ARTIFICIAL INTELLIGENCE FOR EHEALTH

EDITED BY: Deepak Gupta, Joel J. P. C. Rodrigues, Sheng-Lung Peng and Nhu Nguyen

PUBLISHED IN: Frontiers in Public Health

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-976-8
DOI 10.3389/978-2-88974-976-8

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

ARTIFICIAL INTELLIGENCE FOR EHEALTH

Topic Editors:
Deepak Gupta, Maharaja Agrasen Institute of Technology, India
Joel J. P. C. Rodrigues, Federal University of Piauí, Brazil
Sheng-Lung Peng, National Dong Hwa University, Taiwan
Nhu Nguyen, Duy Tan University, Vietnam

Citation: Gupta, D., Rodrigues, J. J. P. C., Peng, S.-L., Nguyen, N., eds. (2022). Artificial Intelligence for Ehealth. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88974-976-8

Table of Contents

frontiers in Public Health

EDITORIAL
published: 04 March 2022
doi: 10.3389/fpubh.2022.852840

Editorial: Artificial Intelligence for eHealth

Deepak Gupta[1]*, *Joel J. P. C. Rodrigues*[2], *Sheng-Lung Peng*[3] and *Nhu Nguyen*[4]

[1] Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, New Delhi, India, [2] Senac Faculty of Ceará, Fortaleza, Brazil, [3] Department of Creative Technologies and Product Design, National Taipei University of Business, Taipei, Taiwan, [4] Duy Tan University, Da Nang, Vietnam

Keywords: artificial intelligence (AI), eHealth, computational intelligence, COVID-19, neural network

Editorial on the Research Topic

Editorial: Artificial Intelligence for eHealth

Artificial intelligence has grown extensively in recent times and is changing the healthcare industry from many perspectives: Clinical Diagnosis, suggests treatment and follow up. Clinical Decision Support (CDS) is a major topic of AI in medicine to assist clinicians at point of care. Existing techniques used for processing health data can be broadly classified into two categories: (a) non-Artificial Intelligence (AI) systems and (b) Artificial Intelligence systems. Even though non-AI techniques are less complex in nature, most of the systems suffer from the drawbacks of inaccuracy and lack of convergence. Hence, these systems are generally replaced by AI based systems which are much superior to the conventional systems. AI techniques are mostly hybrid in nature and include Artificial Neural Networks (ANN), fuzzy theory, and evolutionary algorithms. AI increases the ability for healthcare professionals to better understand the day-to-day patterns and needs of the people they care for, and with that understanding they are able to provide better feedback, guidance and support for staying healthy.

AI-based CDS uses inference and logics, while non-AI-based CDS relies on machine learning to perform the same functions. There are many clinical duties that CDS may assist with, but it is essential that CDS is correctly integrated into the clinical workflow and health records. CDS can be used to assist clinicians in the interpretation of medical pictures through the use of Computer Aided Diagnosis (CAD). CAD incorporates AI as well as computer vision, signal processing, and other components relevant to medicine. Breast cancer, lung cancer, colon cancer, coronary artery disease, and Alzheimer's disease are just a few of the conditions that can benefit from CADs.

There are certain societal concerns about the expanding use of AI in healthcare, including the possibility of bias, lack of transparency for some AI algorithms, privacy problems for data used for AI model training, and security and implementation responsibilities in clinical settings.

All areas of artificial intelligence (AI) in the fields of health informatics, biomedical informatics, and medical image analysis are covered in this special issue. Based on the reviews, eight papers were chosen from a total of fifteen submissions to this special collection. At least two reviewers and at least two rounds of review were required for each paper. Listed below are some of the papers that made important contributions to this discussion.

Using data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and independent latent class analysis (LCA and LTA) and latent transition analysis (LTA) over a three-year period, the authors (Alashwal et al.), describe their findings in the first paper of this special issue. Researchers found that LCA was a better predictor of AD progression than typical clinical cut-off measures on neuropsychological exams when it came to defining and recognizing the disease.

OPEN ACCESS

Edited by:
Bruno Bonnechère, University of Oxford, United Kingdom

Reviewed by:
Dalibor Stanimirovic, National Institute for Public Health, Slovenia

***Correspondence:**
Deepak Gupta
deepakgupta@mait.ac.in

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Public Health

Received: *11 January 2022*
Accepted: *14 February 2022*
Published: *04 March 2022*

Citation:
Gupta D, Rodrigues JJPC, Peng S-L and Nguyen N (2022) Editorial: Artificial Intelligence for eHealth. Front. Public Health 10:852840. doi: 10.3389/fpubh.2022.852840

Researchers (Mishra et al.) recommend data preparation to eliminate the problem of data skewing by employing sample approaches. The K-Nearest Neighbor algorithm is used to classify the data using three different sampling techniques: Resampling, Spread Subsampling, and SMOTE. Evaluation of classification's performance is done using a variety of performance indicators to determine classification's efficiency.

The authors (Iwendi et al.) in this research offer a fine-tuned AdaBoost algorithm-boosted Random Forest model. Data from the COVID-19 patients is used to create a model that can estimate the severity of a patient's condition and whether or not they will recover or die. On the dataset used, the model has an accuracy of 94% and an F1 Score of 0.86. The study of the data shows a link between the gender of the patients and their mortality, as well as the fact that the vast majority of patients are between the ages of 20 and 70.

They are trying to find out if heat massaging of the spinal column can reduce muscle discomfort and increase antioxidant function in this study (Kim et al.). There were 60 people in the study who had lower back discomfort. Both an experimental and a control group were given spinal column heat massage and normal rehabilitative treatment, respectively, as part of their rehabilitation. According to the results of the study, spinal column thermal massage decreases pain more efficiently and improves impairment levels. Because of this, thermal massage may be beneficial in the treatment and prevention of oxidative disorders.

In the next paper (Song et al.) of this special collection shows that the pectolinarin triggers apoptotic cell death in PC12 cells by DNA fragmentation and the production of apoptotic bodies via the activation of ER stress sensors (eIF2 phosphorylation and ATF6 fragmentation) in PC12 cells. The treatment of PC12 cells with 50 μM pectolinarin for 24 h increased the mRNA expression of ATF6, PERK, and IRE1 by up to 1.6, 1.7, and 1.4 times, respectively, compared to the control. Pectolinarin administration enhanced ATF6 fragmentation by roughly twofold compared to the control, and phosphorylation of eIF2 by 2.5 fold. As a result of these findings, future natural medicines and health supplements targeting disorders caused by apoptosis could benefit from a better understanding of the molecular pathways involved.

For the purpose of estimating the number of people who will die from COVID-19-related causes in India over the next decade, the authors (Dhamodharavadhani et al.) conducted an investigation into the suitability of Statistical Neural Network (SNN) models and their hybrid version. These SNN models, including the Probabilistic Neural Network and a Radial Basis Function Neural Network, are used to construct the COVID-19 Mortality Rate Prediction model (MRP) in India. MRP models based on PNN and RBFNN were found to perform better than other models in COVID-19 datasets D2 and D1.

Based on an autoregressive integrated moving average, a model was developed to predict an epidemic of COVID-19 in the world in the next several days (ARIMA) (Dansana et al.). In addition to the 120,000 confirmed fatalities predicted by the ARIMA model until April 1, 2020, we also evaluated the total number of confirmed cases, the total number of fatalities predicted, the autocorrelation function, and the white noise time series for the COVID-19 outbreak's confirmed and fatalities cases.

Patients infected with the COVID-19 virus can have their medical issues diagnosed using a data mining model built on a hybrid deep learning framework (Khadidos et al.). Convolution neural networks (CNNs) and recurrent neural networks (RNNs) combine to form the DeepSense technique, a hybrid deep learning model. In comparison to other deep learning and machine learning classifiers, DeepSense's accuracy was shown to be significantly higher. A patient's prognosis for COVID-19 infections can be improved by knowing the accuracy of the diagnostic approach used.

To summarize, eight of the fifteen papers submitted to this special issue were accepted for publication in this special edition. We, as guest editors, believe that this special issue's research contributions and conclusions will assist readers by expanding their knowledge and inspiring them to work on a variety of elements of Artificial Intelligence for eHealth themselves.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

We would like to thank the editor-in-chief for giving us the opportunity to put together this issue. A special thanks to the editing staffs for their assistance. We would also want to express our gratitude to the authors and reviewers who contributed to this special edition.

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 Public Health

ORIGINAL RESEARCH
published: 03 July 2020
doi: 10.3389/fpubh.2020.00357

COVID-19 Patient Health Prediction Using Boosted Random Forest Algorithm

Celestine Iwendi[1], Ali Kashif Bashir[2], Atharva Peshkar[3], R. Sujatha[4], Jyotir Moy Chatterjee[5], Swetha Pasupuleti[6], Rishita Mishra[7], Sofia Pillai[8] and Ohyun Jo[9]**

[1] BCC of Central South University of Forestry and Technology, Changsha, China, [2] Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, United Kingdom, [3] Department of Information Technology, G H Raisoni College of Engineering, Nagpur, India, [4] School of Information Technology and Engineering, VIT University, Vellore, India, [5] Department of Information Technology, Lord Buddha Education Foundation, Kathmandu, Nepal, [6] School of Civil Engineering, Galgotias University, Greater Noida, India, [7] Department of Electronics and Telecommunications Engineering, G H Raisoni College of Engineering, Nagpur, India, [8] School of Civil Engineering, Galgotias University, Greater Noida, India, [9] Department of Computer Science, College of Electrical and Computer Engineering, Chungbuk National University, Cheongju-si, South Korea

OPEN ACCESS

Edited by:
Deepak Gupta,
Maharaja Agrasen Institute of Technology, India

Reviewed by:
Noor Zaman,
Taylor's University, Malaysia
Dac-Nhuong Le,
Haiphong University, Vietnam
Mumtaz Ali,
Deakin University, Australia

***Correspondence:**
Celestine Iwendi
celestine.iwendi@ieee.org
Ohyun Jo
ohyunjo@chungbuk.ac.kr

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Public Health

Received: 14 May 2020
Accepted: 23 June 2020
Published: 03 July 2020

Citation:
Iwendi C, Bashir AK, Peshkar A, Sujatha R, Chatterjee JM, Pasupuleti S, Mishra R, Pillai S and Jo O (2020) COVID-19 Patient Health Prediction Using Boosted Random Forest Algorithm. Front. Public Health 8:357. doi: 10.3389/fpubh.2020.00357

Integration of artificial intelligence (AI) techniques in wireless infrastructure, real-time collection, and processing of end-user devices is now in high demand. It is now superlative to use AI to detect and predict pandemics of a colossal nature. The Coronavirus disease 2019 (COVID-19) pandemic, which originated in Wuhan China, has had disastrous effects on the global community and has overburdened advanced healthcare systems throughout the world. Globally; over 4,063,525 confirmed cases and 282,244 deaths have been recorded as of 11th May 2020, according to the European Centre for Disease Prevention and Control agency. However, the current rapid and exponential rise in the number of patients has necessitated efficient and quick prediction of the possible outcome of an infected patient for appropriate treatment using AI techniques. This paper proposes a fine-tuned Random Forest model boosted by the AdaBoost algorithm. The model uses the COVID-19 patient's geographical, travel, health, and demographic data to predict the severity of the case and the possible outcome, recovery, or death. The model has an accuracy of 94% and a F1 Score of 0.86 on the dataset used. The data analysis reveals a positive correlation between patients' gender and deaths, and also indicates that the majority of patients are aged between 20 and 70 years.

Keywords: COVID-19, healthcare analytics, patient data, infection, boosting, random forest classification

INTRODUCTION

The healthcare industry is a vast industry that requires real time collection and processing of medical data. Moreover, at the core of this industry lies the problem of data handling which requires real time prediction and dissemination of information to practitioners for quick medical attention. Major actors of this industry, such as physicians, vendors, hospitals, and health-based companies have attempted to collect, manage, and revive data with the aim of using it to enhance medical practices and for technological innovation. However, dealing with healthcare data has, of late, become a complex task due to the massive volume of the data, security issues,

wireless network application incompetence, and the velocity at which it is increasing. Thus, to increase the efficiency, accuracy, and workflow healthcare industries need data analytics tools to manage such complex data.

Coronavirus disease 2019 (COVID-19) is a virus of the Corona virus family and the source of a respiratory illness outbreak throughout the world that originated in Wuhan, China. Studies (1–3) show that Covid-19 has clinical characteristics akin to the SARS-CoV. The dominant symptoms include fever and cough, while gastrointestinal symptoms are uncommon. In COVID-19 infected patients the absence of fever is more frequent than in patients infected by similar viruses, i.e., MERS Corona Virus (2%) and SARS Corona Virus (1%) (4); therefore, there is a possibility of non-febrile patients being missed by a surveillance mechanism with a primary focus on detecting fever (5). The initial patients infected by COVID-19, reportedly indicated an association with a large seafood and animal market in Wuhan that demonstrated an animal-to-person spread. Per contra, a burgeoning number of patients have not displayed any association with the animal markets, revealing the fact of human-to-human transmission of COVID-19. This pandemic has been declared a global health emergency and is spreading at an alarming rate (6). The origin of the virus in Wuhan, China has caused 175,694 deaths globally and has 2,544,792 active patients globally (7). With the stress on medical facilities, it is essential for governments and healthcare facilities to identify and treat cases that are most likely to survive, by so doing, judiciously utilizing the limited stock of medical resources and medications.

Artificial Intelligence (AI) has emerged as the breakthrough technology of the twenty-first century and has found multiple applications in fields from weather prediction, astronomical exploration, to autonomous systems (8). We note a few related works where AI has been applied for detection, prevention, and prediction to combat the COVID-19 pandemic. In Wang and Wong (9) researchers have implemented a Convolutional Neural Network based model to detect COVID-19 patients using CXR images. They used a pre-trained ImageNet and trained the model on an open source dataset of Chest X-Ray images (CXR). While Pal et al. (10) implemented a LSTM model to predict the country-specific risk of COVID-19, that relies on trends and weather data of a particular country to predict the probable spread of COVID-19 in that country. In Liu et al. (11) the AI practitioners applied ML to process internet activity, news reports, health organization reports, and media activity to predict the spread of the outbreak on the providence level in China (12). In Bayes and Valdivieso (13) the authors made use of the Bayesian approach to predict the number of deaths in Peru for 70 days in the future, using the empirical data from China. The authors in Beck et al. (14) applied Artificial intelligence to identify the commercially available drugs that could be used to treat COVID-19 patients. They used Bidirectional Encoder Representations from the Transformers (BERT) framework at the core of their model. In Tang et al. (15) the researchers implemented the random forest algorithm for severity analysis of COVID-19 patients using the Computed Tomography (CT) Scans. In Khalifa et al. (16) the authors proposed a Generative Adversarial Network based fine-tuned model for detecting pneumonia from Chest X-Ray scans, which is one of the symptoms of COVID-19 infection. In Sujatha et al. (17), authors proposed a method which could be helpful in predicting the stretch of COVID-2019, by performing linear regression, and the Multilayer perceptron and Vector autoregression model which could provide an expectation on the COVID-19 Kaggle information, to anticipate the epidemiological pattern of the disease and rate of COVID-2019 cases in India.

Kutia et al. (18) tried to break down client perspectives to eHealth applications in China and the eHealth framework in the Ukraine, which afterwards provided bits of knowledge and proposals for the improvement of an eHealth application (eZdorovya) for mainly health information benefits. Sultan et al. (19) presented a hybrid method that generates and facilitates Alzheimer patients to recall their memories. This egocentric video summary uses important people, objects, and medicines as tools in the realization of their method. Furthermore, an emerging tactile Internet-based nanonetwork that promises a new range of e-health applications has been proposed by Feng et al. (20). The authors use an information based transmit network that goes to an operator via the terahertz band. Finally, the authors in Jain and Chatterjee (21) presented an assortment of strategies intended to speak to, improve, and enable multi-disciplinary and multi-institutional ML to explore in healthcare informatics (22). Khamparia et al. (23) introduced a unique way of an internet of health things (IoHT)-driven deep learning structure for identification and arrangement of cervical cancer in Pap smear pictures, utilizing ideas of transfer learning. Waheed et al. (24) suggested a technique to produce manufactured chest X-ray (CXR) pictures by building up an Auxiliary Classifier Generative Adversarial Network (ACGAN) utilized model called CovidGAN. Sakarkar et al. (25) suggested a profound learning-based mechanized discovery and characterization model for fundus DR pictures.

This paper aims to fill the void of the traditional healthcare system, using machine learning (ML) algorithms to simultaneously process healthcare and travel data along with other parameters of COVID-19 positive patients, in Wuhan, to predict the most likely outcome of a patient, based on their symptoms, travel history, and the delay in reporting the case by identifying patterns from previous patient data. Our contribution includes:

- Processing of healthcare and travel data using machine learning algorithms in place of the traditional healthcare system to identify COVID infected person.
- This work compared multiple algorithms that are available for processing patient data and identified the Boosted Random Forest as the best method for processing data. Further, it executed a grid search to fine-tune the hyper parameters of the Boosted Random Forest algorithm to improve performance.
- Our work obliterates the need to re-compare existing algorithms for processing COVID-19 patient data.

TABLE 1 | Dataset description.

Column	Description	Values (for categorical variables)	Type
id	Patient Id	NA	Numeric
location	The location where the patient belongs to	Multiple cities located throughout the world	String, Categorical
country	Patient's native country	Multiple countries	String, Categorical
gender	Patient's gender	Male, Female	String, Categorical
age	Patient's age	NA	Numeric
sym_on	The date patient started noticing the symptoms	NA	Date
hosp_vis	Date when the patient visited the hospital	NA	Date
vis_wuhan	Whether the patient visited Wuhan, China	Yes (1), No (0)	Numeric, Categorical
from_wuhan	Whether the patient belonged to Wuhan, China	Yes (1), No (0)	Numeric, Categorical
death	Whether the patient passed away due to COVID-19	Yes (1), No (0)	Numeric, Categorical
Recov	Whether the patient recovered	Yes (1), No (0)	Numeric, Categorical
symptom1. symptom2, symptom3, symptom4, symptom5, symptom6	Symptoms noticed by the patients	Multiple symptoms noticed by the patients	String, Categorical

- This work will enable researchers to further work on developing a solution that combines the processing of patient demographics, travel, and subjective health data with image data (scans) for better prediction of COVID-19 patient health outcomes.

The rest of the article is organized as follows: section Materials and Methods discusses the materials and methodology used in detail, along with the dataset description, data pre-processing, and the data analysis of the classification algorithms used. Section Results discusses the result of the experiment followed by further discussion in section Discussion. Section Conclusion and Future Work discusses the results and provides a conclusion and the future direction of the current work.

MATERIALS AND METHODS

The dependencies for the project include the following packages and libraries: Datetime, Numpy, Pandas, SciPy, Scikit Learn, and Matplotlib. The project has been implemented on the Google Colab platform using the CPU runtime. The CPU specifications for Google Colab are; model: 79, CPU Family: 6, model name: Intel(R) Xeon(R) CPU @ 2.20 GHz and cache size: 56,320 KB. The storage used is Google Drive.

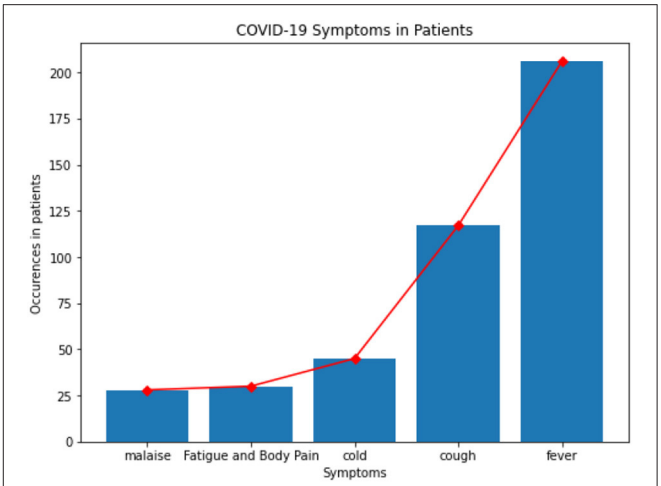

FIGURE 1 | Symptoms in patients.

Dataset

The dataset used in this study was accessed from Kaggle as "Novel Corona Virus 2019 Dataset" (26). The dataset has been compiled from various sources including the World Health Organization and John Hopkins University. However, this dataset has been pre-processed further by us to meet the needs of this study. **Table 1** presents the features of the data.

Data Analysis

Fever, cough, cold, fatigue, body pain, and malaise were the most common symptoms that were noticed in patients whose data is available in this dataset and are shown in **Figure 1**.

Correlation between features of the dataset provides crucial information about the features and the degree of influence they have over the target value. The heat map of Pearson Correlation between the features of the dataset is shown in **Figure 2**, which clearly reveals a relatively stronger positive correlation between age of the patient, whether the patient was native to Wuhan, gap between (in days) when they first felt the symptoms and visited the hospital, and death. However, the country of the patient has a positive correlation with recovery. This implies that foreign patients who visited China had a higher recovery rate. There is also a strong positive correlation between symptom1 and symptom2, and also between symptom2 and symptom3.

Data Pre-processing

The dataset consists of columns with the data being the Date, String, and Numeric type. We also have categorical variables in the dataset. Since the ML model requires all the data that is passed as input to be in the numeric form, we performed label-encoding of the categorical variables. This assigns a number to every unique categorical value in the column.

The dataset consists of multiple missing values which cause an error when passed directly as an input. Thus, we fill the

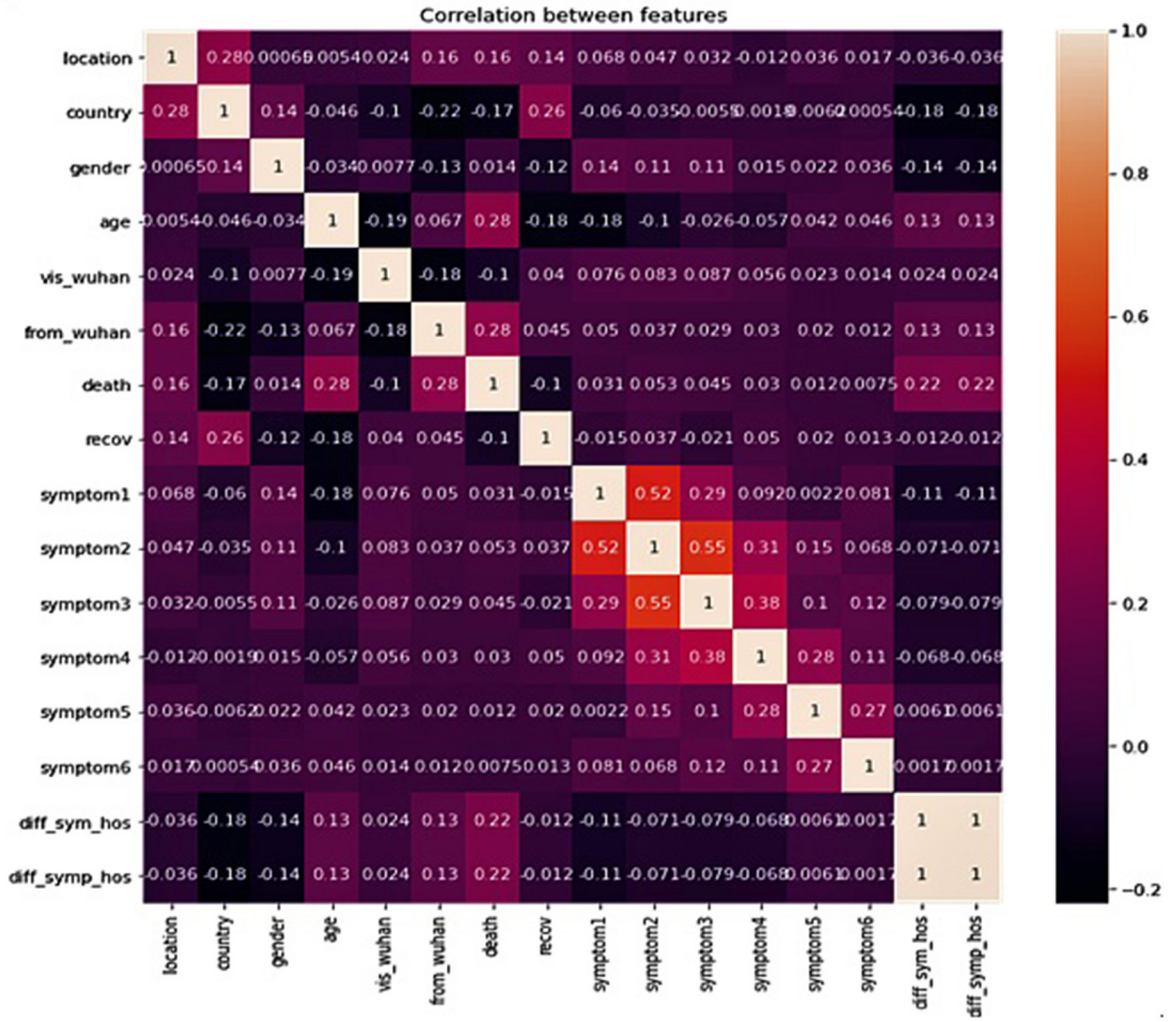

FIGURE 2 | Correlation between data features.

missing values with "NA." Certain patient data records contain missing values for both the "death" and "recov" columns, such patient records have been separated from the main dataset and compiled into the test dataset, while the remaining records have been compiled into the train dataset.

The dataset also consists of columns in the date format. Since the data columns are not directly used, feature engineering has been applied. A new column has been populated with the corresponding (hosp_vis—sym_on) value. This provides us with the number of days that have passed between the symptoms being noticed and the patient visiting the hospital.

Evaluation Metrics

The purpose of the following study is to accurately predict the outcome of a particular patient depending on multiple factors, including but not limited to travel history, demographics etc. Since this is a very crucial prediction, accuracy is very important. Thus, for the purpose of evaluating the model we considered three evaluation metrics for this study.

The following terms are used in the equations: TP, True Positive; TN, True Negative; FP, False Positive; and FN, False Negative.

Accuracy

Given a dataset consisting of $(TP + TN)$ data points, the accuracy is equal to the ratio of total correct predictions $(TP + TN + FP + FN)$ by the classifier to the total data points. Accuracy is an important measure which is used to assess the performance of the classification model. Accuracy is calculated as shown in Equation (1) as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad 0.0 < \text{Accuracy} < 1.0 \tag{1}$$

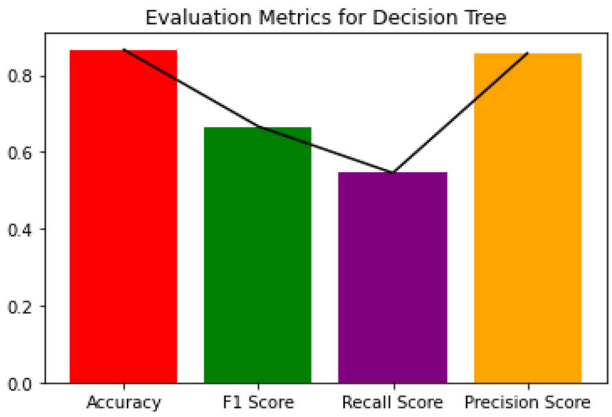

FIGURE 3 | Evaluation metrics for decision tree.

