krusienski et al., 2008), LASSO, GLASSO, and SBL in the case of using the same pre-processing process and classifier. Set 1 includes Fz, Cz, Pz, Oz, PO7, and PO8. Set 2 includes Fz, FCz, Cz, C3, C4, CPz, Pz, P3, P4, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, and O2.

Figure 3 shows the diagram of the data processing framework, which includes three main parts: (1) pre-processing, (2) channel selection, and (3) classification. DS1 and DS2 shared the same pre-processing: bandpass filtering of data from 0.5 to 20 Hz and downsampling by a factor of 5. Then, the sampling rate of the data was 48 Hz. We intercepted 0–667 ms after each stimulus as the primary analysis objective was to obtain 32 sampling points for each stimulus. For the DS3, the 59-channel dataset that went through 0.5–20 Hz bandpass filtering was down-sampled to 50 Hz and the data segment from 0 to 600 ms was taken after stimulation to obtain 30 sampling points for each stimulus. Thus, denoting the number of channels as N_c and number of signal sampling points as N_t, a $1 \times D$ feature matrix was obtained for each stimulus, where $D = N_t N_c$. A feature matrix was labeled "1" only if the corresponding stimulus belongs to the row or column of the target characters. Otherwise, the label was assigned to "0."

The typical classification methods of P300 include traditional machine learning methods and neural network-based methods. Traditional machine learning can achieve outstanding performance with less complexity. This study regarded BLDA as a unified classifier for different channel selection algorithms.

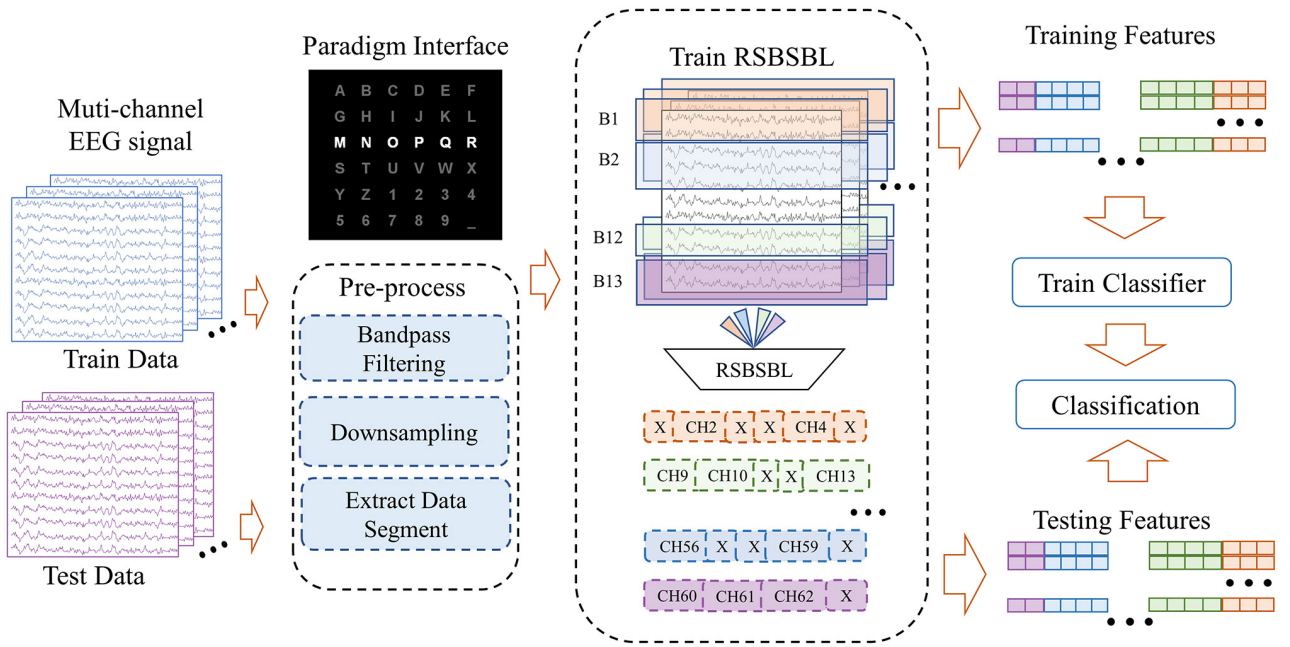


FIGURE 3 | Diagram of the data processing framework, including pre-processing, channel selection, and classification. Using the block sparsity property of RSBSBL, we do pruning on the eligible channels by fitting the training data and labels.

Parameter Setting

The optimal combination of parameters was determined by a 10-fold cross-validation. There were two modes of the selected channel number in the experiment for the channel selections: automatic and fixed. When the channel number was determined automatically, we used a threshold to determine the channel number. For LASSO and SBL, the absolute values of the feature weights in one channel were summed up to represent the importance of the channel. The threshold equaled the mean minus 0.5 times the standard deviation of the channel importance values, and the channels with importance values higher than the threshold were selected. As for GLASSO and RSBSBL, automatic channel selection had been enabled in the methods. When the number of selected channels is fixed (M channels were selected), we used the same way to evaluate each channel. For all the four methods, the absolute values of the feature weights $\mathbf{w}$ of each channel were summed, and the top M channels were selected in descending order.

Evaluation

We used character recognition accuracy to evaluate the performance of a classification. The character recognition accuracy is defined as follows:

$$Acc = \frac{C_{test\_correct}}{C_{test\_total}} \tag{17}$$

where C_{test_total} represents the total number of characters in the test dataset, and $C_{test\_correct}$ is the sum of all the correctly predicted characters. Besides, to evaluate the significance of performance difference, we introduced a non-parametric statistical hypothesis test, the Wilcoxon signed-rank test. The Wilcoxon signed-rank test can be used as an alternative to the paired t-test for matched pairs when the population cannot be assumed to be normally distributed. The significance of the pairs can be confirmed when the corresponding p-value is less than 0.05.

RESULTS

We evaluated the performance of the proposed method on the three datasets. The results covered the experiments of automatic channel selection and the experiments of selecting M channels. For further analysis, we also evaluated the sensitivity of the parameters of the proposed method.

Results of Automatic Channel Selection

Channel selection is supposed to reserve channels with more helpful information and exclude the channels with more noise. According to the data processing, we chose a unified classifier to verify the performance of different methods for a fair comparison. In **Table 2**, we compared the character recognition accuracy of each method on the three datasets, and the number of selected channels was automatically determined as described in section "Parameter Setting." Set 1 and Set 2 are empirical subsets of channels (Set 1 contains 6 channels and Set 2 contains 19 channels). The best results were marked in bold, and the number of channels selected for each participant is presented in the corresponding parentheses.

For DS1, RSBSBL selected the minimum number of channels when the classification accuracy of all the methods was 100%. For DS2, RSBSBL had the highest average accuracy, 97.50%, which

TABLE 2 | Character recognition accuracy (%) (number of channels) and Wilcoxon signed-rank test comparisons for DS1, DS2, and DS3 when each compared method was used for channel selection.

Participant	Methods					
	Set 1	**Set 2**	**LASSO**	**GLASSO**	**SBL**	**RSBSBL**
P1.1	**100.00**	**100.00**	**100.00** (43)	**100.00** (54)	**100.00** (36)	**100.00** (29)
P2.1	80.00	92.00	96.00 (43)	98.00 (64)	97.00 (44)	**99.00** (44)
P2.2	90.00	92.00	93.00 (41)	95.00 (56)	93.00 (39)	**96.00** (45)
Average	85.00	92.00	94.50 (42.00)	96.50 (60.00)	95.00 (41.50)	**97.50** (44.50)
P3.1	55.56	61.11	83.33 (39)	**88.89** (22)	83.33 (40)	**88.89** (15)
P3.2	50.00	61.11	77.78 (39)	72.22 (37)	66.67 (42)	**94.44** (14)
P3.3	72.22	72.22	72.22 (42)	72.22 (38)	72.22 (39)	**94.44** (13)
P3.4	72.22	77.78	77.78 (35)	**83.33 (24)**	77.78 (39)	77.78 (18)
P3.5	55.56	61.11	83.33 (42)	**88.89 (23)**	77.78 (40)	**88.89** (14)
P3.6	44.44	44.44	72.22 (39)	72.22 (24)	83.33 (40)	**88.89** (40)
P3.7	66.67	77.78	**83.33** (41)	**83.33** (31)	72.22 (38)	**83.33** (15)
P3.8	72.22	77.78	72.22 (40)	77.78 (16)	77.78 (40)	**88.89** (13)
P3.9	61.11	66.67	77.78 (42)	83.33 (23)	77.78 (38)	**88.89** (15)
P3.10	72.22	**94.44**	88.89 (42)	88.89 (32)	88.89 (41)	**94.44** (15)
P3.11	61.11	66.67	50.00 (36)	72.22 (20)	50.00 (40)	**77.78** (19)
P3.12	38.89	83.33	83.33 (40)	83.33 (24)	88.89 (39)	**94.44** (13)
Average	60.19	70.37	76.85 (39.75)	80.55 (26.17)	76.39 (39.67)	**88.43** (17.00)
p-value	**0.002**	**0.005**	**0.005**	**0.013**	**0.003**	–

Pi.j represents the jth participant in the ith dataset. The number of selected channels is in parentheses. The highest classification accuracy of each participant of different methods is indicated in bold. p-value is the results of Wilcoxon signed-rank test. Set 1 includes F_z, Cz, Pz, Oz, PO7, and PO8. Set 2 includes Fz, FCz, Cz, C3, C4, CPz, Pz, P3, P4, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, and O2.

was 1.00% higher than the second-ranked GLASSO. Although SBL selected fewer channels than others, the average recognition accuracy was 95.00%.

For DS3, RSBSBL as a channel selection method could bring higher accuracy with BLDA in 11 participants among 12 and got 88.43% average accuracy by eliminating insufficient data than using all channels. It outperformed the second-ranked GSBL on an average by 7.88% and selected the fewest channels as 17. We evaluated the significance of the classification performance of DS3 *via* the Wilcoxon signed-rank test and found that the proposed method performed significantly better than others (RSBSBL vs. LASSO: $p = 0.005 < 0.05$; RSBSBL vs. GLASSO: $p = 0.013 < 0.05$; RSBSBL vs. SBL: $p = 0.003 < 0.05$).

Results of Selecting M Channels

To further compare the effectiveness of the four methods, we compared the recognition results of the algorithms when M channels were selected ($M = [4, 8, 12, 16]$). Top M channels were selected by ranking the corresponding channels according to the sum of the absolute values of the feature weights. The classifiers were retrained with the data with the selected channel. It was supposed that the number of channels M' automatically selected by the method was less than the value of M. In that case, the latest deleted M-M′ channels are added according to the order in which they were deleted during the iteration of the method.

Figure 4 shows the accuracy of each method on DS1, DS2, and DS3, with the horizontal coordinates of the bars indicating the selection of the top M channels. For DS1, the accuracy of all the methods was the same except that the accuracy of SBL was 96.77% when eight channels were selected, and it was lower than others. For DS2, SBL and RSBSBL obtained better performance with 80% average recognition accuracy when four channels were selected. When 8, 12, and 16 channels were selected, GLASSO obtained an average recognition accuracy of 78.5, 84.5, 91, and 92%, respectively, and RSBSBL obtained a better performance of 80, 85.5, 91.5, and 93.5%, respectively. For DS3, GLASSO obtained average recognition accuracy of 73.61, 75.93, 75.46, and 79.63% when 4, 8, 12, and 16 channels were selected, respectively. Moreover, RSBSBL obtained the best performance of 74.07, 82.87, 80.09, and 80.56%, respectively. The average recognition accuracies of LASSO, GLASSO, and RSBSBL on DS3 with $M = 16$ were 77.31, 79.63, and 80.56%, respectively. The results of experiments with the fixed number of selected channels revealed that the feature weights generated by RSBSBL could provide more reasonable guidelines for the channel selection.

We counted the selected channels at the same location and used it to describe the number of times a channel has been selected in the dataset. If 6 of the 12 participants' selected channels contain Pz, then the contribution value of the channel corresponding to the Pz electrode is 6. **Figure 5** indicates the scalp distributions of the contribution value of channels on DS1, DS2, and DS3. The color changes from red to blue, indicating that the channel was selected less often. As shown in **Figure 5**, when the number of selected channels was small ($M = 4, 8$), RSBSBL selected the occipital and parietal electrodes more often. It shows that, in addition to the P300 potential, the early visual components also contribute to a classification in the paradigm (Blankertz et al., 2011).

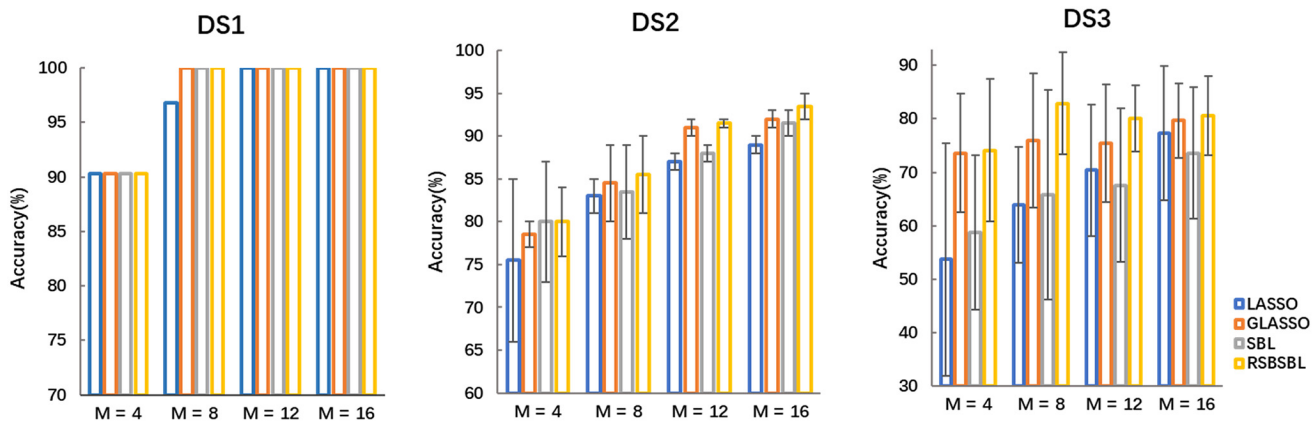


FIGURE 4 | The average recognition accuracy of the four methods on DS1, DS2, and DS3 when M channels are selected, where $M = [4, 8, 12, 16]$. The error bars are the standard deviations for DS2 and DS3.

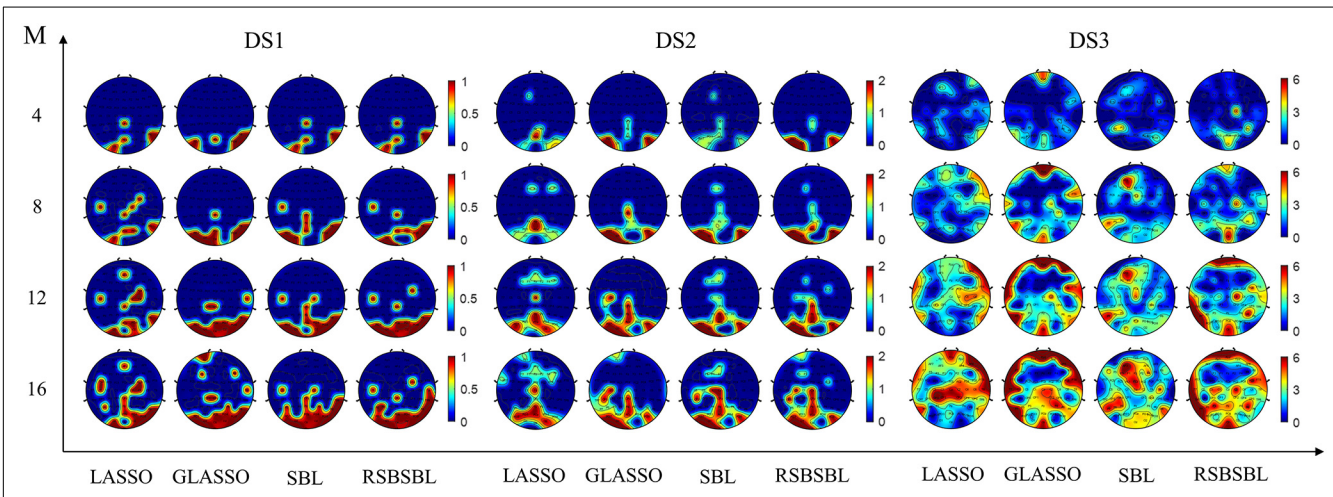


FIGURE 5 | The scalp distribution of the four methods on DS1, DS2, and DS3 when M channels are selected. The contribution value of each channel is equal to the sum of the selected numbers among all participants in the dataset. The color changes from red to blue, indicating that the channel is selected less often.

Parameter Sensitivity

In RSBSBL, γ_b smaller than the threshold τ was set to zero, indicating that τ determines the pruning strength. We analyzed the change in the number of channels selected and the recognition results when τ is assigned different values in the range 10^{-8} to 10^{-1}. The recognition accuracy of each participant varying with τ was normalized to highlight the location of the optimal threshold. **Figure 6** illustrates the effect of the threshold on the proposed method. The x-axis indicates the number of selected channels, the y-axis indicates the value of τ, and the z-axis indicates the participant ID. The color changes from red to blue, indicating that the point corresponds to a higher to lower normalized accuracy.

As shown in **Figure 6**, the number of channels selected by each participant increased as the threshold value decreased. When the threshold was less than or equal to 10^{-6}, the number of selected channels was the original number in the dataset, and the algorithm loses the ability to select the channels automatically. Therefore, 10-fold cross-validation can be used to select the optimal parameter values in the range of 10^{-6} to 10^{-1}. From the curves corresponding to P3.2, P3.3, P3.7, and P3.12, using selected channels can obtain better recognition accuracy than using all the channels, which proves that channel selection can remove weak task-relevant and noisy channels to improve the classification accuracy.

DISCUSSION

The experimental results on the three datasets illustrated that the proposed RSBSBL as a channel selection algorithm could automatically screen out effective channels and get the best overall performance among all the compared methods.

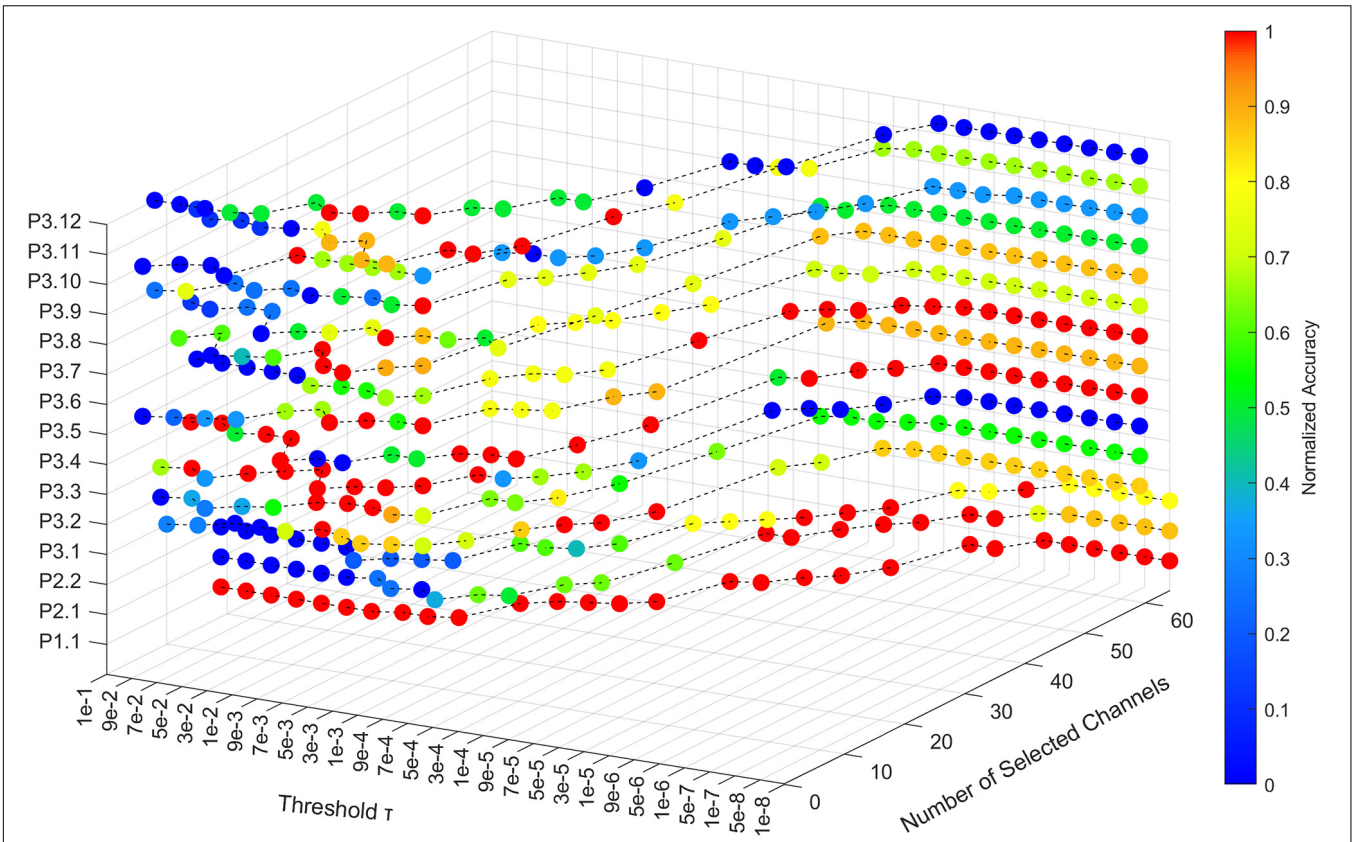


FIGURE 6 | The effect of shear threshold τ in RSBSBL on the number of selected channels and accuracy. The *x*-axis indicates the number of selected channels, the *y*-axis indicates the value of τ, and the *z*-axis indicates the participant ID. The color of the sphere represents the normalized recognition accuracy for each participant with different thresholds.

Effectiveness of Channel Subsets

Fabiani et al. (1987) confirmed that the visual P300 paradigm should at least include Fz, Cz, Pz electrodes signed as the 10–20 international electrode system. Krusienski et al. (2008) and McCann et al. (2015) made sure that Fz, Cz, Pz, Oz, PO7, and PO8 corresponded to the parietal and occipital regions of the brain that take a significant part in the recognition of P300 signals. In **Table 2**, Set 1 and Set 2 represent two empirical channel subsets. Set 1 includes Fz, Cz, Pz, Oz, PO7, and PO8. Set 2 includes Fz, FCz, Cz, C3, C4, CPz, Pz, P3, P4, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, and O2. It can be seen that for many participants (P2.1, P2.2, P3.1, P3.2, P3.5, P3.6), the character recognition accuracy was lower when the empirical channel subsets were used. The empirical selection may not include some channels that contribute to the classification. The channels assumed to reflect visual components and also some frontal channels contribute to the classification for some participants. It also indicates the lower robustness of the empirical channel subset. In **Figure 5**, the scalp mapped according to channel selection of RSBSBL could be observed with high values in Pz, P3, P4, O1, O2, Oz, PO7, PO8, and POz regions. These electrodes are very similar to the abovementioned electrodes, which are closely related to the visually induced ERPs. The P1, N1, and N2 components are mainly concentrated in the parietal and occipital regions. And the central distribution of P2 and P3 is elongated along the midline electrodes (Blankertz et al., 2011). It can be assumed that a multitude of ERP components is affected by attention to the target and utilized by classifiers rather than just the P300 (Treder and Blankertz, 2010). In addition, it can be found from **Figures 5**, **6** that many participants in DS3 had poorer classification using full-channel data compared to DS1 and DS2, and their topographic maps select more frontal channels when M = 8, 12,16. This phenomenon may be due to the effect of eye artifacts and noise during the experiment.

Character Recognition Performance

Table 2 and **Figure 3** show the superiority of RSBSBL in channel selection. When the number of channels was determined automatically, the proposed method achieved the highest average recognition accuracy of 100, 97.5, and 88.43% for DS1, DS2, and DS3, with the lowest average number of channels on DS1 and DS3. The RSBSBL achieved better performance than the compared methods when selecting the channels with the fixed number, and the average accuracies of 90.21, 80, and 74.07% were obtained with the top four selected channels on the three datasets.

To verify the performance of RSBSBL, we compared the proposed method with the state-of-the-art developments in recent years on DS2, as shown in **Table 3**. Most of them are

TABLE 3 | Character recognition accuracy (%) of comparison with state-of-the-art results (DS2).

Author	Channel selection method	Classification method	Accuracy
Kee et al., 2015	NSGA-II	BLDA	94.9%
Khairullah et al., 2020	BPSO	Ensemble LDA	97.0%
Tang et al., 2020	RF-GA	CNN	96.9%
Martinez-Cagigal et al., 2022	BMOPSO	LDA	92.5%
	PEAIL	LDA	94.0%
Our method	RSBSBL	BLDA	97.5%

NSGA-II, Non-dominated sorting genetic algorithm II; BPSO, binary particle swarm optimization; GA, genetic algorithm; BMOPSO, binary multi-objective particle swarm optimization; PEAIL, Pareto Evolutionary Algorithm based on Incremental Learning.

TABLE 4 | The average character recognition accuracy (%) (number of channels) comparisons on three datasets.

Method	DS1	DS2	DS3
Case 1	100.00 (30)	94.50 (35.50)	82.41 (17.75)*
Case 2	100.00 (27)	96.00 (36.50)	82.41 (18.67)*
Case 3	100.00 (27)	96.50 (47.50)	85.65 (16.33)
Our method	100.00 (29)	97.50 (44.50)	88.43 (17.00)

Case 1: B is the unitary matrix. Case 2: All blocks have the same B. Case 3: The B of each block is different. "" represents a significant difference with our method after Wilcoxon signed-rank test ($p < 0.05$).*

based on evolutionary computational algorithms (Kee et al., 2015; Khairullah et al., 2020; Tang et al., 2020; Martinez-Cagigal et al., 2022). The channel selection methods and classifiers used in each study are shown in the table.

The shear threshold τ significantly impacted the final results, so cross-validation was required to determine the optimal parameters. According to the analysis of parameter sensitivity, as shown in **Figure 6**, the recommended threshold selection range was $[10^{-6}, 10^{-1}]$. Besides, **Figure 6** reflects the variation of character recognition accuracy with the shear threshold for each participant. Compared with others, P3.2, P3.3, P3.9, P3.10, and P3.12 cannot achieve the best recognition accuracies with the full channels, which implies that the EEG signals of these participants have more channels with noise, and these channels are not conducive to signal classification. As shown in **Table 2**, when determining the number of channels automatically, RSBSBL can achieve the best recognition accuracies of them with the corresponding number of selected channels of 14, 13, 15, 15, and 13, respectively. It confirms that RSBSBL can remove unfavorable channels and improve the recognition accuracies.

Effectiveness of Regional Smoothing

To verify the effectiveness of regional smoothing, we conducted further controlled experiments on the three datasets, and the results are shown in **Table 4**. Case 1 represents that B is a unit matrix, implying that no temporal correlation is considered. Case 2 has the same B for all blocks, indicating that all channels share the same B. Case 3 has a different B matrix for each block, showing that regional smoothing is no longer done. The comparison between Case 3 and Case 1 in **Table 4** illustrates the improvement of the model due to temporal correlation. The comparison between our algorithm and Cases 3 and 1 indicates the improvement brought by region smoothing. The"*" in **Table 4** represents a significant difference in our method after Wilcoxon signed-rank test (RSBSBL vs. Case 1: $p = 0.015 < 0.05$; RSBSBL vs. Case 2: $p = 0.031 < 0.05$; RSBSBL vs. Case 3: $p = 0.124$).

Time Costs and Limitations

As described in sections "Data Descriptions" and "The Framework of Data Processing," for DS1 and DS2, $N_t = 32$ and$N_c = 64$ after pre-processing, then we can get a $1 \times D$ ($D = N_tN_c = 2048$) vector for each stimulus. As shown in **Table 1**, in the training datasets of DS1 and DS2, the total number of stimuli was 7,560 and 15,300, which is larger than the number of features D. For DS3, $N_t = 30$ and $N_c = 59$ after pre-processing, then the feature is a $1 \times D$ ($D = N_tN_c = 1770$) vector. In **Table 1**, in the training datasets of DS3, the total number of stimuli was 864, which is smaller than the number of its features.