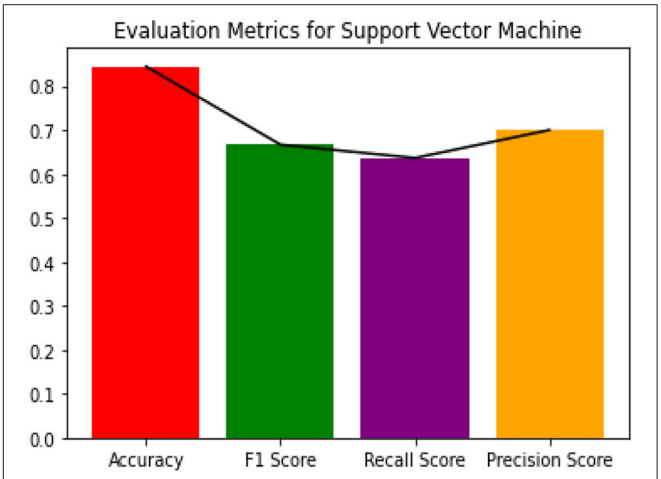

FIGURE 4 | Evaluation metrics for SVM classifier.

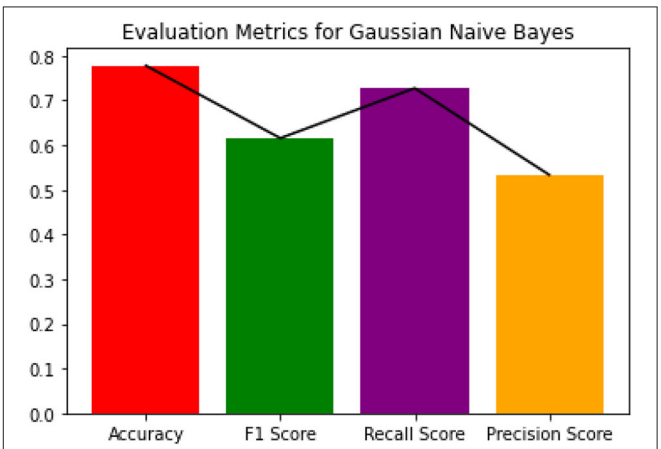

FIGURE 5 | Evaluation metrics for Gaussian NB.

Precision

Precision is equal to the ratio of the True Positive (*TP*) samples to the sum of True Positive (*TP*) and False Positive (*FP*) samples. Precision is also a key metric to identify the number of correctly classified patients in an imbalanced class dataset. Precision is calculated as given in Equation (2) as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

Recall

Recall is equal to the ratio of the True Positive (*TP*) samples to the sum of True Positive (*TP*) and False Negative (*FN*) samples. Recall is a significant metric to identify the number of correctly classified patients in an imbalanced class dataset out of all the patients that could have been correctly predicted. Recall is calculated as given in Equation (3) as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

F1 Score

F1 Score is equal to the harmonic mean of Recall and Precision value. The F1 Score strikes the perfect balance between Precision and Recall thereby providing a correct evaluation of the model's performance in classifying COVID-19 patients. This is the most significant measure that we will be using to evaluate the model. F1 Score can be calculated as shown in Equation (4) as follows:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

RESULTS

We have used the pre-processed dataset to train multiple ML classification models. The models included in this study include: Decision Tree Classifier, Support Vector Classifier, Gaussian Naïve Bayes Classifier, and Boosted Random Forest Classifier.

Since the dataset we used can be an imbalanced dataset, we will be using F1 Score as the primary metric for comparison. **Figures 3–6** shows the model performances for all the models stated above.

The decision tree constructed for estimating the target variable is visualized in **Figure 7**. The decision tree has a depth of 2 and the Gini index of every node is <0.5, which indicates an imbalance in the training data.

Since Boosted Random Forest algorithm is the best performing model, we will fine tune the model for better performance on the dataset.

DISCUSSION

Boosted Random Forest Classification

A Boosted Random Forest is an algorithm, which consists of two parts; the boosting algorithm: AdaBoost and the Random Forest classifier algorithm (27)—which in turn consists of multiple decision trees. A decision tree builds models that are similar to an

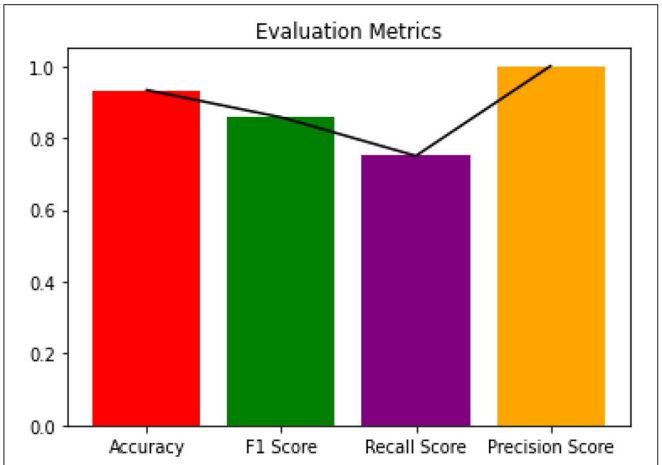

FIGURE 6 | Evaluation metrics for Boosted Random Forest.

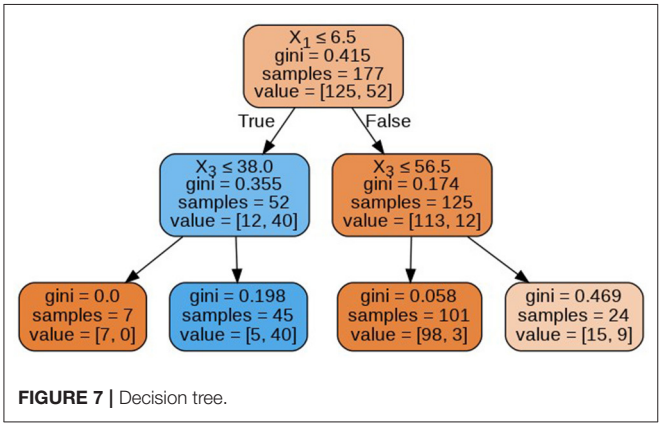

FIGURE 7 | Decision tree.

actual tree. The algorithm divides our data into smaller subsets, simultaneously adding branches to the tree. The outcome is a tree consisting of leaf nodes and decision nodes. A decision node has two or more branches representing the value of each feature (like age, symptom1, etc.) tested and the leaf node holds the result value on the patient's prospective condition (target value).

Multiple classifier decision trees (ensemble of classifiers) eliminate the risk of failure of a single decision tree to correctly predict the target value. Thus, the random forest averages the result provided by multiple trees to provide the final result.

The margin function for the random forest is expressed in Equation (5), the generalization error in Equation (6), and confidence in the prediction in Equation (7). Here $h_1(x), h_2(x), \ldots, h_k(x)$ is the ensemble of classifiers (decision trees) and the training data is drawn from the vectors X, Y.

The margin function is expressed as follows:

$$\mathrm{mg}(\mathrm{X},\mathrm{Y}) = \mathrm{av_k I}\left(\mathrm{h_k}(\mathrm{X}) = \mathrm{Y}\right) - \mathrm{max_{j \neq Y} av_k I}\left(\mathrm{h_k}(\mathrm{X}) = \mathrm{j}\right) \quad (5)$$

where the indicator function is denoted by $I(.)$. The generalization error is given as follows:

$$\mathrm{PE}^* = \mathrm{P_{X,Y}}\left(\mathrm{mg}(\mathrm{X},\mathrm{Y}) < 0\right) \quad (6)$$

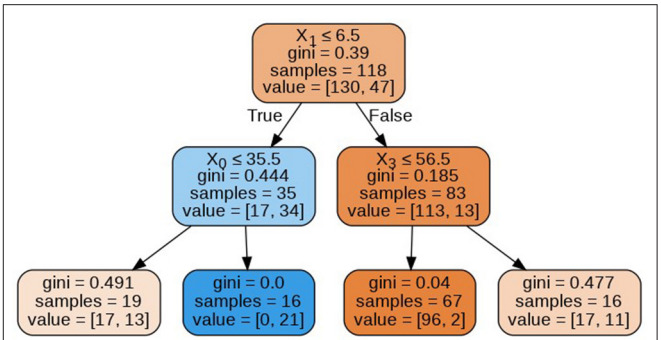

FIGURE 8 | Decision tree 1.

where the probability is expressed over the X, Y space. In random forests, we have $h_k(X) = h(X, \Theta_k)$, therefore the number of classifiers (decision trees) increases, for all the sequences of trees. The probability PE^* converges to Equation (7), from the Strong Law of Large Numbers and tree structure.

$$\mathrm{P_{X,Y}}\left(\mathrm{P_\Theta}\left(\mathrm{h}(\mathrm{X},\Theta) = \mathrm{Y}\right) - \mathrm{max_{j \neq Y} P_\Theta}\left(\mathrm{h}(\mathrm{X},\Theta) = \mathrm{j}\right) < 0\right) \quad (7)$$

Applying the boosting algorithm AdaBoost (28) provides a corrective mechanism to improve the model after every prediction of patient state. Eventually, the decision is a result of summing up of all the base models. It is one of the most efficient techniques in ML.

The corrective mechanism can be expressed as follows Equation (8). Given $(x_1, y_1), \ldots, (x_m, y_m)$, *where* $x_i \in X$, $y_i \in Y = \{-1, +1\}$. For, $t = 1, \ldots, T$. Initialize $D_1(i) = \frac{1}{m}$. After training a weak learner, random forest in our case, using distribution D_t.

Get the hypothesis, $h_t : X \rightarrow \{-1, +1\}$,

With the error $e_t = P_{r_{i \sim D_t}}\left[h_t(x_i) \neq y_i\right]$

After choosing $\alpha_t = \frac{1}{2} ln\, ln\left(\frac{1-e_t}{e_t}\right)$

Update: $D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \{e^{-\alpha t}$ *if* $h_t(x_i) = y_i$ $e^{\alpha t}$ *if* $h_t(x_i) \neq y_i$

$= \frac{D_t(i)\ exp(-\alpha_t y_i h_t(x_i))}{Z_t}$

Here, Z_t is a normalization factor. We get the final hypothesis as follows:

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right) \quad (8)$$

Here the dependent variable was the patient state (dead/recovered) while the independent variables were location, country, vis_wuhan, from_wuhan (hosp_vis—sym_on), age, gender, symptom (1–6). We have used the boosted random forest because of its accurate classification performance on imbalanced datasets (25, 29).

The decision trees visualized in **Figures 8–11** have a depth equal to two. Also, the Gini index in all the leaf nodes of all the trees is <0.5, which indicates the training dataset is imbalanced. Hence, for optimizing the performance of the model

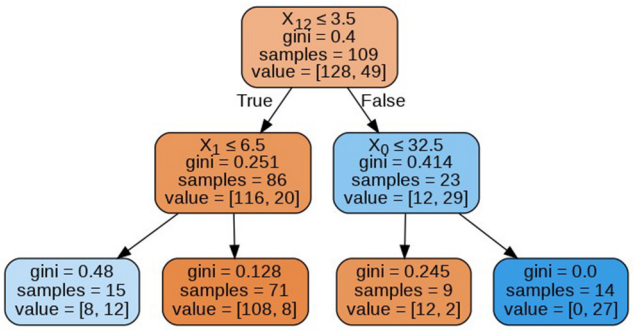

FIGURE 9 | Decision tree 10.

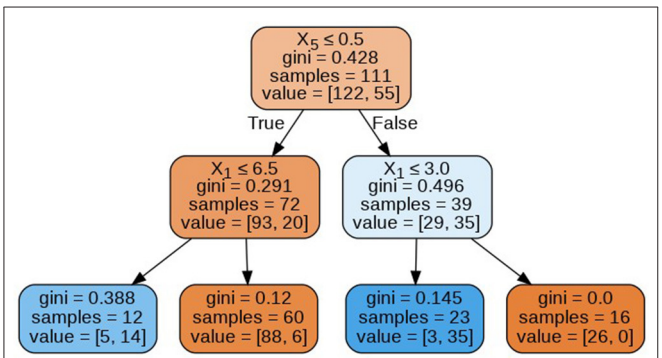

FIGURE 10 | Decision tree 25.

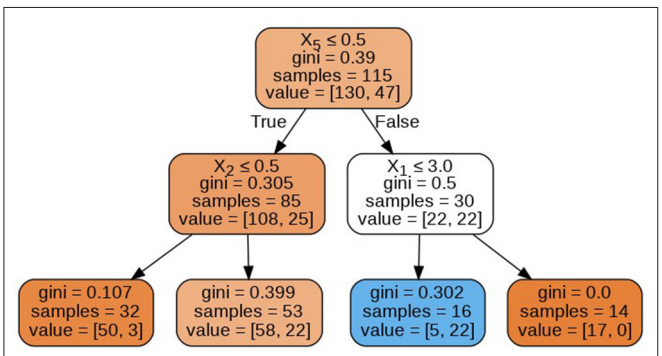

FIGURE 11 | Decision tree 100.

we have reduced the depth of trees to 2 and increased the number of estimators (decision trees) in the random forest to 100. This prevents high variance in the model and provides accurate predictions.

Hyperparameter Optimization

Since the Boosted Random Forest Classifier was implemented using the default parameters, for the optimal performance of the model, we conducted a grid search over a grid of chosen parameters to gain a set of best performing parameters. We implemented the grid search using the GridSearchCV() function from Sklearn library. **Table 2** presents the hyperparameters as returned by the grid search algorithm (30).

TABLE 2 | Optimal hyperparameters returned by grid search.

Parameters	Value
n_estimators	100
max_depth	2
min_samples_leaf	2
min_samples_split	2
criterion	gini

TABLE 3 | Evaluation results.

Metric	Score
Recall score	0.75
Precision score	1.0
F1 score	0.86
Accuracy	0.94

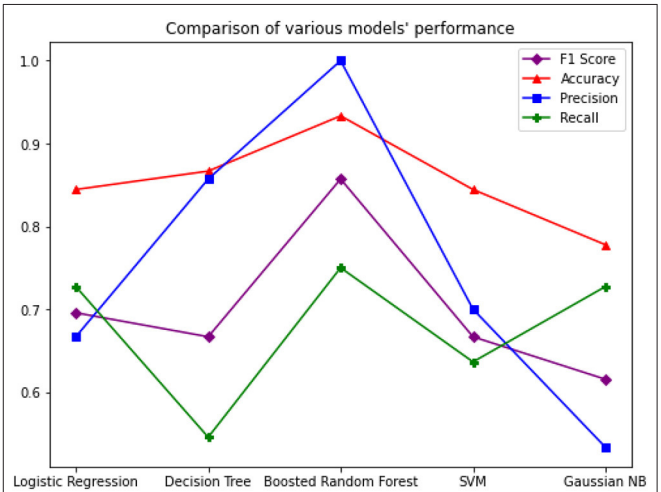

FIGURE 12 | Comparison of Models' performance.

Table 3 presents the evaluation metrics of the Fine Tuned Boosted Random Forest.

The study shows that Boosted Random Forest performs better while predicting COVID-19 patient deaths. **Figure 12** graph compares the performance of all the models including Boosted Random Forest.

CONCLUSION AND FUTURE WORK

The application of Artificial Intelligence is very crucial to process patient data for efficient treatment strategies. In this paper we presented a model that implements the Random Forest algorithm boosted by the AdaBoost algorithm, with a F1 Score of 0.86 on the COVID-19 patient dataset. We have discovered that the Boosted Random Forest algorithm provides accurate predictions even on imbalanced datasets. The data analyzed in this study has revealed that death rates were higher amongst the Wuhan natives

compared to non-natives. Also, male patients had a greater death rate compared to female patients. The majority of affected patients are aged between of 20 and 70 years.

Future work will focus on creating a pipeline that combines CXR scanning computer vision models with these types of demographic and healthcare data processing models. These models will then be integrated into applications that will support the growth of mobile healthcare. This can provide a step toward a semi-autonomous diagnostic system that can provide rapid screening and detection for COVID-19 affected regions and prepare us for future outbreaks.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. For the reproducible code, please check out the GitHub repository: https://github.com/Atharva-Peshkar/Covid-19-Patient-Health-Analytics.

AUTHOR CONTRIBUTIONS

AP, CI, and RM: conceptualization. AP and RM: methodology, investigation, data curation, and writing—original draft preparation. AP, RM, SP, OJ, and NP: software. RS and JC: validation and visualization. CI, RS, and JC: formal analysis. AP, AB, and RM: resources. JC and CI: writing—review and editing, supervision. AB, AP, RM, SP, NP, RS, CI, OJ, and JC: project administration. All authors have read and agreed to the published version of the manuscript.

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2018R1C1B5045013).

REFERENCES

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5
2. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *N Engl J Med.* (2020) 382:1199–207. doi: 10.1056/NEJMoa2001316
3. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet.* (2020) 395:507–13. doi: 10.1016/S0140-6736(20)30211-7
4. *Clinical Management of Severe Acute Respiratory Infection When Novel Coronavirus (2019-nCoV). Infection Is Suspected: Interim Guidance*. (2020). Available online at: https://apps.who.int/iris/handle/10665/330893 (accessed April 31, 2020).
5. Zumla A, Hui DS, Perlman S. Middle East respiratory syndrome. *Lancet.* (2015) 386:995–1007. doi: 10.1016/S0140-6736(15)60454-8
6. Pham QV, Nguyen DC, Hwang WJ, Pathirana PN. Artificial intelligence (AI) and big data for coronavirus (COVID-19) pandemic: a survey on the state-of-the-arts. *Preprints.* (2020) 2020:2020040383. doi: 10.20944/preprints202004.0383.v1
7. *WHO Situation Report-94 Coronavirus disease 2019 (COVID-19).* (2020). Available online at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200423-sitrep-94-covid-19.pdf?sfvrsn=b8304bf0_4 (accessed March 10, 2020).
8. Kathiresan S, Sait ARW, Gupta D, Lakshmanaprabu SK, Khanna A, Pandey HM. Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. *Pattern Recogn Lett.* (2020) 133:210–6. doi: 10.1016/j.patrec.2020.02.026
9. Wang L, Wong A. COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images. *arXiv*. (2020) 2003.09871. Available online at: https://arxiv.org/abs/2003.09871 (accessed May 5, 2020).
10. Pal R, Sekh AA, Kar S, Prasad DK. Neural network-based country wise risk prediction of COVID-19. *arXiv*. (2020) 2004.00959. Available online at: https://arxiv.org/abs/2004.00959 (accessed May 7, 2020).
11. Liu D, Clemente L, Poirier C, Ding X, Chinazzi M, Davis JT, et al. A machine learning methodology for real-time forecasting of the 2019–2020 COVID-19 outbreak using Internet searches, news alerts, and estimates from mechanistic models. *arXiv*. (2020) 2004.04019. Available online at: https://arxiv.org/abs/2004.04019 (accessed May 6, 2020).
12. Cai H. Sex difference and smoking predisposition in patients with COVID-19. *Lancet Respir Med.* (2020) 8:e20. doi: 10.1016/S2213-2600(20)30117-X
13. Bayes C, Valdivieso L. Modelling death rates due to COVID-19: a Bayesian approach. *arXiv.* (2020) 2004.02386. Available online at: https://arxiv.org/abs/2004.02386 (accessed May 5, 2020).
14. Beck BR, Shin B, Choi Y, Park S, Kang K. Predicting commercially available antiviral drugs that may act on the novel coronavirus (2019-nCoV), Wuhan, China through a drug-target interaction deep learning model. *bioRxiv.* (2020). Available online at: https://www.biorxiv.org/content/10.1101/2020.01.31.929547v1.abstract (accessed May 5, 2020).
15. Tang Z, Zhao W, Xie X, Zhong Z, Shi F, Liu J, et al. Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images. *arXiv*. (2020) 2003.11988. Available online at: https://arxiv.org/abs/2003.11988 (accessed May 10, 2020).
16. Khalifa NEM, Taha MHN, Hassanien AE, Elghamrawy S. Detection of coronavirus (COVID-19) associated pneumonia based on generative adversarial networks and a fine-tuned deep transfer learning model using chest X-ray dataset. *arXiv*. (2020) 2004.01184. Available online at: https://arxiv.org/abs/2004.01184 (accessed May 5, 2020).
17. Sujatha R, Chatterjee JM, Hassanien AE. A machine learning forecasting model for COVID-19 pandemic in India. *Stoch Environ Res Risk Assess.* (2020) 34:959–72. doi: 10.1007/s00477-020-01827-8
18. Kutia S, Chauhdary SH, Iwendi C, Liu L, Yong W, Bashir AK. Socio-Technological factors affecting user's adoption of eHealth functionalities: a case study of China and Ukraine eHealth systems. *IEEE Access.* (2019) 7:90777–88. doi: 10.1109/ACCESS.2019.2924584
19. Sultan S, Javed A, Irtaza A, Dawood H, Dawood H, Bashir AK. A hybrid egocentric video summarization method to improve the healthcare for Alzheimer patients. *J Ambient Intell Human Comput.* (2019) 10:4197–206. doi: 10.1007/s12652-019-01444-6
20. Feng L, Ali A, Iqbal M, Bashir AK, Hussain SA, Pack S. Optimal haptic communications over nanonetworks for E-health systems. *IEEE Trans Ind Inform.* (2019) 15:3016–27. doi: 10.1109/TII.2019.2902604
21. Jain V, Chatterjee JM. *Machine Learning with Health Care Perspective.* (2020). Available online at: https://link.springer.com/book/10.1007%2F978-3-030-40850-3 (accessed May 5, 2020).
22. Chatterjee JM. Bioinformatics using machine learning. *Glob J Internet Interv IT Fusion.* (2018) 1:28–35.
23. Khamparia A, Gupta D, de Albuquerque VHC, Sangaiah AK, Jhaveri RH. Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning. *J Supercomput.* (2020) 76:1–19. doi: 10.1007/s11227-020-03159-4

24. Waheed A, Goyal M, Gupta D, Khanna A, Al-Turjman F, Pinheiro PR. Covidgan: data augmentation using auxiliary classifier gan for improved covid-19 detection. *IEEE Access.* (2020) 8:91916–23. doi: 10.1109/ACCESS.2020.2994762
25. Sakarkar G, Pillai S, Rao CV, Peshkar A, Malewar S. Comparative study of ambient air quality prediction system using machine learning to predict air quality in smart city. In *Proceedings of International Conference on IoT Inclusive Life (ICIIL 2019), NITTTR Chandigarh, India*. Singapore: Springer (2020). p. 175–82. doi: 10.1007/978-981-15-3020-3_16
26. *Novel Corona Virus 2019 Dataset*. (2020). Available online at: https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset/ (accessed April 23, 2020).
27. Breiman L. Random forests. *Mach Learn.* (2001) 45:5–32. doi: 10.1023/A:1010933404324
28. Freund Y, Schapire R, Abe N. A short introduction to boosting. *J Jpn Soc Artif Intell.* (1999) 14:1612.
29. Khalilia M, Chakraborty S, Popescu M. Predicting disease risks from highly imbalanced data using random forest. *BMC Med Inform Decis Mak.* (2011) 11:51. doi: 10.1186/1472-6947-11-51
30. Pillai SK, Raghuwanshi MM, Gaikwad M. Hyperparameter tuning and optimization in machine learning for species identification system. In: *Proceedings of International Conference on IoT Inclusive Life (ICIIL 2019), NITTTR Chandigarh, India*. Singapore: Springer (2020). p. 235–41. doi: 10.1007/978-981-15-3020-3_22

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 Public Health

PERSPECTIVE
published: 16 July 2020
doi: 10.3389/fpubh.2020.00274

Optimization of Skewed Data Using Sampling-Based Preprocessing Approach

Sushruta Mishra[1], Pradeep Kumar Mallick[1], Lambodar Jena[2] and Gyoo-Soo Chae[3]*

[1] School of Computer Engineering, Kalinga Institute of Industrial Technology, Deemed to be University, Bhubaneswar, India,
[2] Department of Computer Science and Engineering, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India,
[3] Division of Information & Communication, Baekseok University, ChePonan-si, South Korea

In the past few years, classification has undergone some major evolution. With a constant surge of the amount of data gathered from different sources, efficient processing and analysis of data is becoming difficult. Due to the uneven distribution of data among classes, data classification with machine-learning techniques has become more tedious. While most algorithms focus on major data samples, they ignore the minor class data. Thus, the data-skewing issue is one of the critical problems that need attention of researchers. The paper stresses upon data preprocessing using sampling techniques to overcome the data-skewing problem. Here, three different sampling techniques such as Resampling, SpreadSubSampling, and SMOTE are implemented to reduce this uneven data distribution issue and classified with the K-nearest neighbor algorithm. The performance of classification is evaluated with various performance metrics to determine the efficiency of classification.

Keywords: data skewing problem, machine learning, best first search, KNN algorithm, SMOTE, SpreadSubSampling, F-score

OPEN ACCESS

Edited by:
Deepak Gupta,
Maharaja Agrasen Institute of Technology, India

Reviewed by:
Gyoo Chae,
Baekseok University, South Korea
Shafiz Affendi,
University of Wollongong in Dubai, United Arab Emirates

****Correspondence:***
Sushruta Mishra
mishra.sushruta@gmail.com

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Public Health

Received: *26 April 2020*
Accepted: *26 May 2020*
Published: *16 July 2020*

Citation:
Mishra S, Mallick PK, Jena L and Chae G-S (2020) Optimization of Skewed Data Using Sampling-Based Preprocessing Approach. Front. Public Health 8:274. doi: 10.3389/fpubh.2020.00274

INTRODUCTION

Recently, enormous data is aggregated on a daily basis. Many times, it is observed that such massive data samples are unevenly matched and classified among its classes. As a result, generation of data occurs in a skewed way. This scenario in a data set where samples of data in one class are much higher in comparison to that of the other class is represented as a skewed data set. In this case, the higher data sample class becomes the major class and the class consisting of relatively less data samples is labeled as minor class. As a result of this uneven distribution of data samples, major class is given higher importance than minor class. Hence, the overall performance of machine-learning algorithms is affected, thereby generating inaccurate results. However, in the classification process, both major and minor class samples play a significant role. Subsequently from previous studies, it is observed that this skewed data has a major impact on the performance of machine-learning algorithms (1–3). Besides this, acute knowledge mining from minor class is also a factor considering there are very few data instances (4). This data-skewing concept holds significance, which is applied in several real-time applications such as remote area sensing (5), acute pollution detection (6), risk identification and control (7), and fraud detection and prevention (8). This data skewing can be a critical bottleneck in sensitive applications like in the medical domain in diagnosis of patients where a minor negligence can be dangerous for patients' health. Usually, various classification algorithms that are deployed in predictive analytics consider homogeneity in data partitioning among its classes. It may be a hurdle, thereby degrading the effectiveness of performance of machine-learning

models. Apart from this, overhead causes due to rate of occurrence of error in data skewing are unbalanced, creating inconsistency during classification. It may result in classes overlapping for which noisy instances rise, creating more complexity in prediction.

SAMPLING

The problem of imbalance aggregation of samples of data may be effectively dealt with by the use of a data preprocessing approach called Sampling. It is used to handle problems of uneven data distribution in a given data set. The prime objective of this sampling approach is to identify and choose a sample of data from the raw unstructured data sets gathered, which represent the overall data records. By deploying this mechanism, a smaller data section can be mapped to the entire data set. Two features that govern the selection of a sample include the size of the data sample and the quality of the sample. Several distinct criteria are there for selecting a sample rather than a complete database, which are as follows:

- It is suited in large data sets, which involves handling several constraints.
- The approach of data filtering and preprocessing is cheap.
- Relative loss in data is least.
- It is robust and dynamically applicable to different data sets.

Basically, there are two ways to deal with Sampling which includes Under-Sampling and Over-Sampling. Under-Sampling is performed on the larger training data set while Over-Sampling is done on minority data records.

Under-Sampling: This approach is applicable to a class belonging to majority data samples. Here, data samples of data are selected and eliminated at random so as to create a balance between both class data as shown in **Figure 1**.

Over-Sampling: This procedure is applicable to enhance the size of data sets in the class with fewer instances so as to match it with the class with a larger number of data distributions as shown in **Figure 2**.

RELATED WORK

In (9), the authors discussed a method to determine the features on probability density technique by taking a small data set consisting of uneven classes on the ranking of attributes. Nguwi and Cho (10) present a method of attribute selection that represents a weight vector based on a support vector machine. Its basic aim was to remove the unwanted features,

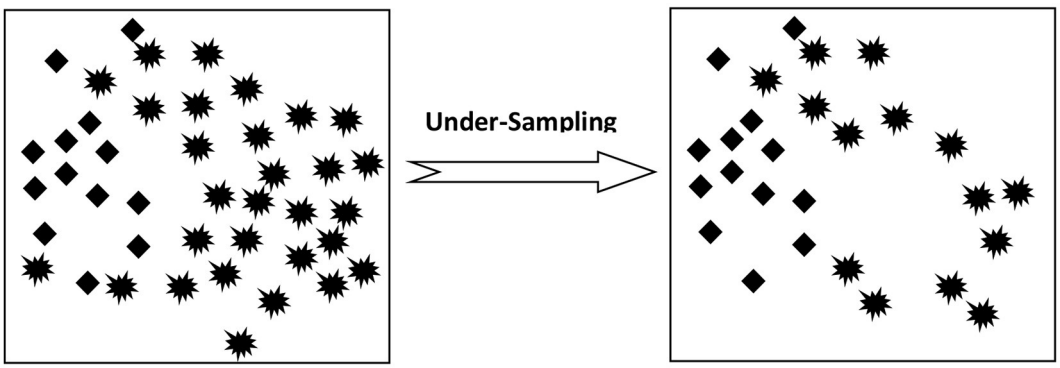

FIGURE 1 | Under-sampling process.

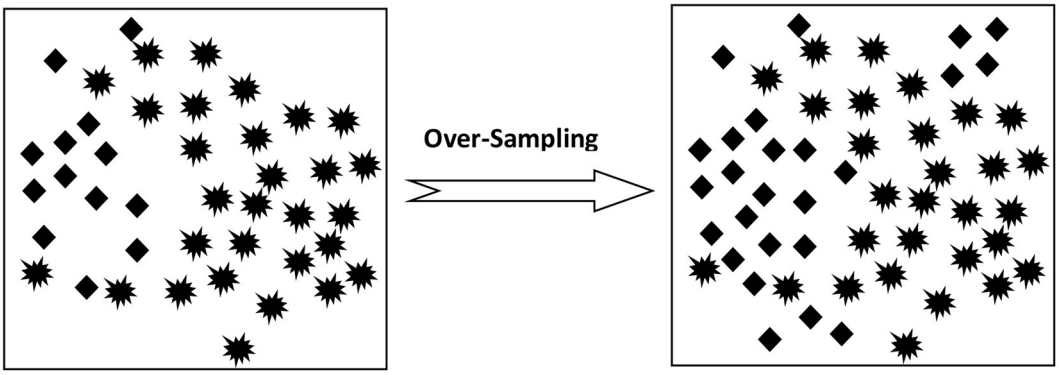

FIGURE 2 | Over-sampling process.

thus enhancing the accuracy rate in classification. In (11), a white blood cell classification model based on the nature-inspired approach is developed, and when a comparison analysis with other existing nature-inspired algorithms is done, it was found that the proposed model was extremely fast and precise in the analysis of hematological issues. An efficient machine-learning framework to identify and classify leaves has been proposed in (12). Later, the proposed model was compared with some classifiers like random forest and KNN algorithm. The accuracy rate of the proposed framework outperformed other classifiers in both training and testing data samples. In (13), the significance of the SMOTE method on sparse and heterogeneous data sets is discussed. Here, the decision tree algorithm and naïve Bayes algorithm are used as classifier. In (14), a hybrid model classifier was developed, which is a combination of SMOTE, particle swarm optimization, and RBF (radial-basis function). This hybrid combination resulted in high prediction performance. In (8), authors analyzed the Undersampling and Oversampling effects on backpropagation neural networks and particle swarm optimization. The result highlighted the sensitivity of the PSO algorithm toward the uneven data distribution and training data with minimum number of instances and many attributes. An enhanced and dynamic adaptation of the crow search algorithm is proposed in (15) to make Parkinson disease diagnosis more effectively and accurately. The proposed optimized algorithm yielded an overall prediction accuracy of 100%. In (16), both undersampling and oversampling use a resampling method that defines several parameters in tuning SVM. The unbalanced data aggregation results in a major data shift to the minor class. A study in (17) developed a resampling-based preprocessing technique to address the skewing of unbalanced data sets and classified various types of tumor in patients. Sharma et al. (18) deals with development of an optimized meta-heuristic model for attribute selection to accurately categorize protein structures. The proposed model was experimentally compared with other meta-heuristic algorithms, and it was found that it yielded optimum results than others. Sahoo et al. (19) used the LVQ technique to illustrate and analyze the clustering deviation issue on the breast cancer data set. In (20), the authors developed a class-based method using ant-colony optimization, which is very effective for major classes. The rate of latency was very high in this method, which is a disadvantage. A novel approach is proposed in (21) on the basis of experimental evaluation on polyster compounds reinforced with fiberglass. The PSO algorithm and genetic algorithm were used to predict global optimum, and it was inferred that the convergence of the PSO algorithm was very fast and needed less execution time. A succinct analysis in (22) presented a cluster analysis under-sampling technique. Here, various clusters were formed for partitioning the entire training set. From every cluster, selected data from the major class were chosen according to the proportion of major data to minor data. The research outcome using clustering on under-sampling enhances the rate of classification accuracy, and the model was more robust than others. Highly accurate with least computational cost evolutionary model-based feature optimization techniques are presented in (23) to detect and diagnose lung disease disorders automatically. Sharma et al. (24) presented an optimized and improved feature selection model that is useful in extracting optimum attributes in Parkinson's disease record samples with enhanced efficiency. It achieved an accuracy rate of 95.91% which is much higher than other related algorithms.

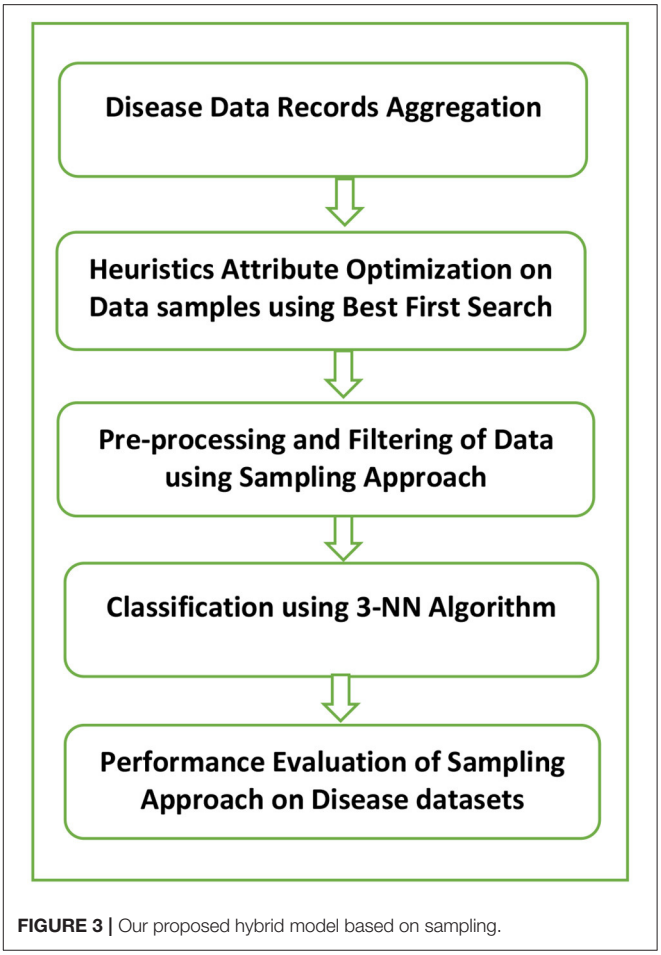

FIGURE 3 | Our proposed hybrid model based on sampling.

PROPOSED WORK

The prime purpose of the analysis is to present disease data samples with minimum uneven data issue so that the required adjustment may be done in data segregation of major and minor classes. The developed framework as shown in **Figure 3** constitutes a five-phase approach of evaluation of performance with the use of data filtering with sampling. The system model proposed in our work is applicable to the healthcare industry. The data sets under consideration include diabetes, breast cancer, and hepatitis. Initially, the sample disease data sets under consideration were gathered from the UCI repository. In the next step, the best-first search algorithm which is based on heuristic search optimization is applied to the original data samples to remove the less relevant features. It is used in the traversal of the graph to determine one or more goal states, which implements priority queue for its operations. This results in a reduced and optimal data set. Eventually, this optimized data

set is implemented with three distinct sampling methods which include SMOTE, Resampling, and SpreadSubSampling. These techniques are useful in varying the sampling data distributions in the already existing data samples where the minority class samples are over-sampled while the majority class samples are under-sampled. This results in a relatively evenly balanced data set. After sampling is performed, the data set undergoes the classification process. The classifier used for this purpose is the K-nearest neighbor (KNN) algorithm where the value of K is 3. It performs classification of a newly arrived unknown data sample based on the Euclidean distance which is used as a similarity measure. It performs classification of new test data on the basis of distance functions used as a similarity metric. During classification, a new unseen data sample is allotted the class label of the most common class among k-closest examples. While classifying a test data, it is allotted a class label of the most common class among 3 closest neighbors. Finally, the effectiveness of the presented model is evaluated by the help of fw vital performance indicators like positive predictive value (PPV), sensitivity, prediction accuracy, F-score, and ROC metric, and the efficiency of the system model is evaluated. The heuristics-based optimization search space technique implemented in our research is best-first search, and the classifier used is the K-NN algorithm where K is taken as 3.

Various sampling techniques used are as follows:

- *SpreadSubSample*: A subsampling filter in which a random subset is processed in order to be fit in memory. In this filter, a maximum spread between the minor and major classes is denoted. For example, the user may specify the class frequency difference to be 2:1. For the subsequent batches, there is no resampling while the batch mode is implemented.
- *SMOTE:* It is a technique specified for Oversampling of the minority class with Random Undersampling of the majority class. The k-nearest minority neighbors are computed in the minority class. Then, some of these neighbors are selected, from which synthetic data samples are extracted which join the minority sample with its chosen neighbors.
- *Resampling:* The objective of Resampling is to add instances to a class. It is done by simply adding instances multiple times to the result data set from the class, which has only a few instances. Suppose one class 2 instance is present then with a resampling with a bias of 1.0, N copies of that instance and N other instances of each other's type for which data is present will be the result. Thus, a random subsample of a data set is produced sampling either with replacement or without replacement.

RESULT AND DISCUSSION

Three sampling approaches which include SMOTE, SpreadSubSample, and Resampling as part of data preprocessing are used in the study. Among the clinical data sets, breast cancer, diabetes, and hepatitis are the data sets under consideration. Various evaluation criteria can be used to determine the effectiveness of machine-learning techniques. However, all evaluation factors may not be useful in dealing with skewed data

TABLE 1 | Parameters of a confusion matrix.

True positives (TP)	Positive instances correctly identified by classifier
True negatives (TN)	Negative instances correctly identified by classifier
False positives (FP)	Negative instances incorrectly identified by classifier
False negatives (FN)	Positive instances incorrectly identified by classifier

		Predicted Class	
		+ve	-ve
Actual Class	+ve	True Positive (TP)	False Negative (FN)
	-ve	False Positive (FP)	True Negative (TN)

FIGURE 4 | Skeleton of a confusion matrix.

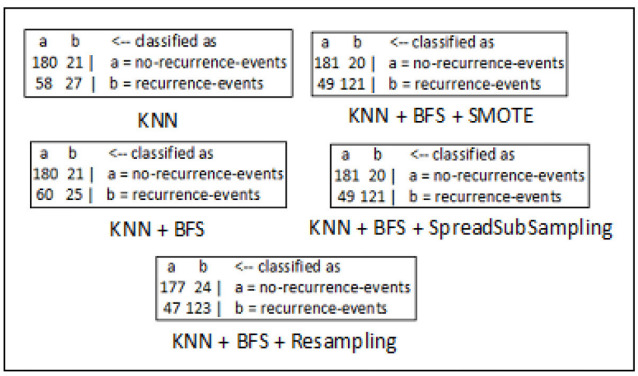

FIGURE 5 | Confusion matrix analysis in breast cancer dataset.

issue due to the presence of unbalanced data samples in complex data sets. Hence, a confusion matrix may be useful to handle such problem in deriving few important metrics to demonstrate the efficiency of classification. Basic parameters of a confusion matrix are illustrated in **Table 1**, and a sample confusion matrix is presented in **Figure 4**.

The confusion matrices for breast cancer, diabetes, and hepatitis disease samples are developed using every feasible combination of the KNN classifier, sampling approaches, and best-first search method. The breast cancer data set consists of 286 data records of different patients, and the entire data samples are spread over in two classes which include "*no-recurrence-events*" and "*recurrence-events*". The confusion matrix of the breast cancer data is seen in **Figure 5**. The confusion matrix for the diabetes data set is depicted in **Figure 6**. It has 786 data samples collected from different patients. "*tested_negative*" and "*tested_positive*" are the two classes under consideration in the diabetes data. As observed in **Figure 7**, "*DIE*" and "*LIVE*" are the two distinct class labels for the hepatitis data set constituting 145 unique samples of patients.

Prediction accuracy forms the basis of classification, which represents the frequency of accurate predictions among all predictions made. Equation 1 denotes the prediction rate of accuracy in terms of confusion matrix.

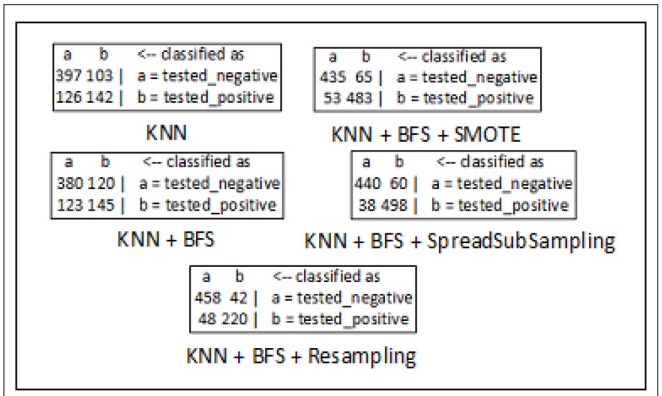

FIGURE 6 | Confusion matrix analysis in diabetes dataset.

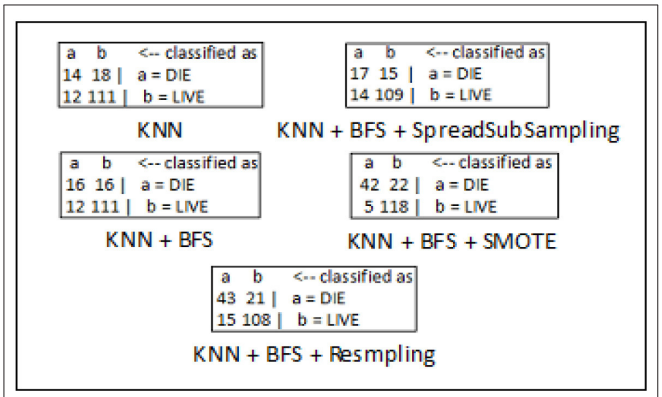

FIGURE 7 | Confusion matrix analysis in hepatitis dataset.

$$\text{Prediction Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

However, at times, prediction accuracy alone is not sufficient to determine the effectiveness of prediction. There are some other equally vital metrics to gauge the prediction performance in skewed data sets such as positive predictive value, sensitivity, and F-score. The positive predictive value denotes the probability of relevance of a randomly chosen data sample from the entire data set.

$$\text{Positive Predictive Value} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

Sensitivity represents the chance of a relevant data sample chosen at random to be extracted in a search.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

These two metrics may not be so useful in determining the superiority of algorithmic performance in the machine-learning task. For example, if one classification model offers high precision value and low recall value than other models, then it is tough to determine the best model among all. In these scenarios, another evaluation metric named F-score is used. This F-score metric computes the balanced mean value in between recall and precision. The efficiency of a classifier is directly proportional to the value of this F-score metric.

$$\text{F} - \text{score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (4)$$

TABLE 2 | Analysis of different sampling techniques on sample disease data sets yielding optimum performance.

Healthcare data set	Evaluation metric	Optimum value	Sampling method
Breast cancer	PPV	0.814	SMOTE and SpreadSubSampling
	Sensitivity	0.820	SMOTE and SpreadSubSampling
	F-score	0.812	SMOTE and SpreadSubSampling
	Prediction accuracy	0.814	SMOTE and SpreadSubSampling
Diabetes	PPV	0.921	SpreadSubSampling
	Sensitivity	0.906	SpreadSubSampling
	F-score	0.905	SpreadSubSampling
	Prediction accuracy	0.884	SMOTE and SpreadSubSampling
Hepatitis	PPV	0.839	Resampling
	Sensitivity	0.837	Resampling
	F-score	0.849	SMOTE
	Prediction accuracy	0.867	SMOTE

The effectiveness of our proposed model can be evaluated by several performance parameters like sensitivity, positive predictive value (PPV), F-score, and prediction accuracy rate. In case of breast cancer, it is sharply observed that data preprocessing with SMOTE and SpreadSubSample methods yields much better results in terms of the performance metrics taken into account. While with the diabetes data set the SpreadSubSample sampling technique performs relatively better than SMOTE and Resampling methods. With hepatitis data, Resampling and SMOTE methods outperform the SpreadSubSample method. Therefore, it can be clearly observed that the data skewing issue is reduced to a large extent by using the sampling approach and thereby fruitful in balancing the uneven data sets. It leads to more optimal performance with data preprocessing using sampling techniques rather than performing classification without sampling methods. The evaluation result analysis is summarized in **Table 2**.

CONCLUSION

The issue of skewed data is a challenging area, which needs to be handled effectively when dealing with time-specific applications like disease diagnosis in the medical domain. Precise disease diagnosis of patients is a very critical task which requires a high level of accuracy. Our paper has presented the data skewing issue and has demonstrated the use of the data-preprocessing approach by implementing some vital sampling

techniques on healthcare disease data sets. Upon implementation of sampling techniques with the KNN classifier on disease data sets, it was observed that the data skewing issue was significantly minimized thereby a more balanced data set is the result.

In this work, sampling techniques like SMOTE, SpreadSubSampling, and Resampling are used. SMOTE was projected as an oversampling method and SpreadSubSampling was used as an under-sampling method for balancing data samples. Though there are no unified norms for class balancing, the study can infer that the classification using the sampling approach generates an optimum result than going in alone. Therefore, it may be concluded that data preprocessing with the sampling approach offers an ideal option to avoid skewing of data samples and is thus beneficial for an effective and accurate disease diagnosis.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://archive.ics.uci.edu/ml/index.php.

ETHICS STATEMENT

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

SM formulated the concept idea, proposed methodology, and helped in implementation. PM helped in prototype design methodology and implementation work. LJ helped in result analysis work and in documentation work. G-SC helped data interpretation, editing and drafting the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to acknowledge our families, friends, and colleagues for their encouragement and support throughout our research work.

REFERENCES

1. Chawla N, Japkowicz N, Kolcz A. *ICML'2003 Workshop on Learning From Imbalanced Data Sets (II)*. (2003). Available online at: http://www.site.uottawa.ca/~nat/Workshop2003/workshop2003.html
2. Japkowicz N. Learning from imbalanced data sets: a comparison of various strategies. In: *AAAI Workshop on Learning From Imbalanced Data Sets*. *Vol. 68.* (2000). p. 10–5.
3. Weiss GM. Mining with rarity: a unifying framework. *ACM Sigkdd Expl News.* (2004) 6:7–19. doi: 10.1145/1007730.1007734
4. Chawla NV, Japkowicz N, Kolcz A. Editorial: special issue on learning from imbalanced data sets. *ACM Sigkdd Expl News.* (2004) 6:1–6. doi: 10.1145/1007730.1007733
5. Lu WZ, Wang D. Ground-level ozone prediction by support vector machine approach with a cost-sensitive classification scheme. *Sci Total Environ.* (2008) 395:109–16. doi: 10.1016/j.scitotenv.2008.01.035
6. Huang YM, Hung CM, Jiau HC. Evaluation of neural networks and data mining methods on a credit assessment task for class imbalance problem. *Nonlin Anal.* (2006) 7:720–47. doi: 10.1016/j.nonrwa.2005.04.006
7. Cieslak DA, Chawla NV, Striegel A. Combating imbalance in network intrusion datasets. In: *IEEE International Conference Granular Computing.* (2006). p. 732–7.
8. Mazurowski MA, Habas PA, Zurada JM, Lo JY, Baker JA, Tourassi GD. Training neural network classifiers for medical decision making: the effects of imbalanced datasets on classification performance. *Neural Netw.* (2008) 21:427–36. doi: 10.1016/j.neunet.2007.12.031
9. Alibeigi M, Hashemi S, Hamzeh A. DBFS: an effective density based feature selection scheme for small sample size and high dimensional imbalanced data sets. *Data Knowl Engin.* (2012) 81:67–103. doi: 10.1016/j.datak.2012.08.001
10. Nguwi YY, Cho SY. An unsupervised self-organizing learning with support vector ranking for imbalanced datasets. *Expert Syst Appl.* (2010) 37:8303–12. doi: 10.1016/j.eswa.2010.05.054
11. Gupta D, Arora J, Agrawal U, Khanna A, de Albuquerque VHC. Optimized binary bat algorithm for classification of white blood cells. *Measurement.* (2019) 143:180–90. doi: 10.1016/j.measurement.2019.01.002
12. Gupta D, Sharma P, Choudhary K, Gupta K, Chawla R, Khanna A, et al. Artificial plant optimization algorithm to detect infected leaves using machine learning. *Exp Syst.* (2020) 2020:e12501. doi: 10.1111/exsy.12501
13. Taft LM, Evans RS, Shyu CR, Egger MJ, Chawla N, Mitchell JA, et al. Countering imbalanced datasets to improve adverse drug event predictive models in labor and delivery. *J Biomed Inform.* (2009) 42:356–64. doi: 10.1016/j.jbi.2008.09.001
14. Gao M, Hong X, Chen S, Harris CJ. A combined SMOTE and PSO based RBF classifier for two-class imbalanced problems. *Neurocomputing.* (2011) 74:3456–66. doi: 10.1016/j.neucom.2011.06.010
15. Gupta D, Sundaram S, Khanna A, Hassanien AE, de Albuquerque VHC. Improved diagnosis of Parkinson's disease based on Optimized Crow Search Algorithm. *Comp Elect Engin.* (2018) 68:412–24. doi: 10.1016/j.compeleceng.2018.04.014
16. Cohen G, Hilario M, Sax H, Hugonnet S, Geissbuhler A. Learning from imbalanced data in surveillance of nosocomial infection. *Artif Intellig Med.* (2006) 37:7–18. doi: 10.1016/j.artmed.2005.03.002
17. Mishra S, Panda A, Tripathy KH. Implementation of re-sampling technique to handle skewed data in tumor prediction. *J Adv Res Dyn Control Syst.* (2018) 10:526–30.
18. Sharma P, Gupta A, Aggarwal A, Gupta D, Khanna A, Hassanien AE, et al. The health of things for classification of protein structures using improved grey wolf optimization. *J Supercomp.* (2018) 76:1226–41. doi: 10.1007/s11227-018-2639-4
19. Sahoo S, Mishra S, Mohapatra SK, Mishra BK. Clustering deviation analysis on breast cancer using linear vector quantization technique. *Int J Control The Appl.* (2016) 9:311–22.
20. Yu H, Ni J, Zhao J. ACOSampling: An ant colony optimization-based undersampling method for classifying imbalanced DNA microarray data. *Neurocomputing.* (2013) 101:309–18. doi: 10.1016/j.neucom.2012.08.018
21. Kalita K, Mallick PK, Bhoi AK, Raichurkar PP, Ramachandran M, Ghadai RK. Optimizing drilling induced delamination in GFRP composites using genetic algorithm and particle swarm optimization. *Adv Comp Lett.* (2013) 27:1–9. doi: 10.1177/096369351802700101

22. Yen SJ, Lee YS. Cluster-based under-sampling approaches for imbalanced data distributions. *Exp Syst Appl.* (2009) 36:5718–27. doi: 10.1016/j.eswa.2008.06.108
23. Gupta N, Gupta D, Khanna A, Filho PR, de Albuquerque VHC. Evolutionary algorithms for automatic lung disease detection. *Measurement*. (2019) 140:590–608. doi: 10.1016/j.measurement.2019.02.042
24. Sharma P, Jain R, Sharma M, Gupta D. Parkinson's diagnosis using AntLion optimization algorithm. *Int J Innov Comp Appl*. (2019) 10:138–46. doi: 10.1504/IJICA.2019.103370

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 Public Health

ORIGINAL RESEARCH
published: 28 August 2020
doi: 10.3389/fpubh.2020.00441

COVID-19 Mortality Rate Prediction for India Using Statistical Neural Network Models

S Dhamodharavadhani[1], *R Rathipriya*[1] *and Jyotir Moy Chatterjee*[2]*

[1] Department of Computer Science, Periyar University, Salem, India, [2] Department of IT, Lord Buddha Education Foundation (LBEF), Kathmandu, Nepal

The primary aim of this study is to investigate suitable Statistical Neural Network (SNN) models and their hybrid version for COVID-19 mortality prediction in Indian populations and is to estimate the future COVID-19 death cases for India. SNN models such as Probabilistic Neural Network (PNN), Radial Basis Function Neural Network (RBFNN), and Generalized Regression Neural Network (GRNN) are applied to develop the COVID-19 Mortality Rate Prediction (MRP) model for India. For this purpose, we have used two datasets as D1 and D2. The performances of these models are evaluated using Root Mean Square Error (RMSE) and "R," a correlation value between actual and predicted value. To improve prediction accuracy, the new hybrid models have been constructed by combining SNN models and the Non-linear Autoregressive Neural Network (NAR-NN). This is to predict the future error of the SNN models, which adds to the predicted value of these models for getting better MRP value. The results showed that the PNN and RBFNN-based MRP model performed better than the other models for COVID-19 datasets D2 and D1, respectively.