In a preliminary study, we found that inappropriate iterations can make the algorithm to have a large time cost [e.g., using equations (10) and (11) on DS1 and DS2]. Therefore, a strategy of automatic selection of the iteration method is used to avoid this problem. In **Figure 7**, we analyze the variation of the matrix

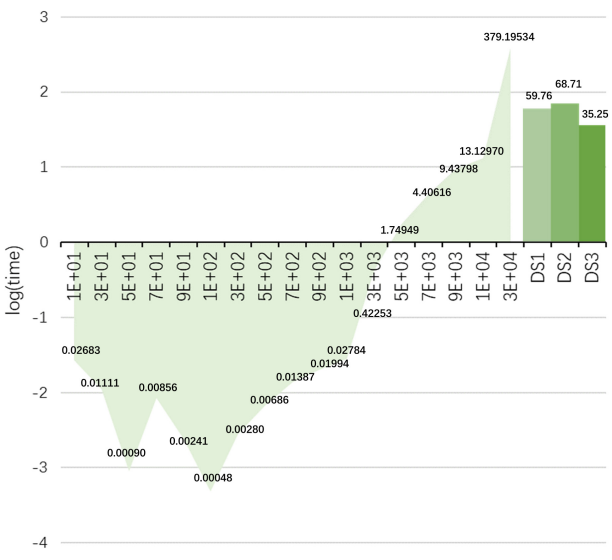


FIGURE 7 | Changes in the run-time (s) of matrix inversion when the size of the matrix increases. In the left part, the horizontal axis represents the size of the square array. The vertical axis is the value after taking the logarithm of the time. The bar chart represents the average time cost of the proposed method on the three datasets.

inversion run-time when the size of the matrix increases (the matrix is a square matrix). In the left part, the horizontal axis represents the size of the square matrix. The vertical axis is the value after taking the logarithm of the time, and the actual time (s) is also indicated in the figure. It can be noticed that the time spent on matrix inversion is more than 1 s when the matrix size is larger than 3,000 × 3,000. Therefore, we consider that the method may not be suitable for data with numbers of features and samples larger than 3,000. Of course, this problem can be solved by reducing the number of features and optimizing the iteration steps. The right bar in **Figure 7** indicates the average time cost of the proposed method on the three data sets, which is acceptable.

Future Work

The sparse Bayesian algorithm can make the sparsity of the algorithm change by changing the prior distribution of **w** (Tipping, 2001). Zhang et al. (2015) used the Laplace distribution instead of the traditional Gaussian distribution for the classification of P300 signals using SBL. Therefore, RSBSBL can change the prior of the weights to make the sparsity stronger in the future, such as the Gamma distribution. The proposed method used the EM algorithm for iteration, and there is still room for improvement in the computational speed. In the future, we will also explore the suitability of the proposed method for other ERPs.

CONCLUSION

This study proposed a novel channel selection method, namely RSBSBL, which improved the original BSBL and obtained the assigned sparse weights. While considering the temporal correlation of sampling points of the same channel, it exploits the spatial distribution characteristics of the electrodes so that channels in adjacent regions share a positive definite matrix to get regional smoothing. Also, we discussed the efficiency of RSBSBL in the channel selection and design an automatic selection iteration strategy model to reduce the time cost caused by the inverse operation of the large-size matrix. The experimental results on three datasets indicate that RSBSBL can select appropriate channels, leading to high recognition accuracy. We will conduct future studies to improve the robustness of this algorithm.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of East China University of Science and Technology. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

XZ was the main author to raise the idea of the manuscript, designed the experimental procedure, and collected the original dataset. JJ made effective suggestions on the manuscript's structure and provided the experimental site. RX has embellished the language of the manuscript and made key suggestions. SL and HS were involved in revising the manuscript's results section. XW and AC provided inputs for optimizing the data processing flow. All authors contributed to the manuscript revision, and read and approved the submitted version.

FUNDING

This work was supported by the Science and Technology Innovation 2030 Major Projects 2022ZD0208900 and the Grant National Natural Science Foundation of China under Grant 62176090; in part by Shanghai Municipal Science and Technology Major Project under Grant 2021SHZDZX; in part by the Program of Introducing Talents of Discipline to Universities through the 111 Project under Grant B17017; in part by the Shu Guang Project supported by the Shanghai Municipal Education Commission and the Shanghai Education Development Foundation under Grant 19SG25; in part by the Ministry of Education and Science of the Russian Federation under Grant 14.756.31.0001, and in part by the Polish National Science Center under Grant UMO-2016/20/W/NZ4/00354.

REFERENCES

Alotaiby, T., Abd El-Samie, F. E., Alshebeili, S. A., and Ahmad, I. (2015). A review of channel selection algorithms for EEG signal processing. *EURASIP J. Adv. Signal Process* 2015, 1–21.

Ando, M., Nobukawa, S., Kikuchi, M., and Takahashi, T. (2021). Identification of Electroencephalogram Signals in Alzheimer's Disease by Multifractal and Multiscale Entropy Analysis. *Front. Neurosci.* 15:667614. doi: 10.3389/fnins.2021.667614

Arican, M., and Polat, K. (2019). ""Comparison of the Performances of Selected EEG Electrodes with Optimization Algorithms in P300 Based Speller Systems,"" in *in 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, (Piscataway: IEEE), 1–4.

Arvaneh, M., Robertson, I. H., and Ward, T. E. (2019).). A p300-based brain-computer interface for improving attention. *Front. Hum. Neurosci.* 12:524. doi: 10.3389/fnhum.2018.00524

Bashashati, A., Fatourechi, M., Ward, R. K., and Birch, G. E. (2007). A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *J. Neural Eng.* 4:R32. doi: 10.1088/1741-2560/4/2/R03

Blankertz, B., Lemm, S., Treder, M., Haufe, S., and Müller, K.-R. (2011). Single-trial analysis and classification of ERP components—a tutorial. *Neuroimage* 56, 814–825. doi: 10.1016/j.neuroimage.2010.06.048

Blankertz, B., Muller, K.-R., Curio, G., Vaughan, T. M., Schalk, G., Wolpaw, J. R., et al. (2004). The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials. *IEEE Trans. Biomed. Eng.* 51, 1044–1051. doi: 10.1109/TBME.2004.826692

Blankertz, B., Muller, K.-R., Krusienski, D. J., Schalk, G., Wolpaw, J. R., Schlogl, A., et al. (2006). The BCI competition III: validating alternative approaches to actual BCI problems. *IEEE Trans. Neural Syst. Rehabil. Eng.* 14, 153–159. doi: 10.1109/TNSRE.2006.875642

Cecotti, H., and Graser, A. (2010). Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Transac. Pattern Anal. Mach. Intell.* 33, 433–445. doi: 10.1109/TPAMI.2010.125

Cecotti, H., Rivet, B., Congedo, M., Jutten, C., Bertrand, O., Maby, E., et al. (2011). A robust sensor-selection method for P300 brain-computer interfaces. *J. Neural Eng.* 8:16001. doi: 10.1088/1741-2560/8/1/016001

Deng, X., Yu, Z. L., Lin, C., Gu, Z., and Li, Y. (2019). A bayesian shared control approach for wheelchair robot with brain machine interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* 28, 328–338. doi: 10.1109/TNSRE.2019.2958076

Dey, M. R., Shiraz, A., Sharif, S., Lota, J., and Demosthenous, A. (2020). Dictionary selection for compressed sensing of EEG signals using sparse binary matrix and spatiotemporal sparse Bayesian learning. *Biomed. Phys. Eng. Express* 6:65024. doi: 10.1088/2057-1976/abc133

Donchin, E., Spencer, K. M., and Wijesinghe, R. (2000). The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE Trans. Biomed. Eng.* 8, 174–179. doi: 10.1109/86.847808

Fabiani, M., Gratton, G., Karis, D., and Donchin, E. (1987). Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. *Adv. Psychophysiol.* 2:78.

Hammer, E. M., Halder, S., Kleih, S. C., and Kübler, A. (2018). Psychological predictors of visual and auditory P300 brain-computer interface performance. *Front. Neurosci.* 12:307. doi: 10.3389/fnins.2018.00307

Hassan, M., and Wendling, F. (2018). Electroencephalography Source Connectivity: aiming for High Resolution of Brain Networks in Time and Space. *IEEE Signal Proc. Magaz.* 35, 81–96. doi: 10.1109/MSP.2017.2777518

He, S., Zhou, Y., Yu, T., Zhang, R., Huang, Q., Chuai, L., et al. (2019). EEG-and EOG-based asynchronous hybrid BCI: a system integrating a speller, a web browser, an e-mail client, and a file explorer. *IEEE Trans. Neural Syst. Rehabil. Eng.* 28, 519–530. doi: 10.1109/TNSRE.2019.2961309

Hoffmann, U., Vesin, J.-M., Ebrahimi, T., and Diserens, K. (2008a). An efficient P300-based brain-computer interface for disabled subjects. *J. Neurosci. Methods* 167, 115–125. doi: 10.1016/j.jneumeth.2007.03.005

Hoffmann, U., Yazdani, A., Vesin, J.-M., and Ebrahimi, T. (2008b). "Bayesian feature selection applied in a P300 brain-computer interface,"," in *2008 16th European Signal Processing Conference*, (Piscataway: IEEE), 1–5.

Huang, Z., Guo, J., Zheng, W., Wu, Y., Lin, Z., and Zheng, H. (2022). A Calibration-free Approach to Implementing P300-based Brain-computer Interface. *Cogn. Comput.* 14, 887–899.

Jin, J., Li, S., Daly, I., Miao, Y., Liu, C., Wang, X., et al. (2019). The study of generic model set for reducing calibration time in P300-based brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* 28, 3–12. doi: 10.1109/TNSRE.2019.2956488

Jin, J., Sellers, E. W., Zhou, S., Zhang, Y., Wang, X., and Cichocki, A. (2015). A P300 brain-computer interface based on a modification of the mismatch negativity paradigm. *Int. J. Neural Syst.* 25:1550011. doi: 10.1142/S0129065715500112

Kee, C.-Y., Ponnambalam, S. G., and Loo, C.-K. (2015). Multi-objective genetic algorithm as channel selection method for P300 and motor imagery data set. *Neurocomputing* 161, 120–131.

Khairullah, E., Arican, M., and Polat, K. (2020). Brain-computer interface speller system design from electroencephalogram signals with channel selection algorithms. *Med. Hypothes.* 141:109690. doi: 10.1016/j.mehy.2020.109690

Kim, K. T., Suk, H. I., and Lee, S. W. (2016). Commanding a brain-controlled wheelchair using steady-state somatosensory evoked potentials. *IEEE Trans. Neural Syst. Rehabil. Eng.* 26, 654–665. doi: 10.1109/TNSRE.2016.2597854

Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2008). Toward enhanced P300 speller performance. *J. Neurosci. Methods* 167, 15–21. doi: 10.1016/j.jneumeth.2007.07.017

Lal, T. N., Schroder, M., Hinterberger, T., Weston, J., Bogdan, M., Birbaumer, N., et al. (2004). Support vector channel selection in BCI. *IEEE Trans. Biomed. Eng.* 51, 1003–1010. doi: 10.1109/TBME.2004.827827

Lei, X., Yang, P., and Yao, D. (2009). An empirical Bayesian framework for brain-computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* 17, 521–529. doi: 10.1109/TNSRE.2009.2027705

Liu, H., and Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Trans. Knowledge Data Eng.* 17, 491–502. doi: 10.1109/TCBB.2009.6

Lopez-Calderon, J., and Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* 8:213. doi: 10.3389/fnhum.2014.00213

Maestú, F., Cuesta, P., Hasan, O., Fernandéz, A., Funke, M., and Schulz, P. E. (2019). The Importance of the Validation of M/EEG With Current Biomarkers in Alzheimer's Disease. *Front. Hum. Neurosci.* 13:17. doi: 10.3389/fnhum.2019.00017

Manyakov, N. V., Chumerin, N., Combaz, A., and van Hulle, M. M. (2011). Comparison of classification methods for P300 brain-computer interface on disabled subjects. *Comput. Intell. Neurosci.* 2011, 1–12. doi: 10.1155/2011/519868

Martinez-Cagigal, V., and Hornero, R. (2017). "Binary Bees Algorithm for P300-Based Brain-Computer Interfaces Channel Selection," in *Advances in Computational Intelligence. IWANN 2017. Lecture Notes in Computer Science()*, Vol. 10306, eds I. Rojas, G. Joya, and A. Catala (Cham: Springer).

Martinez-Cagigal, V., Santamaría-Vázquez, E., and Hornero, R. (2022). Brain-computer interface channel selection optimization using meta-heuristics and evolutionary algorithms. *Martínez-Cagigal. Appl. Soft Comput.* 115:108176.

McCann, M. T., Thompson, D. E., Syed, Z. H., and Huggins, J. E. (2015). Electrode subset selection methods for an EEG-based P300 brain-computer interface. *Disabil. Rehabil.-Assist. Technol.* 10, 216–220. doi: 10.3109/17483107.2014.884174

Nakanishi, M., Wang, Y., Chen, X., Wang, Y.-T., Gao, X., and Jung, T.-P. (2017). Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis. *IEEE Trans. Biomed. Eng.* 65, 104–112. doi: 10.1109/TBME.2017.2694818

Padfield, N., Zabalza, J., Zhao, H., Masero, V., and Ren, J. (2019). EEG-based brain-computer interfaces using motor-imagery: techniques and challenges. *Sensors* 19:1423. doi: 10.3390/s19061423

Rakotomamonjy, A., and Guigue, V. (2008). BCI competition III: dataset II-ensemble of SVMs for BCI P300 speller. *IEEE Trans. Biomed. Eng.* 55, 1147–1154. doi: 10.1109/TBME.2008.915728

Reichert, C., Dürschmid, S., Heinze, H.-J., and Hinrichs, H. (2017). A comparative study on the detection of covert attention in event-related EEG and MEG signals to control a BCI. *Front. Neurosci.* 11:575. doi: 10.3389/fnins.2017.00575

Sorbello, R., Tramonte, S., Giardina, M. E., La Bella, V., Spataro, R., Allison, B., et al. (2017). A human-humanoid interaction through the use of BCI for locked-in ALS patients using neuro-biological feedback fusion. *IEEE Trans. Neural Syst. Rehabil. Eng.* 26, 487–497. doi: 10.1109/TNSRE.2017.2728140

Tang, C., Xu, T., Chen, P., He, Y., Bezerianos, A., and Wang, H. (2020). ""A Channel Selection Method for Event Related Potential Detection based on Random Forest and Genetic Algorithm,"," in *2020 Chinese Automation Congress (CAC), 5419–5424*, (Piscataway: IEEE).

Tekgul, H., Bourgeois, B. F. D., Gauvreau, K., and Bergin, A. M. (2005). Electroencephalography in neonatal seizures: comparison of a reduced and a full 10/20 montage. *Pediatr. Neurol.* 32, 155–161. doi: 10.1016/j.pediatrneurol.2004.09.014

Tibshirani, R. (1996). Regression shrinkage and selection *via* the lasso. *J. R. Stat. Soc. Ser. B-Stat. Methodol.* 58, 267–288.

Tipping, M. E. (2001). Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* 1, 211–244. doi: 10.1016/j.cmpb.2008.05.002

Tomioka, R., and Müller, K.-R. (2010). A regularized discriminative framework for EEG analysis with application to brain-computer interface. *Neuroimage* 49, 415–432. doi: 10.1016/j.neuroimage.2009.07.045

Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C., et al. (2010). A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clin. Neurophysiol.* 121, 1109–1120. doi: 10.1016/j.clinph.2010.01.030

Treder, M. S., and Blankertz, B. (2010). (C) overt attention and visual speller design in an ERP-based brain-computer interface. *Behav. Brain Funct.* 6, 1–13. doi: 10.1186/1744-9081-6-28

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* 113, 767–791.

Wu, W., Wu, C., Gao, S., Liu, B., Li, Y., and Gao, X. (2014). Bayesian estimation of ERP components from multicondition and multichannel EEG. *Neuroimage* 88, 319–339. doi: 10.1016/j.neuroimage.2013.11.028

Yin, L., Wang, K., Tong, T., An, Y., Meng, H., Yang, X., et al. (2020). Improved Block Sparse Bayesian Learning Method Using K-Nearest Neighbor Strategy for Accurate Tumor Morphology Reconstruction in Bioluminescence Tomography. *IEEE Trans. Biomed. Eng.* 67, 2023–2032. doi: 10.1109/TBME.2019.2953732

Yuan, M., and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B* 68, 49–67.

Zhang, Y., Wang, Y., Jin, J., and Wang, X. (2017). Sparse Bayesian learning for obtaining sparsity of EEG frequency bands based feature vectors in motor imagery classification. *Int. J. Neural Syst.* 27:1650032.

Zhang, Y., Zhou, G., Jin, J., Zhao, Q., Wang, X., and Cichocki, A. (2015). Sparse Bayesian classification of EEG for brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* 27, 2256–2267. doi: 10.1109/TNNLS.2015.2476656

Zhang, Z., Jung, T.-P., Makeig, S., and Rao, B. D. (2013). Compressed Sensing for Energy-Efficient Wireless Telemonitoring of Noninvasive Fetal ECG *Via* Block Sparse Bayesian Learning. *IEEE Trans. Biomed. Eng.* 60, 300–309. doi: 10.1109/TBME.2012.2226175

Zhang, Z., and Rao, B. D. (2011). Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning. *IEEE J. Sel. Top. Signal Process.* 5, 912–926.

Conflict of Interest: RX is employed by the company g.tec medical engineering GmbH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 14 June 2022
doi: 10.3389/fnagi.2022.866230

Automated Classification of Mild Cognitive Impairment by Machine Learning With Hippocampus-Related White Matter Network

Yu Zhou[1†], Xiaopeng Si[2,3,4†], Yi-Ping Chao[5,6], Yuanyuan Chen[2,3], Ching-Po Lin[7], Sicheng Li[2,3], Xingjian Zhang[2,3], Yulin Sun[2,3], Dong Ming[2,3*] and Qiang Li[1*]*

[1] School of Microelectronics, Tianjin University, Tianjin, China, [2] Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin, China, [3] Tianjin Key Laboratory of Brain Science and Neural Engineering, Tianjin University, Tianjin, China, [4] Institute of Applied Psychology, Tianjin University, Tianjin, China, [5] Graduate Institute of Biomedical Engineering, Chang Gung University, Taoyuan, Taiwan, [6] Department of Computer Science and Information Engineering, Chang Gung University, Taoyuan, Taiwan, [7] Institute of Neuroscience, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

Background: Detection of mild cognitive impairment (MCI) is essential to screen high risk of Alzheimer's disease (AD). However, subtle changes during MCI make it challenging to classify in machine learning. The previous pathological analysis pointed out that the hippocampus is the critical hub for the white matter (WM) network of MCI. Damage to the white matter pathways around the hippocampus is the main cause of memory decline in MCI. Therefore, it is vital to biologically extract features from the WM network driven by hippocampus-related regions to improve classification performance.

Methods: Our study proposes a method for feature extraction of the whole-brain WM network. First, 42 MCI and 54 normal control (NC) subjects were recruited using diffusion tensor imaging (DTI), resting-state functional magnetic resonance imaging (rs-fMRI), and T1-weighted (T1w) imaging. Second, mean diffusivity (MD) and fractional anisotropy (FA) were calculated from DTI, and the whole-brain WM networks were obtained. Third, regions of interest (ROIs) with significant functional connectivity to the hippocampus were selected for feature extraction, and the hippocampus (HIP)-related WM networks were obtained. Furthermore, the rank sum test with Bonferroni correction was used to retain significantly different connectivity between MCI and NC, and significant HIP-related WM networks were obtained. Finally, the classification performances of these three WM networks were compared to select the optimal feature and classifier.

Results: (1) For the features, the whole-brain WM network, HIP-related WM network, and significant HIP-related WM network are significantly improved in turn. Also, the accuracy of MD networks as features is better than FA. (2) For the classification algorithm, the support vector machine (SVM) classifier with radial basis function, taking the significant HIP-related WM network in MD as a feature, has the optimal classification

OPEN ACCESS

Edited by:
Peng Xu,
University of Electronic Science and Technology of China, China

Reviewed by:
Yong Liu,
Beijing University of Posts and Telecommunications (BUPT), China
Xujun Duan,
University of Electronic Science and Technology of China, China
Dong Wen,
University of Science and Technology Beijing, China

***Correspondence:**
Xiaopeng Si
xiaopengsi@tju.edu.cn
Dong Ming
richardming@tju.edu.cn
Qiang Li
liqiang@tju.edu.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

Received: *31 January 2022*
Accepted: *21 March 2022*
Published: *14 June 2022*

Citation:
Zhou Y, Si X, Chao Y-P, Chen Y, Lin C-P, Li S, Zhang X, Sun Y, Ming D and Li Q (2022) Automated Classification of Mild Cognitive Impairment by Machine Learning With Hippocampus-Related White Matter Network. Front. Aging Neurosci. 14:866230. doi: 10.3389/fnagi.2022.866230

performance (accuracy = 89.4%, AUC = 0.954). (3) For the pathologic mechanism, the hippocampus and thalamus are crucial hubs of the WM network for MCI.

Conclusion: Feature extraction from the WM network driven by hippocampus-related regions provides an effective method for the early diagnosis of AD.

Keywords: mild cognitive impairment, white matter connectivity, Alzheimer's disease, early diagnosis, feature extraction, machine learning

INTRODUCTION

Alzheimer's disease (AD) is a chronic neurodegenerative disease with irreversible progression (Sperling et al., 2011; Hyman et al., 2012). Mild cognitive impairment (MCI) is the prodromal stage of AD (Petersen et al., 1999; Gauthier et al., 2006). The primary clinical manifestation of MCI is memory loss (Petersen et al., 2001). Since the progression of AD is irreversible and the treatment of AD has little effect, the detection of MCI is expected to find out the high risk of AD and further prevent its occurrence (Jacobs et al., 2013; Wang et al., 2013). However, the subtle changes of brain microstructure in MCI make it challenging to distinguish the disease from conventional radiography (Pellegrini et al., 2018). Therefore, establishing reliable biomarkers to diagnose MCI in the early stages remains challenging.

According to the biomarker modeling of AD, white matter demyelination has been proved to appear earlier during AD progression than abnormal changes of gray matter and functional connectivity (Zhuang et al., 2012; Lee et al., 2015). Recent studies have shown that genes and protein molecules ultimately cause microstructure changes in specific white matter fibers (Yu et al., 2021; Zhao et al., 2021). Thus, white matter degeneration is a valid biomarker for MCI. Diffusion tensor imaging (DTI) could detect subtle structural changes in white matter fibers, facilitating large-scale non-invasive screening (Tournier, 2019). The DTI index, mean diffusivity (MD), and fractional anisotropy (FA) describe fiber tracts' state.

Machine learning offers a systematic approach to developing advanced, automatic, and objective classification frameworks for MCI diagnosis. The classification framework mainly includes feature extraction and classification algorithms (Rathore et al., 2017). Although there are plenty of studies on AD classification, there is insufficient research on MCI (Shatte et al., 2019). The classification accuracies for AD were all around 80%, while the accuracies for MCI were only about 60% (Wee et al., 2011; Dyrba et al., 2013, 2015a; Nir et al., 2015; Prasad et al., 2015; Dou et al., 2020). The main reason is that whole-brain changes in AD are so significant that can be classified with high accuracy by the features of white matter, gray matter, and functional connectivity. In contrast, changes in MCI are not obvious for the whole brain. However, the progression of AD is irreversible. Early diagnosis of MCI fascinates high-risk screening of AD in time.

Previous studies focused on the classification algorithm to improve classification performance. Existing researchers found that k-nearest neighbor (KNN) (Ebadi et al., 2017), random forest (RF) (Maggipinto et al., 2017; Wang et al., 2018), and support vector machine (SVM) (Cui et al., 2012; Demirhan et al., 2015; Dyrba et al., 2015a; Nir et al., 2015; Xie et al., 2015; Ahmed et al., 2017) have achieved better classification performance for MCI when taking the white matter as a feature. KNN is based on Euclid's theorem and is classified by measuring the distance between different features. RF integrates many decision trees into a forest and combines them for predicting. SVM is based on statistical theory to solve two classification problems, mainly introducing kernel function to solve the problem of linear inseparability. However, previous studies only attached importance to algorithms while neglecting the features. In fact, selecting the appropriate modality in the data and extracting suitable features are usually more important than the underlying algorithm (Zhang et al., 2021).

The white matter feature extraction methods for MCI classification mainly include specific fiber tracking and the whole-brain white matter network (WM network). The specific fiber tracking method takes features using the fiber tracts that converge between localized brain regions and calculates the DTI index of voxels on a single fiber tract. However, this method focused on a specific fiber tract and had low classification accuracy (Nir et al., 2015; Dou et al., 2020). In the whole-brain white matter connectivity network approach, DTI images were segmented into several anatomical regions and features based on the metrics calculated from the fibers within these regions (Wee et al., 2011; Prasad et al., 2015). Recent studies had started to use multimodality for feature extraction. DTI, T1-weighted (T1w), and resting-state functional MRI (rs-fMRI) images from different modes were used to capture information from different perspectives (Hinrichs et al., 2011; Liu et al., 2014). However, most studies had concatenated the features from different modalities. The disadvantage is that all features are treated equally, and it provides no way to account for the diverse nature of features extracted from different modalities (Zhang et al., 2011, 2012; Dyrba et al., 2015b). Therefore, the urgent problem for MCI classification is to combine with different modalities before training and develop a dedicated feature fusion strategy.

To address this issue, the brain network theory was introduced. First, the brain can be seen as a network of spatially dispersed brain areas that share information continuously in functional connectivity (Sporns and Betzel, 2016). The different brain regions are connected and have synchronization within the same activation pattern. Different brain regions are temporally synchronized under the same activation pattern by the white matter pathway (Gu et al., 2015). Second, the primary clinical manifestation of MCI is memory loss (Rugg and Vilberg, 2013). The hippocampus is the crucial area for memory processing. Therefore, investigating the abnormal white matter network related to the hippocampus could promote the detection of

MCI in time (Benoit and Schacter, 2015). Furthermore, previous studies have also demonstrated that white matter degeneration leads to abnormalities in the functional connectivity of the corresponding brain regions (Kleinschmidt and Vuilleumier, 2013). In the default mode network, damage to the white matter fibers such as fornix and cingulum could weaken the functional connectivity of the medial temporal lobe and precuneus to the hippocampus (Buckner and DiNicola, 2019). Therefore, we hypothesized that white matter connectivity between brain regions associated with the hippocampus could be an important biomarker for MCI recognition. Features extraction of white matter network through hippocampus-related regions would help to improve the MCI classification performance.

This study aimed to extract the effective features of white matter connectivity networks driven by regions related to the hippocampus and improve classification performance. In our study, the elderly people were divided into MCI group and normal control (NC) group. First, the DTI data were obtained, and the orientation distribution function (ODF) was calculated by constrained spherical deconvolution (CSD). Based on the ODF, the fiber tracts of the whole brain were constructed. Second, the DTI index of whole-brain white matter connectivity between each brain region under the automated anatomical labeling (AAL) was calculated to construct the MD and FA brain structure networks as features. Furthermore, the hippocampus was used as the seed point to define the regions of interest (ROIs). Brain regions with high correlation to the hippocampus in functional connectivity were obtained as ROIs compared with the AAL template. Finally, whole-brain white matter connectivity networks were screened by ROIs related to the hippocampus. Different classifiers (e.g., SVM, KNN, and RF) were used to validate the classification performance of the extracted features. The recursive feature elimination (RFE) ranked these features' contribution to analyze the pathological mechanisms of MCI. Our study proposed a method of feature extraction to improve the MCI classification performance. The study also revealed the pathological mechanisms of MCI by the ranked contribution of features. It would provide effective early aid to MCI diagnosis.

MATERIALS AND METHODS

Participants

After excluding the left-handed and other brain injury history, a total of 96 subjects met the criteria for inclusion (48 male and 48 female; age: 80.6 ± 5.4 years, mean ± std). All participants were provided written informed consent based on the Helsinki Declaration. The experimental protocol was approved by the Institutional Review Board of Tianjin University and the Ethics Committee of Chang Gung University.