Keywords: Covid-19, mortality rate prediction (MRP), statistical neural network (SNN), probabilistic neural network (PNN), generalized regression neural network (GRNN), radial basis function neural network (RBFNN), non-linear autoregressive (NAR), root mean square error (RMSE)

OPEN ACCESS

Edited by:
Deepak Gupta,
Maharaja Agrasen Institute of Technology, India

Reviewed by:
Mumtaz Ali,
Deakin University, Australia
Hoang Viet Long,
Ton Duc Thang University, Vietnam
Noor Zaman,
Taylor's University, Malaysia

***Correspondence:**
Jyotir Moy Chatterjee
jyotirchatterjee@gmail.com

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Public Health

Received: *24 June 2020*
Accepted: *17 July 2020*
Published: *28 August 2020*

Citation:
Dhamodharavadhani S, Rathipriya R and Chatterjee JM (2020) COVID-19 Mortality Rate Prediction for India Using Statistical Neural Network Models. Front. Public Health 8:441. doi: 10.3389/fpubh.2020.00441

INTRODUCTION

At the end of December 2019 in Wuhan, China, it was first reported that a human infection was caused by a novel coronavirus (nCov) or Wuhan virus or 2019-nCov (1). One of the biggest challenges of this epidemic is the human-to-human transition of nCov. The coronavirus (COVID-19) infected cases increase at an exponential rate worldwide. On 30 January 2020, the World Health Organization (WHO) issued a worldwide health emergency warning notice (2), labeling that 2019-nCoV is of urgent global concern. The disease and mortality rates for the COVID-19 are uncertain at the early stage (3) especially for young ones and aged people. WHO has estimated the reproduction factor (R0) of nCov is 2.7. In demand to control the extensive and quick spread of the nCov, public health sectors took reliable preventative measures and imposed curfew or lockdown infested cities in China, United States, India, and other countries also (4, 5). This is to limit the social distance between people and to avoid the broadcast of this novel virus via humans to humans.

Since 2000, machine learning techniques have gain momentum and play a vital role in epidemiological data analysis. Machine learning techniques also can be used to develop standard

mortality models. Deprez et al. (6) used machine learning algorithms to fit and assess the mortality model by detecting the weaknesses of different mortality models. Artificial Neural Networks (ANNs) (7) used to track and forecast latent mortality factors with greater predictability. Richman and Wüthrich (8) have used the Lee-Carter model to many population predictions using neural networks.

"Probabilistic Neural Network (PNN)" is used for kernel analysis. PNN makes training faster (9). PNN assimilates statistical concepts with neural networks and thus outcomes in an adjusting classification system in which conventional statistical equivalents have unsuccessful. The PNN used to describe bacterial growth and no growth states and to assess the probability. Evolution as affected by changing working conditions (10).

In (11), the GRNN model was created as another possible instrument for the infectious disease rate expectation field. Han et al. (12) built a GRNN network with one-dimensional input and output layer to forecast the occurrence of blood and sexually transmitted diseases. Hong and Zhou (13) made a comparison study on back propagation neural network (BPNN), GRNN, and RBFNN for evaporation prediction. The results revealed that PNN is a powerful technique than artificial neural network methods.

Montazer et al. (14) carried out a large scale comparison study for the major machine learning models such as multilayer perceptron, Bayesian neural networks, RBFNN, GRNN (also called kernel regression), K-nearest neighbor regression, CART regression trees, support vector regression, and Gaussian processes for time series forecasting. The authors reported that these models have different impacts on the performance purely dependent on the dataset.

The RBF and GRNN (15) have been applied over heart disease patient data for the outcome of the medicine. The results showed that RBF performed well for prescribing medicine for the patient. The RBF and GRNN have been applied over heart disease patient data for the outcome of the medicine (16). The results showed that RBF performed well for prescribing medicine for the patient. Hajmeer and Basheer (17) claimed that the Gaussian process approach performed better than the standard generalized linear model (GLM) for the Phenomenological forecasting of dengue disease incidence. Huber (18) reviewed various learning methods for defining network parameters such as widths, centers, and synaptic weights of the RBF neural network. Williams and Rasmussen (19) general regression neural networks for forecasting time series data was proposed as an automated methodology. This methodology is meant to achieve an effective and fast tool so that a huge amount of time series can be predicted automatically. From these works, one could clearly understand the applications of PNN, GRNN, and RBFNN in various research domains.

In recent years, Artificial neural networks (ANNs) have been used frequently to capture the uncertainty in the time series dataset as they have been proven to be a powerful technique for handling the non-linear data (13). Therefore, the use of these ANN techniques gains huge momentum in recent years in the field of epidemiological predictions for the linear, non-linear, and hybrid data (20–22). Hybrid technique integrating the Autoregressive Integrated Moving Average (ARIMA) with a Non-linear Auto-regressive Neural Network (NAR) yielded better forecasting accuracy for time series data (20) relative to other combinations of ANN models or time series models individually (21) proposed the SARIMA-NARX technique for the prediction of scarlet fever incidence cases in China. Moreover, the authors claimed that this hybrid technique has the promising ability to handle both linearity and non-linearity in the scarlet fever dataset than the other techniques. Wang et al. (21) developed techniques by fusing a seasonal autoregressive integrated moving average (SARIMA) with a neural network non-linear autoregression (NNNAR) for tuberculosis (TB) incidence data in china.

Singh et al. (22) used an advanced ARIMA model for predicting the COVID-19 disease spread using Top 15 COVID-19 affected countries. They forecasted that the recovery and death rates were rose faster in the next 2 months when compared to COVID-19 confirmed cases (23). A fine-tuned Random Forest model was proposed by Iwendi et al. (24) for prediction of the severity of the COVID-19 case using the migration, geographical, demographic, and travel details of COVID-19 patients. Tomar and Gupta (25) used Long Short-Term Memory (LSTM) and curve fitting for forecasting the number of COVID-19 confirmed cases in India 30 days ahead. The main limitation is that the proposed method is accurate only for a short range of values (26).

In accordance with the relevant literature, the error variable has not been considered in the modeling of standard neural network models and hybrid neural network models for the improvisation of epidemiological prediction accuracy. Therefore, this work aimed to propose Statistical Neural Network models and their hybrid version (PNN, GRNN, and RBFNN) with a NAR model, to predict COVID-19 mortality rate prediction in India by considering the error variable. Moreover, to evaluate the performance of these models, a benchmark measure, the RMSE, is used. The results of this study may facilitate the public health officials of the Indian government for better prevention and control measures for COVID-19.

The remaining part of this article is arranged as follows. Section Methods and Materials explains about the methods and materials for forecasting of COVID-19 Mortality for India. Section Proposed Methodology expounds on the proposed methodology for COVID-19 Death case prediction. Section Result and Discussion discusses about the results of this study. Section Conclusion and Future work summarizes this work with possible future work.

METHODS AND MATERIALS

Dataset Description

For experimentation purpose we have used (27) for predicting the Covid-19 death cases for India. This dataset contains India's COVID-19 Confirmed cases and Death cases from January 20, 2020, to May 30, 2020, which is used for training and testing models. First, these data are pre-processed to eliminate missing

values and inappropriate values. These data can be used to create two types of datasets. They are:

- Dataset1 (D1) contains a time series of COVID-19 death cases.
- Dataset2 (D2) contains two attributes such as COVID-19 confirmed cases and death cases. Here, "death case" is a predictive attribute and "confirmed case" is a response attributes or independent attribute.

Probabilistic Neural Network (PNN)

It's a kind of radial basis networks (9). This applies to the Bayesian decision rule and Parzen (estimators of the probability density function), called the Bayes-Parzen classification. PNN contains equally statistical pattern recognition characteristics and BPNN. It applies to various fields including pattern recognition, non-linear mapping, and classification. Equation (2) represents the PNN is a supervised feed-forward neural network.

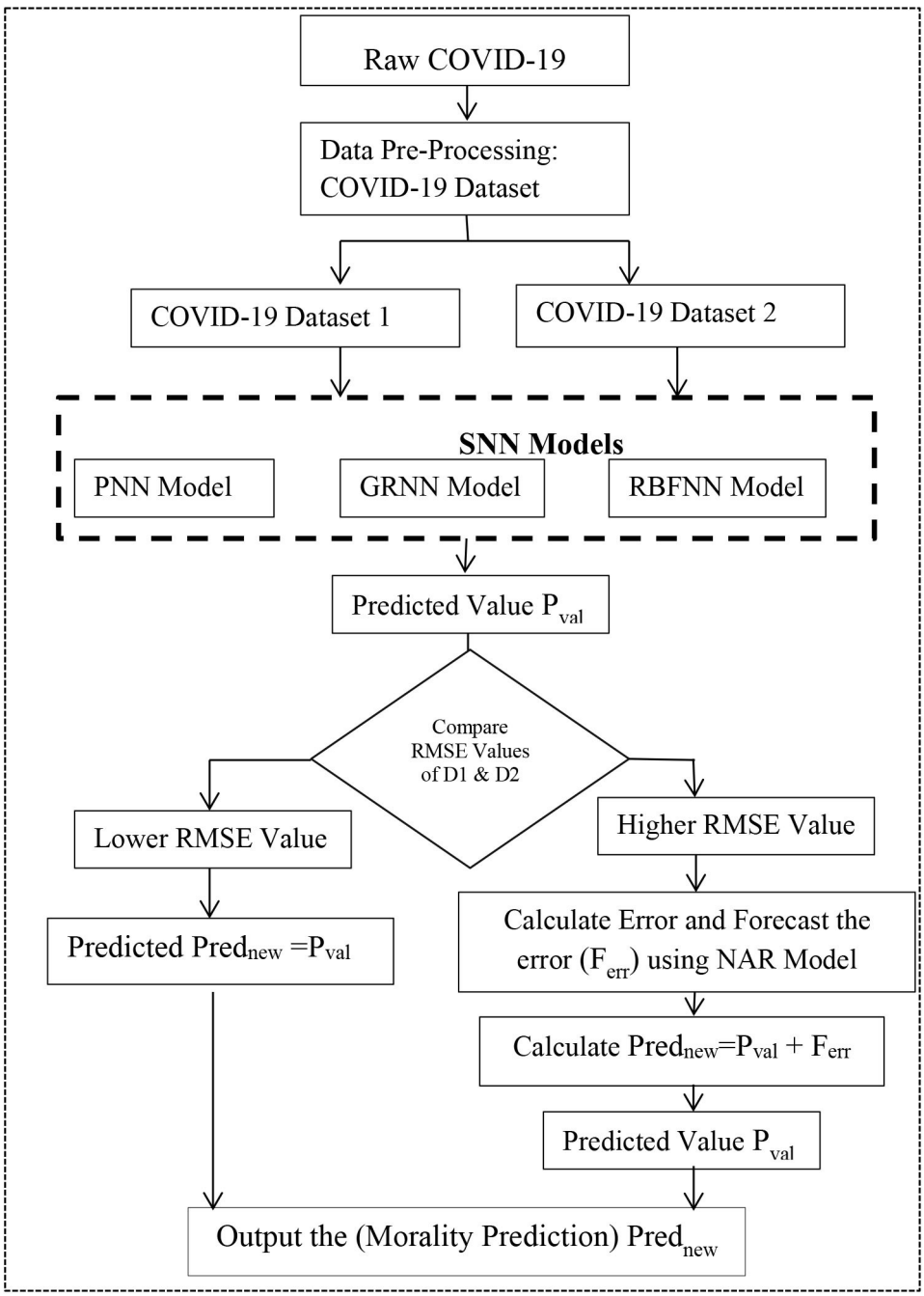

FIGURE 1 | Proposed methodology for COVID-19 MRP model.

This is similarly made of three layers with an algorithm for one-pass training (10). PNN has the capacity of Train on a sparse collection of data. It's also capable of classifying data to different types of outputs (11). There is plenty of usage of PNN aimed at classification advantages. For instance: The PNN processing time is quicker than BPNN and Robust and noisy. The PNN manner of training is Simple and Immediate (15, 28–31).

$$P\left(X_{new} \mid C_i\right) = Pi = \frac{1}{|C_i|} \sum_{j=1}^{|C_i|} w_{i,j} \tag{1}$$

Where P denotes as probability, X as predicted value, w represents as the weight value, C represents as class, where i indexes the input dimension and w_{ji} is a positive parameter signifying the ith weight of the jth hidden unit.

Generalized Regression Neural Network (GRNN)

It's a special case of Radial Basis Networks (RBN) (9). The structure of a GRNN is comparatively easy and fixed with 2 layers. The first layer is the pattern and the second layer are the summation. If each unit in the pattern layer is passed through the input, the input-response relationship will be "memorized" and stored in the unit. As a result, in the training set, no. of units in the pattern layer is equal to the no. of actual values. A Gaussian PDF will be added to the network input in each pattern unit, so that represented as the Equation (2)

$$\boldsymbol{\theta} = EXP[-0.5 * (X - u) \grave{}(X - u) / (\sigma \hat{} 2)] \tag{2}$$

where $\boldsymbol{\theta}$ is the Pattern Unit output, X is the input, u is training vector stored in the unit, and σ is a positive constant known as "spread" or "smooth parameter." If $\boldsymbol{\theta}$ is calculated, computation is moved on to the summation layer

$$Y|X = SUM(Y * \boldsymbol{\theta}) / SUM(\boldsymbol{\theta}) \tag{3}$$

where Y|X is the prediction conditional on X and Y is the response in the training sample (12, 31–39).

Radial Basis Function Neural Network (RBFNN)

It's a ANN (14, 31, 33) that uses functions on a radial basis as activation functions shown in Equation (5). The RBFNN is a neural network with three layers of feed-forwards. The first layer is linear and only the input signal is transmitted, while the next layer is non-linear and uses Gaussian functions (9, 10). The third layer incorporates the Gaussian outputs in linear form. During training, only the tap weights among the hidden layer and the output layer are changed (30–39).

$$f(x) = \frac{1}{\sqrt[\sigma]{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4}$$

The function approximation f(x) is a Gaussian function. x represents as the actual values. The input x, to find the dimensional parameters of the function.

Non-linear Autoregressive Neural Network (NAR-NN)

The NAR (34) is a sort of ANN fitting for evaluating future estimations of the input variable (9, 10). The NAR-NN empowers the forecast of future estimations of a time series. It upheld by its history foundation utilizing a re-feeding care of instrument, in which an anticipated worth may fill in as a contribution for new expectations at further developed focuses in time. In condition (6) speaks to as anticipate arrangement y(t) given d past estimations of y(t).

$$y(t) = f(y(t-1), \ldots, y(t-d)) \tag{5}$$

Where y represents as input parameter, t denotes as time period and d represents as delay.

Root Mean Square Error (RMSE)

RMSE (34, 35) is the square root of the square differences measured between predicted and actual COVID-19 Death cases. Its representation is shown in Equation (6).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(Predicted\ COVID-19\ Death\ Case_i - Actual\ COVID-19\ Death\ Case_i\right)^2} \tag{6}$$

Where n = number of samples.

Correlation Coefficient (R)

It's a measure a linear relationship between the predicted and actual COVID-19 death cases. It represents as in Equation (7)

$$R = \frac{\sum_{i=1}^{n} \left(t_i - \overline{t_l}\right)\left(p_i - \overline{p_l}\right)}{\sqrt{\sum_{i=1}^{n} \left(t_i - \overline{t_l}\right)^2} \sqrt{\sum_{i=1}^{n} \left(p_i - \overline{p_l}\right)^2}} \tag{7}$$

TABLE 1 | Model parameters setup.

Model parameters	PNN model	GRNN model	RBFNN model
Hidden layer (HL)	Fixed architecture	Fixed architecture	Fixed architecture
Number of neurons in HL	10–15	10–15	10–15
Training algorithm	Bayesian regularization	Bayesian regularization	Bayesian regularization
SPREAD (σ)	0–4	0–4	0–4
Performance indicator Measure	RMSE	RMSE	RMSE

where t is the actual COVID-19 death case value, p is the predicted COVID-19 death case value, $\bar{t}$ is the mean of actual COVID-19 death case value $\bar{p}$ is the mean predicted COVID-19 death cases value, and n is the total number of data points.

PROPOSED METHODOLOGY

In this paper, three SNN models (such as PNN, GRNN, and RBFNN) are constructed with the appropriate model parameter values and used in these two datasets to validate the predicted results concerning given the available datasets. **Figure 1** illustrates the proposed methodology. The following steps are used to develop the proposed methodology:

Step 1: Pre-process the raw COVID-19 time series dataset. Create (D1) and (D2).

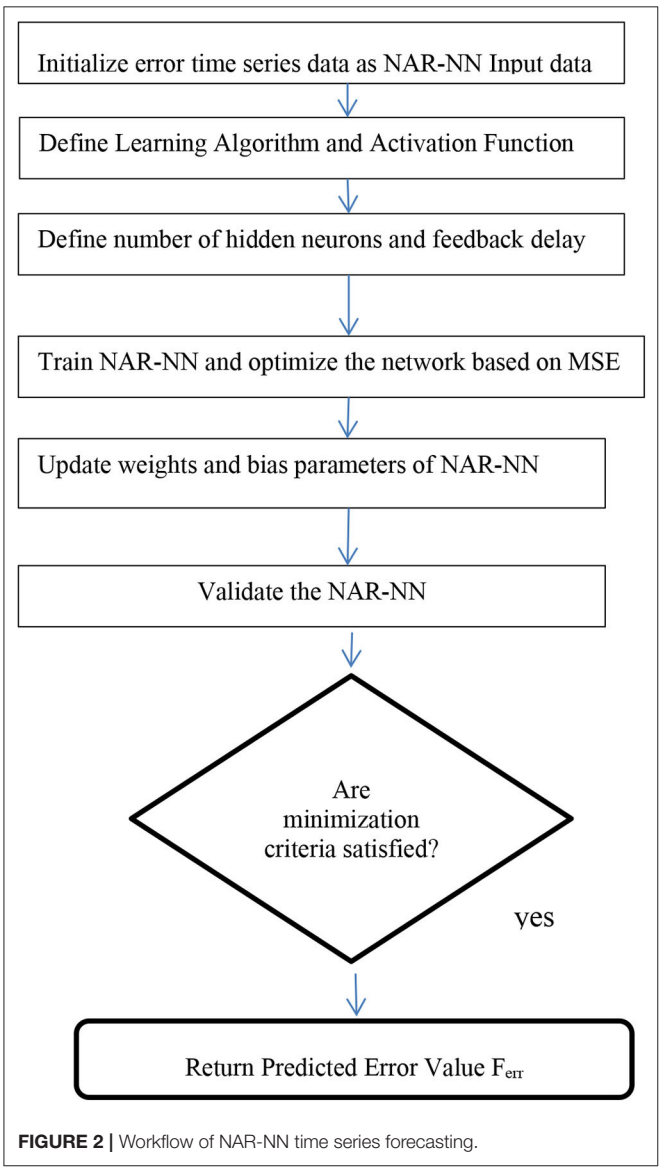

FIGURE 2 | Workflow of NAR-NN time series forecasting.

Step 2: Initialize Model Parameters for PNN, GRNN, and RBFNN. Parameters are shown in **Table 1**.
Step 3: Input D1 and D2 into the PNN model, GRNN model, and RBFNN model, respectively, and predict COVID-19 death cases (Prednew) for "n" period ahead or for given set of confirmed cases.
Step 4: Compare the SNN models of two datasets using RMSE value. Calculate the error or residual of the SNN model with higher RMSE.
Step 5: Input these residuals into the NAR-NN time series forecasting model and predict the residual values (Ferr). It is shown graphically in **Figure 2**.
Step 6: Ferr is added with PredNew to generate an optimized prediction value.
Step 7: Return optimized predicted values as output.

Figure 2 describes the working principle of the NAR-NN model for error forecasting for these models.

RESULT AND DISCUSSION

In this section, the results of three different SNNs: PNN, GRNN, and RBFNN models for D1 and D2 are presented and discussed. The performance of these models was compared. The benchmark key performance indicator metrics such as RMSE and Correlation coefficient (R) is used to estimate the COVID-19 Mortality models for India.

TABLE 2 | Performance metrics for datasets.

Model	RMSE		R	
	D1	D2	D1	D2
PNN	8.889595	7.898071	0.999978	0.999983
GRNN	9.713768	8.388667	0.999975	0.999981
RBFNN	8.528095	9.50462	0.99998	0.999977

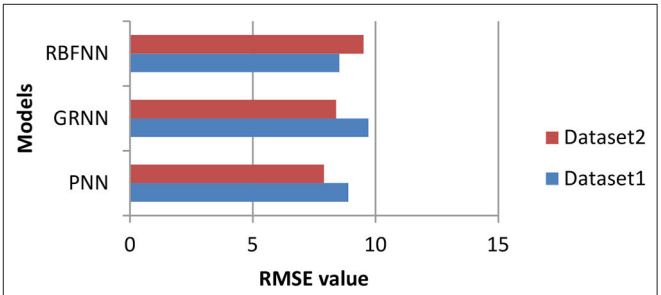

FIGURE 3 | Comparison of RMSE values with a different model.

TABLE 3 | SPREAD value for PNN, GRNN, and RBFNN.

Model	D1 (spread)	D2 (spread)
PNN	0.5	2
GRNN	4	4
RBFNN	1.5	1.68

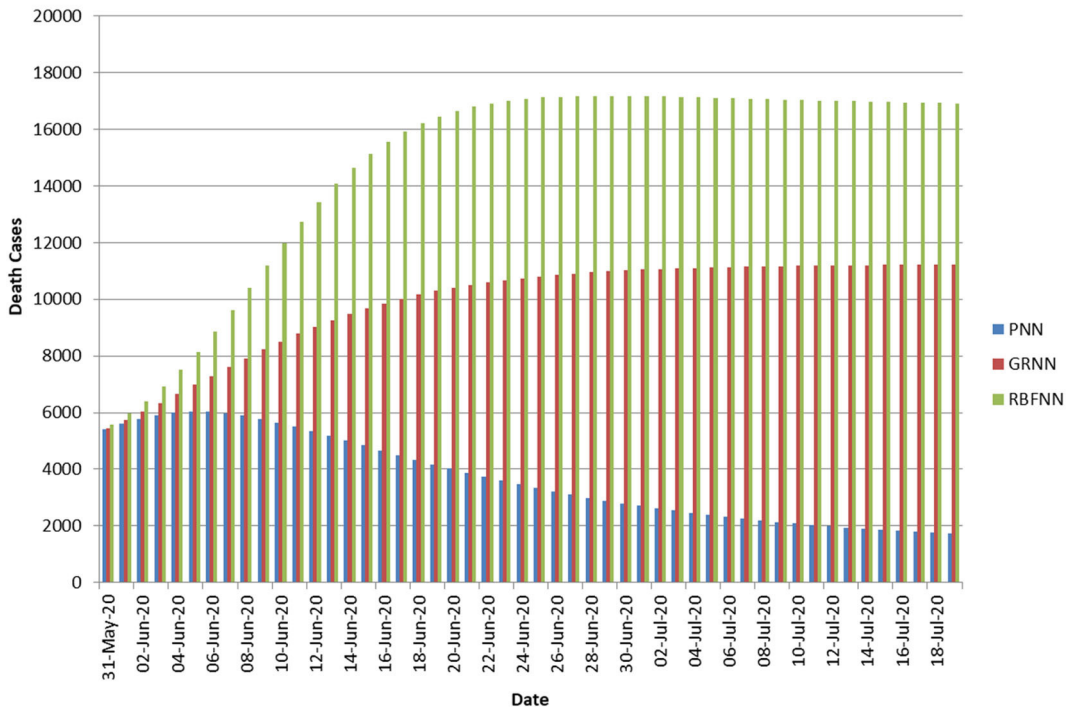

FIGURE 4 | Predicted curve for standard SNN.

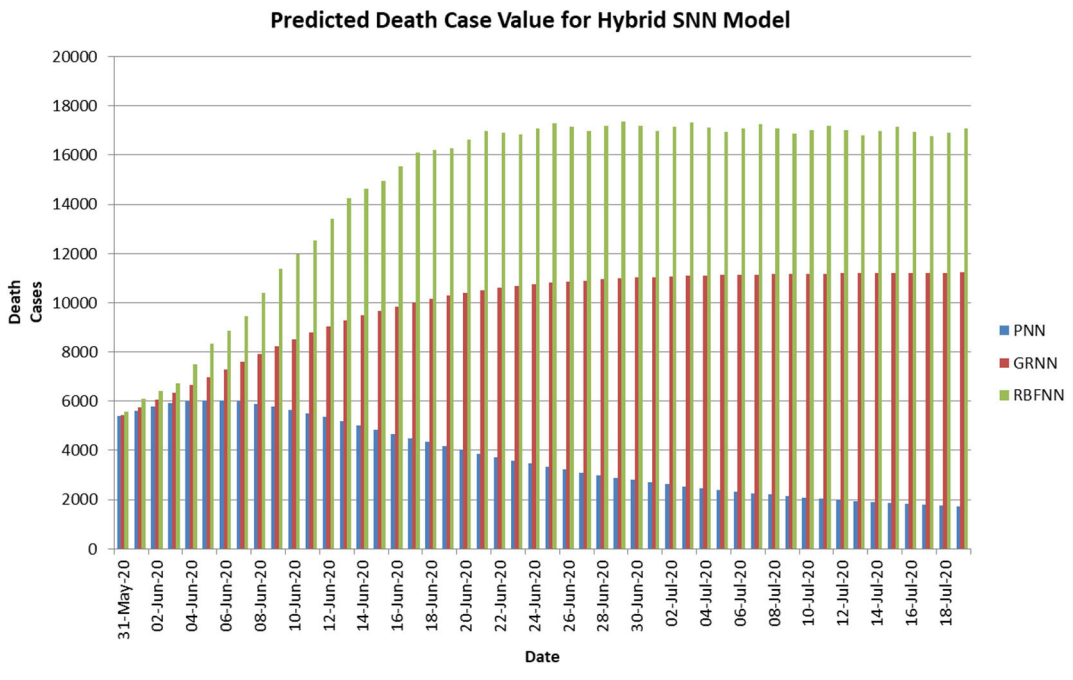

FIGURE 5 | Predicted curve for hybrid SNN.

TABLE 4 | Predicted value Ypred for D1 using standard models.

Date	PNN	GRNN	RBFNN
31-May-20	5,409	5,454	5,574
1-Jun-20	5,608	5,732	5,957
2-Jun-20	5,776	6,026	6,403
3-Jun-20	5,906	6,333	6,917
4-Jun-20	5,990	6,649	7,501
5-Jun-20	6,028	6,970	8,151
6-Jun-20	6,022	7,291	8,859
7-Jun-20	5,976	7,609	9,614
8-Jun-20	5,896	7,920	10,398
9-Jun-20	5,787	8,219	11,191
10-Jun-20	5,656	8,506	11,972
11-Jun-20	5,509	8,776	12,721
12-Jun-20	5,350	9,029	13,422
13-Jun-20	5,183	9,263	14,061
14-Jun-20	5,013	9,479	14,631
15-Jun-20	4,841	9,675	15,129
16-Jun-20	4,670	9,854	15,555
17-Jun-20	4,502	10,015	15,913
18-Jun-20	4,337	10,159	16,209
19-Jun-20	4,177	10,288	16,450
20-Jun-20	4,022	10,402	16,643
21-Jun-20	3,872	10,503	16,795
22-Jun-20	3,729	10,593	16,912
23-Jun-20	3,592	10,672	17,000
24-Jun-20	3,461	10,742	17,066
25-Jun-20	3,336	10,803	17,112
26-Jun-20	3,218	10,856	17,143
27-Jun-20	3,105	10,903	17,163
28-Jun-20	2,998	10,944	17,172
29-Jun-20	2,896	10,980	17,175
30-Jun-20	2,800	11,012	17,171
1-Jul-20	2,709	11,040	17,164
2-Jul-20	2,623	11,064	17,153
3-Jul-20	2,542	11,085	17,140
4-Jul-20	2,465	11,103	17,125
5-Jul-20	2,393	11,120	17,109
6-Jul-20	2,324	11,134	17,093
7-Jul-20	2,260	11,147	17,076
8-Jul-20	2,199	11,158	17,059
9-Jul-20	2,142	11,168	17,043
10-Jul-20	2,088	11,176	17,026
11-Jul-20	2,038	11,184	17,010
12-Jul-20	1,990	11,191	16,995
13-Jul-20	1,945	11,197	16,980
14-Jul-20	1,903	11,202	16,966
15-Jul-20	1,863	11,207	16,952
16-Jul-20	1,826	11,211	16,939
17-Jul-20	1,791	11,215	16,926
18-Jul-20	1,758	11,219	16,915
19-Jul-20	1,727	11,222	16,903

TABLE 5 | Predicted value Ypred for D1 using hybrid models.

Date	PNN	GRNN	RBFNN
31-May-20	5,404	5,434	5,574
1-Jun-20	5,608	5,731	6,111
2-Jun-20	5,781	6,048	6,403
3-Jun-20	5,906	6,333	6,744
4-Jun-20	5,993	6,648	7,500
5-Jun-20	6,029	6,970	8,342
6-Jun-20	6,025	7,291	8,861
7-Jun-20	5,976	7,609	9,441
8-Jun-20	5,898	7,920	10,396
9-Jun-20	5,787	8,220	11,382
10-Jun-20	5,658	8,506	11,975
11-Jun-20	5,509	8,777	12,548
12-Jun-20	5,351	9,029	13,419
13-Jun-20	5,183	9,264	14,252
14-Jun-20	5,014	9,479	14,634
15-Jun-20	4,841	9,676	14,955
16-Jun-20	4,671	9,854	15,552
17-Jun-20	4,502	10,015	16,104
18-Jun-20	4,338	10,160	16,212
19-Jun-20	4,177	10,288	16,277
20-Jun-20	4,022	10,403	16,640
21-Jun-20	3,873	10,504	16,986
22-Jun-20	3,730	10,594	16,915
23-Jun-20	3,592	10,673	16,827
24-Jun-20	3,462	10,742	17,063
25-Jun-20	3,336	10,804	17,303
26-Jun-20	3,218	10,857	17,146
27-Jun-20	3,105	10,904	16,989
28-Jun-20	2,998	10,945	17,170
29-Jun-20	2,896	10,981	17,366
30-Jun-20	2,800	11,013	17,174
1-Jul-20	2,709	11,040	16,990
2-Jul-20	2,623	11,065	17,150
3-Jul-20	2,542	11,086	17,331
4-Jul-20	2,465	11,104	17,128
5-Jul-20	2,393	11,121	16,936
6-Jul-20	2,325	11,135	17,090
7-Jul-20	2,260	11,148	17,267
8-Jul-20	2,200	11,159	17,062
9-Jul-20	2,142	11,169	16,869
10-Jul-20	2,089	11,177	17,024
11-Jul-20	2,038	11,185	17,201
12-Jul-20	1,990	11,192	16,998
13-Jul-20	1,945	11,198	16,807
14-Jul-20	1,903	11,203	16,963
15-Jul-20	1,863	11,208	17,143
16-Jul-20	1,826	11,212	16,942
17-Jul-20	1,791	11,216	16,753
18-Jul-20	1,758	11,220	16,912
19-Jul-20	1,727	11,223	17,094

In general, residues or errors are an inevitable part of any predictive or regression models. Similarly, there are errors in the PNN, GRNN, and RBFNN models. To provide a predictive model with high accuracy, this study explores a hybrid approach, including the NAR-NN time series forecasting model. For hybridization, first is to find out the mean RMSE value of SNN models for D1 and D2. And, then identify which set of SNN models has the highest mean RMSE value. Here, the mean RMSE value of SNN models for D1 is higher than that of D2. Therefore, trends in residues or errors are detected for D1 and predicted by the NAR-NN model. Combining the predicted residual values with predicted COVID-19 death cases of respective SNN models of D1 for higher predictive accuracy.

Table 2 shows the values of the performance metrics such as RMSE and R2 for three SNN models.

The performance of these models is compared based on RMSE value as shown in **Figure 3**. While comparing the mean RMSE value of the three SNN models for COVID19 time series, i.e., D1 data is higher compared to the mean RMSE value of the three SNN models for D2. Therefore, to reduce the RMSE value for D1, the NAR-NN is combined with the SNN models. The purpose of this NAR-NN is to forecast the error of SNN models. Thereafter, this predicted error is included in the predicted COVID19 mortality cases of the respective SNN models for D1.

Table 3 shows the optimum SPREAD values or smoothing factor (σ) of three SNN models. This spread parameter of the SNN models has an important inspiration on the prediction performance. Consequently, in instruction to select the appropriate SPREAD parameter of these models, we run these models with different SPREAD values from 0 to 4 with 0.02 intervals and identified the best SPREAD values of the respective models.

Figure 4 shows the predicted curve of standard SNN model for COVID-19 death cases, respectively. Here, the X-axis represents the dates and Y-axis represents the number of death cases predicted. **Figure 5** shows the predicted curve of hybrid

TABLE 6 | Predicted death cases for D2.

Number of confirmed cases	Models		
	PNN	GRNN	RBFNN
200,000	5465.563	4839.117	6206.483
210,000	5469.28	4300.973	6872.048
220,000	5377.04	3797.288	7468.859
230,000	5211.424	3437.197	7938.901
240,000	4996.182	3216.437	8276.483
250,000	4751.402	3092.094	8504.179
260,000	4492.206	3025.066	8651.558
270,000	4229.201	2989.672	8744.469
280,000	3969.475	2971.118	8802.077
290,000	3717.585	2961.382	8837.423
300,000	3476.337	2956.237	8858.968

TABLE 7 | MRP for D2.

Confirmed cases	Models		
	PNN	GRNN	RBFNN
200,000	2.7	2.4	3.1
210,000	2.6	2.0	3.3
220,000	2.4	1.7	3.4
230,000	2.2	1.5	3.4
240,000	2.1	1.3	3.4
250,000	1.9	1.2	3.4
260,000	1.7	1.2	3.3
270,000	1.6	1.1	3.2
280,000	1.4	1.1	3.1
290,000	1.3	1.0	3.0
300,000	1.1	1.0	2.9

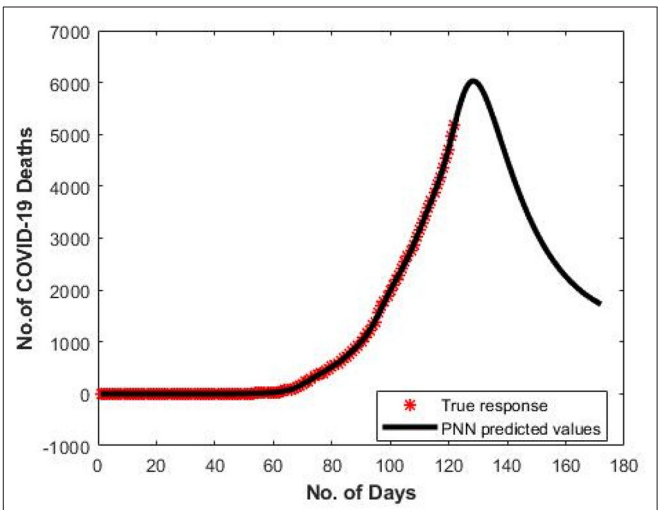

FIGURE 6 | Predicted curve for D1 using PNN.

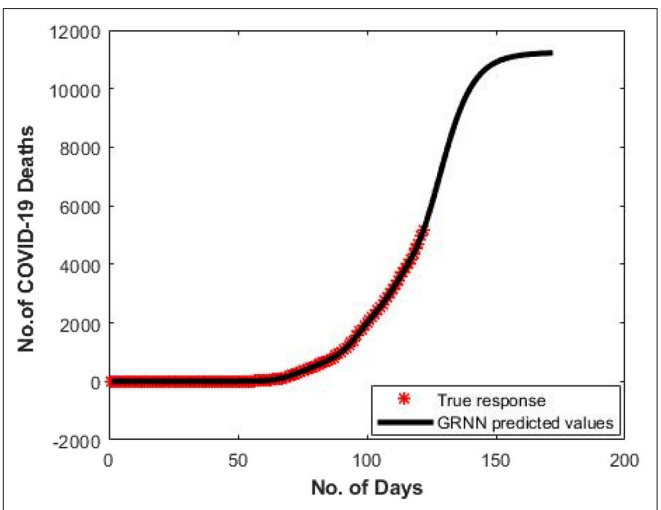

FIGURE 7 | Predicted curve for D1 using GRNN.

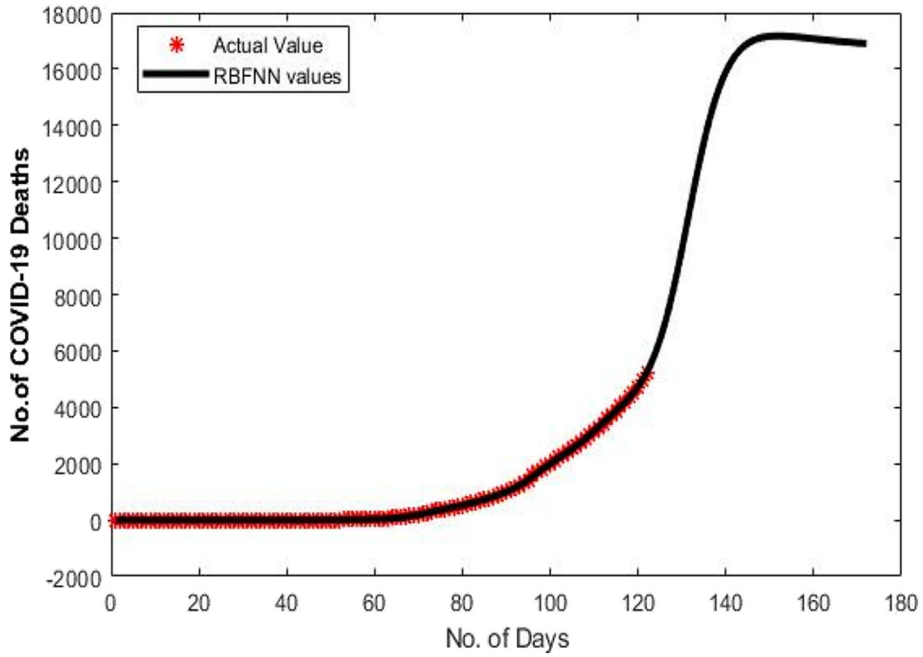

FIGURE 8 | Predicted curve for D1 using RBFNN.

SNN model for COVID-19 death cases, respectively. Here, the X-axis represents the dates and Y-axis represents the number of death cases predicted.

Tables 4, **5** show the predicted number of COVID-19 death cases using time series data (i.e., D1) for three standard and hybrid SNN models. The hybrid model is the combination of standard models and error forecasting model using NAR-NN. There is no difference in the predicted values for the standard and hybrid models since their RMSE value is about 0.2 approximately.

Table 6 shows the predicted number of COVID-19 death cases using time series data (i.e., D2) for three standard and hybrid SNN models.

Table 7 shows the calculated MRP for COVID- 19 predicted death cases using the dataset (D1). MRP is defined as in Equation (9). It is described as the number of predicted death cases divided by the number of confirmed cases and then multiplied by 100. It shows the number of COVID-19 deaths per 100 COVID-19 confirmed cases.

$$MRP = \frac{Number\ of\ Predicted\ Death\ Cases}{Number\ of\ Confirmed\ Cases} * 100 \qquad (8)$$

Figure 6 shows the predicted curve of PNN for COVID-19 death cases vs. the number of days since the first COVID-19 case for India, respectively. Here, the X-axis represents the number of days and Y-axis represents the number of death cases predicted. For the dataset (D1), the PNN model shows a gradual decrease in the number of death cases after the 130th day since 1st COVID-19 in India.

Figure 7 denotes the predicted curve of GRNN for COVID-19 death cases vs. the number of days since the first COVID-19 case for India, respectively. Here, the X-axis signifies the number of days and Y-axis signifies the number of death cases predicted. For the dataset (D1), the GRNN model shows a smoothing means curve in the number of death cases after the 130th day since 1st COVID-19 in India, while the GRNN models show a smoothly increasing pattern.

Figure 8 shows the predicted curve of RBFNN for COVID-19 death cases vs. the number of days since the first COVID-19 case for India, respectively. Here, the X-axis represents the number of days and Y-axis represents the number of death cases predicted. The shape of the curve is bell curve. For the dataset (D1), the RBFNN model shows a increasing in the number of death cases after the 130th day since 1st COVID-19 in India.

Figure 9 illustrates the predicted curve of PNN for D2 death cases vs. the number of confirmed cases for India. Here, the X-axis signifies the number of confirmed cases and Y-axis signifies the number of death cases predicted. The PNN model shows a sharp decrease in the number of death cases after the number of COVID-19 confirmed cases reach 220,000 nearly.

Figure 10 proves the predicted curve of GRNN for COVID-19 death cases vs. the number of confirmed cases for India. Here, the X-axis depicts the number of confirmed cases and Y-axis signifies the number of death cases predicted. The GRNN model shows a decrease pattern in the number of death cases after 245,000 COVID-19 confirmed cases.

Figure 11 illustrates the RBFNN predicted curve for COVID-19 death cases vs. the number of confirmed cases for India. Here, the X-axis signifies the number of confirmed cases and Y-axis signifies the number of death cases predicted. The RBFNN shows

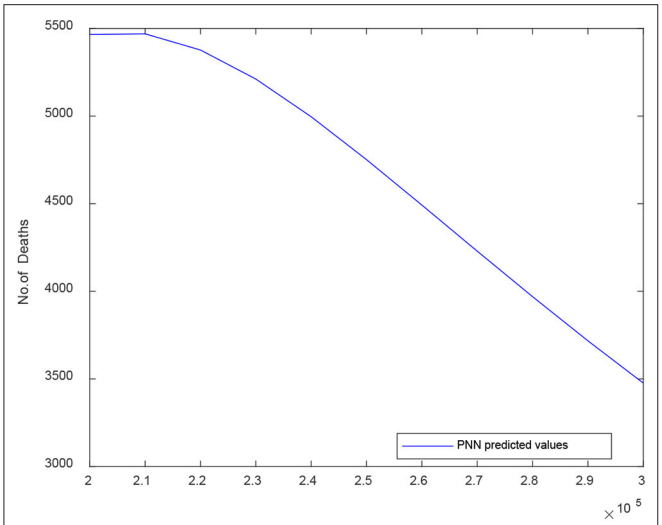

FIGURE 9 | Predicted curve for D2 using PNN.

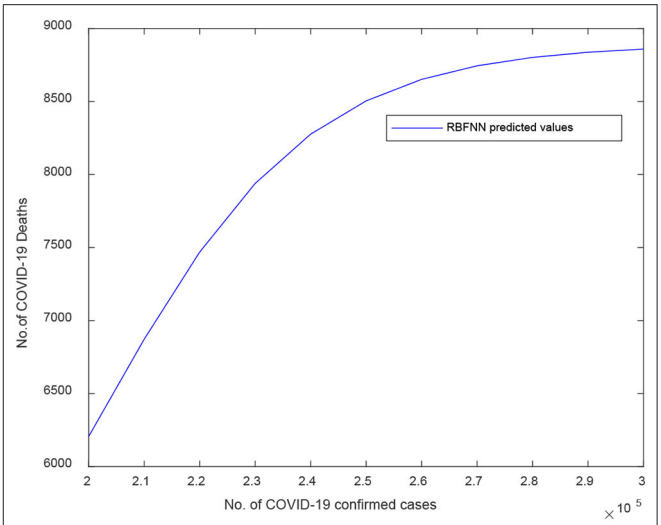

FIGURE 11 | Predicted curve for D2 using RBFNN.

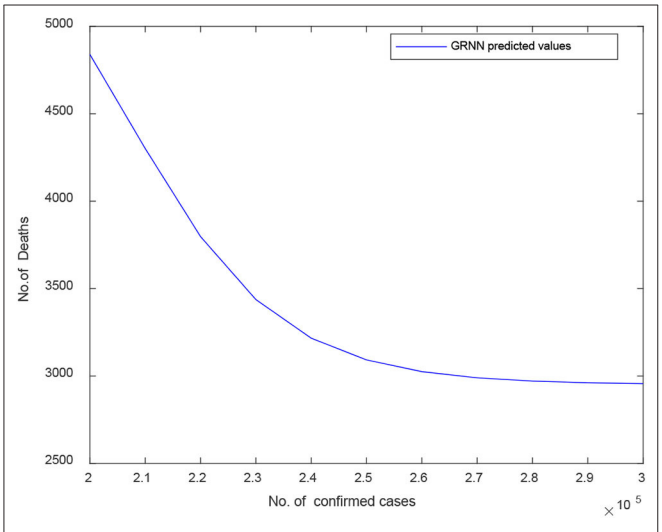

FIGURE 10 | Predicted curve for D2 using GRNN.

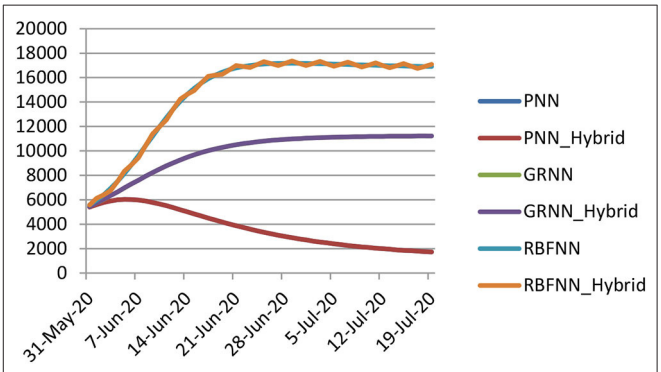

FIGURE 12 | Comparison of SNN models and hybrid models for D1.

an increasing pattern in the number of death cases after 245,000 COVID-19 confirmed cases.

Figure 12 depicts that forecasted error is very less or almost zero in the case of PNN and GRNN whereas RBFNN model shows slightly high error value.

The advantage of this study is that the COVID-19 mortality rate prediction using the SNN model and its hybrid models and gives a profound and solid comprehension of the pattern and qualities of COVID-19. An important observation was made from this investigative performance study of SNN models for the COVID-19 datasets, that no single neural network model can be considered the best model, which depends entirely on the neural network parameters and the characteristics of the data.

Limitations

There is some limitation of our current work, which are as follows:

- First, COVID-19 is initially recognized as mild illness with dry cough, in more cases there are asymptomatic and seldom leads to death. The majority of COVID-19 cases in India is asymptomatic and very mildly infected individuals, which they are not available to human services experts, which resulted in under-reporting.
- Second, other demographical and topographical components related with the event and spread of COVID-19 are excluded from the proposed SNN models; thus, regardless of whether the SNN models consider these factors, encourages the improvement in the prescient exactness will require further confirmation.
- Lastly, the hybrid SNN-NAR-NN model is developed based on the benchmark neural network regression model that is suited for short-term mortality rate prediction very well. Finally,

applicability of these SNN models in other infectious diseases may be carried out as future work.

CONCLUSION AND FUTURE WORK

This research paper proposed a SNN models and their hybrid version with the NAR-NN time series model for the prediction of the COVID-19 mortality rate in India. The performances of these models have evaluated by using RMSE and "R," a correlation value. Based on the comparison of the RMSE values of these models, it was found that the SNN models for D1 are higher than D2. Therefore, in this work, SNN is hybridized with NAR-NN for dataset D1 to predict the future error of the SNN models, which was added to the predicted value of these models for better mortality rate prediction. On the whole, the empirical results were showed that: (i) RBFNN based MRP model performed better than the GRNN and PNN models for D1 dataset, (ii) PNN based MRP model performed better than the GRNN and RBFNN models for D2 dataset. For the both datasets, SNN based MRP models have captured the incremental curve for COVID-19 death cases for India. The proposed method is capable of providing a predictive tool for assessing its current state of infection, severity, and help government and health care workers for better decision making to reduce the mortality rate in India.

In the future, deep learning Recurrent Neural Network time series forecasting model will be used to increase the prediction accuracy for the COVID-19 mortality rate prediction. And also, this study will be enhanced by including many factors or variables like demographical factors, geographical factors, and weather factors (temperature, humidity, wind speed, and rainfall) for modeling the highly accurate prediction model for ongoing COVID-19 pandemic.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset?select=time_series_covid_19_deaths.csv.

AUTHOR CONTRIBUTIONS

SD and RR: conceptualization, implementation, and design. SD, RR, and JC: writing and editing. JC: verification. All authors contributed to the article and approved the submitted version.

REFERENCES

1. WHO. (2020). Available online at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200123-sitrep-3-2019-ncov.pdf?sfvrsn=d6d23643_8 (accessed February 1, 2020).
2. WHO. (2020). Available online at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200131-sitrep-11-ncov.pdf?sfvrsn=de7c0f7_4 (accessed January 30, 2020).
3. Sparrow A. *"How China's Coronavirus Is Spreading-and How to Stop It," January 26, 2020.* (2020). Available online at: https://foreignpolicy.com/2020/01/26/2019-ncov-china-epidemic-pandemic-the-wuhan-coronavirus-a-tentative-clinical-profile/ (accessed February 2, 2020).
4. Sujath R, Chatterjee JM, Hassanien AE. A machine learning forecasting model for COVID-19 pandemic in India. *Stoch Environ Res Risk Assessment.* (2020) 34:1. doi: 10.1007/s00477-020-01827-8
5. Langousis A, Carsteanu AA. Undersampling in action and at scale: application to the COVID-19 pandemic. *Stoch Environ Res Risk Assessment.* (2020) 34:1–3. doi: 10.1007/s00477-020-01821-0
6. Deprez P, Shevchenko PV, Wüthrich MV. Machine learning techniques for mortality modeling. *Eur Actuarial J.* (2017) 7:337–52. doi: 10.1007/s13385-017-0152-4
7. Hainaut D. A neural-network analyzer for mortality forecast. *ASTIN Bull.*(2018) 48:481–508. doi: 10.1017/asb.2017.45
8. Richman R, Wüthrich MV. A neural network extension of the Lee–Carter model to multiple populations. *Annal Actuarial Sci.* (2018) 1–21. doi: 10.1017/S1748499519000071
9. Fekrazad F. A best approach in intrusion detection for computer network PNN/GRNN/RBF. *Int J Comput Sci Issues.* (2014) 11:182.
10. Jeatrakul P, Wong KW. Comparing the performance of different neural networks for binary classification problems. In: *2009 Eighth International Symposium on Natural Language Processing*. Bangkok: IEEE (2009). p. 111–5.
11. Ebden M. *Gaussian Processes for Regression: A Quick Introduction.* (2008) Available online at: http://www.robots.ox.ac.uk/\simmebden/reports.GPtutorial.Pdf
12. Han Q, Su H, Wang CC, Shan XW, Chang WW, Xu ZW. Prediction on the incidence of blood and sexually transmitted diseases with models of ARIMA and GRNN. *Modern Prev Med.* (2012) 6:1337–40.
13. Hong S, Zhou Z. Application of gaussian process regression for bearing degradation assessment. In: *2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining (ISSDM2012).* Taipei: IEEE (2012). p. 644–48.
14. Montazer GA, Giveki D, Karami M, Rastegar H. Radial basis function neural networks: a review. *Comput Rev J.* (2018) 1:52–74.
15. AŽman K, Kocijan J. Dynamical systems identification using Gaussian process models with incorporated local models. *Eng Appl Artif Intell.* (2011) 24:398–408. doi: 10.1016/j.engappai.2010.10.010
16. Aksoy H, Dahamsheh A. Artificial neural network models for forecasting monthly precipitation in Jordan. *Stoch Environ Res Risk Assessment.* (2009) 23:917–31. doi: 10.1007/s00477-008-0267-x
17. Hajmeer M, Basheer I. A probabilistic neural network approach for modeling and classification of bacterial growth/no-growth data. *J Microbiol Methods.* (2002) 51:217–26. doi: 10.1016/S0167-7012(02)00080-5
18. Huber MF. Recursive gaussian process: on-line regression and learning. *Pattern Recogn Lett.* (2014) 45:85–91. doi: 10.1016/j.patrec.2014.03.004
19. Williams CK, Rasmussen CE. *Gaussian Processes for Machine Learning*. Vol. 2. Cambridge, MA: MIT Press (2006).
20. Khamparia A, Singh A, Anand D, Gupta D, Khanna A, Kumar NA, et al. A novel deep learning-based multi-model ensemble method for the prediction of neuromuscular disorders. *Neural Comput Appl.* (2018) 1–13. doi: 10.1007/s00521-018-3896-0
21. Wang Y, Xu C, Li Y, Wu W, Gui L, Ren J, et al. An advanced data-driven hybrid model of SARIMA-NNNAR for tuberculosis incidence time series forecasting in Qinghai Province, China. *Infect Drug Resist.* (2020) 13:867. doi: 10.2147/IDR.S232854
22. Singh RK, Rani M, Bhagavathula AS, Sah R, Rodriguez-Morales AJ, Kalita H, et al. Prediction of the COVID-19 pandemic for the top 15 affected countries: advanced autoregressive integrated moving average (ARIMA) model. *JMIR Public Health Surveil.* (2020) 6:e19115. doi: 10.2196/19115
23. Chouhan V, Singh SK, Khamparia A, Gupta D, Tiwari P, Moreira C, et al. A novel transfer learning-based approach for pneumonia detection in chest X-ray images. *Appl Sci.* (2020) 10:559. doi: 10.3390/app10020559
24. Iwendi C, Bashir AK, Pasupuleti NS, Sujatha R, Chatterjee JM, Peshkar A, et al. COVID-19 Patient health prediction using boosted random forest algorithm. *Front Public Health*. (2020) 8:357. doi: 10.3389/fpubh.2020.00357