Neuropsychological Behavior Testing

All subjects were tested on the clinical dementia rating (CDR) and the mini-mental status examination (MMSE) scale. The entry criteria for MCI diagnosis were as follows: (i) CDR = 0.5; (ii) 24 ≤ MMSE < 30 for well-educated subjects (education years ≥ 6) or 19 ≤ MMSE < 24 for less-educated subjects (education years < 6). The MCI group was 81.3 ± 3.6 (mean ± std) years with education duration of 5.9 ± 5.4 years (mean ± std). There was no difference between groups in age and education duration (**Table 1**).

Magnetic Resonance Imaging Data Acquisition

The DTI, T1w, and rs-fMRI images of each subject were acquired without personal information. The DTI with 30 diffusion encoding directions was acquired using echo planar imaging (EPI) sequence with parameters: *b*-value = 1,000 s/mm^2, TR/TE = 11,000/104 ms, field of view = 192 × 192 mm^2, matrix size = 128 × 128, slice thickness = 2 mm, voxel size = 2 × 2 × 2 mm^3, number of slices = 70, and number of excitations = 1. The rs-fMRI was acquired using gradient EPI sequence with parameters: TR/TE = 2,500/27 ms, flip angle = 77°, band width = 2,400 Hz/pixel, field of view = 260 × 260 mm^2, matrix size = 64 × 64, voxel size = 3.4 × 3.4 × 3.4 mm^3, number of slices = 43, scan time = 360 s, and time points = 180. The T1w imaging was acquired using a 3D magnetization prepared rapid gradient EPI sequence with parameters: TR/TE = 2,530/3.5 ms; TI = 1,100 ms; field of view = 260 × 260 mm^2; matrix size = 256 × 256; slice thickness = 1 mm; voxel size = 1 × 1 × 1 mm^3; and number of slices = 192.

White Matter Structure Network Establishment and Feature Extraction

The feature extraction of the white matter structure network for MCI classification consists of the following steps (**Figure 1**):

1. MRI data preprocessing.
2. Reconstruction for white matter fiber tracts by orientation distribution function (ODF). Based on DTI, the ODF was calculated by CSD.
3. Establishment for whole-brain structure network. DTI indexes (MD and FA) of white matter connectivity between the AAL 90 brain region were calculated to establish the whole-brain structure network as the full feature for MCI classification.
4. Feature selection driven by hippocampus-related ROIs. Brain regions with significant functional connectivity to the hippocampus in rs-fMRI were selected as ROIs. The

TABLE 1 | Demographic and neuropsychological information of the MCI and NC groups.

	NC (*n* = 54)	MCI (*n* = 42)	*p*-value
Gender (m/f)	27/27	21/21	1.0000[a]
Age (year)	80.0 ± 6.3[b]	81.3 ± 3.6[b]	0.6353[c]
Education (year)	6.5 ± 4.1[b]	5.9 ± 5.4[b]	0.1530[c]
CDR	0	0.5	<0.0001[c]
MMSE	28.0 ± 1.8[b]	24.9 ± 2.8[b]	<0.0001[c]

[a]*Chi-square test.* [b]*Mean ± std.* [c]*Rank sum test.*

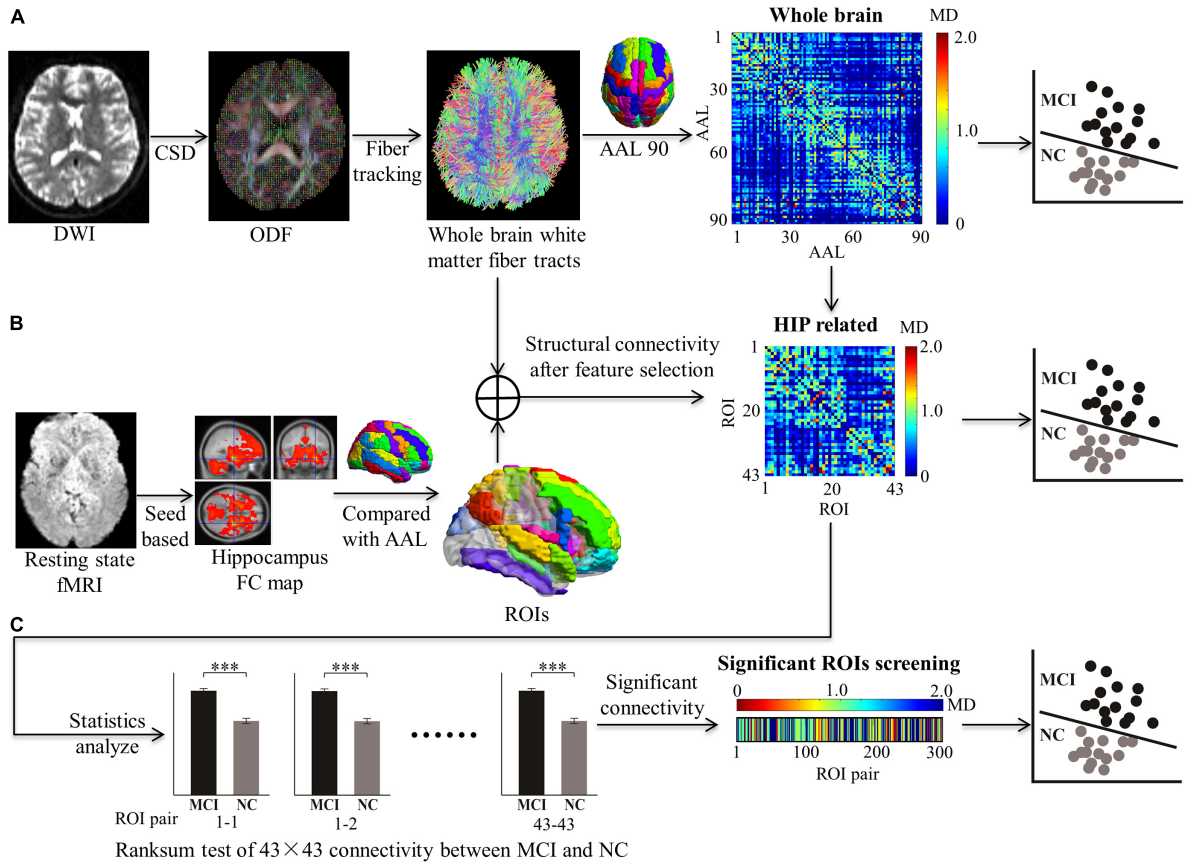


FIGURE 1 | Feature extraction in structural connectivity map driven by hippocampus related ROIs. Three kinds of features were extracted for MCI classification, including: **(A)** The whole brain WM network was acquired from diffusion MRI with all 90 × 90 AAL regions. **(B)** The HIP related WM network was selected by 43 × 43 ROIs from the whole brain WM network. **(C)** The significant HIP related WM network was acquired (*$p < 0.05$, **$p < 0.005$, and ***$p < 0.0005$; ranksum test with bonferroni correction).

ROIs are used to screen features for the white matter structural network.

5. Machine learning for MCI recognition. Different classifiers (e.g., SVM, KNN, and RF) were used to test the features for MCI recognition.
6. Performance comparison for searching optimal features and classifiers. Classification performances were validated to prove that our feature extraction method was effective. Pathological mechanisms were analyzed with feature sort.

Magnetic Resonance Imaging Data Preprocessing

The DWI data were denoised and corrected for Gibb's ringing using MRtrix3[1] and then motion-corrected, and the eddy current distortion was corrected using the eddy tool in FSL (v5.0.11).[2] Next, a brain mask was constructed using the Brain Extraction Tool in FSL, and the diffusion tensor fitting was performed at each voxel within the brain mask to generate DTI index maps. The rs-fMRI data were temporal band-pass filtered (0.01–0.10 Hz) and detrended (both linear and quadratic trends). 3D geometrical displacement was used to correct for head motion. Spatial smoothing was performed with a Gaussian filter kernel (FWHM = 6 mm). The entire process was performed using the Statistical Parametric Mapping (SPM) software package,[3] in which the head motion parameters, global signals, white matter signals, and cerebrospinal fluid signals were obtained and combined into a complete covariate. Covariate from functional signals was then removed using the Resting-State fMRI Data Analysis Toolkit (REST).[4]

Reconstruction for White Matter Fiber Tracts

To reconstruct the whole-brain white matter fiber tracts of each subject, fibers were tracked in DTI data (**Supplementary Figure 1**). (1) The ODF necessary for the fiber tracking algorithm can be obtained using the MRtrix software (see text footnote 1). The CSD was used with maximum spherical harmonic degree = 6 during this process. (2) The ODF was integrated into DSI Studio[5] to achieve a 3D reconstruction

[1]http://www.brain.org.au/software/mrtrix/

[2]https://fsl.fmrib.ox.ac.uk/fsl

[3]http://www.fil.ion.ucl.ac.uk/spm

[4]http://restfmri.net/forum/REST

[5]http://dsi-studio.labsolver.org/

of the whole brain's white matter fiber connectivity. The tracking parameters were as follows: Number of tracts = 100,000, Qa_threshold = 0.04, max angle = 60, length constraint = 20.0–450.0 mm, step size = 1.00 mm, and smoothness = 0.5 mm. Seed direction was set to random, and seed position was subvoxel.

Establishment for Whole-Brain Structure Network

Since the AAL is located in the MNI152 standard space, the transfer matrixes between the standard space and native space (**Supplementary Figure 2**) were required to obtain the AAL brain region for each subject space. (1) The b0 image from the DTI image served as the native space. The T1w image after skull stripping in FSL served as the structure space. The MNI152 template from FSL served as the standard space. (2) With the high-resolution T1w image as the reference, FMRIB's Linear Image Registration Tool (FLIRT) was used to obtain the transition matrix from native space to structural space (T_{NS}) and the transition matrix from standard space to structural space (T_{MS}). The transition matrix from native space to standard space was $T_{NM} = T_{NS} \times T_{MS}^{-1}$. (3) The transition matrix from standard space to native space was $T_{MN} = T_{MS} \times T_{NS}^{-1}$.

To establish the whole-brain WM network, a 90 × 90 matrix for the whole brain divided by the AAL template was calculated (**Supplementary Figure 1**). (1) The AAL 90 brain regions in the MNI152 standard space were transformed to each subject's space separately by applying transfer matrix T_{MN}. (2) For the 90 × 90 matrix, every element records the mean DTI index (MD or FA) of fiber tracts between every two of the AAL 90 brain regions. Also, the whole-brain WM networks were established, including the MD connectivity 90 × 90 matrix and the FA connectivity 90 × 90 matrix. (3) Whole-brain WM networks of 54 NC and 42 MCI were established separately as entire features for MCI recognition.

Feature Selection Driven by Hippocampus-Related Regions of Interest

The primary clinical symptom of MCI is memory loss, and the hippocampus is closely associated with memory. So, the hippocampus was chosen as a seed point to calculate the brain regions that had significant functional connectivity (**Supplementary Figure 3**). The concrete steps were as follows: (1) All participants' rs-fMRI images were transferred from the native space to the standard space. (2) The left hippocampus (30, −16, −14) and right hippocampus (−30, −16, −14) with a radius of 3 mm served as the seed points to create the FC map of each participant in the MNI152 standard space. With the same time series, brain regions that activated correlation with hippocampus were calculated in software REST. The Fisher Z transformation was applied so that the results of each participant followed normal distributions. (3) After individual-level analysis, group analysis for MCI and NC was conducted separately by the second-level analysis in SPM. A one-sample *T*-test and a familywise error correction were employed to revise the statistics of the group analysis results. (4) Significant area ($p < 0.05$) with the threshold (*T*-value > 10, size > 27) was selected to obtain the brain regions associated with the functional connectivity of the hippocampus at the group level.

To confirm the locations of the hippocampus-related ROIs, the following steps were conducted: (1) The functional connectivity maps of the MCI and NC groups were compared with AAL templates. The percentage of overlapping voxels taken in the AAL was calculated (**Supplementary Figure 3**). (2) To select as many ROIs as possible in the AAL and prevent false positives, a volume percentage threshold of 10% was set. If one of the 90 regions in the AAL had a volume fraction greater than the threshold value, the brain region in the AAL was selected as an ROI (**Supplementary Figure 4**). The selected ROIs are listed in **Supplementary Table 1**, including their abbreviations. (3) Totally 43 ROIs related to the hippocampus were used to extract features of the whole-brain WM network (90 × 90 matrices), and the HIP-related WM network (43 × 43 matrices) was established. The rank sum test with Bonferroni correction of HIP-related WM network between MCI and NC was used to extract significant features. The significant HIP-related WM network was acquired (300 vectors) (**Figure 1**).

Machine Learning for Mild Cognitive Impairment Recognition

To train features of different WM networks for MCI classification, the process of machine learning was as follows (**Figure 2**): (1) The whole-brain WM network, the hippocampus-related WM network (HIP-related WM network), and the significant HIP-related WM network were separately used as features for machine learning. The recursive feature elimination (RFE) algorithm was used for both dimension reduction and features ranking. (2) The machine learning

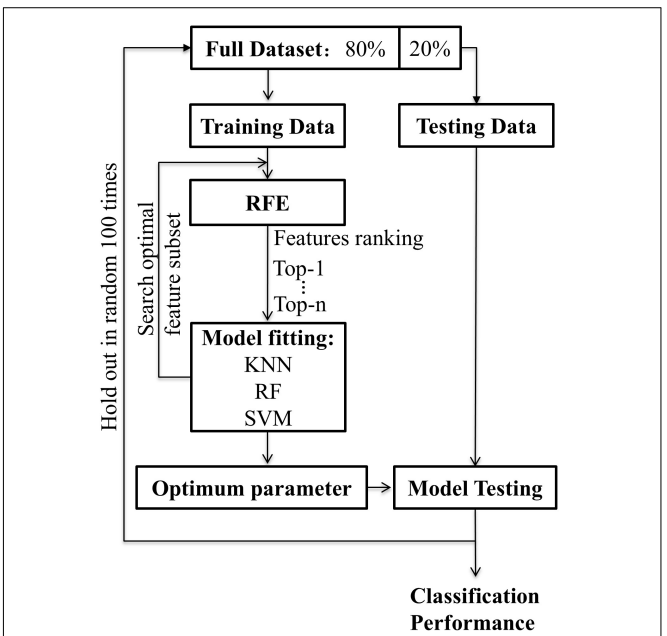


FIGURE 2 | The flowchart for machine learning. Recursive feature elimination was used to search for the optimal feature subset. The classification algorithm was respectively used KNN, RF, SVM (linear, poly, rbf, sigmoid). The hold-out method was repeated 100 times randomly with 80% of the data for training and 20% for testing.

methods included SVM, KNN, and RF. For SVM, four different kernel functions were tried, namely, linear kernel (linear), polynomial kernel (polynomial), radial basis function kernel (rbf), and sigmoid kernel (sigmoid). For KNN, neighbors of 1, 3, 5, and 7 were used. For RF, 50, 100, 150, and 200 decision trees were used. (3) The hold-out method was used to evaluate classification performance; 80% of the data set was selected as the training set and the remaining 20% as the test set. This process was repeated 100 times randomly to obtain the average classification accuracy and the area under the curve (AUC). All the above experiments were based on the Scikit-Learn library.[6]

Performance Comparison for Searching Optimal Feature and Classifier

The following steps were conducted to search for optimal features and classifiers. (1) To demonstrate the effectiveness of the feature extraction method, the classification accuracies (100 times hold-out method) were statistically analyzed in the rank sum test between three features (i.e., whole-brain WM network, HIP-related WM network, and significant HIP-related WM network), and the AUC values were compared in the identical coordinate system (**Figure 3**). (2) To obtain a better classifier for MCI recognition, the classification accuracies (100 times hold-out method) were statistically analyzed in the rank sum test between different classifiers, and the AUC values were compared in the identical coordinate system (**Figure 4**). (3) To compare which biomarker is more effective for MCI recognition, the classification accuracies (100 times hold-out method) were statistically analyzed in the rank sum test between MD and FA features, and the AUC values were compared in the identical coordinate system (**Figure 5**). (4) To rank all features' contribution to the classification, recursive feature elimination (RFE) was performed. Feature ranking order of the most contributing connectivity was selected based on their classification performance (**Figure 6A** and **Table 4**). A schematic illustration of degenerated white matter in MCI was made according to the ranking features (**Figure 6B**).

RESULTS

Participant Characteristics

There were no significant differences in demographic information between the MCI and NC. There were only significant differences in the cognitive scales (CDR and MMSE, **Table 1**). This indicates that irrelevant variables' influence was eliminated between the MCI and NC, and the results were credible.

Regions of Interest Related to Hippocampus in Resting-State Functional Magnetic Resonance Imaging

A total of 43 ROIs related to hippocampus in rs-fMRI were obtained by comparing with AAL (**Supplementary Figure 4**). Full names and abbreviations of the 43 ROIs were shown in **Supplementary Table 1**. Among them, 15 ROIs belong to the limbic lobe, 16 ROIs belong to the frontal lobe, 5 ROIs belong to the temporal lobe, 3 ROIs belong to the central region, 3 ROIs belong to the parietal lobe, and 1 ROI belongs to the occipital lobe. Most ROIs are concentrated in the limbic, frontal, and temporal lobes.

Classification Performance Promoted by Screening Features in Hippocampus-Related Regions of Interest

For the WM network of MD in the same classifier, the performances of progressively optimized features (whole-brain WM network, HIP-related WM network, and significant HIP related WM network) were significantly promoted ($p < 0.05$, rank sum test with Bonferroni correction) (**Figure 3** and **Table 2**). It proved that our feature extraction method in this experiment significantly improved the performance of the MCI classification.

Comparison of Mild Cognitive Impairment Classification Performance Under Different Algorithms

The three machine learning classifiers (of which SVM contains four kernel functions) were compared separately (**Figure 4**). The performance in order of average classification accuracy (100 times hold-out method) was as follows: SVM rbf (ACC = 89.4%, AUC = 0.954), SVM sigmoid (ACC = 88.2%, AUC = 0.950), KNN (ACC = 86.9%, AUC = 0.920), SVM linear (ACC = 86.2%, AUC = 0.937), RF (ACC = 84.8%, AUC = 0.935), and SVM ploy (ACC = 78.5%, AUC = 0.951). The performance of the SVM rbf was significantly better than other classifiers. Therefore, training with the significant HIP-related WM network as features using SVM rbf can better recognize MCI.

Comparison of Classification Performance Under Mean Diffusivity and Fractional Anisotropy Features

For the same machine learning classifier, the performance of the feature MD was significantly better than FA ($p < 0.05$, rank sum test with Bonferroni correction) in the case of both using optimal features for training (**Figure 5** and **Table 3**). Therefore, the biomarker MD is superior to FA for MCI diagnosing from a machine learning perspective.

Ranking the Contribution of White Matter Structural Connectivity to Mild Cognitive Impairment Classification

According to the significant HIP-related WM network's feature ranking order through RFE (**Figure 6A** and **Table 4**), schematic illustrations of degenerated white matter in MCI were made (**Figure 6B**). The most common connections between MCI and NC are HIP-temporal connectivity, limbic connectivity, and THA-frontal connectivity.

[6]https://scikit-learn.org/

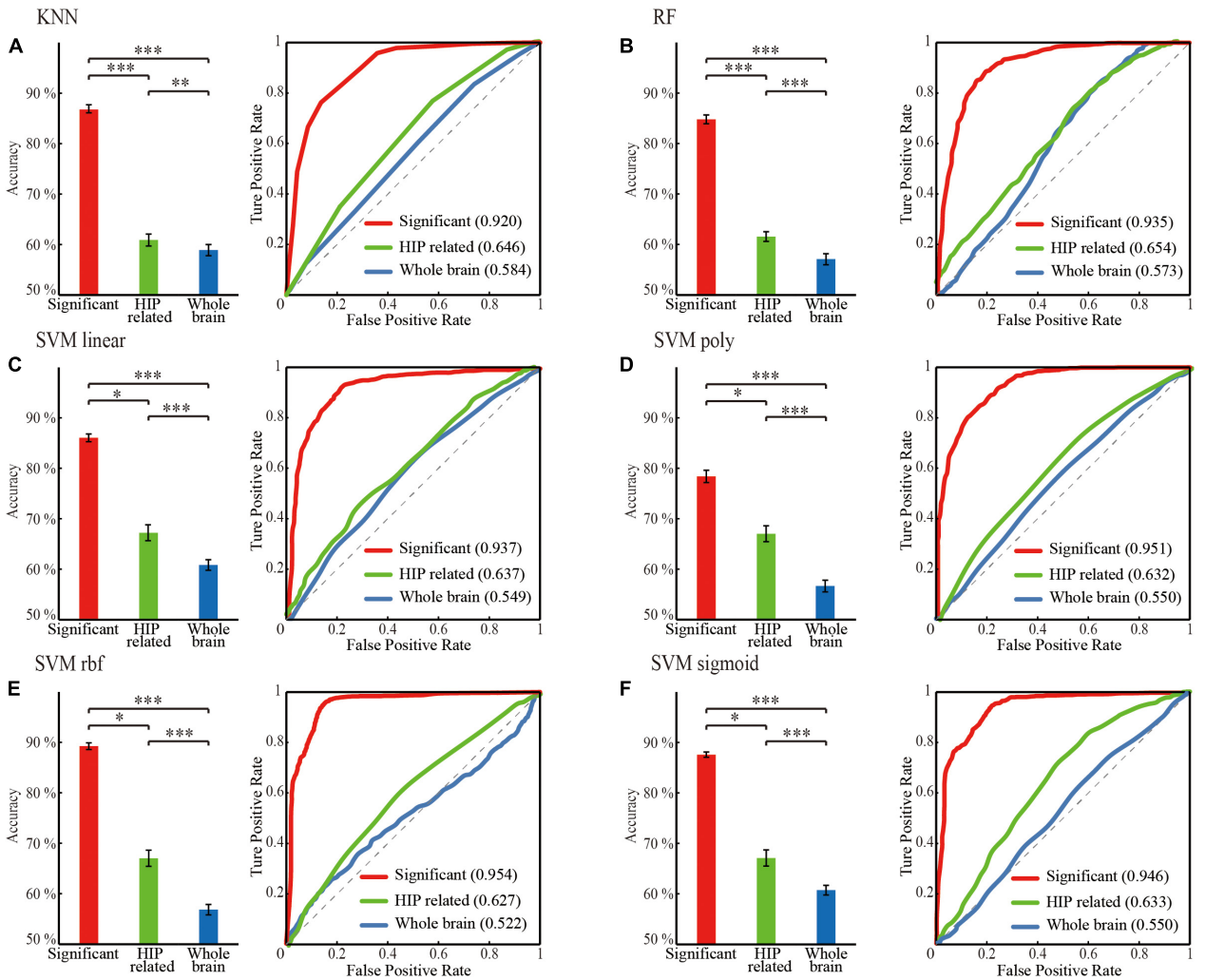


FIGURE 3 | Classification performance comparison between different mean diffusivity (MD) feature sets for different classifier. **(A)** KNN. **(B)** RF. **(C)** SVM linear. **(D)** SVM poly. **(E)** SVM rbf. **(F)** SVM sigmoid. 'red': the significant HIP related WM network; 'green': HIP related WM network; 'blue': the whole brain WM network. Classification performance including mean accuracy (mean ± std; $*p < 0.05$, $**p < 0.005$, and $***p < 0.0005$; ranksum test with bonferroni correction) and AUC of randomly 100 times hold out method.

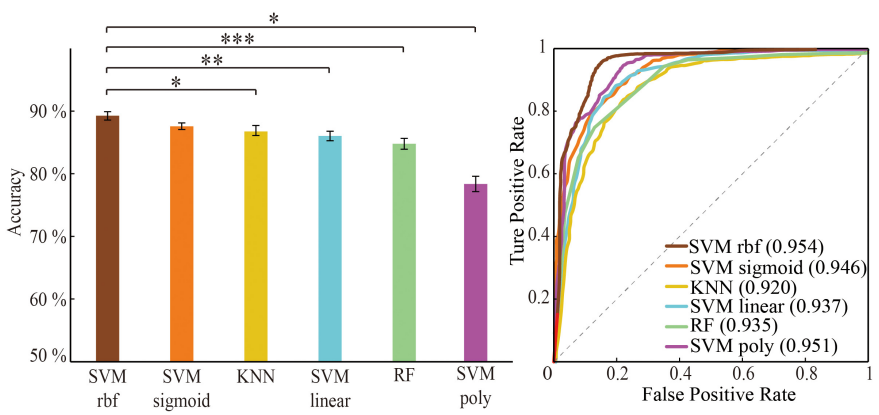


FIGURE 4 | Classification performance comparison between different algorithm through optimal feature set. Classification performance including mean accuracy (mean ± std; $*p < 0.05$, $**p < 0.005$, and $***p < 0.0005$; ranksum test with Bonferroni correction) and AUC of randomly 100 times hold-out method.

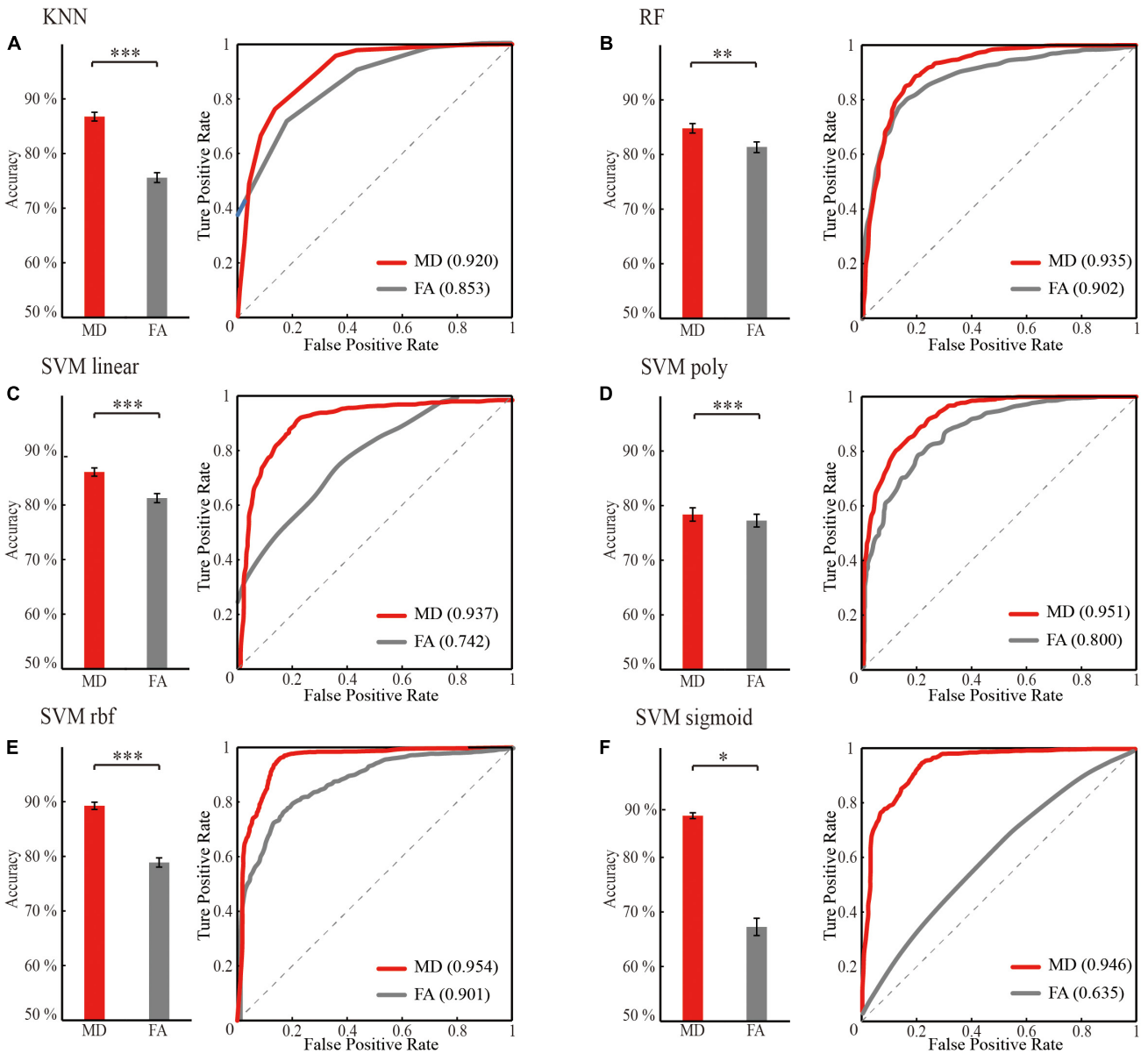


FIGURE 5 | Optimal classification performance comparison between mean diffusivity (MD) and fractional anisotropy (FA) feature sets (the significant HIP related WM network) for different classifier. **(A)** KNN. **(B)** RF. **(C)** SVM linear. **(D)** SVM poly. **(E)** SVM rbf. **(F)** SVM sigmoid. 'red': classification performance of mean diffusivity feature sets; 'gray': classification performance of fractional anisotropy feature set. Classification performance including mean accuracy (mean ± std; $^{*}p < 0.05$, $^{**}p < 0.005$, and $^{***}p < 0.0005$; rank sum test with Bonferroni correction) and AUC of randomly 100 times hold out method.