25. Tomar A, Gupta N. Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *Sci Total Environ.* (2020) 728:138762. doi: 10.1016/j.scitotenv.2020.138762
26. Waheed A, Goyal M, Gupta D, Khanna A, Al-Turjman F, Pinheiro PR. Covidgan: data augmentation using auxiliary classifier gan for improved covid-19 detection. *IEEE Access.* (2020) 8:91916–23. doi: 10.1109/ACCESS.2020.2994762
27. Novel Corona Virus 2019 Dataset. Kaggle (2020). Available online at: https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset?select=time_series_covid_19_deaths.csv (accessed June 27, 2020)
28. Bagtzoglou AC, Hossain F. Radial basis function neural network for hydrologic inversion: an appraisal with classical and spatio-temporal geostatistical techniques in the context of site characterization. *Stoch Environ Res Risk Assessment.* (2009) 23:933–45. doi: 10.1007/s00477-008-0262-2
29. Seyedzadeh S, Rahimian FP, Glesk I, Roper M. Machine learning for estimation of building energy consumption and performance: a review. *Visual Eng.* (2018) 6:5. doi: 10.1186/s40327-018-0064-7
30. Saha S, Saha B, Saxena A, Goebel K. Distributed prognostic health management with Gaussian process regression. In: *2010 IEEE Aerospace Conference*. Big Sky, MT: IEEE (2010). p. 1–8.
31. Ahmed NK, Atiya AF, Gayar NE, El-Shishiny H. An empirical comparison of machine learning models for time series forecasting. *Econometr Rev.* (2010) 29:594–621. doi: 10.1080/07474938.2010.481556
32. Sharma N, Om H. Usage of probabilistic and general regression neural network for early detection and prevention of oral cancer. *Sci World J.* (2015) 2015:234191. doi: 10.1155/2015/234191
33. Martínez F, Charte F, Rivera AJ, Frías MP. Automatic time series forecasting with GRNN: a comparison with other models. In: *International Work-Conference on Artificial Neural Networks.* Cham: Springer (2019). p. 198–209.
34. Dhamodharavadhani S, Rathipriya R. Enhanced logistic regression (ELR) model for big data. In: Fausto Pedro Garcia Marquez, editor. *Handbook of Research on Big Data Clustering and Machine Learning.* IGI Global (2020). p. 152–76.
35. Dhamodharavadhani S, Rathipriya R. Variable selection method for regression models using computational intelligence techniques. In: Padmavathi G, and Shanmugapriya D, editors. *Handbook of Research on Machine and Deep Learning Applications for Cyber Security*. IGI Global (2020). p. 416–36.
36. Asante-Okyere S, Shen C, Yevenyo Ziggah Y, Moses Rulegeya M, Zhu X. Investigating the predictive performance of Gaussian process regression in evaluating reservoir porosity and permeability. *Energies.* (2018) 11:3261. doi: 10.3390/en11123261
37. Al-Mahasneh AJ, Anavatti SG, Garratt MA. Review of applications of generalized regression neural networks in identification and control of dynamic systems. *arXiv.* (2018).
38. Rasmussen CE. Gaussian processes in machine learning. In: Bousquet O, von Luxburg U, and Rätsch G, editors. *Summer School on Machine Learning.* Berlin: Springer (2003). p. 63–71.
39. Hannan SA, Manza RR, Ramteke RJ. Generalized regression neural network and radial basis function for heart disease diagnosis. *Int J Comput Appl.* (2010) 7:7–13. doi: 10.5120/1325-1799

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 Public Health

BRIEF RESEARCH REPORT
published: 09 September 2020
doi: 10.3389/fpubh.2020.00478

Impact of Endoplasmic Reticulum Stress Sensors on Pectolinarin Induced Apoptosis

*Ji-Hye Song[1], Kisang Kwon[1], O-Yu Kwon[1], Eun-Ryeong Lee[2], Seung-Whan Kim[3] and Kyung-Hee Kang[4]**

[1] Departments of Anatomy and Cell Biology, College of Medicine, Chungnam National University, Daejeon, South Korea,
[2] Department of Biomedical Laboratory Science, College of Health and Welfare, Kyungwoon University, Gumi, South Korea,
[3] Department of Emergency Medicine, College of Medicine, Chungnam National University, Daejeon, South Korea,
[4] Department of Dental Hygiene, College of Medical Science, Konyang University, Daejeon, South Korea

Pectolinarin, [5,7-Dihydroxy 4′,6-dimethoxyflavone 7-rutinoside, 7-[[6-O-(6-Deoxy-α-L-mannopyranosyl)-β-D-glucopyranosyl] oxy]-5-hydroxy-6-methoxy-2-(4-ethoxyphenyl)-4H-1-benzopyran-4-one], has been stated one of the major compounds in *Cirsium nipponicum* (Maxim.) Makino. It is characterized by biological functions of hepatoprotective, anti-inflammatory and antiobesity activities. In this research, it was explained that pectolinarin causes apoptosis in PC12 cells conducted by DNA fragmentation and formation on apoptotic bodies through the activation of ER stress sensors (ATF6 fragmentation and eIF2α phosphorylation). The result of treating the PC12 cells with 50 μM pectolinarin for 24 h has come to increase ATF6 mRNA expression up to 1.6 times, PERK expression up to 1.7 times and IRE1 expression up to 1.4 times, respectively, compared to those of the control. ATF6 fragmentation by pectolinarin treatment was increased about 2 times compared with its control, and phosphorylation of eIF2α was increased 2.5 times. The results proposed that the perception of the molecular mechanisms underlying pectolinarin-caused apoptosis may be useful in new natural medicinal products and health supplements for the apoptosis-related diseases.

Keywords: pectolinarin, PC12 cells, endoplasmic reticulum (ER), stress, sensor

OPEN ACCESS

Edited by:
Deepak Gupta,
Maharaja Agrasen Institute of Technology, India

Reviewed by:
Sandeep Kumar Satapathy,
VIT University, India
Gyoo-Soo Chae,
Baekseok Culture University, South Korea
Shruti Mishra,
VIT-AP University, India

***Correspondence:**
Kyung-Hee Kang
dhkhkang@konyang.ac.kr

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Public Health

Received: *11 June 2020*
Accepted: *27 July 2020*
Published: *09 September 2020*

Citation:
Song J-H, Kwon K, Kwon O-Y, Lee E-R, Kim S-W and Kang K-H (2020) Impact of Endoplasmic Reticulum Stress Sensors on Pectolinarin Induced Apoptosis. Front. Public Health 8:478. doi: 10.3389/fpubh.2020.00478

INTRODUCTION

Cirsium nipponicum (Maxim). Makino is an asteraceae perennial herb, called Island thistle (in English) and Mul-eong-gcong-kwi (in Korean), distributed widely throughout Ul Leung Do (an island located at the east of Korean Peninsula), and it is used as a medicinal or edible plant (1). In Oriental medicine, roots are generally used for disease treatment, therefore the leaves and stems are collected when flowers are in full bloom (2). Traditionally, it has been widely used as a traditional medicinal product for the treatment of bleeding, hepatitis, hypertension and blood circulation diseases (3, 4). Recently, pharmacological studies have shown that its extract has antitumor (5) and antidiabetic (6), antioxidant (7), anti-inflammatory (8), and antifungal functions (9). This contains a significant amount of flavonoid mixtures, among which pectolinarin has been reported to be the major compound (10). The important biological activity of pectolinarin reported so far is as follows; anti-inflammatory, hepatoprotective, antiobesity activities and analgesic effect (11). However, the biological function of pectolinarin is not precisely defined.

The endoplasmic reticulum (ER) is an organelle spotted in eukaryotic cells. It is a very important manufacturing site for the post-translational step that mediates the synthesis, folding, modification and transport of secretory proteins. Some kinds of stressors that disrupt the endoplasmic reticulum function lead to accumulation of un-, mis-, misfolded proteins in the endoplasmic reticulum lumen (12). The endoplasmic reticulum stress induces ER-stress adaptable signal called the unfolded protein response (UPR) to maintain endoplasmic reticulum homeostasis *via* activation of endoplasmic reticulum chaperones, such as binding immunoglobulin protein (BiP), glucose-regulated protein 94 (GRP94), calnexin, calreticulin, endoplasmic reticulum protein 29 (ERp29), heat shock protein 47 (HSP47) and protein disulfide isomerase (PDI), which regulates three types of endoplasmic reticulum stress sensors, containing IRE1(Inositol Requiring Enzyme 1), PERK(PKR-like ER kinase) and initiating ATF6(transcription factor 6) (13–15). Though this study did not fully document the specific significance of endoplasmic reticulum stress protein expression, it demonstrated that pectolinarin controls the expression of endoplasmic reticulum stress sensors associated with apoptosis using the PC12 cells, which is widely used as a classical neuronal cell model.

MATERIALS AND METHODS

Sample, Cell Culture, and MTT Assay

Pectolinarin (chemical formula, $C_{29}H_{34}O_{15}$; molar mass, 622.57 g/mol) purified at a purity of >95.0% (HPLC) derived from *Cirsium nipponicum* (Maxim.) Makino was gifted by National Development Institute of Korean Medicine (NIKOM). The PC12 cells were cultured in collagen-coated plates or flasks containing 85% RPMI-1640 medium, augmented with 25 mM HEPES buffer, horse serum 10% heat inactivated, fetal bovine serum 5% heat inactivated, 1 mM sodium pyruvate, 1 g/l d-(+) -glucose, 2 mM L-glutamine, 25 μg/ml streptomycin and 25 U/ml penicillin (all Gibco; Thermo Fisher Scientific, Inc., USA). The cells were preserved in a humidified incubator at 37°C at 5% CO_2 and the medium was changed every 48 h. The effects of pectolinarin on cell survival of the PC12 cells were made using an MTT kit (Sigma-Aldrich, USA). Color development was observed at 595 nm with a reference wavelength of 650 nm using the Sunrise™ microplate reader (Tecan Trading AG, Switzerland).

RT-PCR Analysis

Each gene expression was mainly determined by RT-PCR as described below. RT-PCR conditions included 30 cycles comprising each of the following: 94°C for 30 s, 58°C for 30 s and 72°C for 1 min (10 min in the final cycle) employing the primers with *Taq* DNA polymerase (Solgent Co., Ltd., Korea). The RT-PCR primers were provided by Bioneer Corporation, Korea. The RT-PCR primers were as follows: IRE1 forward, 5′-ACC ACC AGT CCA TCG CCA TT-3′ and reverse, 5′-CCA CCC TGG ACG GAA GTT TG-3′; ATF6 forward, 5′-CTA GGC CTG GAG GCC AGG TT-3′ and reverse, 5′-ACC CTG GAG TAT GCG GGT TT-3′; PERK forward, 5′-GGT CTG GTT CCT TGG TTT CA-3′ and reverse, 5′-TTC GCT GGC TGT GTA ACT TG-3′; BiP forward, 5′-AGT GGT GGC CAC TAA TGG AG-3′ and reverse, 5′-TCT TTT GTC AGG GGT CGT TC-3′. Bcl-xl forward, 5′-CCC CAG AAG AAA CTG AAC CA-3′ and reverse, 5′-GCA GAA CTA CAC CAG CCA CA-3′; Bax forward, 5′-AGG GGC CTT TTT GTT ACA GG-3′ and reverse, 5′-GAT CAG CTC GGG CAC TTT AG-3′ Bcl2 forward, 5′-AAG CTG CAC AGC GGG GCT A-3′ and reverse, 5′-CAG ATG CCG GTT CAG GTA CT-3′ Bak1 forward, 5′-TTA CCT CCA GCA GGA AC-3′ and reverse, 5′-ACC ACC TCT CTG TGC AAT CC-3′ LC3a forward, 5′-GCC TGT CCT GGA TAA GAC CA-3′ and reverse, 5′-GTT CAC CAG GAA GG-3′ Beclin forward, 5′-GTG CTC CTG TGG AAT GGA AT-3′ and reverse, 5′-GCT GCA CAC AGT CCA GAA AA-3′ Cal forward, 5′-GGC ATC TTC ATC CCA GTC AT-3′ and reverse, 5′-CTC CTC TCT GCT CCT CAT GG-3′ PDI forward, 5′-CAG AGT TCT GCC ACC GCT TC-3′ and reverse, 5′-TCC TCG AGA TCG TCA TC-3′ ERp29 forward, 5′-CTC CTC TCT GCT CCT CAT GG-3′ and reverse, 5′-GCT CCA TGT TCA GCT TGT CA-3′ Xbp1 forward, 5′-AAA CAG AGT AGC TCA GAC TGC-3′ and reverse, 5′-TCC TCC TGG GTA GAC CTC TGG GAG-3′. All chemicals were acquired from Sigma-Aldrich; Merck KGaA. The figures show the outputs of a representative experiment in triplicate with different sampling units.

Western Blotting

Immunoblotting was carried out according to the standard procedure. The PC12 cells treated with pectolinarin were dissolute by the addition of SDS sample buffer [62.5 mM Tris-HCl, pH 6.8, 6% (w/v) SDS, 30% glycerol, 125 mM DTT, 0.03% (w/v) bromphenol blue] and parted by SDS-PAGE. The proteins were moved to a nitrocellulose membrane, and the membrane was developed with the main antibodies. The rabbit anti-eIF2α antibody, eIF2α-P antibody and goat anti-actin antibody were acquired from Santa Cruz Biotechnology, Inc., USA. The mouse anti-ATF6 antibody was acquired from Novus Biologicals, LLC, USA. The horseradish peroxidase-conjugated anti-rabbit, anti-goat and anti-mouse IgG secondary antibodies were acquired from Santa Cruz Biotechnology, Inc. Goat anti-actin antibody was used to systematize the quantity of sample proteins. The blots were developed by utilizing an upgraded chemiluminescence Western blotting detection system kit (Amersham, Sweden). The experiments were carried out in triplicate, and the protein bands were defined employing Image J software (version 1.48; https://imagej.nih.gov/ij/).

DNA Fragmentation Assay and Hoechst 33342 Staining

Following the treatment with pectolinarin, the PC12 cells were dissolute in 100 μl 10 mM Tris-HCl buffer (pH 7.4) having 10 mM EDTA and 0.5% Triton X-100. After centrifugation for 5 min at 16,000 x g, the supernatant dealt with RNase A and proteinase K (Promega Corporation, Madison, WI, USA). Eventually, 20 μl of 5 M NaCl and 120 μl isopropanol were put together and preserved with ice for 1 h. Following centrifugation for 15 min at 16,000 x g, the DNA pellets were dissolved in

20 μl TE buffer. The DNA samples were stuffed onto a 0.7% agarose gel and noted using a UV source after ethidium bromide (Sigma-Aldrich; Merck KGaA) staining. After treatment with pectolinarin, the PC12 cells were developed in incubator for 30 min with Hoechst 33342 (Molecular Probes; Thermo Fisher Scientific, Inc.) loading dye and washed in ice-cold 1X PBS. After following staining for 10 min, the stained cells were screened by employing a fluorescence microscope (Axio Scope A1; Zeiss GmbH, Germany) at 340 nm.

RESULTS AND DISCUSSION

Cell Viability and ER Chaperone Expression

Pectolinarin is a representative component of flavonoid mixtures isolated from *Cirsium nipponicum* (Maxim.) Makino (10). It has anti-inflammatory activities and is similar in chemical structure to linarin (8, 16, 17). It is well-known that one of the remarkable features of cellular injury is leakage of soluble lactate dehydrogenase (LDH) from the cells stimulated by any stimulant. As a result of assaying, pectolinarin did not induce LDH leakage, so it was proven not to cause cellular injury within a certain concentration (18). In this study, the PC12 cell line used here is extracted from a pheochromocytoma cell of the rat adrenal medulla having a mixture of neuroblastic and eosinophilic cells. This cell line has been extensively used in *in vitro* studies to examine various neuronal diseases (19, 20). At the beginning of this study, we tested the effects of pectolinarin on cell viability in the MTT assay following at 1, 5, 10, 50, and 100 μg/ml of pectolinarin treatment for 24 h. As shown in **Figure 1A**, the result of the MTT assay has shown that no morphological differences were shown at concentrations below 100 μg/ml pectolinarin treatment and control cells were observed.

The endoplasmic reticulum which has newly-synthesized secretory and cell membrane proteins are post-translationally modified and correctly folded. Protein modification and folding in the endoplasmic reticulum are impaired by the endoplasmic reticulum chaperones, such as BiP, GRP94, calnexin, calreticulin, ERp29 and PDI. One of the most abundant endoplasmic reticulum chaperones in the endoplasmic reticulum lumen is BiP, a member of the Hsp70 family of proteins, which perceives newly synthesized proteins as they are transferred in the endoplasmic reticulum and sustains them in a state of competent for following folding and oligomerization. Like BiP, each endoplasmic reticulum chaperone exerts a unique function in the endoplasmic reticulum lumen to complete the correct protein folding (13–15). Although the effect of pectolinarin on cell viability was not confirmed in the MTT assay, the expression of endoplasmic reticulum chaperones (PDI, BiP, calnexin and ERp29) which indicate the degree of cell stress, was examined by pectolinarin treatment in the PC12 cells. As presented in **Figure 1B**, notable endoplasmic reticulum chaperone expression was not observed in the PC12 cells treated with pectolinarin. In summary, no remarkable cell viability and endoplasmic reticulum chaperone expression were observed by pectolinarin treatment on the PC 12 cells.

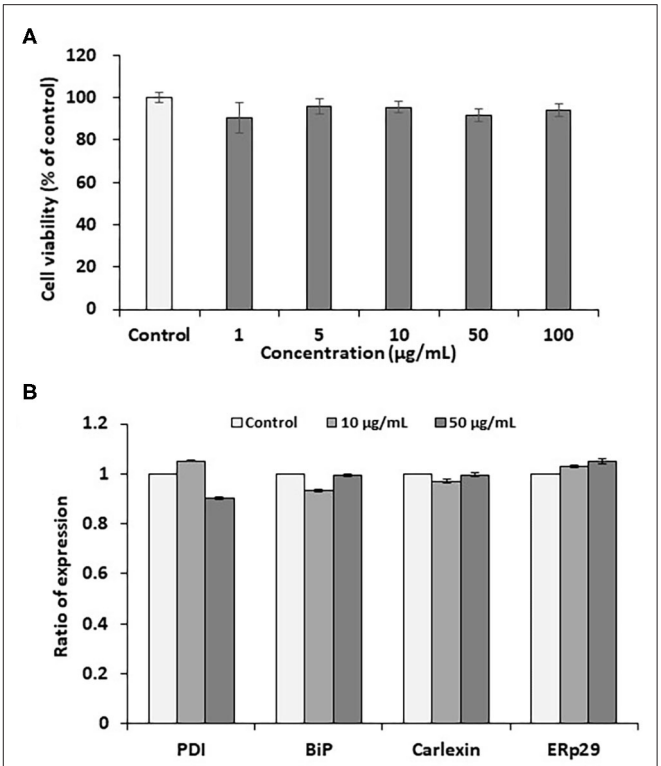

FIGURE 1 | Effects of pectolinarin on cell viability and gene expression of endoplasmic reticulum chaperones in the PC12 cells. **(A)** The cells were under the incubator for 24 h with diverse concentrations of pectolinarin (0–0.1 mg/ml). Then cell viability assay was performed using MTT kit (Sigma-Aldrich). **(B)** The cells were treated with pectolinarin (10 and 50 μg/ml) for 24 h. The results of gene expression of endoplasmic reticulum chaperones of disulfide isomerase (PDI), binding immunoglobulin protein (BiP), calexin and ERp29. The RT-PCR outcomes were checked three times.

Expression and Activation of Endoplasmic Reticulum Sensors

Although pectolinarin induces unremarkable cell viability and endoplasmic reticulum chaperone expression, next we have examined the gene expression of endoplasmic reticulum stress sensors (ATF6, PERK and IRE1) and the activations by pectolinarin treatment on the PC 12 cells. Under altered endoplasmic reticulum homeostasis, the endoplasmic reticulum stress signal transduction pathway is mediated *via* being active of ER(endoplasmic reticulum)- stress sensors. IRE1 activates XBP mRNA cleaving, producing an activated form of the XBP1 protein. PERK induces phosphorylation of the eIF2α, which hinders translation initiation. Active ATF6 is cleaved at the cytosolic face and the resulting N-terminal cytoplasmic domain ties to the endoplasmic reticulum stress-responding element, which enhances endoplasmic reticulum chaperone gene expression (21).

We have determined the expression of endoplasmic reticulum stress sensors under same experimental conditions as described in **Figure 1B**. As showed in **Figure 2A**, dealing the PC12 cells with 50 μM pectolinarin for 24 h increased ATF6 mRNA

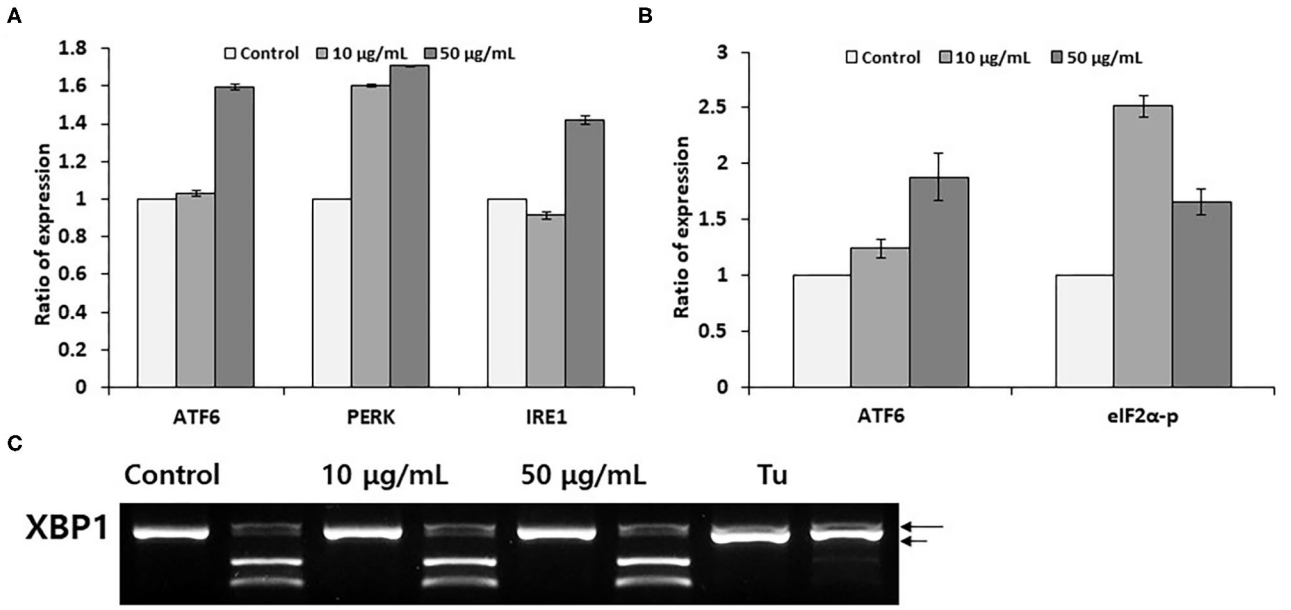

FIGURE 2 | Pectolinarin commands ERSS(Endoplasmic Reticulum-Stress Sensors). **(A)** The cells were dealt with pectolinarin (10 and 50 μg/ml) for 24 h. The gene expression of endoplasmic reticulum stress sensors of ATF6, PERK and IRE1 was performed by RT-PCR. (B) Western blotting was performed using anti-eIF2-α antibody/eIF2-α-P antibody, anti-ATF6 monoclonal antibody against cells treated with pectolinarin (10 and 50 μg/ml) for 24 h, respectively. The resulting bands (fragmented ATF6 and phosphorylated eIF2-α) were estimated by Image J software (version 1.48). **(C)** RT-PCR was performed against mRNAs of XBP1. The following PCR product was moreover digested by *Pst*I to disclose a restriction site that was lost following XBP1 splicing under endoplasmic reticulum stress. The following XBP1 cDNA products were disclosed on a 2% agarose gel. Unspliced XBP1 mRNA made the two lower bands which were indicated by arrows (upper, 290 bp and lower, 183 bp). The spliced XBP1 mRNA pointed out by a bold arrow. ATF6, activating transcription factor; PERK, PKR-like ER kinase (PERK); IRE1, inositol requiring enzyme 1 (IRE1); eIF2-alpha-P, phosphorylated form of translation initiation factor eIF2α; XBP1, X-box tying protein 1, Tu, tunicamycin, respectively.

expression up to 1.6 times, PERK expression is 1.7 times and IRE1 expression is 1.4 times compared to those of the control, respectively. This result suggests that although pectolinarin does not regulate endoplasmic reticulum chaperone expression directly, it regulates the gene expression of endoplasmic reticulum stress sensors. ATF6 fragmentation by pectolinarin treatment was increased about 2 times compared with its control, and phosphorylation of eIF2α was increased 2.5 times (**Figure 2B**). However, there was a little change in the XBP1 mRNA uncleavaging that means endoplasmic reticulum stress level (**Figure 2C**). In summary, it has been shown that pectolinarin treatment on PC 12 cells actively regulates endoplasmic reticulum stress sensor activity through both the ATF6 fragmentation and eIF2α phosphorylation rather than the regulation of ER chaperone gene expression.

Induction of Apoptosis

If early cellular responses fail to maintain endoplasmic reticulum homeostasis, endoplasmic reticulum stress that activates UPR signal to stimulate the apoptosis as well as this autophagy for cell survival or local cell death. Already, there are already some reports that UPR strongly associated with both apoptosis and autophagy by several regulators, such as for apoptosis (Bax, Bak, Bcl2 and Bcl-xl) and autophagy (LC3a and Beclin) (22). We found that based on the results of **Figure 2**, nonetheless, pectolinarin induces mild endoplasmic reticulum chaperone expression, the ER stress sensors are actively induced. These findings provide new insights that the mild endoplasmic reticulum stress through PERK-eIF2α-p or/and ATF6 fragmentation indicating pathway has a main role in saving cellular damage from pectolinarin. We therefore tested the role of instantaneous endoplasmic reticulum stress in both apoptosis and autophagy induction of bystander cells treated with pectolinarin. **Figure 3A** showed that the pectolinarin induces only the expression of pro-apoptosis (Bax and Bak), no meaningful expression of anti-apoptosis (Bcl2 and Bcl-xl) and autophagy (LC3a and Beclin) induction. The result may provide an evidence indicating that pectolinarin induces apoptosis through endoplasmic reticulum stress signaling but not autophagy.

Apoptosis shows two typical cell changes, it morphologically makes apoptotic bodies by cell shrinkage and chromosomal DNA fragmentation. It is being used as a useful marker for identification of apoptotic cells detection of both apoptotic bodies by microscope and apoptotic DNA fragmentation *via* the DNA laddering assay (23). In this study, to confirm that pectolinarin induces apoptosis, we investigated PC12 cells treated by pectolinarin that shows apoptotic bodies and DNA fragmentation or not. As a result, inter-nucleosomal DNA fragmentation increased in cells treated with pectolinarin dose-dependently (**Figure 3B**). Moreover, apoptotic bodies were observed following Hoechst 33342 staining (**Figure 3C**). The

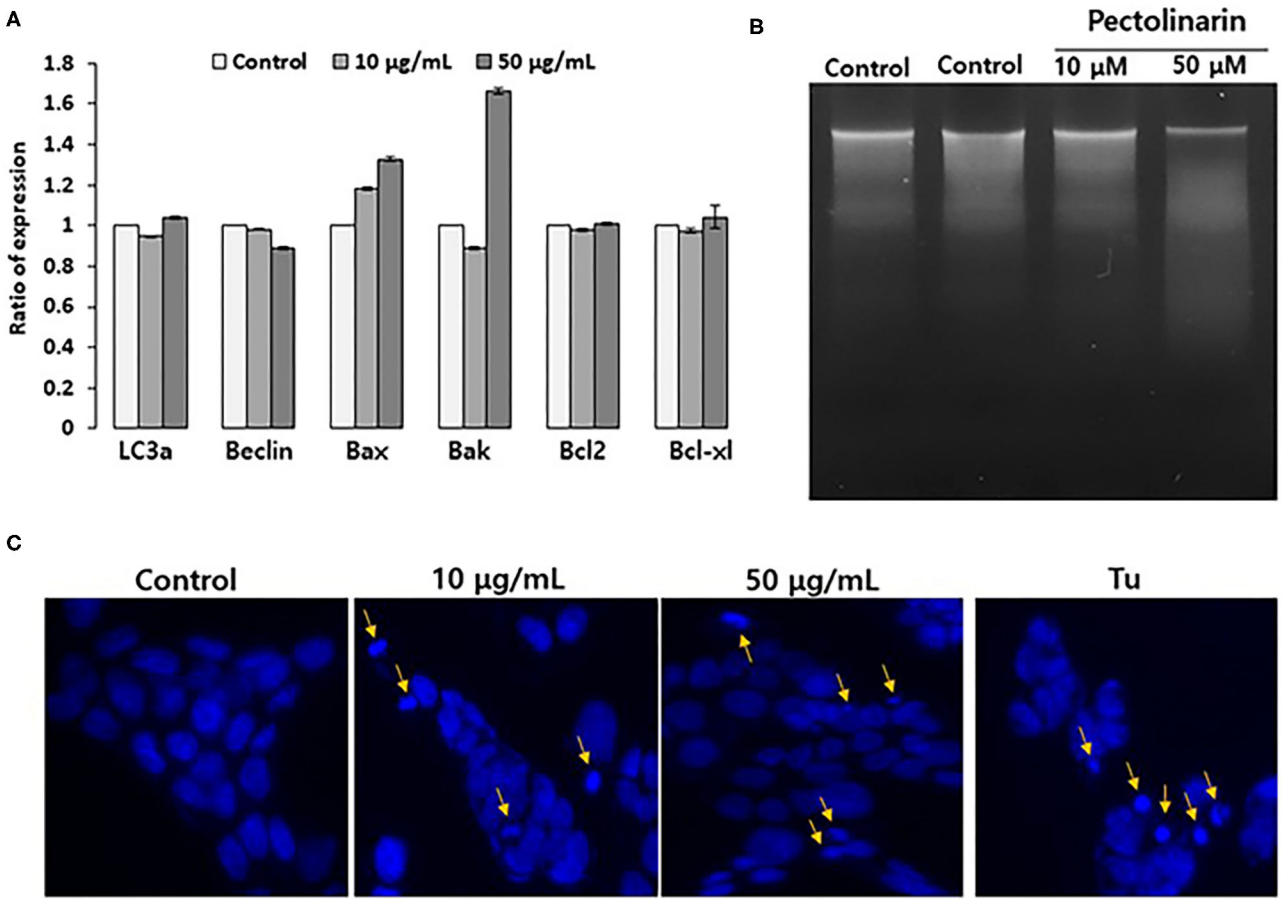

FIGURE 3 | Pectolinarin causes apoptosis in the PC12 cells. **(A)** The cells were treated with pectolinarin (10 and 50 μg/ml) for 24 h. RT-PCR was performed against various mRNAs; LC3a and Beclin-1 LC3a, microtubule-associated protein light chain 3; Beclin, coiled-coil moesin-like BCL2 interacting protein; Bax, Bcl-2-associated X; Bad, Bcl-2-associated death promoter; Bcl-2, B-cell lymphoma 2; Bcl-xl, B-cell lymphoma/leukemia-x long, respectively. **(B)** To make sure DNA fragmentation, the cells were dealt with pectolinarin (10 and 50 μg/ml) for 24 h, respectively. The DNA was lysed on a 1.5% agarose gel and pictured with Ethidium bromide, an intercalating agent. **(C)** The cells were dealt with the same conditions of the above **(B)** and mixed with Hoechst 33342 solution to perceive the formation of apoptotic bodies indicated by arrows. Mixed nuclei were detected from a fluorescent microscope employing a blue filter.

above results clearly revealed that pectolinarin participates in the induction of PC12 cells apoptosis.

In conclusion, pectolinarin dealings obviously caused apoptosis through the expression of ERSS(Endoplasmic Reticulum-Stress Sensors) in the PC12 cells. However, at this time, it is difficult to adequately explain the signaling through endoplasmic reticulum stress triggering several regulators connected with the apoptosis pathway for the extended endoplasmic reticulum stress. It is thought that better understanding of the biological mechanisms focusing pectolinarin-caused apoptosis may be helpful and useful in the diagnosis and treatment of apoptosis-related diseases based on new natural medicinal products.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

AUTHOR CONTRIBUTIONS

J-HS gave the conceptual idea. KK developed the methodology work. O-YK helped in implementation work. E-RL helped in result analysis work. S-WK helped in documentation work. K-HK helped in final editing and drafting of manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(Ministry of Education) (No. NRF-2017R1D1A1B03035544).

REFERENCES

1. Lee YN. *Flora of Korea. Seoul: Kyohaksa*. (2002) *p*. 843.
2. Kim JG. *Illustrated Natural Drugs Encyclopedia*. Seoul: *Namsandang*. (1984) *p*. 37.
3. Hertog MGL, Kromhout D, Aravanis C, Blackburn H, Buzina R, Ridanza F, et al. Flavonoid intake and long-term risk of coronary heart disease and cancer in the seven countries study. *Arch Int Med*. (1995) 155:381–6. doi: 10.1001/archinte.155.4.381
4. Heo J. *Dongui Bogam*. Seoul: Bubin Publishers Co. (1999). p. 1942.
5. Jung JH, Kim Y, Lee CO, Kang SS, Park JH, Im KS. Cytotoxic constituents of Saussurea lappa. *Arch Pharm Res*. (1998) 21:153–6. doi: 10.1007/BF02974020
6. Jeong DM, Jung HA, Choi JS. Comparative antioxidant activity and HPLC profiles of some selected Korean thistles. *Arch Pharm Res*. (2008) 31:28–33 doi: 10.1007/s12272-008-1116-7
7. Yoo YM, Nam JH, Kim MY, Choi J, Park HJ. Pectolinarin and pectolinarigenin of cirsium setidens prevent the hepatic injury in rats caused by D-galactosamine *via* an antioxidant mechanism. *Biol Pharm Bull*. (2008) 31:760–4 doi: 10.1248/bpb.31.760
8. Cho S, Lee J, Lee YK, Chung MJ, Kwon KH, Lee S. Determination of pectolinarin in Cirsium spp. using HPLC/UV analysis. *J Appl Biol Chem*. (2016) 59:107–12. doi: 10.3839/jabc.2016.020
9. Jung JE, Sin SM, Kim HM, Lee S, Choi K, Park KW, et al. Screening of radical scavenging activity and gastric cancer prevention activity from Korean folk plants. *Cancer Prev Res*. (2011) 16:65–73 Available online at: http://www.jcpjournal.org/journal/view.html?volume=16&number=1&spage=65
10. Thao NT, Cuong TD, Hung TM, Lee JH, Na M, Son JK, et al. Simultaneous determination of bioactive flavonoids in some selected Korean thistles by high-performance liquid chromatography. *Arch Pharm Res*. (2011) 34:455–61. doi: 10.1007/s12272-011-0314-x
11. Lee S, Lee DH, Kim JC, Um BH, Sung SH, Jeong LS, et al. Pectolinarigenin, an aglycone of pectolinarin, has more potent inhibitory activities on melanogenesis than pectolinarin. *Biochem Biophys Res Commun*. (2017) 493:765–72. doi: 10.1016/j.bbrc.2017.08.106
12. Hwang J, Qi L. Quality control in the endoplasmic reticulum: crosstalk between ERAD and UPR pathways. *Trends Biochem Sci*. (2018) 43:593–605. doi: 10.1016/j.tibs.2018.06.005
13. Back SH, Kaufman RJ. Endoplasmic reticulum stress and type 2 diabetes. *Annu Rev Biochem*. (2012) 81:767–93. doi: 10.1146/annurev-biochem-072909-095555
14. Schröder M, Kaufman RJ. ER stress and the unfolded protein response. *Mutat Res*. (2005) 569:29–63. doi: 10.1016/j.mrfmmm.2004.06.056
15. Schröder M, Kaufman RJ. The mammalian unfolded protein response. *Annu Rev Biochem*. (2005) 74:739–89. doi: 10.1146/annurev.biochem.73.011303.074134
16. Chouhan V, Kumar SS, Khamparia A, Gupta D, Tiwari P, Moreira C, et al. A novel transfer learning based approach for pneumonia detection in chest x-ray images. *Appl Sci*. (2020) 10:559. doi: 10.3390/app10020559
17. Khamparia A, Saini G, Gupta D, Khanna A, Tiwari S, Albuquerque VHC. Seasonal crops disease prediction and classification using deep convolutional encoder network. *Circuits Syst Signal Process*. (2020) 39:818–36. doi: 10.1007/s00034-019-01041-0
18. Lee YJ, Lee JH, Kim YH, Kim JH, Yu SY, Kim DB, et al. Assessment of the pectolinarin content and the radical scavenging-linked antiobesity activity of Cirsium setidens Nakai extracts. *Food Sci Biotechnol*. (2015) 24:2235–43. doi: 10.1007/s10068-015-0298-2
19. Roth JA, Horbinski C, Higgins D, Lein P, Garrick MD. Mechanisms of manganese-induced rat pheochromocytoma (PC12) cell death and cell differentiation. *Neurotoxicology*. (2002) 23:147–57. doi: 10.1016/S0161-813X(01)00077-8
20. Khamparia A, Singh A, Anand D, Gupta D Khanna A, Kumar NA, et al. A novel deep learning-based multi-model ensemble method for the prediction of neuromuscular disorders. *Neural Comput Appl*. (2020) 32:11083–95. doi: 10.1007/s00521-018-3896-0
21. Parmar VM, Schröder M. Sensing endoplasmic reticulum stress. *Adv Exp Med Biol*. (2012) 738:153–68. doi: 10.1007/978-1-4614-1680-7_10
22. Song S, Tan J, Miao Y, Li M, Zhang Q. Crosstalk of autophagy and apoptosis: involvement of the dual role of autophagy under ER stress. *J Cell Physiol*. (2017) 232:2977–84. doi: 10.1002/jcp.25785
23. Elmore S. Apoptosis: a review of programmed cell death. *Toxicol Pathol*. (2007) 35:495–516. doi: 10.1080/01926230701320337

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 Public Health

ORIGINAL RESEARCH
published: 16 September 2020
doi: 10.3389/fpubh.2020.00493

Use of a Spinal Thermal Massage Device for Anti-oxidative Function and Pain Alleviation

Ka-Eun Kim[1], Jeong-Sook Park[2], Il-Young Cho[3], Yong-Soon Yoon[4], Soon-Kwon Park[5] and Sang-Yun Nam[6]*

[1] College of Medical Sciences, Jeonju University, Jeonju-si, South Korea, [2] Department of Nursing, Nambu University, Gwangju, South Korea, [3] College of Medical Sciences, Jeonju University, Jeonju-si, South Korea, [4] Department of Rehabilitation Medicine, Presbyterian (Jesus) Medical Center, Jeonju-si, South Korea, [5] College of Social Sciences, Jeonju University, Jeonju-si, South Korea, [6] College of Medical Sciences, Jeonju University, Jeonju-si, South Korea

Background: Elderly people are vulnerable to a variety of diseases, including chronic pain, which reduces their levels of physical fitness. Thermal massage has been shown to relieve pain and activate antioxidant enzymes. The objective of this study was to determine whether thermal massaging of the spinal column can reduce muscle pain and induce antioxidant function.

Methods: This study included participants aged ≥60 years with lower back pain. The participants were assigned to either an experimental group who received spinal column thermal massage and standard rehabilitative treatment or a control group who received standard rehabilitative treatment only. Data from a total of 116 participants (61 and 55 in the control and experimental groups, respectively) were used for analysis. Participants were assessed before treatment and at 4 (POST1) and 8 weeks (POST2) post-treatment, using a pain numeric rating scale (PNRS) and the Roland and Morris Disability Questionnaire (RMDQ), and by measuring the serum levels of superoxide dismutase (SOD), serum glutathione-peroxidase (GPx), and serum catalase (CAT).

Results: The extent of pain reduction, as measured by the PNRS, was greater in the experimental group. The RMDQ score in the control group decreased at POST1, but the decrease was not maintained at POST2, whereas the decrease in POST1 in the experimental group continued until POST2. SOD concentrations were significantly higher in the experimental group at POST1 and POST2, and GPx levels were significantly higher in the experimental group at POST2; however, there were no changes in CAT concentrations. Incidentally, there was a significant correlation between antioxidant activity and pain perception in the experimental group.

Conclusions: The study findings suggest that spinal column thermal massage reduces pain more effectively, improves self-reported levels of disability, and increases the antioxidant enzyme levels. Thermal massage may, therefore, be useful in the prevention and treatment of diseases associated with oxidation.

Keywords: spinal thermal massage device, pain, superoxide dismutase, glutathione peroxidase, catalase

OPEN ACCESS

Edited by:
Deepak Gupta, Maharaja Agrasen Institute of Technology, India

Reviewed by:
Pradeep Kumar Mallick, KIIT University, India
Akash Kumar Bhoi, Sikkim Manipal University, India
Kyung Tae Chung, Dong-Eui University, South Korea

***Correspondence:**
Jeong-Sook Park
pk0207@nambu.ac.kr

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Public Health

Received: *02 June 2020*
Accepted: *31 July 2020*
Published: *16 September 2020*

Citation:
Kim K-E, Park J-S, Cho I-Y, Yoon Y-S, Park S-K and Nam S-Y (2020) Use of a Spinal Thermal Massage Device for Anti-oxidative Function and Pain Alleviation. Front. Public Health 8:493. doi: 10.3389/fpubh.2020.00493

INTRODUCTION

Elderly people are vulnerable to various diseases, including chronic pain, and often experience a sudden decline in fitness, limiting them to low-level physical activity (1). Among various theories regarding the causes of aging, the theory involving oxidative stress is a widely accepted one. With aging, the levels of antioxidant enzymes in the body decrease and the total blood antioxidant activity that can respond to oxidative stress decreases gradually (2, 3). This subsequently leads to an increase in the production of reactive oxygen species (ROS), which is a major cause of geriatric diseases and other diseases responsible for chronic pain (4, 5).

While various factors can cause pain, inflammation by ROS is of particular importance. Excessive amounts of ROS, which are produced naturally by metabolic processes, may cause cell damage by oxidation (6, 7). In such events, the cell membrane, DNA, and other cellular structures may become damaged. Consequently, the cells become dysfunctional or mutated depending on the extent of damage, and pain may occur due to the substances generated during this process (8).

Recently, studies have identified the intracellular mechanisms and physiological effects triggered by massage. Specifically, studies have reported that the levels of cytokines, including interleukin-6 (IL-6) and tumor necrosis factor-alpha (TNF-α), which increase following the inflammatory response and micro-rupture of tissues due to muscle injury or strenuous exercise, and nuclear factor-κB (NF-κB), a transcription factor that activates these cytokines, were reduced significantly by massage (9, 10). NF-κB is especially sensitive to oxidation and is known to activate inflammatory cytokines when it is stimulated by ROS (11). Taken together, the results from previous studies suggest that massage activates antioxidant functions and alleviates inflammation and pain through the reduction of NF-κB, IL-6, and TNF-α levels. When ROS activate the signaling systems, such as NF-κB or mitogen-activated protein kinase, or disturb their homeostasis, cells may proliferate abnormally or mutate (12). Cells possess various antioxidant enzymes and chemicals for eliminating ROS (13). The antioxidant enzymes that protect the body by removing free radicals generated by oxidative stress include superoxide dismutase (SOD), catalase (CAT), glutathione-peroxidase (GPx), and glutathione S-transferase (GST) (14, 15). Karabulut et al. reported that the level of antioxidant enzymes was increased after a massage intervention was applied to obese elderly women (16). Moreover, results from a randomized double-blind study revealed that the pain experienced by patients with chronic pancreatitis was reduced after they were treated with antioxidants (17).

When heat is applied to human tissues, the increase in tissue temperature is accompanied by physiological responses (18), including an increase in collagen fibril extensibility, blood flow, and cell permeability, and a reduction in muscle spasms and pain (19, 20).

Recently, home healthcare devices, such as spinal thermal massage (STM) devices, with massage and heating functions, have been developed and are commercially available. Among these devices, some products have already received food and drug administration (FDA) approval from various countries, to be used as medical devices for reducing muscle pain. However, few studies have evaluated the pain-alleviating effects of these medical devices. Accordingly, the present study aimed to clinically identify the pain-alleviating effects of a STM program, and to determine whether these programs induce changes in the antioxidant enzyme activity.

METHODS

Participants

This study followed the principles and recommendations of the Helsinki Declaration, and the protocol was reviewed and approved by the Institutional Review Board (IRBN. 2017-06-022). A total of 140 adults (≥60 years of age) who had experienced lower back, shoulder, knee, hip, and/or neck pain for at least 3 months were recruited for this study, and we obtained the written informed consent from them. We evaluated the data obtained from their medical records and interviews prior to the study to determine if the candidates were eligible to participate in this study.

Candidates with any one of the following conditions were excluded from the study: diseases of the immune system, pacemaker or an electronic implanted device, malignant tumor, spinal infection, thrombosis, skin disease or skin hypersensitivity, spinal deformity or scoliosis with Cobb's angle $\geq 20^{\circ}$, osteoporosis or history of spinal fracture due to osteoporosis, high risk of spinal fracture, myopathy, or spinal instability (failed back syndrome) following spinal surgery. We also excluded anyone who was determined to be unfit for the study by the researcher. Ten candidates were eliminated during this screening phase.

Finally, a total of 130 participants were enrolled in the study. Sixty-five participants were assigned to the control group who received the standard rehabilitative treatment (SRT), whereas the remaining 65 patients were assigned to the experimental group who received SRT and a STM intervention. Participants were restricted from taking medications or nutraceutical foods that could affect the SOD, GPx, and CAT concentrations in the blood, which are used to measure the antioxidant enzyme activity. Given that some patients dropped out from the study, data of 116 participants (61 and 55 in the control and experimental groups, respectively) were ultimately used for the final analysis (**Figure 1**). There was no difference in the age distribution and pain duration between the two groups (**Table 1**). Although the sex distribution was different between the two groups ($\chi^2 = 10.13$, $p < 0.05$), a previous study showed that the effect of massage is not affected by sex difference (21).

Massage Device

The present study used a thermal massage device that was approved as a personal medical device (Model: CGM MB-1401, Ceragem, South Korea). It consists of a bed covered with a tough synthetic fabric and a massage roller (projector) moving beneath the fabric. This device is operated with a remote control and allows the user to receive massage and heat treatment while lying down (**Figure 2**). The temperature of the projector is adjustable

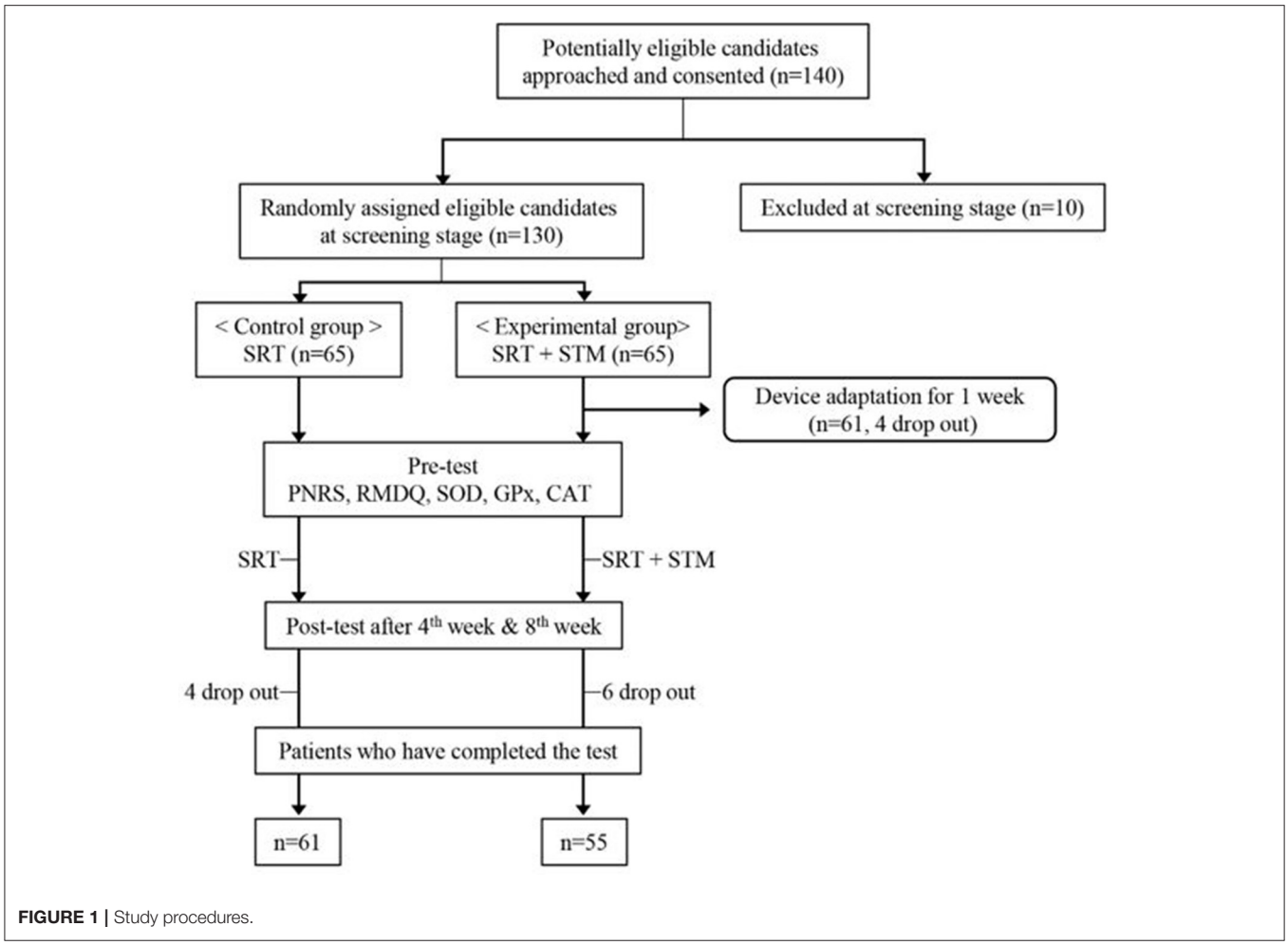

FIGURE 1 | Study procedures.

TABLE 1 | General characteristics of the participants.

Variable	CON	EXP	Chi^2 test
Sex			
Male	23	37	$\chi^2 = 10.13$
Female	38	18	$p < 0.05$
Age (Years)			
–69	27	32	$\chi^2 = 2.38$
70–79	30	21	$p > 0.05$
80–	4	2	
Pain Duration (Months)			
1–3	7	9	
4–6	10	16	$\chi^2 = 4.24$
7–12	16	11	$p > 0.05$
13–24	10	11	
25–	13	9	

within the range of 30–60°C, and the strength of the massage ranged from level 1–6. The duration of the massage session set in this device is 36.5 min.

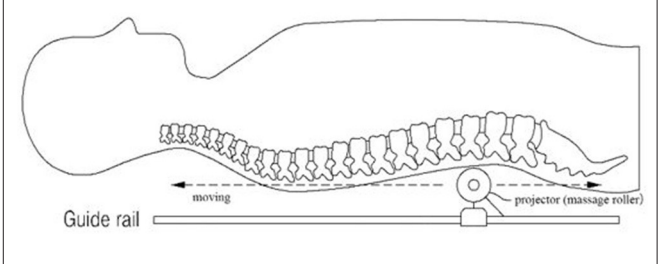

FIGURE 2 | Schematic diagram of the spinal thermal massage device.

Massage Intervention

Participants were instructed to lie on the device and operate it using the remote control. Before the actual massage, scanning is performed, wherein the projector moves from the cervical vertebrae to the sacrum, to measure the length and curvature of the spinal column. The participants were allowed to use the device 3–4 times for 1 week to become acclimatized. The STM intervention was applied five times per week for 8 weeks. Participants who experienced severe pain or discomfort during the massage session were excluded from the study.

Measured Variables and Tools

The present study used self-report questionnaires to determine the level of disability due to the lower back pain and other pains. Blood samples were also collected from the participants to determine the antioxidant effects of the massage intervention. A questionnaire was administered and blood samples were collected before the intervention (PRE), immediately after 4 weeks of intervention (POST1), and immediately after 8 weeks of intervention (POST2).

Pain-Related Indices

The following two different clinical pain-related indices were measured: (1) the pain numeric rating scale (PNRS) and (2) the Roland and Morris disability questionnaire (RMDQ), which is used to determine the level of disability due to lower back pain. The PNRS is a segmented numeric version of the visual analog scale (VAS), in which a participant selects a whole number (0–10 integers) that best reflects the intensity of his/her pain (22). The common format is a horizontal bar or line. Similar to the VAS, the PNRS is anchored by terms describing pain severity extremes (23). The RMDQ is a 24-items self-report questionnaire about how low back pain affects functional activities. Each question is worth one point; thus, scores can range from 0 (no disability) to 24 (severe disability) (24). Higher scores for the PNRS and RMDQ indicate higher levels of pain and disability, respectively.

Measurement of Antioxidant Enzyme Activity

SOD, GPx, and CAT are representative antioxidants in the body that deal with free radicals that cause aging and disease. In the present study, blood samples were collected from the participants to determine the effects of the STM intervention on the antioxidant function based on the activities of SOD, GPx, and CAT.

Blood Sampling

Blood was collected before and immediately after 4 and 8 weeks of the massage intervention. The collected blood samples were immediately divided into serum collection tubes (BD Vacutainer® SST™ II Advance Plus Blood Collection Tube; Becton Dickinson UK Ltd., Oxford, UK) and heparinized tubes (BD Vacutainer® Sodium Heparin 75 USP Units Blood Collection Tube; Becton Dickinson UK Ltd.) for immediate transport to the laboratory.