DISCUSSION

This study proposes a feature extraction method for whole-brain white matter connectivity networks driven by ROIs related to the hippocampus. The whole-brain WM network, HIP-related WM network, and significant HIP-related WM network were obtained in the process of feature extraction and optimization. Different classification algorithms, such as KNN, RF, and SVM (linear, poly, rbf, sigmoid), were used to test the classification performance. The pathological mechanisms of MCI were also revealed by RFE. Our study found that feature extraction of whole-brain white matter connectivity by hippocampus-related brain regions can significantly improve MCI classification performance. It can be summarized in three points:

1. In the feature, compared with the whole-brain WM network, the HIP-related WM network can significantly promote the MCI classification performance in machine learning.
2. In the algorithm, the classification performance of the SVM rbf is optimum when taking significant HIP-related WM networks as features.

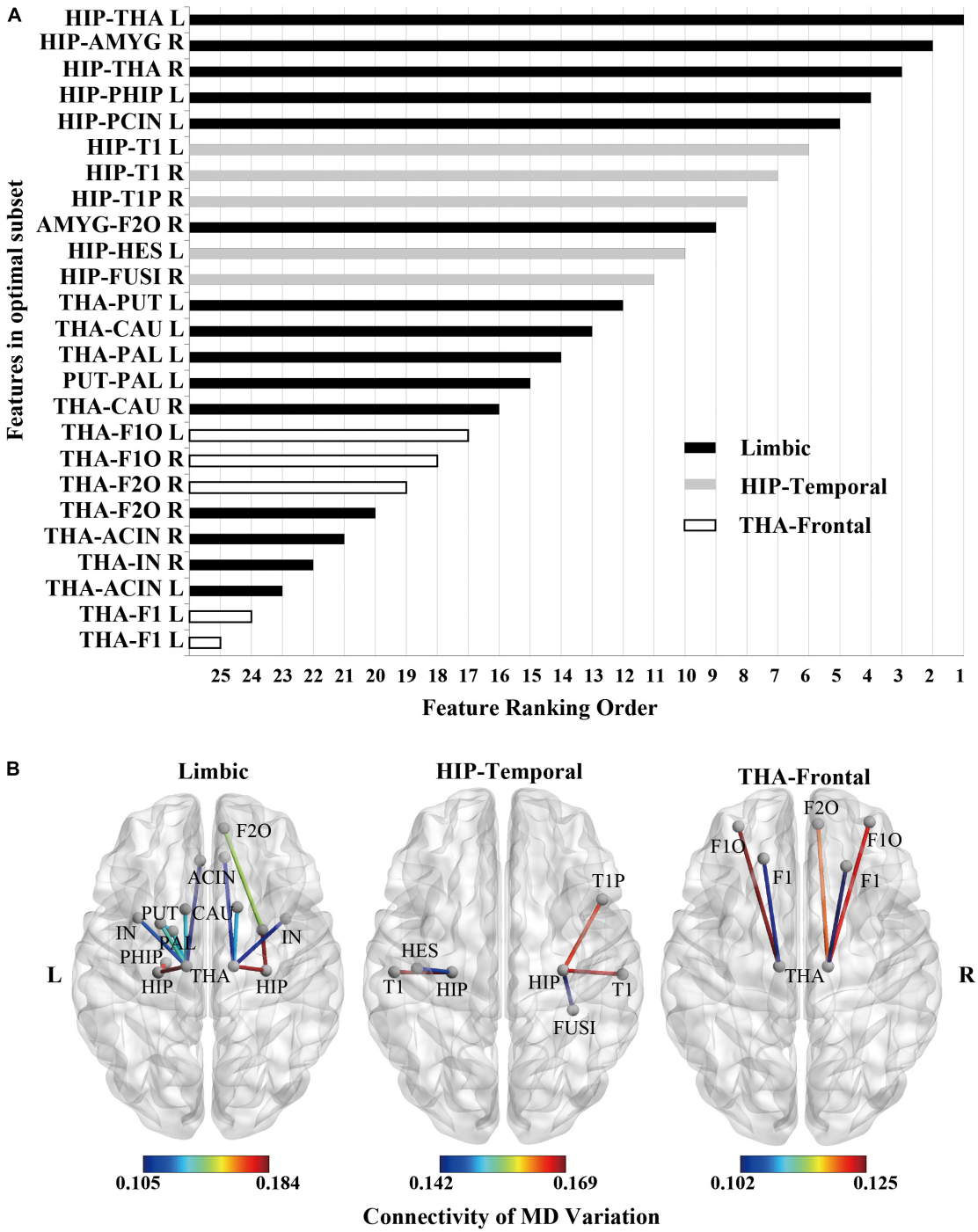


FIGURE 6 | Feature ranking of MD feature (the significant HIP-related WM network) as the result of RFE in SVM rbf model. **(A)** Feature ranking order of the most contributing connectivity (25 features). The bar represents connectivity in different groups. "black": limbic connectivity (15 features). "gray": Hip-temporal connectivity (5 features). "white": THA-frontal connectivity (5 features). **(B)** Schematic illustration of degenerated white matter in MCI. Nodes represented ROIs from AAL templates. Edges represented connectivity; the value represents MD variation, MCI vs. NC.

3. In the pathology, the hippocampus- and thalamus-related white matter connectivity greatly contributed to MCI recognition. So, the method that combines MCI pathology and uses a suitable classifier can significantly improve the classification performance, while the ranking of features contribution can reveal MCI pathology in turn.

Mild Cognitive Impairment Classification Performance Significantly Promoted by Feature Extraction of Hippocampus-Related Regions of Interest

For the same classifier on features, the performance on MCI classification was significantly higher for features screening by HIP-related ROIs than the whole brain. First, in our results, the whole-brain WM network, HIP-related WM network, and significant HIP-related WM network had sequentially significantly higher classification performance (**Figure 3**). Second, in terms of AD pathogenesis, the degeneration of MCI first appears in specific fiber tracts and gradually spreads to the whole brain when developing to the AD stage (Daianu et al., 2015; Jones et al., 2015; Wang et al., 2016). Furthermore, in either specific fiber tracking (Nir et al., 2015; Dou et al., 2020) or whole-brain white matter connectivity network measures (Wee et al., 2011; Prasad et al., 2015) in previous studies, the accuracies of MCI classification were around 60%, while the accuracies of AD can be about 80%. It also corroborates the pathogenesis of AD from an engineering perspective. Finally, for MCI recognition, a single use of whole-brain white matter would weaken the features considering AD's pathogenesis. Based on the pathology of MCI, our study improves the MCI classification performance effectively by feature extraction of HIP-related ROIs.

TABLE 2 | Classification performance based on mean diffusivity (MD) feature sets.

Classifier	Feature	ACC (mean ± sem)	Sen	Spe	AUC
A. KNN	Significant	86.90% ± 0.80%	0.938	0.778	0.920
	HIP related	60.90% ± 1.20%	0.97	0.142	0.646
	Whole brain	58.90% ± 1.10%	0.837	0.265	0.584
B. RF	Significant	84.80% ± 0.90%	0.916	0.781	0.935
	HIP related	61.60% ± 1.00%	0.958	0.144	0.654
	Whole brain	57.10% ± 1.10%	0.935	0.117	0.573
C. SVM linear	Significant	86.20% ± 0.80%	0.905	0.808	0.937
	HIP related	67.30% ± 1.60%	0.886	0.419	0.637
	Whole brain	60.90% ± 1.00%	0.791	0.406	0.549
D. SVM poly	Significant	78.50% ± 1.20%	0.965	0.609	0.951
	HIP related	67.10% ± 1.60%	0.886	0.414	0.632
	Whole brain	56.70% ± 1.20%	0.972	0.029	0.550
E. SVM rbf	Significant	89.40% ± 0.70%	0.938	0.849	0.954
	HIP related	67.10% ± 1.60%	0.887	0.413	0.627
	Whole brain	56.90% ± 1.10%	0.991	0.011	0.522
F. SVM sigmoid	Significant	88.20% ± 0.60%	0.945	0.806	0.946
	HIP related	67.20% ± 1.60%	0.888	0.414	0.633
	Whole brain	60.80% ± 1.00%	0.708	0.501	0.550

ACC, accuracy; Sen, sensitivity; Spe, specificity; AUC, area under the curve; Ave, average; Std, standard error; Sem, standard error mean.

TABLE 3 | Classification performance of MD and FA feature sets.

Classifier	Feature	ACC (mean ± sem)	Sen	Spe	AUC
A.KNN	MD	86.90% ± 0.80%	0.938	0.778	0.92
	FA	75.70% ± 0.90%	0.905	0.571	0.853
B. RF	MD	84.80% ± 0.90%	0.916	0.781	0.935
	FA	81.40% ± 1.00%	0.878	0.743	0.902
C. SVM linear	MD	86.20% ± 0.80%	0.905	0.808	0.937
	FA	81.40% ± 0.80%	0.891	0.728	0.742
D. SVM poly	MD	78.50% ± 1.20%	0.965	0.609	0.951
	FA	77.50% ± 1.20%	0.877	0.672	0.8
E. SVM rbf	MD	89.40% ± 0.70%	0.938	0.849	0.954
	FA	79.00% ± 0.80%	0.841	0.738	0.901
F. SVM sigmoid	MD	88.20% ± 0.60%	0.945	0.806	0.946
	FA	67.30% ± 1.60%	0.886	0.418	0.635

ACC, accuracy; Sen, sensitivity; Spe, specificity; AUC, area under the curve; Ave, average; Std, standard error; Sem, standard error mean.

Mean Diffusivity Can Be a Valid Biomarker for Mild Cognitive Impairment Recognition

For MCI, the MD is more sensitive than FA values to reflect white matter degeneration. First, our results showed that the MD of WM networks as features outperformed FA in all the machine learning classifiers (**Figure 5**). Second, research showed the sensitivity of the MD index to MCI (Yu et al., 2017), while

TABLE 4 | Feature ranking order for WM connectivity.

Feature ranking order	ROI pairs for connectivity	
1 HIP-THA L	Hippocampus	Thalamus
2 HIP-AMYG R	Hippocampus	Amygdala
3 HIP-THA R	Hippocampus	Thalamus
4 HIP-PHIP L	Hippocampus	Parahippocampal gyrus
5 HIP-PCIN L	Hippocampus	Posterior cingulate
6 HIP-T1 L	Hippocampus	Superior temporal gyrus
7 HIP-T1 R	Hippocampus	Superior temporal gyrus
8 HIP-T1P R	Hippocampus	Temporal pole: superior temporal gyrus
9 AMYG-F2O R	Amygdala	Middle frontal gyrus, orbital part
10 HIP-HES L	Hippocampus	Heschl gyrus
11 HIP-FUSI R	Hippocampus	Fusiform gyrus
12 THA-PUT L	Thalamus	Lenticular nucleus, putamen
13 THA-CAU L	Thalamus	Caudate nucleus
14 THA-PAL L	Thalamus	Lenticular nucleus, pallidum
15 PUT-PAL L	Lenticular nucleus, putamen	Lenticular nucleus, pallidum
16 THA-CAU R	Thalamus	Caudate nucleus
17 THA-F1O L	Thalamus	Superior frontal gyrus, orbital part
18 THA-F1O R	Thalamus	Superior frontal gyrus, orbital part
19 THA-F2O R	Thalamus	Middle frontal gyrus, orbital part
20 THA-IN L	Thalamus	Insula
21 THA-ACIN R	Thalamus	Anterior cingulate and paracingulate gyri
22 THA-IN R	Thalamus	Insula
23 THA-ACIN L	Thalamus	Anterior cingulate and paracingulate gyri
24 THA-F1 L	Thalamus	Superior frontal gyrus, dorsolateral
25 THA-F1 R	Thalamus	Superior frontal gyrus, dorsolateral

FA only begins to have statistical differences in multiple fiber tracts during AD (Mito et al., 2018). Our experiment confirmed that the MD index was more sensitive in the MCI stage from classification performance. Furthermore, FA reflects the density of white matter fibers and is influenced by the axon diameter of fiber tracts. MD represents the integrity of the myelin outer the fiber tracts. Demyelination can cause MD to increase (Tournier, 2019). It can be inferred that disintegrated myelin appeared more obvious than axon loss during MCI. Therefore, the MD index can be an effective biomarker for MCI compared with FA.

Comparison With Previous Studies

Compared with previous studies that used white matter as the feature to recognize MCI, our method using HIP-related ROIs to extract white matter features significantly improved MCI classification performance (**Table 5**). By comparing the classification performance between all of our classifiers, the SVM classifier with rbf kernel significantly outperformed the other classifiers (**Figure 4**). It can also be seen from previous studies that the SVM classifier with rbf kernel for MCI recognition performed well when taking the white matter as features (**Table 5**). Therefore, training with the white matter as features after HIP-related ROIs extraction can better recognize MCI when using the SVM rbf classifier.

Important Hubs of the White Matter Connectivity Network for Mild Cognitive Impairment: Hippocampus and Thalamus

White matter connections related to hippocampal and thalamic contribute most to MCI classification. First, previous studies have shown that the white matter associated with the hippocampus and thalamus degenerates during MCI. The white matter microstructure between the hippocampus and medial temporal lobe changes during MCI (Teipel et al., 2016; Zhuo et al., 2016; Dumont et al., 2019). The fornix connects the hippocampus to the thalamus, and its white matter degeneration leads to decreased memory function (Christiansen et al., 2016). Degeneration of the projection tracts connecting the thalamus to the prefrontal lobes leads to a decrease in executive function (Gu and Zhang, 2019; Liu et al., 2021).

Second, in our results, features connected with the hippocampus or thalamus in the white matter connectivity network ranked high in contributions for MCI classification, and features connected with the hippocampus ranked more advanced than the thalamus (**Figure 6**). This result also confirmed our previous findings (Zhou et al., 2021). In our previous study, the MD of all voxels in a single fiber tract was taken as features. The fiber tract with high separability for MCI recognition passed through the hippocampus or thalamus. So, whether single fiber tract or whole-brain network had corroborated that hippocampus and thalamus were important hubs of the white matter connectivity network for MCI.

Furthermore, in our previous study, the highest average accuracy for MCI classification reached 71.0% when taking MD of all voxels in the left inferior longitudinal fasciculus. In this study, the highest average accuracy reached 89.4% by taking white matter networks related to the hippocampus. Compared with the single fiber tract, white matter networks related to the hippocampus as features can improve classification performance. Finally, the contribution of every white matter connectivity can be sorted by RFE. Thus, the hippocampus and thalamus are important hubs for MCI. Features of combinational white matter connectivity networks outperform single fiber tracts.

Limitations

This study mainly focused on feature extraction for MCI recognition. However, the valuable contributions of this study must be considered within the context of certain limitations. First, a certain amount of new data will be added as the test set alone. Data in this article will be used as the training set and

TABLE 5 | Summary of the studies using dMRI features for MCI classification.

Comparison with the previous studies	Classifier	Subjects	Feature	Database	Performance	
		MCI/NC			ACC	AUC
Ebadi et al. (2017)	KNN	15/15	Network	Local	60.0%	0.560
Our study	KNN	42/54	Network	Local	86.9%	0.920
Maggipinto et al. (2017)	RF	90/89	MD/FA voxel	ADNI	54.0%	0.600
Wang et al. (2018)	RF	169/379	Network	ADNI/NACC	75.0%	0.850
Our study	RF	42/54	Network	Local	84.8%	0.935
Xie et al. (2015)	SVM linear	64/64	MD/FA voxel	Local	78.9%	0.856
Our study	SVM linear	42/54	Network	Local	86.2%	0.937
Cui et al. (2012)	SVM rbf	79/204	FA voxel	SMA	71.1%	0.700
Dyrba et al. (2015a)	SVM rbf	35/42	MD/FA voxel	EDSD	77.0%	0.680
Demirhan et al. (2015)	SVM rbf	43/70	FA voxel	ADNI	78.5%	0.758
Nir et al. (2015)	SVM rbf	113/50	Network	ADNI	79.0%	-
Ahmed et al. (2017)	SVM rbf	58/52	MD voxel	ADNI	79.4%	0.788
Our study	SVM rbf	42/54	Network	Local	89.4%	0.954

local, collect by hospital; ADNI, Alzheimer's Disease Neuroimaging Initiative; NACC, National Alzheimer's Coordinating Center; SMA, Sydney Memory and Aging; EDSD, European DTI Study on Dementia; -, not applicable.

validation set. Second, cross-validation will be added to obtain more accurate parameters for the classifier further to improve performance in the future. Third, deep learning will be used based on our existing findings, and generalization could be improved in the future. More subjects and multiple datasets will be acquired from different hospitals to test the generalization of the classifier. Furthermore, future studies should recruit participants with both MCI and AD. The NC, MCI, and AD should be divided to investigate the pathological mechanisms underlying AD development. Finally, the behavior scale would be added, and the support vector regression (SVR) will be used to predict MCI patients' behavior.

CONCLUSION

Our study proposes a feature extraction method driven by hippocampus-related ROIs for white matter connectivity networks. In the feature extraction process, the whole-brain WM network, the HIP-related network, and the significant HIP-related network continuously optimize the performance of MCI classification. By recursive feature elimination, the pathological mechanism revealed that the hippocampus and thalamus are important hubs in white matter networks for MCI. Our results provide a valid basis for the early diagnosis of AD.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. The main codes that support the findings of this study are available on GitHub (https://github.com/ThreePoundUniverse/MCI-network-classification). Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board of Tianjin University; Ethics Committee of Chang Gung University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XS and YZ designed and conceptualized the research and wrote the manuscript. YZ, Y-PC, and C-PL collected the data. YZ, YC, SL, XZ, YS, and XS analyzed the data. XS, DM, and QL supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (grant nos. 61906132 to XS, to DM, 61471263 to QL), Key Project and Team Program of Tianjin City (grant no. XC202020 to XS), Natural Science Foundation of Tianjin City (grant no. 16JCZDJC31100 to QL), Chang Gung University (grant no. BMRPC78 to Y-PC), and Tianjin University (grant no. 2020XRY-0015 to XS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.866230/full#supplementary-material

REFERENCES

Ahmed, O. B., Benois-Pineau, J., Allard, M., Catheline, G., and Amar, C. B. (2017). Recognition of Alzheimer's disease and Mild Cognitive Impairment with multimodal image-derived biomarkers and Multiple Kernel Learning. *Neurocomputing* 220, 98–110. doi: 10.1016/j.neucom.2016.08.041

Benoit, R. G., and Schacter, D. L. (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia* 75, 450–457. doi: 10.1016/j.neuropsychologia.2015.06.034

Buckner, R. L., and DiNicola, L. M. (2019). The brain's default network: updated anatomy, physiology and evolving insights. *Nat. Rev. Neurosci.* 20, 593–608. doi: 10.1038/s41583-019-0212-7

Christiansen, K., Aggleton, J. P., Parker, G. D., O'Sullivan, M. J., Vann, S. D., and Metzler-Baddeley, C. (2016). The status of the precommissural and postcommissural fornix in normal ageing and mild cognitive impairment: an MRI tractography study. *Neuroimage* 130, 35–47. doi: 10.1016/j.neuroimage.2015.12.055

Cui, Y., Wen, W., Lipnicki, D. M., Beg, M. F., Jin, J. S., Luo, S., et al. (2012). Automated detection of amnestic mild cognitive impairment in community-dwelling elderly adults: a combined spatial atrophy and white matter alteration approach. *Neuroimage* 59, 1209–1217. doi: 10.1016/j.neuroimage.2011.08.013

Daianu, M., Jahanshad, N., Nir, T. M., Jack, C. R. Jr., Weiner, M. W., Bernstein, M. A., et al. (2015). Rich club analysis in the Alzheimer's disease connectome reveals a relatively undisturbed structural core network. *Hum. Brain Mapp.* 36, 3087–3103. doi: 10.1002/hbm.22830

Demirhan, A., Nir, T. M., Zavaliangos-Petropulu, A., Jack, C. R., Weiner, M. W., Bernstein, M. A., et al. (2015). "Feature selection improves the accuracy of classifying Alzheimer disease using diffusion tensor images," in *Proceedings of the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, New York, NY, 126–130. doi: 10.1109/ISBI.2015.7163832

Dou, X., Yao, H., Feng, F., Wang, P., Zhou, B., Jin, D., et al. (2020). Characterizing white matter connectivity in Alzheimer's disease and mild cognitive impairment: an automated fiber quantification analysis with two independent datasets. *Cortex* 129, 390–405. doi: 10.1016/j.cortex.2020.03.032

Dumont, M., Roy, M., Jodoin, P. M., Morency, F. C., Houde, J. C., Xie, Z., et al. (2019). Free Water in White Matter Differentiates MCI and AD From Control Subjects. *Front. Aging Neurosci.* 11:270. doi: 10.3389/fnagi.2019.00270

Dyrba, M., Barkhof, F., Fellgiebel, A., Filippi, M., Hausner, L., Hauenstein, K., et al. (2015a). Predicting Prodromal Alzheimer's disease in subjects with mild cognitive impairment using machine learning classification of multimodal multicenter diffusion-tensor and magnetic resonance imaging data. *J. Neuroimaging* 25, 738–747. doi: 10.1111/jon.12214

Dyrba, M., Grothe, M., Kirste, T., and Teipel, S. J. (2015b). Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM. *Hum. Brain Mapp.* 36, 2118–2131. doi: 10.1002/hbm.22759

Dyrba, M., Ewers, M., Wegrzyn, M., Kilimann, I., Plant, C., Oswald, A., et al. (2013). Robust automated detection of microstructural white matter degeneration in Alzheimer's disease using machine learning classification of multicenter DTI Data. *PLoS One* 8:e64925. doi: 10.1371/journal.pone.0064925

Ebadi, A., Dalboni da Rocha, J. L., Nagaraju, D. B., Tovar-Moll, F., Bramati, I., Coutinho, G., et al. (2017). Ensemble Classification of Alzheimer's disease and mild cognitive impairment based on complex graph measures from diffusion tensor images. *Front. Neurosci.* 11:56. doi: 10.3389/fnins.2017.00056

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., et al. (2006). Mild cognitive impairment. *Lancet* 367, 1262–1270.

Gu, L., and Zhang, Z. (2019). Exploring structural and functional brain changes in mild cognitive impairment: a whole brain ALE meta-analysis for multimodal MRI. *ACS Chem. Neurosci.* 10, 2823–2829. doi: 10.1021/acschemneuro.9b00045

Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q. K., Yu, A. B., Kahn, A. E., et al. (2015). Controllability of structural brain networks. *Nat. Commun.* 6:8414.

Hinrichs, C., Singh, V., Xu, G. F., Johnson, S. C., and Neuroimaging, A. D. (2011). Predictive markers for AD in a multi-modality framework: An analysis of MCI progression in the ADNI population. *Neuroimage* 55, 574–589. doi: 10.1016/j.neuroimage.2010.10.081

Hyman, B. T., Phelps, C. H., Beach, T. G., Bigio, E. H., Cairns, N. J., Carrillo, M. C., et al. (2012). National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement.* 123, 1–11.

Jacobs, H. I., van Boxtel, M. P., Gronenschild, E. H., Uylings, H. B., Jolles, J., and Verhey, F. R. (2013). Decreased gray matter diffusivity: A potential early Alzheimer's disease biomarker? *Alzheimers Dement.* 9, 93–97. doi: 10.1016/j.jalz.2011.11.004

Jones, D. T., Knopman, D. S., Gunter, J. L., Graff-Radford, J., Vemuri, P., Boeve, B. F., et al. (2015). Cascading network failure across the Alzheimer's disease spectrum. *Brain* 139, 547–562. doi: 10.1093/brain/awv338

Kleinschmidt, A., and Vuilleumier, P. (2013). Disconnecting cognition. *Curr. Opin. Neurol.* 26, 333–338. doi: 10.1097/wco.0b013e328363393b

Lee, S. H., Coutu, J. P., Wilkens, P., Yendiki, A., Rosas, H. D., and Salat, D. H. (2015). TRACT-BASED ANALYSIS OF WHITE MATTER DEGENERATION IN ALZHEIMER'S DISEASE. *Neuroscience* 301, 79–89. doi: 10.1016/j.neuroscience.2015.05.049

Liu, F. Y., Zhou, L. P., Shen, C. H., and Yin, J. P. (2014). Multiple Kernel learning in the primal for multimodal Alzheimer's disease classification. *IEEE J. Biomed. Health* 18, 984–990. doi: 10.1109/JBHI.2013.2285378

Liu, L., Jiang, H., Wang, D., and Zhao, X. F. (2021). A study of regional homogeneity of resting-state Functional Magnetic Resonance Imaging in mild cognitive impairment. *Behav. Brain Res.* 402, 113103. doi: 10.1016/j.bbr.2020.113103

Maggipinto, T., Bellotti, R., Amoroso, N., Diacono, D., Donvito, G., Lella, E., et al. (2017). DTI measurements for Alzheimer's classification. *Phys. Med. Biol.* 62, 2361–2375. doi: 10.1088/1361-6560/aa5dbe

Mito, R., Raffelt, D., Dhollander, T., Vaughan, D. N., Tournier, J., Salvado, O., et al. (2018). Fibre-specific white matter reductions in Alzheimer's disease and mild cognitive impairment. *Brain* 141, 888–902. doi: 10.1093/brain/awx355

Nir, T. M., Villalon-Reina, J. E., Prasad, G., Jahanshad, N., Joshi, S. H., Toga, A. W., et al. (2015). Diffusion weighted imaging-based maximum density path analysis and classification of Alzheimer's disease. *Neurobiol. Aging* 36, S132–S140. doi: 10.1016/j.neurobiolaging.2014.05.037

Pellegrini, E., Ballerini, L., Hernandez, M., Chappell, F. M., Gonzalez-Castro, V., Anblagan, D., et al. (2018). Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review. *Alzheimers Dement.* 10, 519–535. doi: 10.1016/j.dadm.2018.07.004

Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., et al. (2001). Current concepts in mild cognitive impairment. *Arch. Neurol.* 58, 1985–1992.

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., and Kokmen, E. (1999). Mild cognitive impairment - Clinical characterization and outcome. *Arch. Neurol.* 56, 303–308. doi: 10.1001/archneur.56.3.303

Prasad, G., Joshi, S. H., Nir, T. M., Toga, A. W., and Thompson, P. M. (2015). Brain connectivity and novel network measures for Alzheimer's disease classification. *Neurobiol. Aging* 36, S121–S131. doi: 10.1016/j.neurobiolaging.2014.04.037

Rathore, S., Habes, M., Iftikhar, M. A., Shacklett, A., and Davatzikos, C. (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *Neuroimage* 155, 530–548. doi: 10.1016/j.neuroimage.2017.03.057

Rugg, M. D., and Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Curr. Opin. Neurobiol.* 23, 255–260. doi: 10.1016/j.conb.2012.11.005

Shatte, A. B. R., Hutchinson, D. M., and Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. *Psychol. Med.* 49, 1426–1448. doi: 10.1017/S0033291719000151

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., et al. (2011). Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 280–292. doi: 10.1016/j.jalz.2011.03.003

Sporns, O., and Betzel, R. F. (2016). Modular brain networks. *Annu. Rev. Psychol.* 67, 613–640. doi: 10.1146/annurev-psych-122414-033634

Teipel, S., Grothe, M. J., and Alzheimer 's Disease Neuroimaging Initiative (2016). Does posterior cingulate hypometabolism result from disconnection or local pathology across preclinical and clinical stages of Alzheimer's disease? *Eur. J. Nucl. Med. Mol. Imaging* 43, 526–536. doi: 10.1007/s00259-015-3222-3

Tournier, J. D. (2019). Diffusion MRI in the brain - Theory and concepts. *Prog. Nucl. Magn. Reson. Spectrosc.* 112-113, 1–16. doi: 10.1016/j.pnmrs.2019.03.001

Wang, J., Zuo, X., Dai, Z., Xia, M., Zhao, Z., Zhao, X., et al. (2013). Disrupted functional brain connectome in individuals at risk for Alzheimer's disease. *Biol. Psychiatry* 73, 472–481. doi: 10.1016/j.biopsych.2012.03.026

Wang, Q., Guo, L., Thompson, P. M., Jack, C. R. Jr., Dodge, H., Zhan, L., et al. (2018). The added value of diffusion-weighted mri-derived structural connectome in evaluating mild cognitive impairment: a multi-cohort validation. *J. Alzheimers Dis.* 64, 149–169. doi: 10.3233/JAD-171048

Wang, T., Shi, F., Jin, Y., Yap, P.-T., Wee, C.-Y., Zhang, J., et al. (2016). Multilevel deficiency of white matter connectivity networks in Alzheimer's disease: a diffusion MRI study with DTI and HARDI models. *Neural Plast.* 2016:2947136. doi: 10.1155/2016/2947136

Wee, C. Y., Yap, P. T., Li, W. B., Denny, K., Browndyke, J. N., Potter, G. G., et al. (2011). Enriched white matter connectivity networks for accurate identification of MCI patients. *Neuroimage* 54, 1812–1822. doi: 10.1016/j.neuroimage.2010.10.026

Xie, Y., Cui, Z., Zhang, Z., Sun, Y., Sheng, C., Li, K., et al. (2015). Identification of amnestic mild cognitive impairment using multi-modal brain features: a combined structural MRI and Diffusion Tensor Imaging Study. *J. Alzheimers Dis.* 47, 509–522. doi: 10.3233/JAD-150184

Yu, J., Lam, C. L. M., and Lee, T. M. C. (2017). White matter microstructural abnormalities in amnestic mild cognitive impairment: a meta-analysis of whole-brain and ROI-based studies. *Neurosci. Biobehav. Rev.* 83, 405–416. doi: 10.1016/j.neubiorev.2017.10.026

Yu, M., Sporns, O., and Saykin, A. J. (2021). The human connectome in Alzheimer disease - relationship to biomarkers and genetics. *Nat. Rev. Neurol.* 17, 545–563. doi: 10.1038/s41582-021-00529-1

Zhang, D. Q., Shen, D. G., and Alzheimer's Disease Neuroimaging Initiative (2012). Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease (vol 59, pg 895, 2012). *Neuroimage* 62, 2179–2179. doi: 10.1016/j.neuroimage.2011.09.069

Zhang, D. Q., Wang, Y. P., Zhou, L. P., Yuan, H., and Shen, D. G. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage* 55, 856–867. doi: 10.1016/j.neuroimage.2011.01.008

Zhang, Y. P., Wang, S. H., Xia, K. J., Jiang, Y. Z., and Qian, P. J. (2021). Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion. *Inf. Fusion* 66, 170–183. doi: 10.1016/j.inffus.2020.09.002

Zhao, B., Li, T., Yang, Y., Wang, X., Luo, T., Shan, Y., et al. (2021). Common genetic variation influencing human white matter microstructure. *Science* 372:eabf3736. doi: 10.1126/science.abf3736

Zhou, Y., Si, X., Chen, Y., Chao, Y., Lin, C. P., Li, S., et al. (2021). Hippocampus- and thalamus-related fiber-specific white matter reductions in mild cognitive impairment. *Cereb. Cortex*. [Epub ahead of print]. doi: 10.1093/cercor/bhab407

Zhuang, L., Sachdev, P. S., Trollor, J. N., Kochan, N. A., Reppermund, S., Brodaty, H., et al. (2012). Microstructural white matter changes in cognitively normal individuals at risk of amnestic MCI. *Neurology* 79, 748–754. doi: 10.1212/WNL.0b013e3182661f4d

Zhuo, J. J., Fan, L. Z., Liu, Y., Zhang, Y. C., Yu, C. S., and Jiang, T. Z. (2016). Connectivity profiles reveal a transition subarea in the parahippocampal region that integrates the anterior temporal-posterior medial systems. *J. Neurosci.* 36, 2782–2795. doi: 10.1523/JNEUROSCI.1975-15.2016

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 30 June 2022
doi: 10.3389/fnagi.2022.892178

Research on Rehabilitation Training Strategies Using Multimodal Virtual Scene Stimulation

*Ping Xie[1], Zihao Wang[1], Zengyong Li[2], Ying Wang[1], Nianwen Wang[1], Zhenhu Liang[1], Juan Wang[1] and Xiaoling Chen[1]**

[1] Key Laboratory of Measurement Technology and Instrumentation of Hebei Province, Institute of Electric Engineering, Yanshan University, Qinhuangdao, China, [2] National Research Center for Rehabilitation Technical Aids, Beijing, China

It is difficult for stroke patients with flaccid paralysis to receive passive rehabilitation training. Therefore, virtual rehabilitation technology that integrates the motor imagery brain-computer interface and virtual reality technology has been applied to the field of stroke rehabilitation and has evolved into a physical rehabilitation training method. This virtual rehabilitation technology can enhance the initiative and adaptability of patient rehabilitation. To maximize the deep activation of the subjects motor nerves and accelerate the remodeling mechanism of motor nerve function, this study designed a brain-computer interface rehabilitation training strategy using different virtual scenes, including static scenes, dynamic scenes, and VR scenes. Including static scenes, dynamic scenes, and VR scenes. We compared and analyzed the degree of neural activation and the recognition rate of motor imagery in stroke patients after motor imagery training using stimulation of different virtual scenes, The results show that under the three scenarios, The order of degree of neural activation and the recognition rate of motor imagery from high to low is: VR scenes, dynamic scenes, static scenes. This paper provided the research basis for a virtual rehabilitation strategy that could integrate the motor imagery brain-computer interface and virtual reality technology.

Keywords: brain-computer interface, motor imagery, virtual reality, neural activation, virtual rehabilitation

OPEN ACCESS

Edited by:
Peng Xu,
University of Electronic Science and Technology of China, China

Reviewed by:
Yunfa Fu,
Kunming University of Science and Technology, China
Banghua Yang,
Shanghai University, China

****Correspondence:***
Xiaoling Chen
xlchen@ysu.edu.cn

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *08 March 2022*
Accepted: *21 April 2022*
Published: *30 June 2022*

Citation:
Xie P, Wang Z, Li Z, Wang Y, Wang N, Liang Z, Wang J and Chen X (2022) Research on Rehabilitation Training Strategies Using Multimodal Virtual Scene Stimulation. Front. Aging Neurosci. 14:892178. doi: 10.3389/fnagi.2022.892178

INTRODUCTION

The brain-computer interface (BCI) is a direct communication and control channel established between the human brain and computer or other electronic devices (Wolpaw et al., 2002). Through this channel, people can express ideas or manipulate equipment directly through the brain without any language or action, which can effectively enhance the ability of patients with severe physical disabilities to communicate with the outside world or to control the external environment to improve the quality of life of patients (Ren et al., 2004). Virtual reality (VR) technology creates a virtual world through the use of a new computer system, allowing the experiencer to integrate into the virtual environment and achieve mutual interaction. In addition to being used in the gaming industry, VR also can help stroke patients with flaccid paralysis perform purposeful training in a virtual environment, thereby improving the effect of rehabilitation training. Motor imagery (MI) is defined as the cognitive activity in which a subject imagines a movement without

actually performing the movement (Vries and Mulder, 2007), and it is a common application paradigm in the field of brain-computer interface research (Mattia et al., 2012). This method realizes communication and control with external devices by imaging body movements (Xu et al., 2013).

Motor imagery brain-computer interface (MI-BCI) is a type of BCI that recognizes the patient's motor imagery intention by guiding the patient to perform motor imagery based on motor imagery therapy (Clark et al., 2019). It can be used to effectively remodel the central nervous system in patients with motor dysfunction (Guger et al., 2017). VR technology can provide patients with a more immersive training environment (Ren et al., 2020), help patients perform motor imagery more accurately (Bayliss and Ballard, 2000), and generate more easily identifiable electroencephalography (EEG) signals (Li et al., 2018). Remsik et al. (2016) reviewed 17 independent MI-BCI stroke rehabilitation studies, and they found that 16 produced significant treatment effects. Studies have shown that both motor imagery and actual movement can activate bilateral premotor areas (Pfurtscheller and Neuper, 2001), parietal lobes, basal ganglia, and cerebellum. Studies (Pfurtscheller and Neuper, 1997) also have shown that stroke patients can perform motor imagery to partially activate the damaged motor network (Fumanal-Idocin et al., 2021). In addition, studies have found that the rehabilitation treatment model combining BCI and VR technology is suitable for people suffering from stroke (Varsehi and Firoozabadi, 2021), depression, addiction and other diseases (Takenaka et al., 2021). VR-guided action offers the advantages of intuitive and specific actions and a strong sense of substitution (Velasquez-Martinez et al., 2020). Well-designed scene feedback can produce neural activation (Vourvopoulos et al., 2015).

Rehabilitation training strategies based on VR and MI-BCI have the following limitations (Benitez-Andonegui et al., 2020). First, the current rehabilitation strategy based on MI-BCI improves accuracy primarily by improving the algorithm without using a scene stimulation to improve the neural activation of the subject (Xiao and Fang, 2021), and thereby improving the quality of the EEG signals to improve accuracy (Bagarinao et al., 2020). Second, the virtual rehabilitation scene is singular (Vidaurre et al., 2020), the individual adaptability is poor, and few studies (Li et al., 2021) have compared the neural region activation and enhanced EEG signals feature mechanisms in different scenes (Barsotti et al., 2015). Third, most of the existing virtual rehabilitation training strategies are performed by observing virtual scenes on a computer screen (Moctezuma and Molinas, 2020). This training mode is not only less immersive, but also easily disturbed by the external environment, which introduces difficulties to the rehabilitation training (Gomez-Pilar et al., 2016). Finally, visual feedback training is lacking during rehabilitation. At the neural mechanism level, visual feedback training promotes brain plasticity changes and functional reorganization through the activation of the mirror neuron system, thereby promoting the recovery of motor function (Birbaumer et al., 2006).

In this study, to examine ways to improve the deep activation mechanism of motor nerves in stroke patients with flaccid paralysis, we designed different virtual scenes and training tasks to stimulate motor imagery from different angles. The virtual scenes include the following: limb MI in static scenes, such as text and pictures; limb MI in dynamic scenes of three-dimensional (3D) life and virtual games; and limb MI in the VR environment. This study compared and analyzed energy changes in the motor areas of the brain and the recognition rate of motor imagery before and after rehabilitation training using three-scene stimulation. This study verified the positive activation effect of rehabilitation training strategy on stroke patients and explored the brain activation mechanism using stimulation with different virtual scenes.

MATERIALS AND METHODS

Experimental Paradigm Under Virtual Scenes Stimulation

Design of Virtual Scenes and Training Tasks

As shown in **Figure 1**, in this study, we designed a rehabilitation training strategy based on MI-BCI and VR. Multiple-evoked MI stimulation scenes (e.g., static scenes, dynamic scenes, and VR scenes) induced subjects to perform limb motor imagery. The EEG signals of the brain motor area were collected in real time and then the signals were subjected to preprocessing, intention feature extraction, and intention recognition. Finally, the results were output to the scene for interactive control.

The virtual scenes were interesting and immersive, and the patients had a strong awareness of active training. In this study, we first built the scenes. Then, we added the image of the virtual left and right hands into the scenes and set the training tasks and audio-visual feedback. Finally, through human-computer interaction control, the subjects could feel that they were performing actual body movements. Then we guided the subjects to perform active limb motor imagery.

In addition, we observed the brain electrical activity mapping of the subject and analyzed the degree of neural activation during the subject's brain electrical activity mapping. We observed the activation changes in the brain areas through brain electrical activity mapping and analyzed the effect of training on the subjects. According to the current neural activity in the motor area of the brain, we adaptively adjusted the training scenes to ensure that the subjects could continuously achieve maximum activation of nerves and to accelerate the remodeling of nerve functions.

Different research protocols can lead to different promotion effects of the BCI rehabilitation training system on either the unaffected or affected side of the cerebral hemisphere (Dodd et al., 2017). Different subjects respond differently to different scenes, and the activation areas and intensity of the cerebral cortex and the EEG signal features also are different when the same subject performs different tasks. Considering these differences, this study used the Unity 3D platform to design the training scenes (static and dynamic scenes) in the computer screen as well as the training scenes in the VR environment. We also designed some life-skills training in the scene to feature rehabilitation tasks (e.g., holding goods, pouring water, and picking food).

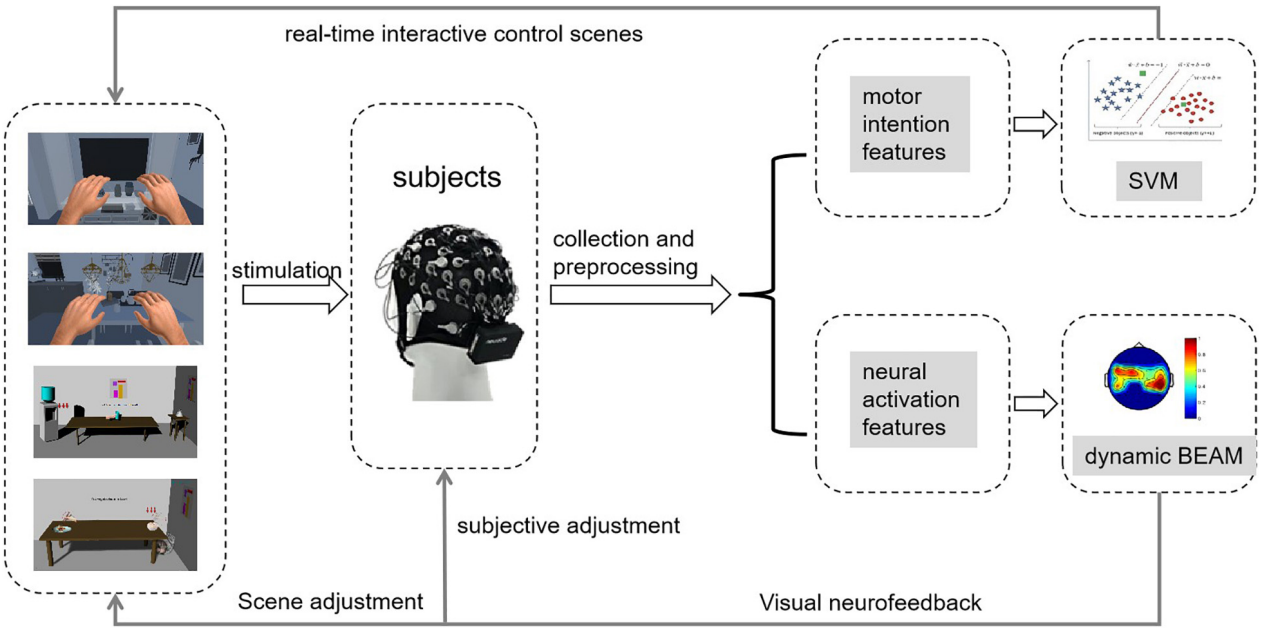


FIGURE 1 | Virtual rehabilitation training strategy map.

As shown in **Figure 2**, this study designed four different static and dynamic scenes without using VR. They included static text scenes, static picture scenes, dynamic 3D life scenes, and dynamic virtual game scenes.

As shown in **Figure 3**, we also designed scenes in the VR environment to enhance the immersion and realism of the subject's training in this study. The control method of the training tasks required the patient to perform motor imagery according to the prompt. In other words, when the patient's use of left (right) motor imagery was recognized, the corresponding limb in the scenes would perform the corresponding movement.

The following describes the scene in detail.

Static scene: Mainly contains text, pictures, etc., the stimulation effect is poor, but the stimulation is more direct. Directly prompt the subjects to perform motor imagery through left, right, etc.

Dynamic scenes: Mainly include videos, games, etc., with general stimulation effects and strong interactivity. Realize motor imagery by playing videos or playing games.

VR scene: The main stimulation mode is the same as that under the computer monitor, but after VR rendering, the sense of immersion is strong, making the patient feel immersive, and the stimulation effect is the best.

Experimental Design

Experimental Principle and Process

The purpose of this study was to verify the effect of the rehabilitation training strategy to improve the subjects' limb control ability during motor imagery. At the same time, we studied the differences in brain activation when subjects were stimulated to perform motor imagery in different scenes. This enabled us to analyze the effect of neural activation when subjects performed motor imagery using stimulation with different virtual scenes.

In this study, we designed a controlled experiment to compare and analyze the changes in neural activity of the brain and to determine the recognition rate of motor imagery using stimulations with different virtual scenes to find the mechanism of neural deep activation. **Figure 4** shows a schematic diagram of the experiment. We assessed the subjects' motor imagery ability before training and then subjects performed multiple motor imagery training. After completing the training, participants finished a post-training assessment.

We selected nine healthy college students in good mental condition as subjects (all male; average age: 24 $\pm$ 2 years old). All subjects are right handedness. The experiment required that all the subjects had not completed similar experiments before and had no history of neurological diseases. All the subjects were informed of the research intention of the trial, the details of the study, and the potential dangers associated with the experiment. Additionally, all of the experiments were conducted within 3 h of the subject having eaten at noon, and each subject had to close his eyes and rest for 5 min before starting motor imagery. Doing so could relieve tension and anxiety and ensured that subjects were in a good mental state.

We grouped the nine subjects equally into three groups: static scene control group, called S1–S3; dynamic scene experimental group, called S4–S6; and VR scene experimental group, called S7–S9. Except for the different virtual scenes for stimulating motor imagery training, all of the other conditions were the same. All the subjects performed a total of 17 days of experiments (14 days of motor imagery task training experiments and 3 days of motor imagery assessment experiments). The specific experimental process was as follows: Subjects completed three

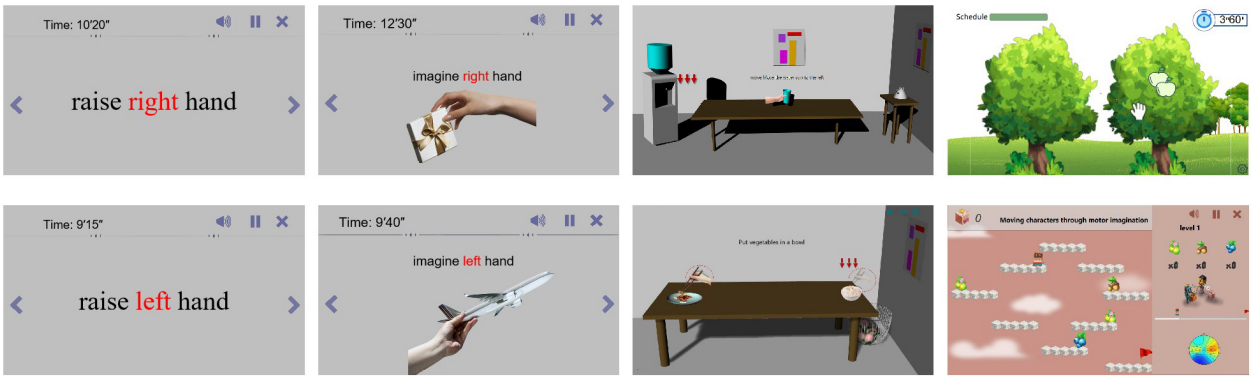


FIGURE 2 | Static and dynamic scenes without VR.

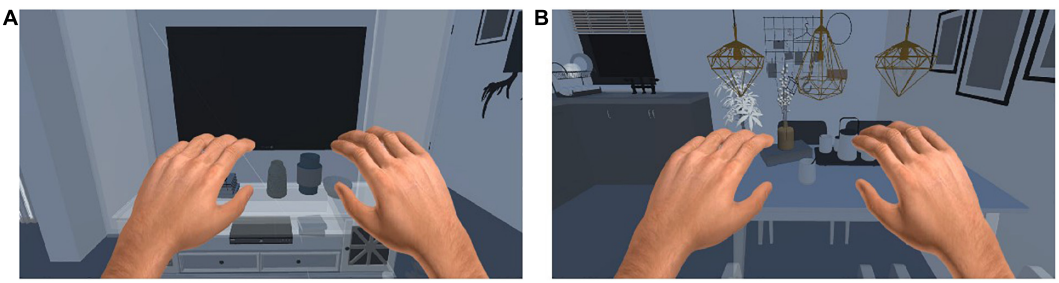


FIGURE 3 | Training scenes in VR enviroment. **(A)** One scene for computer. **(B)** One scence for dinner.

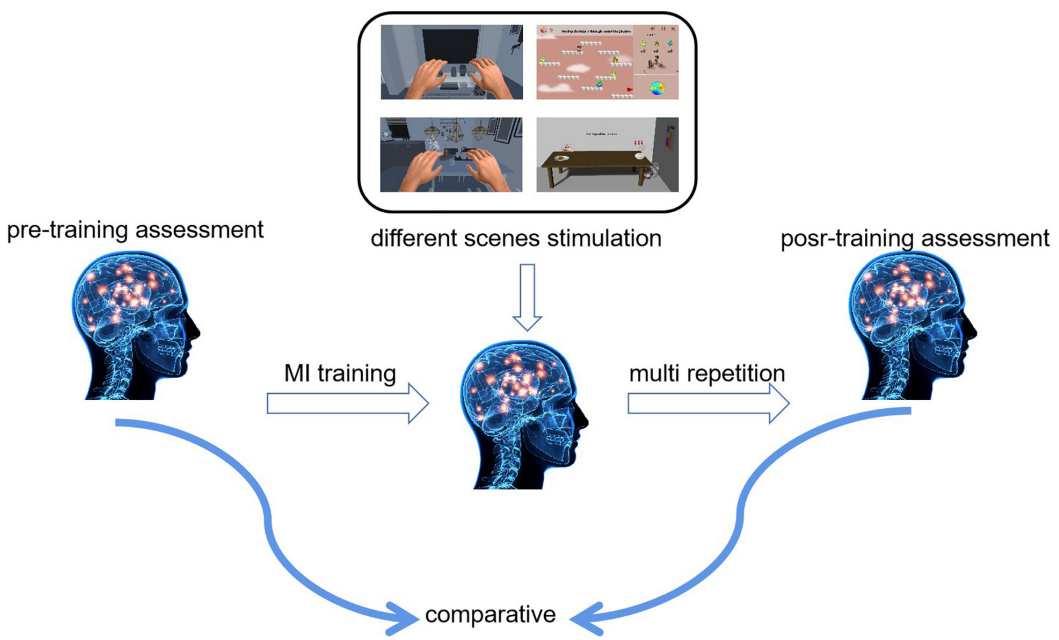


FIGURE 4 | Experimental schematics.

sets of enhanced motor imagery paradigm training per day, and each group included 40 trials. In other words, subjects had to complete a total of 120 limb motor imagery trials per day. In the 14-day motor imagery training using virtual scene stimulation, we obtained the correct rate of 1,680 limb motor images for each subject.

Subjects also performed three motor imagery assessment experiments. Three experiments were conducted on the day before the motor imagery training, the day after the training was completed for 7 days, and the day after the training was completed for 14 days. The subjects completed three sets of motor imagery tasks according to the assessment paradigm, and each group contained 30 motor imagery trials. We collected 19 channels of EEG data during the trials and obtained a total of 270 EEG data points for each subject in the motor imagery trials.

Experimental Paradigm

Figure 5 shows the experimental paradigm for the pre-assessment and post-assessment in the enhanced motor imagery training. A single trial contained three periods totaling 10 s. From 0 to 2 s, a red dot appeared in the center of the screen and then shrunk to remind the subjects to concentrate to start the next motor imagery. From 2 to 7 s, the subjects had to focus on the direction of the left/right movement of the red dot on the screen and performed left/right hand (or left/right limb) motor imagery. From 7 to 10 s, a plus sign (+) appeared in the center of the screen to remind the subjects that this trial had ended. During this period, subjects would rest for 3 s and then they would enter the next trial to repeat this experimental process.

The experimental paradigm of enhanced motor imagery training in different scenes is shown in **Figure 6**. We divided the training process into three groups: the enhanced static scenes control group, the enhanced dynamic scenes experimental group, and the VR scenes experimental group. The motor imagery training process adopted an interactive mode of real-time motor imagery feedback. A single trial contained three periods totaling 10 s. From 0 to 1 s, a red dot appeared in the center of the screen to remind the subjects to concentrate on the start of the motor imagery. At the end of 1 s, the red dot disappeared, which reminded the subjects to start the limb motor imagery. From 1 to 5 s, subjects had to focus on the prompt information provided on the screen and performed left/right hand (or left/right limb) motor imagery. Different prompt information was given to different groups. From 5 to 7 s, the results of the classification model in the motor imagery appeared on the screen, and the virtual characters and limbs in the scenes were controlled to move accordingly. During this period, subjects took a short rest until the end of 7 s. Then, the subjects would enter the next trial and repeat this experimental process.

Electroencephalography Signals Acquisition and Preprocessing

The electrode distribution of EEG data acquisition adopted the international standard 10- to 20-electrode lead positioning. We set the reference electrode in the central area at the top of the

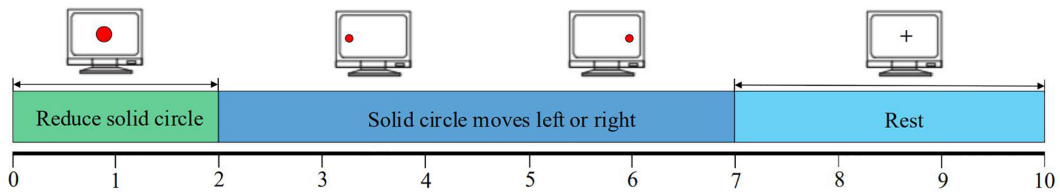


FIGURE 5 | Assessment paradigm before and after enhanced MI training.

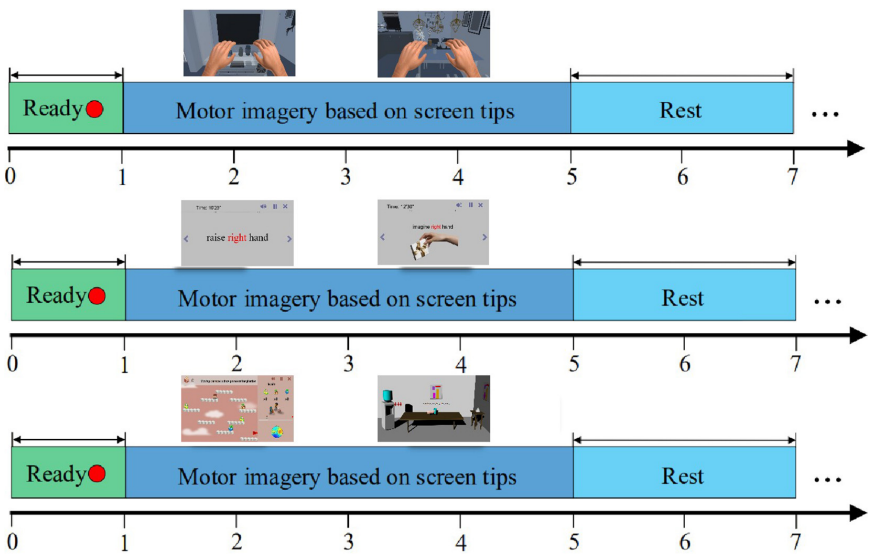


FIGURE 6 | Experimental paradigm of enhanced MI training in different scenes.

head and set the sampling frequency to 1,000 Hz. The acquisition channels included 19 channels in the motor nerve-related area, namely, FC1, FC3, FC5, C1, C3, C5, CP1, CP3, CP5, Cz, FC2, FC4, FC6, C2, C4, C6, CP2, CP4, and CP6.

Most of the EEG acquisition devices were non-invasive dry electrodes. Although EEG acquisition was simpler and more convenient, it was unavoidable that various noises would appear in the acquisition process, such as EOG, ECG, and EMG. This seriously interfered with the subsequent analysis of EEG signals and affected the results of EEG signal analysis. Therefore, it was necessary to preprocess the collected raw EEG signals to remove noises.

The main noises included baseline drift caused by the device, 50 Hz power frequency interference, and EOG artifacts formed by blinking. We used the moving average method to remove baseline drift and used an adaptive 50 Hz filter to remove power frequency interference. We used an independent component analysis to remove EOG artifacts and used the sixth-order Butterworth filter as the band-pass filter. Finally, we acquired 8–32 Hz EEG signals.