Superoxide Dismutase

The SOD activity was measured using a SOD assay kit (Bioassay Systems, Hayward, CA, USA) according to the manufacturer's protocol. Briefly, the removal of superoxide by SOD was indirectly evaluated by measuring the reduction in cytochrome c levels. The protocol is as follows: First, the standards and samples were prepared to measure the SOD activity. Then, 160-μL assay buffer, 5-μL xanthine, and 5-μL WST-1 were mixed in each well. Subsequently, 160-μL working reagent was transferred to each well, the tap plate was mixed, XO enzyme was immediately added after reading OD440nm (ODo), and the plates were incubated for 60 min. Finally, we read the OD440nm (OD_{60}) again to obtain the changed SOD value.

Glutathione Peroxidase

The GPx activity was measured using a GPx assay kit (Bioassay Systems) according to the manufacturer's protocol and was calculated based on the reduction of nicotinamide adenine dinucleotide phosphate. For GPx measurements, we prepared enough working reagents for all sample and control wells by mixing, for each well, 90-μL assay buffer, 5-μL glutathione, 3-μL 35-mM nicotinamide adenine dinucleotide phosphate, and 2-μL GR enzyme. We quickly added 90 μL of a working reagent to the sample/control wells. Then, this solution was diluted in dH2O with a 1:10 ratio to generate the substrate solution. The diluted solutions were used within 1 h. Using a multi-channel pipettor, we added a 100-μL substrate solution to the sample and control wells and mixed the contents well. We immediately read OD340nm (time 0, OD_0) and read again in 4 min (OD_4).

Catalase

CAT catalyzes the decomposition of hydrogen peroxide (produced by SOD) to oxygen and water. The CAT activity was measured using a CAT assay kit (Bioassay Systems) according to the manufacturer's instructions. The protocol is as follows. We prepared enough detection reagent by mixing, for each reaction well (sample, control, and standard wells), 102-μL assay buffer, 1-μL dye reagent, and 1-μL HRP enzyme. After completion of incubation for 30 min, a 100-μL detection reagent was added and mixed, followed by incubation for 10 min. Then, we read at 570 nm.

Statistical Analysis

PNRS and RMDQ scores were analyzed by non-parametric methods, using the Mann–Whitney U and Freedman tests, because they are measured with ordinal scales. We used a mixed design with three-repeated measurements (PRE, POST1, and POST2). Accordingly, a two-way repeated measure analysis of variance was conducted for the statistical analysis of the data (SOD, GPx, and CAT). After omnibus testing, we analyzed the differences in the simple main effects of the interventions within the two groups and between the two groups at each time point. A $p < 0.05$ was considered statistically significant.

RESULTS

Pain-Related Indices

The medians and interquartile ranges of PNRS and RMDQ scores are presented in **Table 2**. The Mann–Whitney U-test and Freedman test were applied to analyze the intergroup differences at each measurement time and intergroup differences in each group, respectively. The Freedman test showed that the pain levels decreased significantly in both the control ($\chi^2 = 8.74$, df $= 2$, $p < 0.05$) and experimental ($\chi^2 = 68.88$, df $= 2$, $p < 0.01$) groups. The Mann–Whitney U-test showed that the pain level was higher in the experimental group than in the control group before the intervention ($U = 973.00$, $p < 0.01$) and at POST1 ($U = 1282.50$, $p < 0.05$). Pain levels were reversed at POST2 with higher pain levels in the control group than in the experimental group ($U = 1269.50$, $p < 0.05$). These analyses indicate that

TABLE 2 | Effect of spinal thermal massage on PNRS and RMDQ scores.

	Group	PRE	POST1	POST2
PNRS	CON ($n = 61$)	3.00 (1)	3.00 (2)	3.00 (2)
	EXP ($n = 55$)	4.00 (2)$^{\#\#}$	4.00 (2)$^{\#}$	2.00 (2)$^{\#}$
RMDQ	CON ($n = 61$)	6.00 (3)	6.00 (3)	6.00 (4)
	EXP ($n = 55$)	11.00 (6)$^{\#\#}$	9.00 (7)$^{\#\#}$	8.00 (9)$^{\#\#}$

Data are expressed as median (interquartile range).
$^{\#}(p < 0.05)$ and $^{\#\#}(p < 0.01)$ indicate mean significant differences (Mann–Whitney U-test).
PNRS, pain numeric rating scale; RMDQ, the Roland and Morris Disability Questionnaire; PRE, pre-intervention; POST1, 4 weeks post-thermal massage treatment; POST2, 8 weeks post-thermal massage treatment; CON, control; EXP, experimental.

TABLE 3 | Effect of spinal thermal massage on antioxidant function.

	Group	PRE	POST1	POST2
SOD	CON ($n = 61$)	0.48 (0.02)	0.45 (0.02)	0.47 (0.02)
(U/mL)	EXP ($n = 55$)	0.42 (0.03)	0.53 (0.02)$^{**\#}$	0.59 (0.02)$^{**\#\#}$
GPx	CON ($n = 61$)	2502.87 (114.01)	2333.36 (117.87)	2174.84 (132.16)*
(U/L)	EXP ($n = 55$)	2404.44 (116.13)	2270.84 (111.53)	2946.71 (159.05)**
CAT	CON ($n = 61$)	1.46 (.012)	1.32 (0.01)	1.28 (0.01)
(U/L)	EXP ($n = 55$)	1.18 (0.02)$^{\#}$	1.21 (0.02)$^{\#\#}$	1.21 (0.02)$^{\#\#}$

Values are expressed as the mean and standard error (SE).
**Significant difference from "PRE" values ($p < 0.05$).*
***Significant difference from "PRE" values ($p < 0.01$).*
$^{\#}$Significant difference from the control group ($p < 0.05$).
$^{\#\#}$Significant difference from the control group ($p < 0.01$).
PRE, pre-intervention; POST1, 4 weeks post-thermal massage treatment; POST2, 8 weeks post-thermal massage treatment; SOD, superoxide dismutase; GPx, glutathione-peroxidase; CAT, catalase; CON, control; EXP, experimental.

pain reduction was greater in the experimental group than in the control group.

The Mann–Whitney *U*-test showed that the RMDQ scores were higher in the experimental group than in the control group before the intervention ($U = 538.00$, $p < 0.01$), at POST1 ($U = 948.00$, $p < 0.01$), and at POST2 ($U = 1162.00$, $p < 0.01$). The Freedman test showed that the RMDQ scores decreased in the experimental group as time progressed ($\chi^2 = 43.19$, df $= 2$, $p < 0.01$); however, the RMDQ scores in the control group did not. The results of these analyses indicate that the disability improved in both groups, but the degree of improvement was higher in the experimental group.

Antioxidant Enzyme Activity

The SOD activity was significantly affected by the intervention [$F_{(1,114)} = 3.97$, $p < 0.05$] and measurement time [$F_{(2,228)} = 8.56$, $p < 0.01$]. There was a significant interaction effect on SOD activity [$F_{(2,228)} = 10.04$, $p < 0.01$]. When simple main effects were analyzed to examine these differences in more detail, there were no significant differences in the SOD activity between pre-intervention and at POST1 and POST2 in the control group. However, the SOD activity was significantly higher at POST1 [$F_{(1,54)} = 13.76$, $p < 0.01$] and POST2 [$F_{(1,54)} = 32.53$, $p < 0.01$) than before the treatment in the experimental group. There were no significant differences in the SOD activity between the two groups before the intervention, but the SOD activity was significantly higher in the experimental group at POST1 [$F_{(1,114)} = 5.78$, $p < 0.05$) and POST2 [$F_{(1,114)} = 18.11$, $p < 0.01$).

The intervention and measurement time had no significant effects on GPx activity, but there was a significant interaction effect observed for time and intervention on GPx activity [$F_{(2,228)} = 9.58$, $p < 0.01$]. The GPx activity was significantly lower at POST2 in the control group [$F_{(1,60)} = 4.316$, $p < 0.05$], whereas the GPx activity was significantly higher at POST2 in the experimental group [$F_{(1,54)} = 13.68$, $p < 0.01$]. The levels of GPx activity were not significantly different between the two groups before the intervention and at POST1, but the GPx activity was significantly higher in the experimental group at POST2 [$F_{(1,114)} = 14.13$, $p < 0.01$].

Overall, the CAT activity was higher in the experimental group than in the control group [$F_{(1,114)} = 10.88$, $p < 0.01$]; however, the interaction effect of time and effect was not significant. When the simple main effects were analyzed to examine these differences in more detail, there were no changes in the CAT activity before and after treatment in both groups. However, the CAT activity was significantly higher in the control group than in the experimental group before the intervention [$F_{(1,114)} = 4.51$, $p < 0.05$] and at POST1 [$F_{(1,114)} = 41.04$, $p < 0.01$] and POST2 [$F_{(1,114)} = 9.09$, $p < 0.01$] (**Table 3**, **Figure 3**). These results demonstrated that the CAT activity, which was higher in the control group than in the experimental group before the intervention, was maintained without any changes after 4 and 8 weeks of treatment.

Correlations Between Pain and Antioxidant Activity

To investigate the correlations between changes in antioxidant activity and pain-related measurements following the massage intervention, we performed a correlation analysis using the results from the antioxidant and pain analyses of the experimental group. The rate of change in values measured before and after the massage intervention was calculated and applied to the correlation analysis [rate of change (%) = 100 * (measured value after massage – measured value before massage) / measured value before massage]. With respect to the rate of change at POST1, the analysis revealed that the correlations between changes in pain and the values of the antioxidation indices were all significant ($p < 0.05$). Additionally, with respect to the rate of change at POST2, the correlations between changes in PNRS scores and the values of the antioxidation indices, except for CAT activity, were significant ($p < 0.05$) (**Table 4**).

DISCUSSION

To date, only a few studies have attempted to monitor the changes in antioxidant enzyme levels and pain by simultaneously using massage and heat as sources of stimuli. This is the first study wherein two stimuli were used to confirm the increase in the antioxidant enzyme levels and the decrease in pain intensity in the elderly with reduced antioxidant enzyme levels. The subjective levels of pain (PNRS) and disability in performing activities of daily living due to lower back pain (RMDQ) were

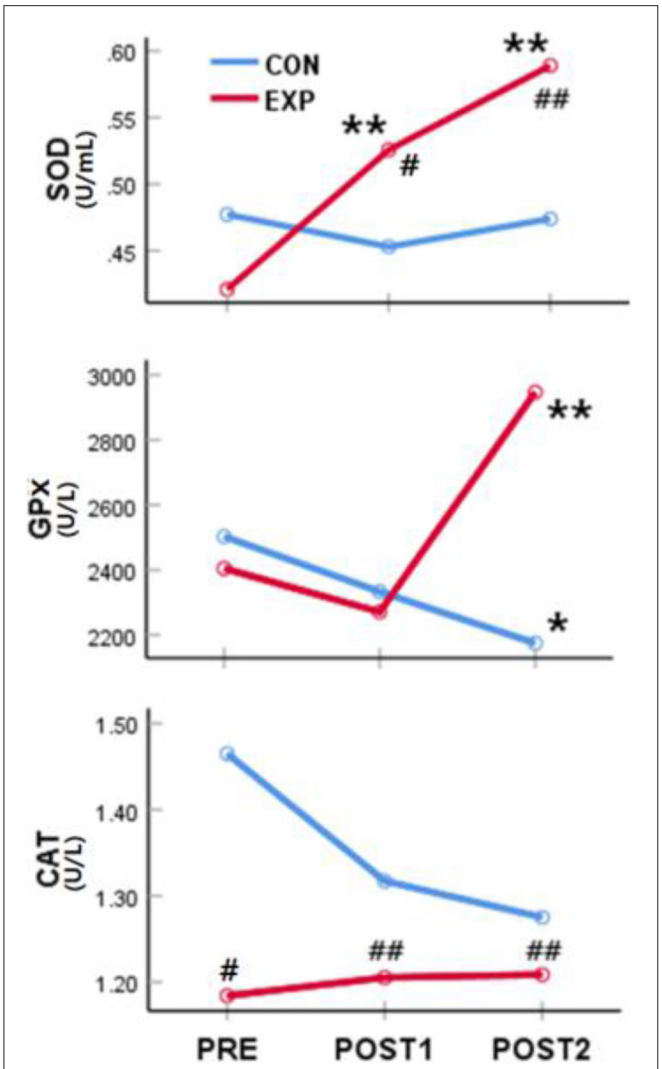

FIGURE 3 | Effect of the spinal thermal massage on antioxidant function. *significant difference from "PRE" values ($p < 0.05$); **significant difference from "PRE" values ($p < 0.01$); #significant difference from the control group ($p < 0.05$). ##significant difference from the control group ($p < 0.01$). PRE, pre-intervention; POST1, 4 weeks post-thermal massage treatment; POST2, 8 weeks post-thermal massage treatment; SOD, superoxide dismutase; GPx, glutathione-peroxidase; CAT, catalase; CON, control; EXP, experimental.

TABLE 4 | Correlations between the change rates in pain-related indices and antioxidant activity.

		PNRS	RMDQ
Change rate at POST1	SOD	0.870*	0.880*
	GPx	0.440*	0.322*
	CAT	0.229*	0.346*
Change rate at POST2	SOD	0.988*	0.987*
	GPx	0.517*	0.598*
	CAT	0.162	0.371*

*$p < 0.05$.
PNRS, pain numeric rating scale; RMDQ, Roland and Morris disability questionnaire; SOD, superoxide dismutase; GPx, glutathione peroxidase; CAT, catalase.

significantly lower (improved) in the experimental group than in the control group after SRT with thermal massage. Antioxidant function, as measured by the activities of SOD and GPx, was also improved following SRT with thermal massage compared to that observed after SRT only. This finding is consistent with the results of other studies that reported that the production of SOD and the level of glutathione (GSH), which is involved in GPx activity, were increased following massage (16), and that antioxidant enzyme levels were increased after thermal stimulation (25). We also observed a positive correlation between the extent of improvement in pain-related indices (PNRS and RMDQ) and antioxidant function, which suggests that the thermal massage intervention affected the physiological processes associated with pain and antioxidant enzyme activities.

Muscle pain is a complex progressive health problem that causes muscle weakness and fatigue (26). There are various causes of pain, including strenuous exercise or musculoskeletal injury, inflammatory response to oxidative stress, including ROS, and fatigue due to lactic acid build-up (27). Watkins et al. suggested that inflammatory cytokines, including IL-1ß, which are released from the microglial cells in the peripheral and central nervous system (CNS), are associated with pain, and that ROS can cause chronic pain by activating the glial cells in the CNS (28). Perez et al. reported that the symptoms of complex regional pain syndrome type I (CRPS-I) were reduced after the administration of ROS scavengers (29). This finding also serves as evidence that antioxidant activity is associated with pain. Based on the results of these studies and the correlation analysis in the present study, we believe that antioxidant function is improved by thermal massage. Consequently, pain alleviation is also achieved with improved antioxidant function, which protects against ROS, and this antioxidant function appears to decrease with aging. The concentrations of hepatic GSH and glutathione reductase, and the activity of SOD have been found to be lower in older animals than in younger animals (30). Even in humans, compared to the levels of antioxidant enzymes in 25-years-old individuals, these levels decrease by 30, 40, and 60% in people aged 40–49, 60–69, and 70–79 years, respectively (31). The SOD activity has been found to decrease with age, with the metabolism of antioxidants consumed through food also decreasing (32, 33). The level of antioxidant enzyme, SOD, which has the ability to remove ROS, increases when ROS and lipid peroxide are generated in the body. This ability rapidly decreases after the age of 40 years, which can lead to various geriatric diseases (34, 35). Diabetes and lung cancer related to smoking rarely develop in people with high SOD activity. However, if the SOD activity decreases, the probability of developing brain and cardiovascular disorders, such as cancer, stroke, and myocardial infarction, increases (34).

GPx catalyzes the decomposition of hydrogen peroxide formed by SOD into oxygen and water, and promotes the activity of GSH, which catalyzes the decomposition and detoxification of lipid peroxides (36). GPx plays a secondary role in removing the remaining hydrogen peroxide after the action of CAT on

hydrogen peroxide (37). Low GPx activity can lead to exposure to several diseases, including multiple sclerosis and diabetic kidney disease. Various reports have indicated that the ability to synthesize antioxidant enzymes, efficiency of antioxidant enzymes, and relevant physiological metabolic actions all decrease with age, clearly suggesting that antioxidant function should be augmented in elderly populations. The mean age of the participants in the present study was 65 years, and improvement in their antioxidant function was achieved through an intervention involving a thermal massage device.

We recognize some limitations associated with the study presented here. First, although we identified that the use of thermal massage improved the antioxidant function, we did not evaluate the specific mechanisms involved in this improvement. Second, as the role of antioxidant function in aging and various diseases has already been established, additional studies should investigate the effects of thermal massage on improving other symptoms or other diseases associated with antioxidant function. Finally, future studies should investigate whether a thermal massage program affects the improvement in antioxidant function for various causes of back pain, such as stress, exercise, injury, or inflammatory responses.

CONCLUSION

While additional systematic studies are required, the findings of this study suggest that thermal massage may be an effective strategy for reducing pain and preventing a decrease in antioxidant enzyme activity in elderly individuals. Thermal massage may, therefore, be useful in the prevention and treatment of diseases associated with oxidation. No side effects due to the intensity of massage and heat stimulation were reported during this trial.

DATA AVAILABILITY STATEMENT

All data generated for this study are included in the article.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board at the Presbyterian (Jesus) Medical Center (IRBN. 2017-06-022). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

K-EK: conception, design, material preparation, data collection, analysis, writing—original draft, review, and editing. I-YC: conception, design, material preparation, data collection, analysis, review writing, editing, and project administration. Y-SY and S-YN: conception, design, material preparation, data collection, and analysis. S-KP: conception, design, material preparation, data collection, analysis, review writing, and editing. J-SP: conception, design, material preparation, data collection, analysis, review writing, editing, and supervision. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Frisard M, Ravussin E. Energy metabolism and oxidative stress: impact on the metabolic syndrome and the aging process. *Endocrine.* (2006) 29:27–32. doi: 10.1385/ENDO:29:1:27
2. Andriollo-Sanchez M, Hininger-Favier I, Meunier N, Venneria E, O'Connor JM, Maiani G, et al. Age-related oxidative stress and antioxidant parameters in middle-aged and older European subject: the ZENITH study. *Eur J Clin Nutr.* (2005) 59:S58–62. doi: 10.1038/sj.ejcn.1602300
3. Franzoni F, Ghiadoni L, Galetta F, Plantinga Y, Lubrano V, Huang Y, et al. Physical activity, plasma antioxidant capacity, and endothelium-dependent vasodilation in young and older men. *Am J Hypertens.* (2005) 18:510–6. doi: 10.1016/j.amjhyper.2004.11.006
4. Choi I. Reactive oxygen species and cancer. *Hanyang Med Rev.* (2013) 33:118–22. doi: 10.7599/hmr.2013.33.2.118
5. Ryu SE. Do reactive oxygen species cause aging? *Hanyang Med Rev.* (2013) 33:75–6. doi: 10.7599/hmr.2013.33.2.75
6. Duru R, Njoku O, Maduka I. Oxidative stress indicators in patients with prostate disorders in Enugu, South-East Nigeria. *Biomed Res Int.* (2014) 2014:313015. doi: 10.1155/2014/313015
7. Forrester SJ, Kikuchi DS, Hernandes MS, Xu Q, Griendlin KK. Reactive oxygen species in metabolic and inflammatory signaling. *Circ Res.* (2018) 122:877–902. doi: 10.1161/CIRCRESAHA.117.311401
8. Rimessi A, Previati M, Nigro F, Wieckowski MR, Pinton P. Mitochondrial reactive oxygen species and inflammation: molecular mechanisms, diseases and promising therapies. *Int J Biochem Cell Biol.* (2016) 81:281–93. doi: 10.1016/j.biocel.2016.06.015
9. Crane JD, Ogborn DI, Cupido C, Melov S, Hubbard A, Bourgeois JM, et al. Massage therapy attenuates inflammatory signaling after exercise-induced muscle damage. *Sci Transl Med.* (2012) 4:119ra13. doi: 10.1126/scitranslmed.3002882
10. Sweeney G, Song J. The association between PGC-1α and Alzheimer's disease. *Anat Cell Biol.* (2016) 49:1–6. doi: 10.5115/acb.2016.49.1.1
11. Nathan C, Cunningham-Bussel A. Beyond oxidative stress: an immunologist's guide to reactive oxygen species. *Nat Rev Immunol.* (2013) 13:349–61. doi: 10.1038/nri3423
12. Panday A, Sahoo MK, Osorio D, Batra S. NADPH oxidases: an overview from structure to innate immunity-associated pathologies. *Cell Mol Immunol.* (2015) 12:5–23. doi: 10.1038/cmi.2014.89
13. Mailloux RJ, Gardiner D, O'Brien M. 2-oxoglutarate dehydrogenase is a more significant source of O2(·-)/H2O2 than pyruvate dehydrogenase in cardiac and liver tissue. *Free Radic Biol Med.* (2016) 97:501–12. doi: 10.1016/j.freeradbiomed.2016.06.014
14. Touyz RM. Reactive oxygen species in vascular biology: role in arterial hypertension. *Exp Rev Cardiovasc Ther.* (2003) 1:91–106. doi: 10.1586/14779072.1.1.91
15. Kurz DJ, Decary S, Hong Y, Trivier E, Akhmedov A, Erusalimsky JD. Chronic oxidative stress compromises telomere integrity and accelerates the onset of senescence in human endothelial cells. *J Cell Sci.* (2004) 117:2417–26. doi: 10.1242/jcs.01097
16. Karabulut AB, Kafkas ME, Kafkas AS, Onal Y, Kiran TR. The effect of regular exercise and massage on oxidant and antioxidant parameters. *Indian J Physiol Pharmacol.* (2013) 57:3 78–83.
17. Cai GH, Huang J, Zhao Y, Chen J, Wu HH, Dong YL, et al. Antioxidant therapy for pain relief in patients with chronic pancreatitis: systematic review and meta-analysis. *Pain Physician.* (2013) 16:521–32.

18. Ide T, Tsutsui H, Hayashidani S, Kang D, Suematsu N, Nakamura K, et al. Mitochondrial DNA damage and dysfunction associated with oxidative stress in failing hearts after myocardial infarction. *Circ Res.* (2001) 88:529–35. doi: 10.1161/01.RES.88.5.529
19. Toyokawa H, Matsui Y, Uhara J, Tsuchiya H, Teshima S, Nakanishi H, et al. Promotive effects of far-infrared ray on full-thickness skin wound healing in rats. *Exp Biol Med.* (2003) 228:724–9. doi: 10.1177/153537020322800612
20. Skillgate E, Bill AS, Côté P, Viklund P, Peterson A, Holm LW. The effect of massage therapy and/or exercise therapy on subacute or long-lasting neck pain-the Stockholm neck trial (STONE): study protocol for a randomized controlled trial. *Trials*. (2015) 16:414. doi: 10.1186/s13063-015-0926-4
21. Smith TE, Valentine P, Thyer BA. Gender effects in massage therapy. *J Altern Med Res*. (2012) 4:201–6.
22. Rodriguez CS. Pain measurement in the elderly: a review. *Pain Manag Nurs*. (2001) 2:38–46. doi: 10.1053/jpmn.2001.23746
23. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: visual analog scale for pain (VAS pain), numeric rating scale for pain (NRS pain), McGill pain questionnaire (MPQ), short-form McGill pain questionnaire (SF-MPQ), chronic pain grade scale (CPGS), short form-36 bodily pain scale (SF-36 BPS), and measure of intermittent and constant osteoarthritis pain (ICOAP). *Arthritis Care Res*. (2011) 63:S240–52. doi: 10.1002/acr.20543
24. Siebenga J, Leferink VJM, Segers MJM, Elzinga MJ, Bakker FC, Ten DH, et al. Prospective cohort study comparing the VAS spine score and Roland–Morris disability questionnaire in patients with a type A traumatic thoracolumbar spinal fracture. *Eur Spine J.* (2008) 17:1096–100. doi: 10.1007/s00586-008-0705-y
25. Currie RW, Tanguay RM. Analysis of RNA for transcripts for catalase and SP71 in rat hearts after *in vivo* hyperthermia. *Biochem Cell Biol.* (1991) 69:375–82. doi: 10.1139/o91-057
26. Khamparia A, Singh A, Anand D, Gupta D, Khanna A, Kumar NA, et al. A novel deep learning based multi-model ensemble methods for prediction of neuromuscular disorders. *Neural Comput Appl.* (2018) 32:11083–95. doi: 10.1007/s00521-018-3896-0
27. Brosseau L, Yonge KA, Welch V, Marchand S, Judd M, Wells GA, et al. Thermotherapy for treatment of osteoarthritis. *Cochrane Database Syst Rev*. (2003) 2003:CD004522. doi: 10.1002/14651858.CD004522
28. Watkins LR, Wiertelak EP, Goehler LE, Smith KP, Martin D, Maier SF. Characterization of cytokine-induced hyperalgesia. *Brain Res*. (1994) 654:15–26. doi: 10.1016/0006-8993(94)91566-0
29. Perez RSGM, Zuurmond WWA, Bezemer PD, Kuik DJ, van Loenen AC, de Lange JJ, et al. The treatment of complex regional pain syndrome type I with free radical scavengers: a randomized controlled study. *Pain.* (2003) 102:297–307. doi: 10.1016/S0304-3959(02)00414-1
30. Rinaldi B, Corbi G, Boccuti S, Filippelli W, Rengo G, Leosco D, et al. Exercise training affects age-induced changes in SOD and heat shock protein expression in rat heart. *Exp Gerontol.* (2006) 41:764–70. doi: 10.1016/j.exger.2006.05.008
31. Reddy Thavanati PK, Kanala KR, de Dios AE, Cantu Garza JM. Age-related correlation between antioxidant enzymes and DNA damage with smoking and body mass index. *J Gerontol A Biol Sci Med Sci.* (2008) 63:360–4. doi: 10.1093/gerona/63.4.360
32. Choi KM, Lee EJ, Kim YH, Baik SH, Kim KY, Choi DS. Effects of red ginseng on the lipid peroxidation of erythrocyte and antioxidant superoxide dismutase (SOD) activity in NIDDM patients. *J Ginseng Res.* (1997) 21:153–9.
33. Kim HK, Park SK, Zhou JL, Taglialatela G, Chung K, Coggeshall RE, et al. Reactive oxygen species (ROS) play an important role in a rat model of neuropathic pain. *Pain.* (2004) 111:116–24. doi: 10.1016/j.pain.2004.06.008
34. Briggs ON, Brown H, Elechi-amadi K, Ezeiruaku F, Nduka N. Superoxide dismutase and glutathione peroxidase levels in patients with long standing type 2 diabetes in Port Harcourt, Rivers State, Nigeria. *Int J Sci Res*. (2016) 5:1282–8. doi: 10.21275/