RESULTS

Analysis of Energy Changes in Brain Motor Areas

Brain Electrical Activity Mapping

Studies have shown that when subjects perform motor imagery, the energy features of EEG signals in the motor areas of the brain must change. In this study, we used brain electrical activity mapping to analyze the neural activation of multiple channels in the motor area, and visual neurofeedback was provided.

Brain electrical activity mapping (BEAM) is a commonly used method for multichannel EEG signal analysis. This method collects EEG signals from multiple channels of the head, performs fourth-order energy extraction, and then uses color bands and digitization to represent different gray levels. Finally, a dynamic BEAM can be drawn.

Brain Motor Area Energy Analysis Based on Brain Electrical Activity Mapping

On the basis of these experiments, we obtained the EEG data of nine subjects when they conducted motor imagery assessment on the day before the motor imagery training and on the day after the training had been completed for 14 days. We obtained and normalized fourth-order cumulative energy value of each channel for a single trial for each subject. Then, we used the energy value of each channel for multiple trials of nine subjects, which then was used as the input parameter for drawing the BEAM. We obtained the spatial distribution map of the neural activation in different lead positions of the left and right limb motor imagery, as shown in **Figure 7**.

Figure 7 shows the spatial distribution of neural activation in the brain regions of the subjects: S1, S4, and S8 before and after 10 trials of enhanced motor imagery training. **Figure 7A** shows that after the subjects were trained in the rehabilitation strategy designed in this study, the activation range of motor nerves in the brain region was significantly expanded (i.e., the activation breadth of motor imagery nerves increased). **Figure 7B** shows that after the subjects were trained in the rehabilitation strategy, the color of the activation area was darker (i.e., the ERD/ERS phenomenon was more obvious), which indicated that the activation depth of the motor imagery nerves increased. **Figure 7C** shows that compared with static scenes and dynamic scenes, this strategy had a more obvious effect on the activation breadth and depth of motor imagery nerves in VR scenes.

Electroencephalography Signals Feature Extraction, the Modeling of Electroencephalography Signals, and Analysis of the Recognition Rate of Motor Imagery

Electroencephalography Signal Feature Extraction

Feature extraction is the mathematical transformation or mapping of the input signal to obtain eigenvalues that are easier to observe and monitor. The widely used EEG signals that feature the extraction methods used during motor imagery mainly include the following: power spectrum estimation method, wavelet transform, independent component analysis, and common spatial pattern (Xin and Wang, 2017). In this study, we selected three feature combinations of mean square error, power spectral density, and common spatial pattern from the time domain, frequency domain, and spatial domain of EEG signal as the feature quantities for motor imagery intention recognition (Yang et al., 2012).

Furthermore, we used the existing particle swarm optimization (PSO)-support vector machine (SVM) as a classifier to perform motion imagery intention recognition.

Analysis of the Recognition Rate of Motor Imagery

Studies have shown that there is a positive correlation between the degree of active participation, the activity degree of motor nerves, the performance of classification models, and the recognition rate of motor imagery when subjects perform motor imagery. The higher the recognition rate of motor imagery, the higher the active participation of the subjects in motor imagery and the stronger the control ability of motor nerves. The recognition rate of motor imagery can be used as an indicator to stimulate motor imagery ability in different scenes. In this study, the analysis of the recognition rate of motor imagery intention included the following two aspects: online MI training and offline MI assessment.

Online Motor Imagery Training Recognition Rate Analysis

We counted the recognition rate of the nine subjects who completed online MI training in virtual scenes within 14 days (**Table 1**). The recognition rate was the online recognition result obtained statistically using the experimental paradigm designed in this study. The feature quantities used in the classification model were mean square error, power spectral density, and common spatial pattern features. The classification model was PSO-SVM.

Each subject performed 120 motor imagery trials per day. We counted the number of correct motor imagery classifications

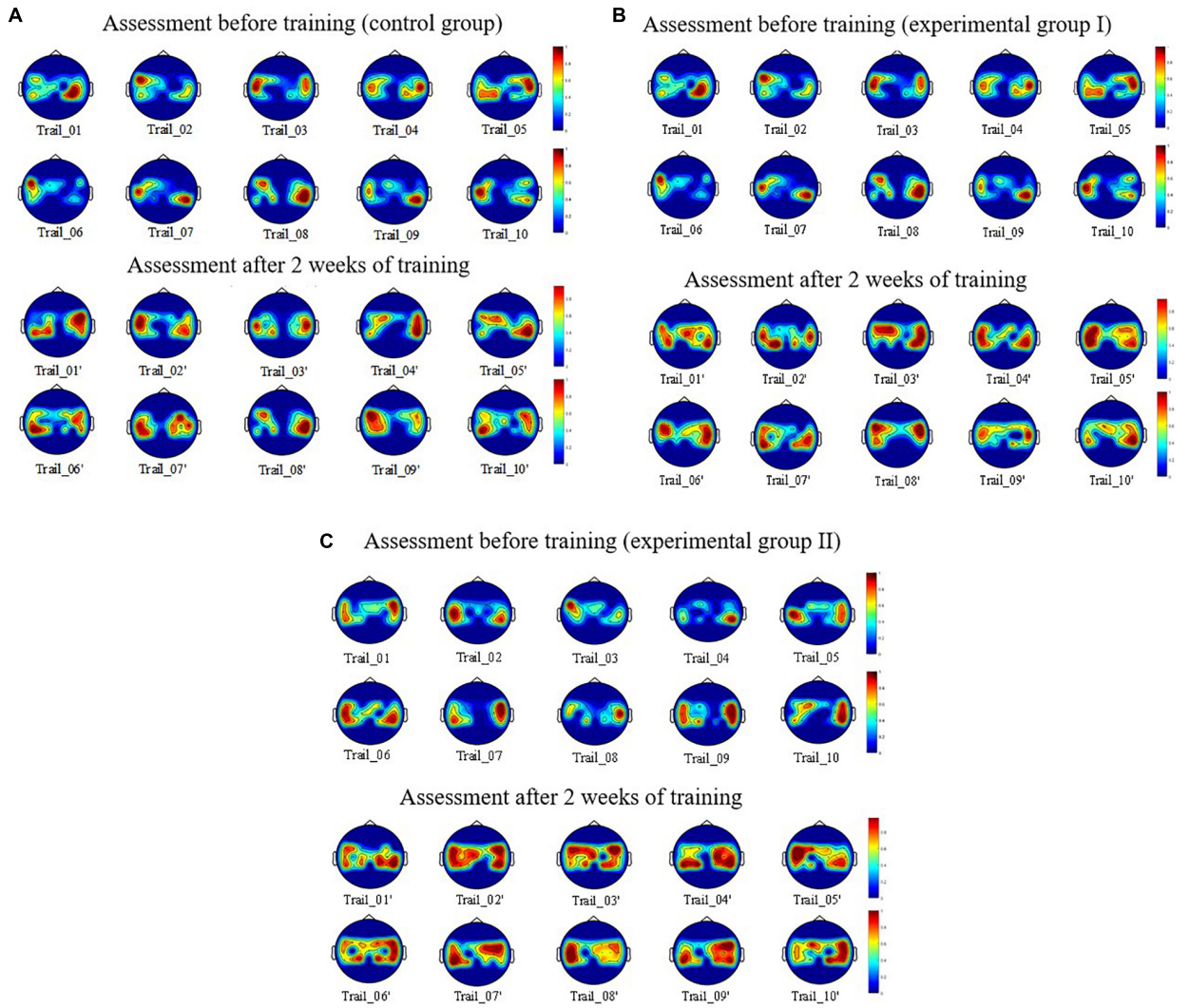


FIGURE 7 | Spatial distribution of the brain activation between the before and after enhanced MI training. **(A)** For control group. **(B)** For experimental group. **(C)** For experimental group II.

per day and calculated each subject's single-day recognition rate, average recognition rate, and standard deviation within 14 days. The data marked with an asterisk (*) in the table indicate that the subject had the highest recognition rate on this day.

According to **Table 1**, by completing the enhanced motor imagery training task, the subjects' motor imagery recognition rates improved to varying degrees. After training, the average recognition rate of motor imagery for three subjects (S1–S3) in static scenes (control group) was between 64.5 and 67.5%, the average recognition rate of motor imagery for three subjects (S4–S6) in dynamic scenes (experimental group I) was between 67.5 and 70.5%, and the average recognition rate of motor imagery for three subjects (S7–S9) in VR scenes (experimental group II) was between 71.0 and 75.0%.

These data showed that the rehabilitation training strategy designed in this study could improve subjects' limb motor imagery ability. The subjects' motor imagination ability in these three scenarios is from weak to strong: static scene, dynamic scene, and VR scene.

Offline Motor Imagery Assessment Recognition Rate Analysis

By analyzing the improvement range of the subjects' motor imagery recognition rate under different training periods, we used an assessment method to evaluate whether the rehabilitation training strategy could improve the subjects' motor imagery ability. We collected the EEG data of nine subjects when they conducted motor imagery assessments 1 day before the motor imagery training, the day after the training was completed for 7 days, and the day after the training was completed for 14 days. Each subject completed three sets of motor imagery tasks per day, and each set consisted of 30 random left and right hand (limb) motor imagery trials.

We subjected the collected data to offline recognition rate analysis. To obtain accurate results, we used a

TABLE 1 | Statistical results of online recognition rate during MI training in different virtual scenes (%).

Group	Static scenes (control group)			Dynamic scenes (experimental group I)			VR scenes (experimental group II)		
Training day	S1	S2	S3	S4	S5	S6	S7	S8	S9
Day_01	60.9%	55.9%	57.6%	61.8%	58.5%	62.6%	62.6%	59.4%	60.3%
Day_02	65.7%	57.4%	60.7%	60.7%	64.8%	59.1%	64.1%	63.1%	64.7%
Day_03	63.4%	61.6%	65.2%	65.1%	67.4%	67.4%	67.4%	64.8%	71.1%
Day_04	69.1%	59.3%	61.5%	69.1%	59.3%	65.9%	71.8%	71.0%	73.9%
Day_05	67.4%	62.7%	70.7%	72.4%	62.6%	70.7%	65.7%	67.3%	74.5%
Day_06	65.9%	66.5%	65.9%	76.8%	65.7%	72.6%	76.8%	73.5%	72.2%
Day_07	56.5%	65.2%	67.6%	70.7%	70.7%	70.7%	72.6%	74.0%	70.7%
Day_08	67.7%	69.8%	56.6%	69.3%	69.3%	65.1%	73.2%	72.7%	75.9%
Day_09	69.3%	67.3%	63.5%	78.2%*	74.3%	74.3%	79.3%	79.2%*	80.8%
Day_10	72.4%	72.7%*	65.6%	72.6%	77.4%*	67.4%	76.6%	76.7%	74.2%
Day_11	70.1%	62.1%	69.3%	67.6%	69.1%	70.0%	82.4%*	72.1%	79.5%
Day_12	74.9%*	69.3%	70.7%	76.6%	68.4%	67.5%	73.4%	70.4%	83.0%
Day_13	68.2%	67.1%	71.6%*	72.5%	70.0%	73.5%	78.3%	76.8%	75.3%
Day_14	70.9%	67.6%	67.6%	74.2%	71.7%	74.0%*	80.0%	78.2%	88.8%*
Average recognition rate	67.3%	64.6%	65.3%	70.5%	67.8%	68.6%	73.2%	71.4%	74.6%
Standard deviation	0.045	0.042	0.046	0.052	0.051	0.043	0.060	0.057	0.070

**denotes the highest MI recognition rate for each subject across the 14 days.*

TABLE 2 | Offline evaluation of recognition rate on different virtual scenes before, during and after enhanced MI training (%).

Group	Subjects	Recognition rate before the training	Recognition rate after 7-day training	Increasing range (%)	Recognition rate after 14-day training	Increasing range (%)
Static scenes	S1	63.6%	69.7%	6.1	72.8%	3.1
	S2	61.6%	66.1%	4.5	68.9%	2.8
	S3	64.1%	68.4%	4.3	72.1%	3.7
Dynamic scenes	S4	63.8%	69.6%	5.8	74.7%	5.1
	S5	64.1%	71.4%	7.3	76.0%	4.6
	S6	60.7%	66.4%	5.7	71.3%	4.9
VR scenes	S7	64.6%	74.4%	9.8	81.6%	7.2
	S8	59.6%	72.5%	12.9	79.3%	6.8
	S9	62.7%	73.4%	10.7	80.5%	7.1

three-fold cross-validation method, as follows: First, we divided the 90 motor imagery datasets into three groups on average. One group was taken out each time as the test data, and the remaining data were used as the training data to complete the PSO-SVM model training. Then, we used the trained model to classify the test data. Finally, we averaged the obtained three groups of classification results as the final recognition rate. The statistical results are given in **Table 2**.

Table 2 shows that after training, the motor imagery recognition rate of the nine subjects was improved. In addition, the overall increase in the recognition rate in the static scenes (control group) and the dynamic scenes (experimental group I) was roughly the same, and the overall increase of the recognition rate in the VR scenes (experimental group II) was significantly higher than that of the other two groups. This result indicated that the rehabilitation training strategy designed in this study could improve the subjects' EEG signal identification and ability to control EEG signals. Compared with static scenes and dynamic scenes, VR scenes had a more significant effect on improving subjects' control of their motor nerves.

DISCUSSION

The experiments designed in this study compared the degree of neural activation as well as the classification and recognition rate of motor imagery when subjects performed motor imagery using stimulation with different virtual scenes. The rehabilitation training strategy designed in this paper greatly improved the motor nerve activation of the subjects, accelerated the remodeling of the subjects' neural functions, and improved the motor imagery ability of the subjects.

The experiments further investigated in which scene the subjects' motor nerve activation layers were wider and deeper.

The virtual scenes designed in this study included static scenes (e.g., pictures, text), dynamic scenes (e.g., animation, games) displayed on a computer screen, and VR scenes. After completing the motor imagery training, we compared and analyzed the energy changes in the brain motor areas and the motor imagery recognition rate changes for the different subjects to explore the neural activation and motor imagery ability changes of subjects using stimulation of different virtual scenes.

First, we analyzed the energy changes in the brain motor areas of subjects under different scene simulation and analyzed the BEAM of one subject randomly selected from each of the three groups. By comparing the distribution of BEAM of each subject before and after training, we found that the range of energy distribution in the motor area after training expanded and the color depth deepened. This proved that after training with the rehabilitation strategy described in this paper, the neural activity intensity and energy in the motor areas of subjects increased, which could stimulate the remodeling of damaged motor nerves to a certain extent.

We further analyzed motor imagery recognition rates. First, the recognition rates of online training (**Table 1**) and offline assessment (**Table 2**) of the nine subjects under different virtual scene simulation remained at around 60%. The main reason that the recognition rate of the latter was slightly higher than that of the former was that the signal interference was less formed in the offline state. Second, from the data in **Table 1**, we found that with an increase in the number of training days, the motor imagery recognition rate of the nine subjects followed an overall upward trend. The fluctuation of the recognition rate in the short period was affected mainly by the mental state of the subjects during the experiment and the interference differences of the EEG signals in each acquisition. Third, by calculating the average recognition rate of 14 days, we found that the recognition rate of subjects in the VR scenes was significantly higher than that of the static scenes and dynamic scenes displayed on a computer screen. This result showed that the subjects could mobilize more nerve cells in the motor area to pursue regular physiological activities in the VR scenes, thereby improving the recognition of the subjects' EEG signals. As shown in **Table 2**, to verify whether the improved recognition rate could be formed by the transient stimulation of the virtual scenes or by the change of the motor nerve activity mechanism, we adopted the same experimental paradigm of motor imagery for offline assessment. We found that the improvement of the recognition rate in the VR scenes was significantly better than that in the static and the dynamic scenes, which was consistent with previous results, and further indicated that the VR scenes had a better effect on improving the recognition of subjects' EEG signals.

CONCLUSION

Today, the virtual rehabilitation that integrates MI-BCI and VR holds significant potential in the field of stroke rehabilitation, and this technology can greatly improve the rehabilitation effect for patients. To investigate how MI-BCI therapy can maximize the deep activation of subjects' motor nerves and accelerate the remodeling mechanism of motor nerve function, we designed the following rehabilitation training strategy: we enhanced motor imagery training under different virtual scenes and compared and analyzed the degree of neural activation and the recognition rate of motor imagery in stroke patients after enhanced motor imagery training using stimulation with different virtual scenes. The experimental results showed that the motor imagery training using virtual scene stimulation could improve the motor nerve activation and motor imagery ability of the subjects. Compared with the static and dynamic scenes displayed on a computer screen, the VR scenes had a more significant effect in improving neural activation intensity and recognition rate.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Human Ethics and Administrative Council of Yanshan University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

PX, XC, and ZW designed this study. ZW wrote the manuscript. ZL, YW, NW, and ZL contributed to conception of the study. JW and XC wrote sections of the manuscript. All authors contributed to manuscript revision, and approved the submitted version.

FUNDING

This work was supported by National Natural Science Foundation of China (Grant Nos. U20A20192, 61901407, and 62076216) and the Key Research and Development Program of Hebei Province of China (Grant Nos. 21372001D and 21372005D).

ACKNOWLEDGMENTS

We would like to thank editor and reviewers for helpful comments to improve the manuscript.

REFERENCES

Bagarinao, E., Yoshida, A., Terabe, K., Kato, S., and Nakai, T. (2020). Improving Real-Time Brain State Classification of Motor Imagery Tasks During Neurofeedback Training. *Front. Neurosci.* 14:623. doi: 10.3389/fnins.2020.00623

Barsotti, M., Leonardis, D., Loconsole, C., Solazzi, M., Sotgiu, E., Procopio, C., et al. (2015). "A full upper limb robotic exoskeleton for reaching and grasping rehabilitation triggered by MI-BCI," in *IEEE International Conference on Rehabilitation Robotics*, (Piscataway: IEEE), doi: 10.1109/ICORR.2015.7281174

Bayliss, J. D., and Ballard, D. H. (2000). A virtual reality testbed for brain-computer interface research. *IEEE Trans. Rehabil. Eng.* 8:188. doi: 10.1109/86.847811

Benitez-Andonegui, A., Burden, R., Benning, R., Mckel, R., and Sorger, B. (2020). An Augmented-Reality fNIRS-Based Brain-Computer Interface: A Proof-of-Concept Study. *Front. Neurosci.* 14:346. doi: 10.3389/fnins.2020.00346

Birbaumer, N., Weber, C., Neuper, C., Buch, E., and Cohen, L. (2006). Physiological regulation of thinking: brain-computer interface (BCI) research. *Prog. Brain Res.* 159, 369–391. doi: 10.1016/S0079-6123(06)59024-7

Clark, E., Czaplewski, A., Dourney, S., Gadelha, A., and Korostenskaja, M. (2019). "Brain-Computer Interface for Motor Rehabilitation," in *HCI International 2019 - Posters: 21st International Conference*, (Orlando, FL: HCII 2019). doi: 10.1088/1741-2552/aae4b8

Dodd, K. C., Nair, V. A., and Prabhakaran, V. (2017). Role of the Contralesional vs. Ipsilesional Hemisphere in Stroke Recovery. *Front. Hum. Neurosci.* 11:469. doi: 10.3389/fnhum.2017.00469

Fumanal-Idocin, J., Takac, Z., Fernandez, J., Sanz, J. A., and Bustince, H. (2021). "Interval-valued aggregation functions based on Moderate deviations applied to Motor-Imagery-Based Brain Computer Interface," in *IEEE Transactions on Fuzzy Systems*, (Piscataway: IEEE), doi: 10.1109/TFUZZ.2021.3092824

Gomez-Pilar, J., Corralejo, R., Nicolas-Alonso, L. F., Álvarez, D., and Hornero, R. (2016). Neurofeedback training with a motor imagery-based BCI: neurocognitive improvements and EEG changes in the elderly. *Med. Biol. Eng. Comput.* 54, 1655–1666. doi: 10.1007/s11517-016-1454-4

Guger, C., Allison, B., Cao, F., and Edlinger, G. (2017). A Brain-Computer Interface for Motor Rehabilitation With Functional Electrical Stimulation and Virtual Reality. *Arch. Phys. Med. Rehabilit.* 98:e24. doi: 10.1016/j.apmr.2017.08.074

Li, F., Fei, W., Zhang, L., Tao, Z., and Xu, P. (2018). The Dynamic Brain Networks of Motor Imagery: time-Varying Causality Analysis of Scalp EEG. *Int. J. Neural Syst.* 29:1850016. doi: 10.1142/S0129065718500168

Li, Y., Wei, Q., Chen, Y., and Zhou, X. (2021). Transfer learning based on hybrid Riemannian and Euclidean space data alignment and subject selection in brain-computer interfaces. *IEEE Access* 9, 6201–6212. doi: 10.1109/ACCESS.2020.3048683

Mattia, D., Pichiorri, F., Molinari, M., and Rupp, R. (2012). "Brain Computer Interface for Hand Motor Function Restoration and Rehabilitation," in *Towards Practical Brain-Computer Interfaces*, eds Z. A. Brendan, D. Stephen, L. Robert, D. R. M. José, and N. Anton (Berlin: Springer Nature), doi: 10.1007/978-3-642-29746-5_7

Moctezuma, L. A., and Molinas, M. (2020). EEG Channel-Selection Method for Epileptic-Seizure Classification Based on Multi-Objective Optimization. *Front. Neurosci.* 14:593. doi: 10.3389/fnins.2020.00593

Pfurtscheller, G., and Neuper, C. (1997). Motor imagery activates primary sensorimotor area in humans. *Neurosci. Lett.* 239, 65–68. doi: 10.1016/S0304-3940(97)00889-6

Pfurtscheller, G., and Neuper, C. (2001). Motor imagery and direct brain-computer communication. *Proc. IEEE* 89, 1123–1134. doi: 10.1109/5.939829

Remsik, A., Young, B., Vermilyea, R., Kiekoefer, L., Abrams, J., Evander Elmore, S., et al. (2016). A review of the progression and future implications of brain-computer interface therapies for restoration of distal upper extremity motor function after stroke. *Expert Rev. Med. Devices* 13, 445–454. doi: 10.1080/17434440.2016.1174572

Ren, S., Wang, W., Hou, Z. G., Liang, X., and Shi, W. (2020). "Enhanced Motor Imagery Based Brain-Computer Interface via FES and VR for Lower Limbs," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, (Piscataway: IEEE). doi: 10.1109/TNSRE.2020.3001990

Ren, Y., Wang, G., and Cheng, M. (2004). The control on the artificial limb based on brain-computer interface. *CHIN. J. Rehabilit. Med.* 19, 330–333.

Takenaka, Y., Suzuki, T., and Sugawara, K. (2021). Time course effect of corticospinal excitability for motor imagery. *Eur. J. Neurosci.* 54, 6123–6134. doi: 10.1111/ejn.15404

Varsehi, H., and Firoozabadi, S. (2021). An EEG channel selection method for motor imagery based brain–computer interface and neurofeedback using Granger causality. *Neural Netw.* 133, 193–206. doi: 10.1016/j.neunet.2020.11.002

Velasquez-Martinez, L. F., Zapata-Castano, F., and Castellanos-Dominguez, G. (2020). Dynamic Modeling of Common Brain Neural Activity in Motor Imagery Tasks. *Front. Neurosci.* 14:714. doi: 10.3389/fnins.2020.00714

Vidaurre, C., Haufe, S., Jorajuría, T., Müller, K., and Nikulin, V. V. (2020). Sensorimotor functional connectivity: a neurophysiological factor related to BCI performance. *Front. Neurosci.* 14:575081. doi: 10.3389/fnins.2020.575081

Vourvopoulos, A., Cardona, J., and Badia, S. (2015). "Optimizing Motor Imagery Neurofeedback through the Use of Multimodal Immersive Virtual Reality and Motor Priming," in *International Conference on Virtual Rehabilitation*, (Valencia: IEEE), doi: 10.1109/ICVR.2015.7358592

Vries, S. D., and Mulder, T. (2007). Motor imagery and stroke rehabilitation: a critical discussion. *J. Rehabilit. Med.* 39, 5–13. doi: 10.2340/16501977-0020

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clin. Neurophysiol.* 113, 767–791. doi: 10.1016/S1388-2457(02)00057-3

Xiao, X., and Fang, Y. (2021). Motor Imagery EEG Signal Recognition Using Deep Convolution Neural Network. *Front. Neurosci.* 15:655599. doi: 10.3389/fnins.2021.655599

Xin, X. U., and Wang, N. (2017). Feature extraction and classification of EEG signals in four kinds of motion imagination. *J. Nanjing Univ. Posts Telecommun.* 37, 18–22. doi: 10.14132/j.cnki.1673-5439.2017.06.003

Xu, B. L., Fu, Y. F., Shi, G., Yin, X. X., Miao, L., Wang, Z. D., et al. (2013). Comparison of optical and concentration feature used for fNIRS-based BCI system using HMM. *Appl. Mech. Mat.* 385-386, 1443–1448. doi: 10.4028/www.scientific.net/amm.385-386.1443

Yang, B., Lu, W., He, M., and Liu, L. (2012). Novel feature extraction method for BCI based on WPD and CSP. *CHIN. J. Sci. Instr.* 33, 2560–2565.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Aging Neuroscience

TYPE Original Research
PUBLISHED 05 August 2022
DOI 10.3389/fnagi.2022.898940

OPEN ACCESS

EDITED BY
Xiuqin Jia,
Capital Medical University, China

REVIEWED BY
Lin Chen,
Chongqing Institute of Green and Intelligent Technology (CAS), China
Jin San Lee,
Kyung Hee University, South Korea

*CORRESPONDENCE
Kyunga Kim
kyunga.j.kim@samsung.com
Sang Won Seo
sangwonseo@empas.com

†These authors have contributed equally to this work and share first authorship

‡These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 18 March 2022
ACCEPTED 18 July 2022
PUBLISHED 05 August 2022

CITATION
Chun MY, Park CJ, Kim J, Jeong JH, Jang H, Kim K and Seo SW (2022) Prediction of conversion to dementia using interpretable machine learning in patients with amnestic mild cognitive impairment. *Front. Aging Neurosci.* 14:898940. doi: 10.3389/fnagi.2022.898940

Prediction of conversion to dementia using interpretable machine learning in patients with amnestic mild cognitive impairment

Min Young Chun [1†], Chae Jung Park [2,3†], Jonghyuk Kim [2†], Jee Hyang Jeong [4], Hyemin Jang [1,3], Kyunga Kim [2,5*‡] and Sang Won Seo [1,2,3,6,7*‡]

[1]Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea, [2]Department of Digital Health, Samsung Advanced Institute for Health Sciences and Technology, Sungkyunkwan University, Seoul, South Korea, [3]Alzheimer's Disease Convergence Research Center, Samsung Medical Center, Seoul, South Korea, [4]Department of Neurology, Ewha Womans University Seoul Hospital, Ewha Womans University College of Medicine, Seoul, South Korea, [5]Biomedical Statistics Center, Data Science Research Institute, Research Institute for Future Medicine, Samsung Medical Center, Seoul, South Korea, [6]Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences and Technology, Sungkyunkwan University, Seoul, South Korea, [7]Department of Intelligent Precision Healthcare Convergence, Sungkyunkwan University, Suwon, South Korea

Purpose: Amnestic mild cognitive impairment (aMCI) is a transitional state between normal aging and Alzheimer's disease (AD). However, not all aMCI patients are observed to convert to AD dementia. Therefore, developing a predictive algorithm for the conversion of aMCI to AD dementia is important. Parametric methods, such as logistic regression, have been developed; however, it is difficult to reflect complex patterns, such as non-linear relationships and interactions between variables. Therefore, this study aimed to improve the predictive power of aMCI patients' conversion to dementia by using an interpretable machine learning (IML) algorithm and to identify the factors that increase the risk of individual conversion to dementia in each patient.

Methods: We prospectively recruited 705 patients with aMCI who had been followed-up for at least 3 years after undergoing baseline neuropsychological tests at the Samsung Medical Center between 2007 and 2019. We used neuropsychological tests and apolipoprotein E (*APOE*) genotype data to develop a predictive algorithm. The model-building and validation datasets were composed of data of 565 and 140 patients, respectively. For global interpretation, four algorithms (logistic regression, random forest, support vector machine, and extreme gradient boosting) were compared. For local interpretation, individual conditional expectations (ICE) and SHapley Additive exPlanations (SHAP) were used to analyze individual patients.

Results: Among the four algorithms, the extreme gradient boost model showed the best performance, with an area under the receiver operating characteristic curve of 0.852 and an accuracy of 0.807. Variables, such as age,

education, the scores of visuospatial and memory domains, the sum of boxes of the Clinical Dementia Rating scale, Mini-Mental State Examination, and *APOE* genotype were important features for creating the algorithm. Through ICE and SHAP analyses, it was also possible to interpret which variables acted as strong factors for each patient.

Conclusion: We were able to propose a predictive algorithm for each aMCI individual's conversion to dementia using the IML technique. This algorithm is expected to be useful in clinical practice and the research field, as it can suggest conversion with high accuracy and identify the degree of influence of risk factors for each patient.

KEYWORDS

Alzheimer's disease, amnestic mild cognitive impairment, prediction algorithm, interpretable machine learning, artificial intelligence, clinical decision-support system, SHapley Additive exPlanations (SHAP)

Introduction

Amnestic mild cognitive impairment (aMCI) refers to a transitional state between normal aging and dementia (Flicker et al., 1991; Petersen et al., 2001; Sarazin et al., 2007). Previous studies showed that within 3 years, approximately 50% of aMCI patients converted to dementia (Fischer et al., 2007; Espinosa et al., 2013), with an annual conversion rate of 5–25% (Larrieu et al., 2002; Mitchell and Shiri-Feshki, 2009; Alegret et al., 2014). However, some aMCI patients maintain a stable state of cognitive function or reverted to normal cognition (Busse et al., 2006; Mitchell and Shiri-Feshki, 2009). Several factors, including age, sex, neuropsychological test results, and apolipoprotein E (*APOE*) genotype were found to affect the rate of conversion to dementia (Petersen et al., 1995; Daly et al., 2000; DeCarli et al., 2004; Yaffe et al., 2006). Thus, as the clinical outcomes of aMCI patients are heterogeneous, it is important to consider the risk factors of each patient individually while predicting their conversion to dementia.

Several studies have been conducted to create algorithms that predict the conversion of aMCI to dementia (Ravaglia et al., 2006; Tabert et al., 2006; De Simone et al., 2019). Specifically, Jang et al. developed a dementia risk prediction algorithm by using traditional statistical methods, such as multivariate logistic regression (LR) and the nomogram (Jang et al., 2017). However, when the LR is applied to complex multivariate non-linear relationships, it may have low robustness because of the multicollinearity between the variables (Tu, 1996).

Machine learning (ML) techniques, a form of artificial intelligence that is increasingly used in the medical research field, have also been considered in developing prediction algorithms for conversion to dementia (Chen and Herskovits, 2010; Mattila et al., 2012; Hall et al., 2015; So et al., 2017; Zhu et al., 2020; Lian et al., 2021; Qiao et al., 2021). These prediction algorithms are based on computer algorithms that help ML to learn complex relationships with empirical data and to make more accurate decisions (Bishop, 2006; Waljee et al., 2014). Compared to the traditional statistical methods, ML has a lower possibility of overlooking unexpected predictors and potential interactions between variables (Waljee et al., 2014). However, unlike nomograms, ML techniques are not able to show which factors play a major role in the conversion. Thus, interpretable ML (IML) was developed to provide understandable explanations for learning complex outputs with predictive accuracy, descriptive accuracy, and relevancy (Murdoch et al., 2019).

Therefore, in the present study, we aimed to develop an IML algorithm with a higher predictive power than that of LR, which predicts conversion to dementia in aMCI participants in an accurate manner. We used clinical demographics, *APOE* genotype, and neuropsychological results as features that are easily accessible in clinical practice. We also attempted to develop a graphic-based interpretable method to show which risk factors influence conversion to dementia, and to what extent, in individual aMCI participants.

Materials and methods

Participants

We conducted a cohort study among participants with aMCI who visited the Samsung Medical Center (SMC) in South Korea from June 2007 to December 2019 and were followed-up for at least 3 years after baseline neuropsychological tests. In total, 705 participants with aMCI were enrolled in this study. All aMCI subjects met the following criteria for aMCI (Albert et al., 2011): (1) subjective memory complaints by participants or caregivers; (2) objective memory decline

below −1.0, standard deviation (SD) on either verbal or visual memory tests; (3) normal activities of daily living (ADL), as judged clinically; and (4) not demented.

All the subjects underwent neurological examination, laboratory tests, including *APOE* genotype, and neuropsychological tests. We excluded participants with secondary causes of cognitive impairment through laboratory tests, such as vitamin B_{12}/folate determination, syphilis serology, and thyroid function tests. In addition, participants with structural lesions, such as territorial infarction, intracranial hemorrhage, brain tumor, traumatic brain injury, hydrocephalus, or severe white matter hyperintensities on brain magnetic resonance imaging (MRI), were excluded.

The study was approved by the Institutional Review Board of SMC, and informed consent was obtained from all participants and caregivers.

TABLE 1 Demographics of the study.

Feature	Training set (N = 565)		Validation set (N = 140)	
	Mean	SD (%)	Mean	SD (%)
Conversion to dementia	204	(36.1%)	50	(35.7%)
Age (years)	71.6	7.8	72.2	7.6
Sex – Women	348	(61.6%)	84	(60.0%)
Education (years)	11.1	5.2	11.1	4.8
APOE ε4 carrier	214	(37.9%)	45	(32.1%)
APOE ε2 carrier	46	(8.1%)	9	(6.4%)
K-BNT	39.9	10.1	39.6	10.3
Ideomotor praxis	4.2	1.2	4.2	1.2
Calculation total score	10.9	2.0	10.6	2.1
RCFT copy score	29.7	6.3	29.7	5.7
RCFT copy time (seconds)	258.5	124.3	273.5	139.4
SVLT delayed recall	2.6	2.5	2.5	2.4
SVLT recognition score	18.3	2.8	18.4	2.4
RCFT delayed recall	6.9	5.4	6.8	4.8
RCFT recognition score	18.2	2.3	18.3	2.3
Contrasting program	19.1	2.8	19.0	2.9
Go/no-go	16.9	5.0	16.8	4.9
COWAT animal	12.5	4.2	12.6	4.3
K-MMSE	25.9	3.2	25.6	3.2
CDR-SOB	1.5	0.9	1.5	0.9

The numbers are mean and standard deviation (or percentage in parenthesis) of the training and validation sets.
APOE, apolipoprotein E; K-BNT, Korean version of the Boston Naming Test; RCFT, Rey–Osterrieth Complex Figure Test; SVLT, Seoul Verbal Learning Test; COWAT, Controlled Oral Word Association; K-MMSE, Korean version of the Mini-Mental State Examination; SD, standard deviation; CDR-SOB, clinical dementia rating-sum of boxes.

TABLE 2 Performance of classifiers on validation set.

Classifier	Accuracy	AUC
Logistic regression	0.743	0.813
Random forest	0.771	0.834
Support vector machine	0.800	0.830
Artificial neural network	0.757	0.841
Extreme gradient boost	0.807	0.852

Each classifier's accuracy, area under the receiver operating characteristic curve, and optimized hyperparameters as presented.
AUC, area under the receiver operating characteristic curve.

Neuropsychological assessments

All the participants underwent the Seoul Neuropsychological Screening Battery (SNSB), a standardized neuropsychological battery widely used in South Korea (Kang and Na, 2003; Kang et al., 2016). Four major cognitive domains were evaluated: memory, language, visuospatial, and frontal/executive function. If the z-score of SNSB was below −1.0 SD of age and education, it was considered impaired.

The scorable tests are comprised of the Korean version of the Boston Naming Test (Kim and Na, 1999), Rey-Osterrieth Complex Figure Test (RCFT) (Kang and Na, 2003), which involves copying, immediate and 20-min delayed recall, and recognition, the Seoul Verbal Learning Test (SVLT) (Kang and Na, 2003), which includes three learning-free recall trials of 12 words, a 20-min delayed recall trial of these 12 items, and a recognition test, the contrasting program (instructing the patient to raise the second and third fingers when the examiner raises the second finger, and to raise the second finger when the examiner raises the second and third fingers), go/no-go test (changing the initial rule as follows: instructing the patient to make a fist in respond to examiner's raising the second and third fingers) (Dubois et al., 2000), and phonemic and semantic Controlled Oral Word Association Tests (COWAT) (Kang et al., 2000). In addition, the ideomotor praxis and the total calculation score were evaluated. The Korean version of the Mini-Mental State Examination (K-MMSE) and clinical dementia rating-sum of boxes (CDR-SOB) of all the participants were investigated (Kang et al., 2016).

Follow-up

All the participants underwent two or more SNSB during a follow-up period of at least 3 years. Dementia was diagnosed on the basis of the criteria of the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders and required evidence of cognitive deficits (confirmed by neuropsychological testing) and social and/or occupational dysfunction (confirmed by ADL impairment). The criteria of the National Institute of Neurological and Communicative Disorders and Stroke and

the Alzheimer's Disease and Related Disorders Association were used for the diagnosis of probable AD (McKhann et al., 2011). A consensus panel and an experienced neurologist reviewed the interview records and neuropsychological results of each aMCI patient and confirmed the conversion to dementia in the SMC cohort.

The primary outcome was defined as conversion to dementia within 3 years of the baseline neuropsychological test. The predictive algorithm used variables, such as age, gender, years of education, neuropsychological features, *APOE* ε2, and *APOE* ε4 status as the potential predictors.

Feature selection

Three major steps were performed to select variables: First, domain knowledge was used to remove the unnecessary variables from the results of neuropsychological tests; second, the remaining variables were used to confirm the significance of the variables through LR analysis for a single variable and remove the insignificant variables; and third, one of the variables suspected of multicollinearity was removed or integrated through the correlation coefficient. We specified the primary outcome as 3-year dementia conversion and included features, such as demographics, *APOE* genotypes, and neuropsychological features (including K-MMSE and CDR-SOB) selected using the above process. The selected features were used as inputs for predictive model building, and as potential predictors for model interpretation.

Algorithm constructions

Eighty percent of the total data was randomly selected by the matching class imbalance and used it to develop the predictive algorithm, and the remaining 20% was used for the algorithm test. Stratified 5-fold cross-validation was repeated five times by random dataset splitting, and Bayesian optimization was used for hyperparameter tuning. Five types of ML models were developed: multivariable LR, random forest (RF), support vector machine (SVM), artificial neural network (ANN) and extreme gradient boost (XGB).

Statistical analyses

The performance of the model was compared by using areas under the receiver operating characteristic curve (AUCs) with DeLong test (P-value < 0.05 indicated statistical significance) (DeLong et al., 1988). Statistical analyses were performed using the Daim (v1.1.0) package in R 4.1.2 (R Core Team, 2021).

Interpretation methods

The interpretation of the developed ML models was based on both global and local perspectives. IML analysis was carried out using R 4.1.2 (R Core Team, 2021), the caret (v6.0-90), the iml (v0.10.1), the vip (v0.3.2), the pdp (v0.7.0), the breakDown (v0.2.1), SHAPforxgboost (v0.1.1), the caret (v6.0-90), the DALEX (v2.3.0), and the modelStudio (v3.0.0) packages.

Global interpretation

The global analysis method was used to evaluate the overall performance of the developed model, which we evaluated through the model performance, feature importance (Breiman, 2001; Fisher et al., 2019), and partial dependence (Friedman, 2001). The ML model performance of the four groups divided by gender and age was measured by accuracy and AUC. The feature importance is to observe a lowered performance change

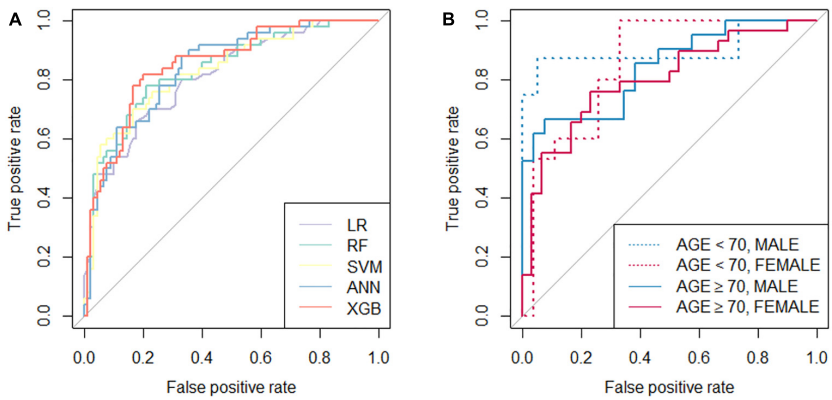


FIGURE 1
Receiver operation characteristic (ROC) curves of the classifiers. **(A)** ROC curves of five developed classifiers; **(B)** ROC curves of the extreme gradient boost classifier tested with validation set divided by age (threshold of 70 years old) and gender. LR, logistic regression; RF, random forest; SVM, support vector machine; ANN, artificial neural network; XGB, extreme gradient boost.

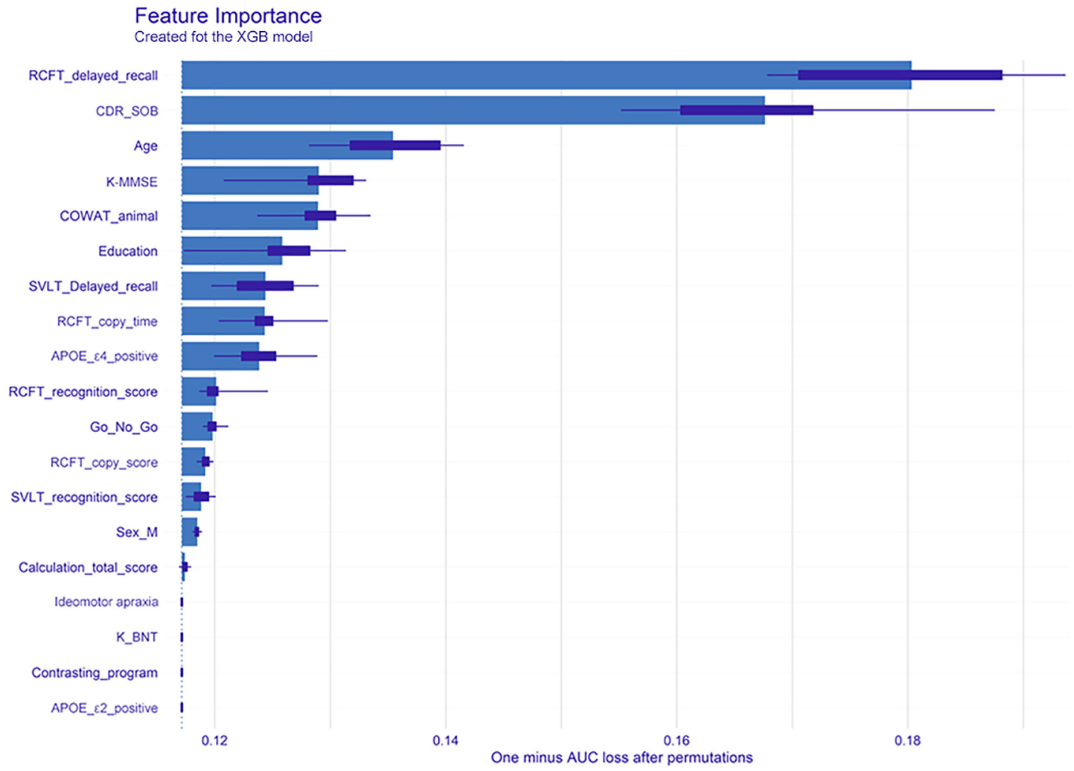


FIGURE 2
Feature importance of the extreme gradient boost model. The bars indicate the feature importance, while the interval bands indicate difference due to random permutations. From the model, clinical neuropsychological features of RCFT delayed recall, clinical dementia rating-sum of boxes, and age were noted as important factors to the global performance. XGB, extreme gradient boost; RCFT, Rey–Osterrieth Complex Figure Test; CDR-SOB, clinical dementia rating-sum of boxes; K-MMSE, Korean version of the Mini-Mental State Examination; COWAT, Controlled Oral Word Association; SVLT, Seoul Verbal Learning Test; APOE, apolipoprotein E; K-BNT, Korean version of the Boston Naming Test; AUC, area under the receiver operating characteristic curve.

by randomly mixing a specific feature. The partial dependence plot (PDP) is a global interpretation method in the ML model that shows the marginal effect of one or two features on the prediction result (Friedman, 2001).

Local interpretation

The local analysis method interpreted the prediction results for individual participants. In this study, we implemented Individual Conditional Expectations (ICE) (Goldstein et al., 2015), Break-down (Robnik-Šikonja and Kononenko, 2008), and SHapley Additive exPlanations (SHAP) (Lundberg and Lee, 2017). First, ICE (or Ceteris-paribus) plots display one line per individual that shows how the individual's prediction changes when a feature changes (Goldstein et al., 2015). Other feature values are fixed with the individual's data. Second, Break-down plots show feature attributions; that is, the prediction is decomposed into contributions that can be attributed to different interpretive features (Robnik-Šikonja and Kononenko, 2008). A plot can be drawn by adding or subtracting each feature contribution one by one on the basis of the average predicted value for all datasets. Finally, SHAP explains individual predictions by computing the contribution of each feature to the prediction. This is based on the game theoretically optimal Shapley values (Lundberg and Lee, 2017). Unlike break-down plots, the order of adding features is calculated by numerous trials; therefore, the mean and SD is estimated.

We plotted three local interpretations above with the XGB model using six exemplary patients. **Supplementary Table 1** shows demographic and dementia conversion information. Also, we collected all IML results and developed dashboards with a graphical view of each patient's analysis results.

Results

Demographics and clinical characteristics

Table 1 shows the patient demographics and clinical characteristics. The model-building and validation datasets were composed of 565 and 140 participants, respectively. Among the aMCI participants of the development set, 36.1% (204/565) of

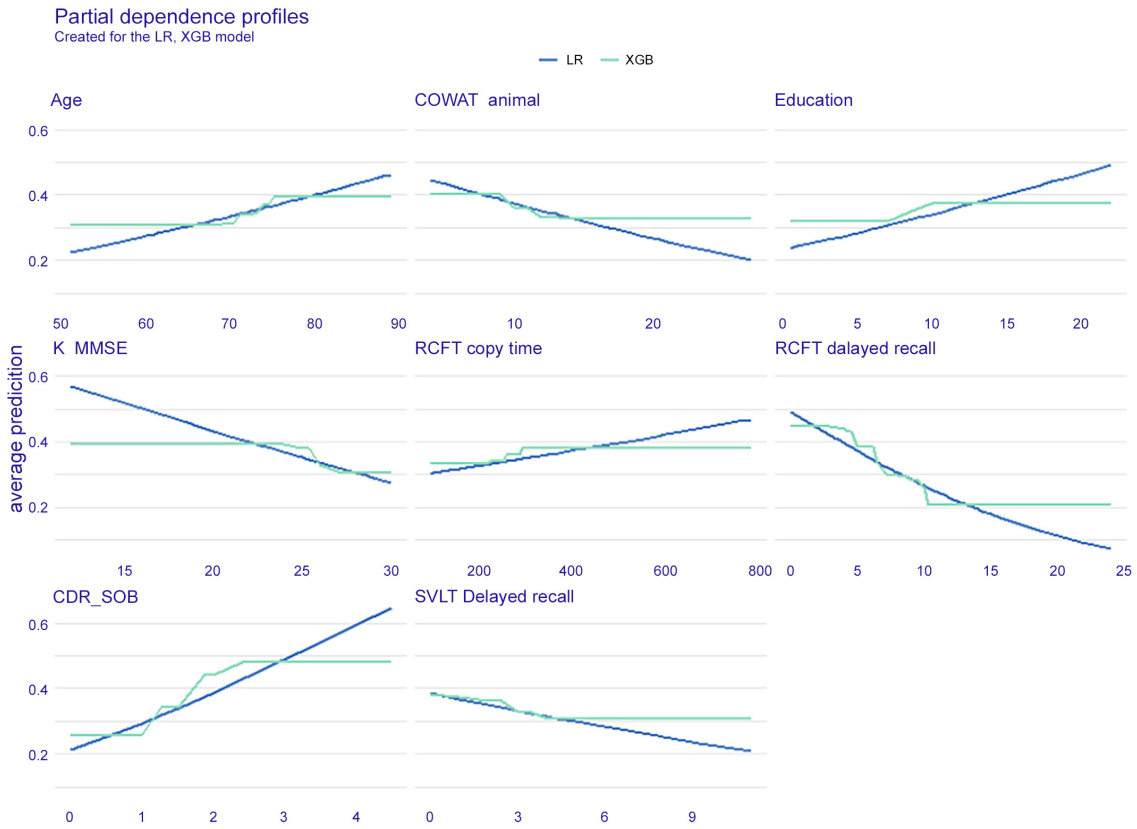


FIGURE 3
Partial dependence plot of six features. The extreme gradient boost (blue) model and logistic regression (green) model are presented. LR, logistic regression; XGB, extreme gradient boost; COWAT, Controlled Oral Word Association; K-MMSE, Korean version of the Mini-Mental State Examination; RCFT, Rey–Osterrieth Complex Figure Test; CDR-SOB, clinical dementia rating-sum of boxes; SVLT, Seoul Verbal Learning Test.

the participants were observed to convert to dementia within 3 years. In the validation set, 50 out of 140 participants (35.7%) converted to dementia, which is similar to the conversion rate in the development set. Among participants who converted to dementia, 90.2% (n = 229) progressed to clinical AD–type dementia by meeting the core clinical criteria for probable AD (McKhann et al., 2011), and 9.8% to other types of dementia including subcortical vascular dementia (n = 12, 4.7%), frontotemporal dementia (n = 2, 0.8%), dementia with Lewy bodies (n = 2, 0.8%), and others (n = 9, 3.5%).

The following 19 features were used for model building: age, gender, education, *APOE* ε2, *APOE* ε4, K-BNT, ideomotor apraxia, calculation total score, RCFT copy score, RCFT copy time, SVLT delayed recall, SVLT recognition score, RCFT delayed recall, RCFT recognition score, contrasting program, go/no-go test, COWAT animal, K-MMSE, and CDR-SOB.

Global interpretation

The global interpretation results on the three methods are as follows:

Algorithm performance

The performance of the developed classifiers on validation set and the optimized hyperparameters is shown in **Table 2**. The XGB model showed the highest performance (accuracy 0.807, AUC 0.852) compared to the other models. **Figure 1A** shows the receiver operating characteristic curve of the developed classifiers. Statistical tests showed that the AUCs of the XGB and the LR models were significantly different (P-value $<$ 0.05). The hyperparameters of best performed XGB model was as follows: booster = gbtree, eta = 0.1, max_depth = 6, min_child_weight = 17, subsample = 0.81, colsample_bytree = 0.66. The hyperparameters of other models were as follows: mtry = 4 for RF, sigma = 0.020 and C = 0.849 for SVM, and size = 4 and decay = 0.32 for ANN. We determined the XGB to be the best-performing classifer and proceeded with the model interpretation. Also, we divided test set into 4 groups by gender and age: (1) age $<$ 70 and male (n = 20), (2) age $<$ 70 and female (n = 29), (3) age $\geq$ 70 and male (n = 36), (4) age $\geq$ 70 and female (n = 55). The prediction result from XGB model of each group was (1) 0.902, (2) 0.838, (3) 0.865, and (4) 0.828, respectively (**Figure 1B**).

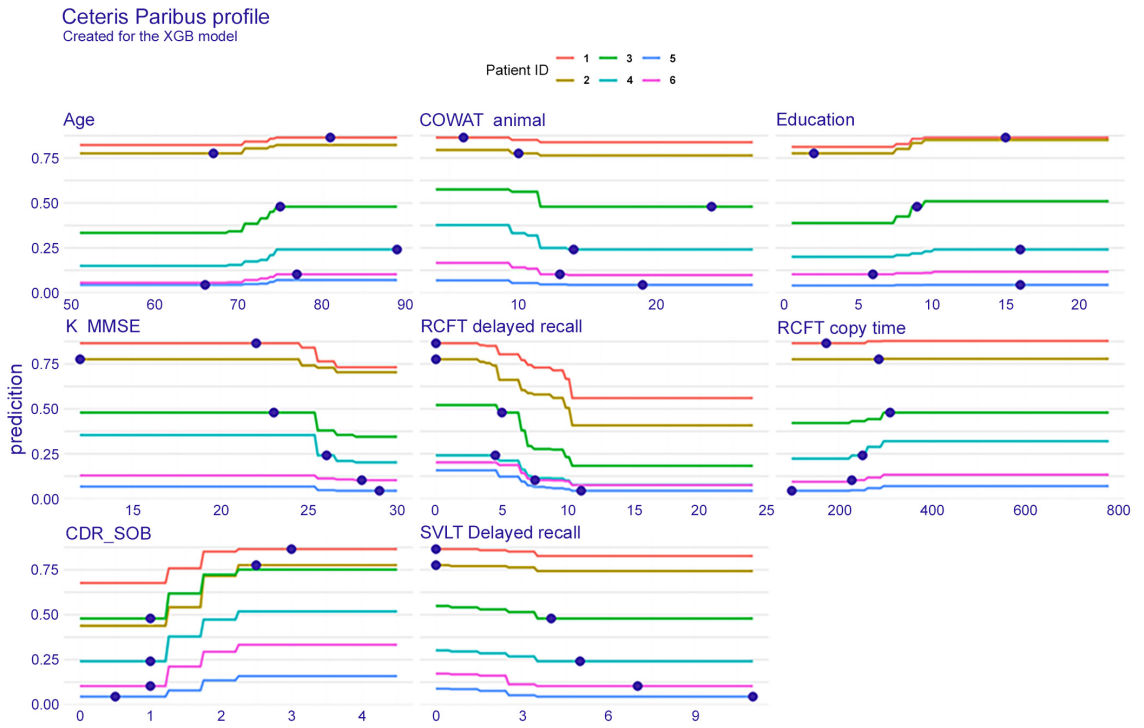


FIGURE 4
Individual conditional expectation on eight features when predicted with the extreme gradient boost model. The results of a total of six patients are plotted in different line colors. XGB, extreme gradient boost; COWAT, Controlled Oral Word Association; K-MMSE, Korean version of the Mini-Mental State Examination; RCFT, Rey–Osterrieth Complex Figure Test; CDR-SOB, clinical dementia rating-sum of boxes; SVLT, Seoul Verbal Learning Test.

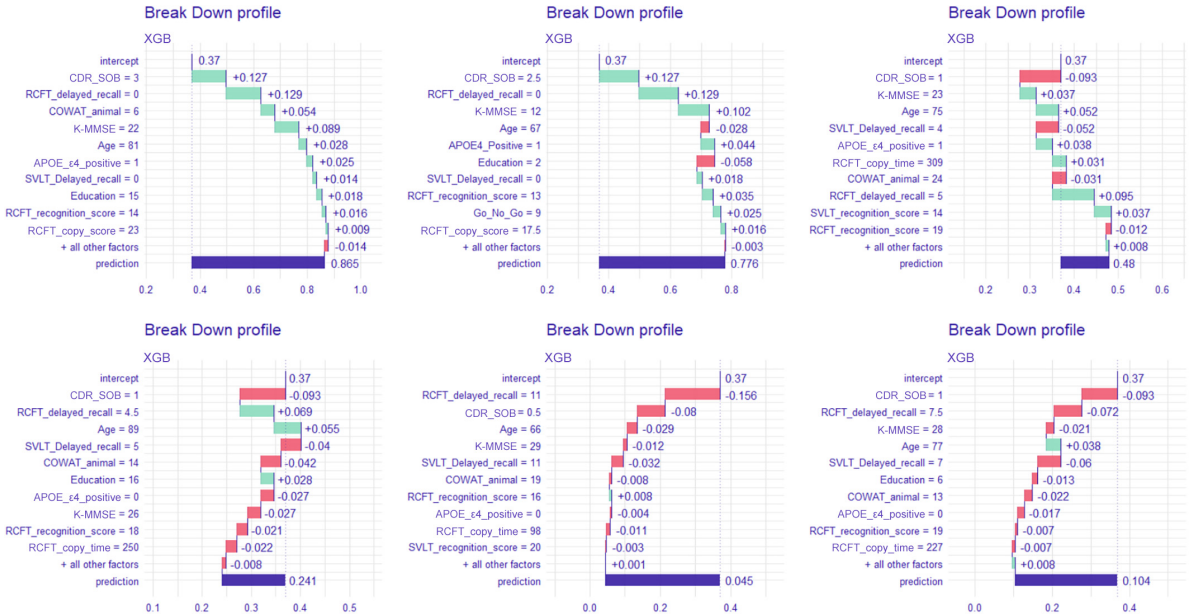


FIGURE 5
Break-down plot on six patients when predicted with the extreme gradient boost model. XGB, extreme gradient boost; CDR-SOB, clinical dementia rating-sum of boxes; RCFT, Rey–Osterrieth Complex Figure Test; K-MMSE, Korean version of the Mini-Mental State Examination; COWAT, Controlled Oral Word Association; SVLT, Seoul Verbal Learning Test; APOE, apolipoprotein E.

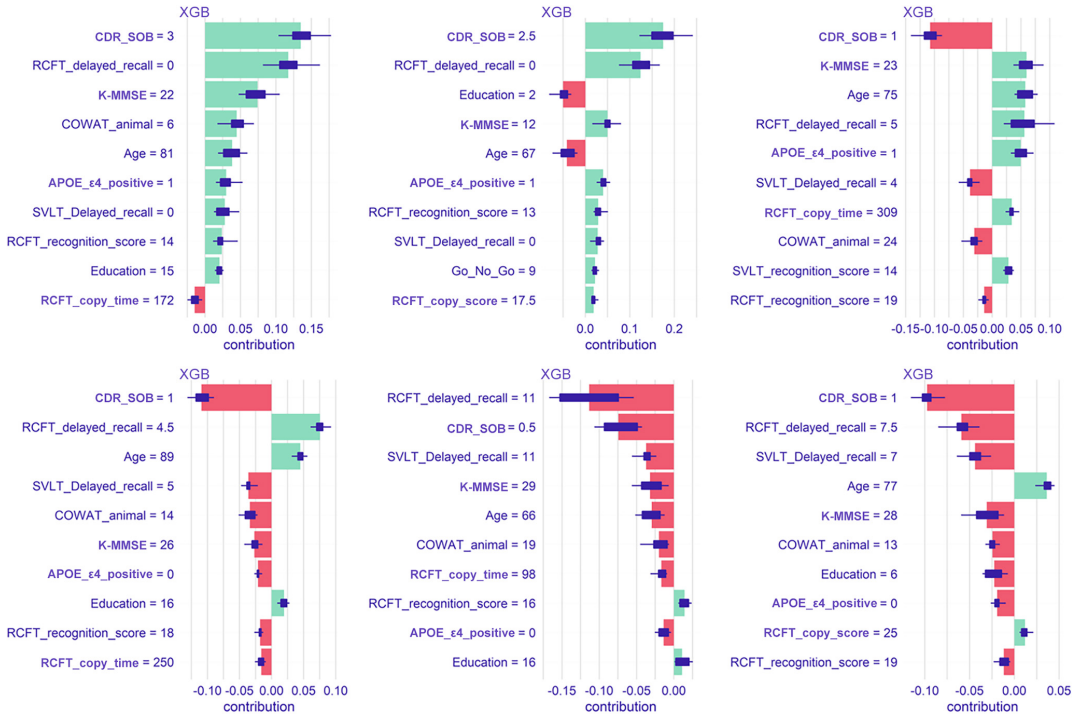


FIGURE 6
Shapley values plot of six patients when predicted with the extreme gradient boost model. XGB, extreme gradient boost; CDR-SOB, clinical dementia rating-sum of boxes; RCFT, Rey–Osterrieth Complex Figure Test; K-MMSE, Korean version of the Mini-Mental State Examination; COWAT, Controlled Oral Word Association; SVLT, Seoul Verbal Learning Test; APOE, apolipoprotein E.

Feature importance

Figure 2 shows feature importance of XGB, where the bars indicate feature importance, and the interval bands indicate difference due to random permutations. According to the result, clinical neuropsychological features of RCFT, CDR-SOB, as well as age were important factors to the global performance.

Partial independence

In Figure 3, the PDP of six features is shown with the XGB and LR models. It can be explained that under the condition that other features are fixed, the possibility of dementia conversion increases with age, while it decreases when the RCFT delayed recall score increases. The slope patterns of the XGB and LR were similar.

Local interpretation

The local interpretation results on three methods are as follows.

Individual conditional expectations

Figure 4 shows the ICE plot, which presents eight features for six individuals. To explain the result on patient number 3 (green line), the probability of dementia conversion increases between the ages of 70 and 75 years. The age of this patient is 75 years as seen in a blue dot on the green line, the interpretation plot shows the prediction value (y-axis), that is, the conversion probability, indicating approximately 0.5 within 3 years. Likewise, regarding RCFT delayed recall, this subject scored 5; therefore, the conversion possibility was approximately 0.5. If the patient had performed the test better and obtained a higher score, the conversion probability would be reduced.

Break-down plots

Figure 5 shows the break-down plots in six individuals, with the XGB model. In patient number 1, the most upper left plot, the subject had a sum of box value of 3, which attributes as much as 0.127 to the baseline mean prediction value of 0.36. In the same way, the RCFT delayed recall value of 0 contributes as much as 0.127 to the prediction.

SHapley Additive exPlanations

Figure 6 shows Shapley values plot of six individuals. In patient number 1 (the most upper left plot), the feature that contributed the most to predicting dementia conversion is the CDR-SOB. In patient number 5 (lower middle plot), RCFT delayed recall contributed most to the conversion.

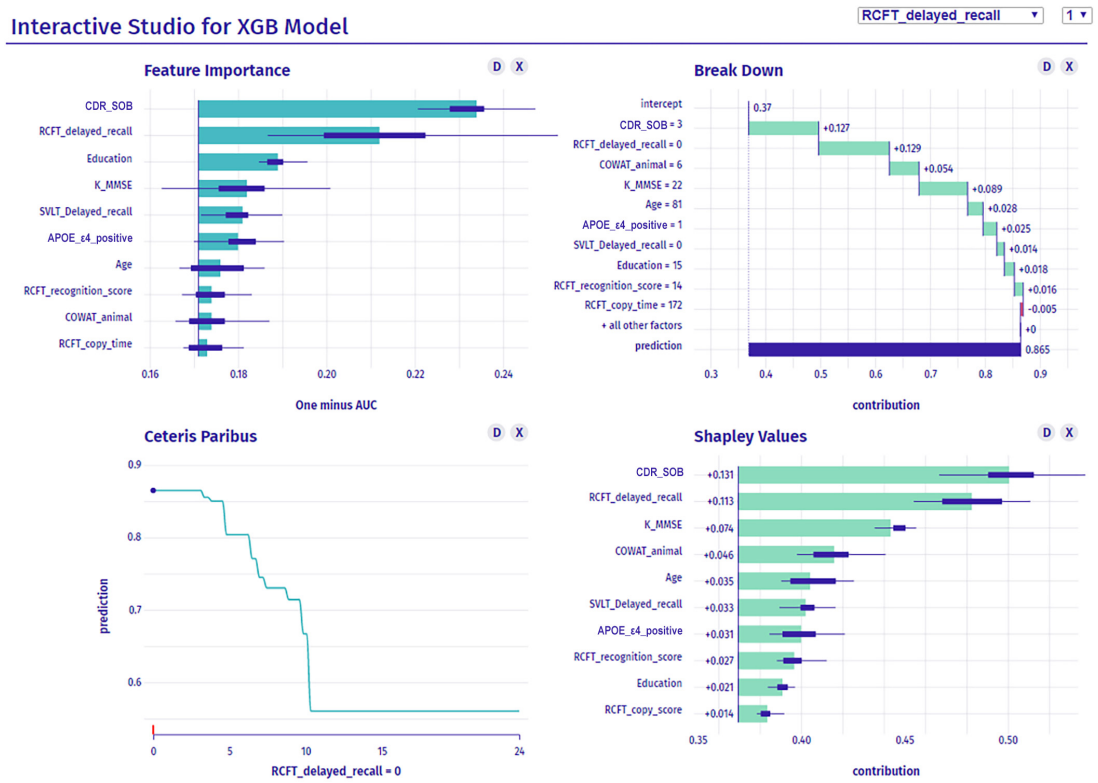


FIGURE 7
Dashboard for a patient's interpretation for predicting dementia conversion. XGB, extreme gradient boost; CDR-SOB, clinical dementia rating-sum of boxes; RCFT, Rey–Osterrieth Complex Figure Test; K-MMSE, Korean version of the Mini-Mental State Examination; SVLT, Seoul Verbal Learning Test; APOE, apolipoprotein E; COWAT, Controlled Oral Word Association.

Graphic-based overall interpretation on individuals

Figure 7 shows the dashboard displaying the global and the local interpretation of patient 1. We collected all the IML results above and developed a dashboard that provides a graphical view of each patient's analysis results by displaying them on a screen (Figure 7). It not only provides the probability of aMCI to dementia conversion, but also presents quantitative information on the risk factors attributed to the conversion.

Discussion

In the present study, using the clinical and neuropsychological features of carefully phenotyped aMCI patients, we developed an algorithm to predict conversion to dementia by applying the IML technique. Our major findings are as follows. First, among the ML techniques, the XGB model showed the best accuracy, which was superior to that of LR. Second, variables, such as visual memory delayed recall, CDR-SOB, age, K-MMSE score, frontal executive function, education, verbal memory delayed recall, visuospatial function, and *APOE* genotype were important features for creating the algorithm. Finally, ICE and SHAP analyses allowed for the interpretation of variables acted as important factors in the conversion to dementia of each aMCI patient. Taken together, our findings suggest that an algorithm using the IML technique enables us to individually predict the conversion of patients with aMCI to dementia within 3 years in clinical practice and the research field. Using our newly developed IML algorithm, we predict that, with the aid of visualized graphs, patients will be able to more easily understand the neuropsychological factors that are at risk, which would become a further step toward precision medicine.

In the present study, when compared with other algorithms including LR, the XGB model showed the best performance with an AUC of 0.852 and an accuracy of 0.807. Thus, these findings suggest that our newly developed algorithm with the XGB model overcomes this limitation and results in better AUC and accuracy than LR. If the predictive algorithm is applied to the electronic medical record system, the conversion rate would be readily calculated in clinical practice with more accuracy.

The second major finding was that RCFT delayed recall, CDR-SOB, age, K-MMSE, COWAT-animal, education, SVLT

delayed recall, RCFT copy time, and *APOE* genotype were the important factors in the IML algorithm, which is consistent with previous studies. Consistent with our findings, MMSE (Hou et al., 2019), CDR-SOB (Daly et al., 2000; Dickerson et al., 2007; Montano et al., 2013; Woolf et al., 2016), and frontal/executive dysfunction, which can be examined by the COWAT-animal test (Lezak et al., 2004), were found to be the predictors of conversion to dementia in other studies (Tabert et al., 2006; Jung et al., 2020). The *APOE* ε4 genotype was also found to play an important role in conversion to dementia, which was again consistent with previous studies (Petersen et al., 1995; Mosconi et al., 2004; Elias-Sonnenschein et al., 2011).

In our previous studies (Ye et al., 2015; Jang et al., 2017), the odds ratio of conversion to dementia was higher in Verbal-aMCI patients than in Visual-aMCI patients. However, our global interpretation results showed that the RCFT delayed recall score (visual memory) had higher feature importance than the SVLT delayed recall score (verbal memory), which is thought to be due to differences in the classification of participants. The previous studies defined Visual-aMCI as only visual memory impairment, Verbal-aMCI as only verbal memory impairment, and Both-aMCI as visual and verbal memory impairment, and then analyzed the odds ratio compared to Visual-aMCI. On the other hand, we analyzed the variables of the RCFT delayed recall score and SVLT delayed recall score together with other neuropsychological test scores of all participants without classification.

There are also some debates on the educational effects in participants with aMCI among studies. Specifically, a previous study (Cooper et al., 2015) did not show that high educational levels predict conversion to dementia in participants with aMCI. However, another study from our group showed that highly educated aMCI participants were at a higher risk of conversion to AD dementia than less educated aMCI participants (Ye et al., 2013). Furthermore, early stage aMCI participants with higher levels of education showed a slower cognitive decline while late-stage aMCI participants with higher levels of education showed a more rapid cognitive decline. Thus, our present findings that aMCI patients with higher education levels were more likely to convert to dementia should be replicated in the future studies with larger MCI participants.

Some studies have proposed an algorithm for differentiating cognitive decline using ML methods, including the Disease State Index, naïve Bayes, Bayesian network classifier with inverse tree structure, decision tree, SVM, multiple-layer perceptrons, Begging, RF, and rule-based classifier (Chen and Herskovits, 2010; Hall et al., 2015; So et al., 2017; Bansal et al., 2018; Bhagyashree et al., 2018; Zhu et al., 2020). Beheshti et al. also developed a predictive algorithm with feature ranking and a genetic algorithm, which can predict the conversion rate to dementia after 3 years (Beheshti et al., 2017). However, compared to previous studies, the present study is meaningful in that we predicted the conversion of aMCI to dementia with IML, especially by presenting the attribution of each feature to the prediction. Thus, the IML predictive algorithm used in our study might be more useful in clinical practice because it is composed of clinical data that are widely and commonly used for evaluating cognition status.

Our final major finding was that our IML, which consisted of the ICE and SHAP analyses, allowed for the interpretation of variables that acted as important factors in the conversion to dementia in each patient. Therefore, we suggest that our IML is an improved predictive algorithm that has both the high accuracy of ML and the advantage of the nomogram. Identifying the specific factors that influence conversion to dementia for each aMCI patient will be helpful for the development of personalized intervention strategies in the future.

To our knowledge, our study is the first to develop an IML algorithm to predict conversion to dementia within a large sample size of well-phenotyped aMCI patients. Another strength of this study is that the IML algorithm was based on variables that are most commonly used in clinical practice, specifically neuropsychological test results and *APOE* genotype. However, this study has some limitations. First, MRI volumetry and cortical thickness, which are highly correlated with neurodegenerative dementia, were not used in this algorithm. Future studies incorporating structural brain MRI information are required to achieve higher predictive power. Second, since we did not perform amyloid and tau positron emission tomography in all participants, we could not determine the biomarker guided diagnosis in our participants. Third, the number of samples to train the model might not be large enough because of the limited number of subjects of 3-year followed-up. Finally, since this study was conducted only at SMC, there is a limitation regarding the generalizability of the outcomes. External validation in an independent cohort should be conducted in the future. Nevertheless, our study is noteworthy in demonstrating that the IML algorithm is able to estimate the individual risk of conversion to dementia in each aMCI patient.

Conclusion

This study was able to develop an IML algorithm to predict conversion to dementia in aMCI patients. This IML algorithm is expected to be useful in clinical practice and the research field as it can identify the degree to which individual risk factors influence each patient.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board of Samsung Medical Center. The patients/participants provided their written informed consent to participate in this study.

Author contributions

MC, CP, and JK: conceptualization and formal analysis and investigation. CP and JK: methodology. MC and CP: writing – original draft preparation. JJ, HJ, KK, and SS: writing – review and editing. SS: funding acquisition. KK and SS: supervision. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This research was supported by a grant of the Korean Health Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (HI19C1132); a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health and Welfare and Ministry of science and ICT, Republic of Korea (grant numbers: HU20C0111 and HU22C0170); the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2019R1A5A2027340); Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-02068, Artificial Intelligence Innovation Hub); Future Medicine 20*30 Project of the Samsung Medical Center (#SMX1220021); and the "National Institute of Health" research project (2021-ER1006-01).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.898940/full#supplementary-material

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., et al. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 270–279. doi: 10.1016/j.jalz.2011.03.008

Alegret, M., Cuberas-Borrós, G., Espinosa, A., Valero, S., Hernández, I., Ruíz, A., et al. (2014). Cognitive, genetic, and brain perfusion factors associated with four year incidence of Alzheimer's disease from mild cognitive impairment. *J. Alzheimer's Dis.* 41, 739–748. doi: 10.3233/JAD-132516

Bansal, D., Chhikara, R., Khanna, K., and Gupta, P. (2018). Comparative analysis of various machine learning algorithms for detecting dementia. *Procedia Comput. Sci.* 132, 1497–1502. doi: 10.1016/j.procs.2018.05.102

Beheshti, I., Demirel, H., and Matsuda, H. (2017). Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. *Comput. Biol. Med.* 83, 109–119. doi: 10.1016/j.compbiomed.2017.02.011

Bhagyashree, S. I. R., Nagaraj, K., Prince, M., Fall, C. H. D., and Krishna, M. (2018). Diagnosis of dementia by machine learning methods in epidemiological studies: a pilot exploratory study from south India. *Soc. Psychiatry Psychiatr. Epidemiol.* 53, 77–86. doi: 10.1007/s00127-017-1410-0

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York, NY: Springer-Verlag.

Breiman, L. (2001). Statistical modeling: the two cultures (with comments and a rejoinder by the author). *Stat. Sci.* 16, 199–231.

Busse, A., Hensel, A., Gühne, U., Angermeyer, M. C., and Riedel-Heller, S. G. (2006). Mild cognitive impairment: long-term course of four clinical subtypes. *Neurology* 67, 2176–2185. doi: 10.1212/01.wnl.0000249117.23318.e1

Chen, R., and Herskovits, E. H. (2010). Machine-learning techniques for building a diagnostic model for very mild dementia. *Neuroimage* 52, 234–244. doi: 10.1016/j.neuroimage.2010.03.084

Cooper, C., Sommerlad, A., Lyketsos, C. G., and Livingston, G. (2015). Modifiable predictors of dementia in mild cognitive impairment: a systematic review and meta-analysis. *Am. J. Psychiatry* 172, 323–334.

Daly, E., Zaitchik, D., Copeland, M., Schmahmann, J., Gunther, J., and Albert, M. (2000). Predicting conversion to Alzheimer disease using standardized clinical information. *Arch. Neurol.* 57, 675–680. doi: 10.1001/archneur.57.5.675

De Simone, M. S., Perri, R., Fadda, L., Caltagirone, C., and Carlesimo, G. A. (2019). Predicting progression to Alzheimer's disease in subjects with amnestic mild cognitive impairment using performance on recall and recognition tests. *J. Neurol.* 266, 102–111. doi: 10.1007/s00415-018-9108-0

DeCarli, C., Mungas, D., Harvey, D., Reed, B., Weiner, M., Chui, H., et al. (2004). Memory impairment, but not cerebrovascular disease, predicts progression of MCI to dementia. *Neurology* 63, 220–227. doi: 10.1212/01.wnl.0000130531.90205.ef

DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44, 837–845.

Dickerson, B. C., Sperling, R. A., Hyman, B. T., Albert, M. S., and Blacker, D. (2007). Clinical prediction of Alzheimer disease dementia across the spectrum of mild cognitive impairment. *Arch. Gen. Psychiatry* 64, 1443–1450. doi: 10.1001/archpsyc.64.12.1443

Dubois, B., Slachevsky, A., Litvan, I., and Pillon, B. (2000). The FAB: a frontal assessment battery at bedside. *Neurology* 55, 1621–1626.

Elias-Sonnenschein, L. S., Viechtbauer, W., Ramakers, I. H., Verhey, F. R., and Visser, P. J. (2011). Predictive value of APOE-epsilon4 allele for progression from MCI to AD-type dementia: a meta-analysis. *J. Neurol. Neurosurg. Psychiatry* 82, 1149–1156. doi: 10.1136/jnnp.2010.231555

Espinosa, A., Alegret, M., Valero, S., Vinyes-Junqué, G., Hernández, I., Mauleón, A., et al. (2013). A longitudinal follow-up of 550 mild cognitive impairment patients: evidence for large conversion to dementia rates and detection of major risk factors involved. *J. Alzheimer's Dis.* 34, 769–780.

Fischer, P., Jungwirth, S., Zehetmayer, S., Weissgram, S., Hoenigschnabl, S., Gelpi, E., et al. (2007). Conversion from subtypes of mild cognitive impairment to Alzheimer dementia. *Neurology* 68, 288–291. doi: 10.1212/01.wnl.0000252358.03285.9d

Fisher, A., Rudin, C., and Dominici, F. (2019). All Models are wrong, but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously. *J. Mach. Learn. Res.* 20, 177.

Flicker, C., Ferris, S. H., and Reisberg, B. (1991). Mild cognitive impairment in the elderly: predictors of dementia. *Neurology* 41, 1006–1009. doi: 10.1212/wnl.41.7.1006

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Ann. Statist.* 29, 1189–1232.

Goldstein, A., Kapelner, A., Bleich, J., and Pitkin, E. (2015). Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. *J. Comput. Graph. Stat.* 24, 44–65.

Hall, A., Muñoz-Ruiz, M., Mattila, J., Koikkalainen, J., Tsolaki, M., Mecocci, P., et al. (2015). Generalizability of the disease state index prediction model for identifying patients progressing from mild cognitive impairment to Alzheimer's disease. *J. Alzheimers Dis.* 44, 79–92. doi: 10.3233/jad-140942

Hou, X. H., Feng, L., Zhang, C., Cao, X. P., Tan, L., and Yu, J. T. (2019). Models for predicting risk of dementia: a systematic review. *J. Neurol. Neurosurg. Psychiatry* 90, 373–379. doi: 10.1136/jnnp-2018-318212

Jang, H., Ye, B. S., Woo, S., Kim, S. W., Chin, J., Choi, S. H., et al. (2017). Prediction model of conversion to dementia risk in subjects with amnestic mild cognitive impairment: a longitudinal. Multi-Center Clinic-Based Study. *J. Alzheimers Dis.* 60, 1579–1587. doi: 10.3233/JAD-170507

Jung, Y. H., Park, S., Jang, H., Cho, S. H., Kim, S. J., Kim, J. P., et al. (2020). Frontal-executive dysfunction affects dementia conversion in patients with amnestic mild cognitive impairment. *Sci. Rep.* 10:772. doi: 10.1038/s41598-020-57525-6

Kang, I. W., Beom, I. G., Cho, J. Y., and Son, H. R. (2016). Accuracy of Korean-mini-mental status examination based on seoul neuro-psychological screening battery ii results. *Korean J. Fam. Med.* 37, 177–181. doi: 10.4082/kjfm.2016.37.3.177

Kang, Y., Chin, J., Na, D. L., Lee, J., and PArk, J. (2000). Brief Report: a normative study of the korean version of controlled oral word association test (COWAT) in the elderly. *Korean J. Clin. Psychol.* 19, 385–392.

Kang, Y., and Na, D. L. (2003). *Seoul Neuropsychological Screening Battery*. Seoul: Human Brain Research & Consulting Co.

Kim, H., and Na, D. L. (1999). BRIEF REPORT normative data on the Korean version of the Boston naming test. *J. Clin Exp. Neuropsychol.* 21, 127–133. doi: 10.1076/jcen.21.1.127.942

Larrieu, S., Letenneur, L., Orgogozo, J. M., Fabrigoule, C., Amieva, H., Le Carret, N., et al. (2002). Incidence and outcome of mild cognitive impairment in a population-based prospective cohort. *Neurology* 59, 1594–1599. doi: 10.1212/01.wnl.0000034176.07159.f8

Lezak, M. D., Howieson, D. B., and Loring, D. W. (2004). *Neuropsychological Assessment*. New York, NY: Oxford University Press.

Lian, C., Liu, M., Wang, L., and Shen, D. (2021). Multi-Task weakly-supervised attention network for dementia status estimation with structural MRI. *IEEE Trans Neural Netw Learn Syst* *, doi: 10.1109/tnnls.2021.3055772

Lundberg, S. M., and Lee, S. I. (2017). A unified approach to interpreting model predictions. *Adv. Neural Inf. Proc. Syst* 30, 4768–4777.

Mattila, J., Soininen, H., Koikkalainen, J., Rueckert, D., Wolz, R., Waldemar, G., et al. (2012). Optimizing the diagnosis of early Alzheimer's disease in mild cognitive impairment subjects. *J. Alzheimers Dis.* 32, 969–979. doi: 10.3233/JAD-2012-120934

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. Jr., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005

Mitchell, A. J., and Shiri-Feshki, M. (2009). Rate of progression of mild cognitive impairment to dementia–meta-analysis of 41 robust inception cohort studies. *Acta. Psychiatr. Scand.* 119, 252–265. doi: 10.1111/j.1600-0447.2008.01326.x

Montano, M. B., Andreoni, S., and Ramos, L. R. (2013). Clinical dementia rating independently predicted conversion to dementia in a cohort of urban elderly in Brazil. *Int. Psychogeriatr.* 25, 245–251. doi: 10.1017/S1041610212001615

Mosconi, L., Perani, D., Sorbi, S., Herholz, K., Nacmias, B., Holthoff, V., et al. (2004). MCI conversion to dementia and the APOE genotype: a prediction study with FDG-PET. *Neurology* 63, 2332–2340. doi: 10.1212/01.wnl.0000147469.18313.3b

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., and Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proc. Natl. Acad. Sci. U.S.A.* 116, 22071–22080. doi: 10.1073/pnas.1900654116

Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., et al. (2001). Current concepts in mild cognitive impairment. *Arch. Neurol.* 58, 1985–1992. doi: 10.1001/archneur.58.12.1985

Petersen, R. C., Smith, G. E., Ivnik, R. J., Tangalos, E. G., Schaid, D. J., Thibodeau, S. N., et al. (1995). Apolipoprotein E status as a predictor of the development of Alzheimer's disease in memory-impaired individuals. *JAMA* 273, 1274–1278.

Qiao, H., Chen, L., Ye, Z., and Zhu, F. (2021). Early Alzheimer's disease diagnosis with the contrastive loss using paired structural MRIs. *Comput Methods Prog. Biomed.* 208:106282. doi: 10.1016/j.cmpb.2021.106282

Ravaglia, G., Forti, P., Maioli, F., Martelli, M., Servadei, L., Brunetti, N., et al. (2006). Conversion of mild cognitive impairment to dementia: predictive role of mild cognitive impairment subtypes and vascular risk factors. *Dement. Geriatr. Cogn. Disord.* 21, 51–58. doi: 10.1159/000089515

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/

Robnik-Šikonja, M., and Kononenko, I. (2008). Explaining classifications for individual instances. *IEEE Trans. Knowl. Data Eng.* 20, 589–600. doi: 10.1109/TKDE.2007.190734

Sarazin, M., Berr, C., De Rotrou, J., Fabrigoule, C., Pasquier, F., Legrain, S., et al. (2007). Amnestic syndrome of the medial temporal type identifies prodromal AD: a longitudinal study. *Neurology* 69, 1859–1867. doi: 10.1212/01.wnl.0000279336.36610.f7

So, A., Hooshyar, D., Park, K. W., and Lim, H. S. (2017). Early Diagnosis of Dementia from Clinical Data by Machine Learning Techniques. *Appl. Sci.* 7:651. doi: 10.3390/app7070651

Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., et al. (2006). Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Arch. Gen. Psychiatry* 63, 916–924. doi: 10.1001/archpsyc.63.8.916

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J. Clin. Epidemiol.* 49, 1225–1231. doi: 10.1016/s0895-4356(96)00002-9

Waljee, A. K., Higgins, P. D., and Singal, A. G. (2014). A primer on predictive models. *Clin. Transl. Gastroenterol.* 5:e44. doi: 10.1038/ctg.2013.19

Woolf, C., Slavin, M. J., Draper, B., Thomassen, F., Kochan, N. A., Reppermund, S., et al. (2016). Can the clinical dementia rating scale identify mild cognitive impairment and predict cognitive and functional decline? *Dement Geriatr. Cogn. Disord.* 41, 292–302. doi: 10.1159/000447057

Yaffe, K., Petersen, R. C., Lindquist, K., Kramer, J., and Miller, B. (2006). Subtype of mild cognitive impairment and progression to dementia and death. *Dement Geriatr. Cogn. Disord.* 22, 312–319. doi: 10.1159/000095427

Ye, B. S., Chin, J., Kim, S. Y., Lee, J. S., Kim, E. J., Lee, Y., et al. (2015). The heterogeneity and natural history of mild cognitive impairment of visual memory predominant type. *J. Alzheimers Dis.* 43, 143–152. doi: 10.3233/JAD-140318

Ye, B. S., Seo, S. W., Cho, H., Kim, S. Y., Lee, J.-S., Kim, E.-J., et al. (2013). Effects of education on the progression of early-versus late-stage mild cognitive impairment. *Int. Psychogeriatr.* 25, 597–606. doi: 10.1017/S1041610212002001

Zhu, F., Li, X., Tang, H., He, Z., Zhang, C., Hung, G.-U., et al. (2020). Machine learning for the preliminary diagnosis of dementia. *Sci. Prog.* 2020:5629090. doi: 10.1155/2020/5629090

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership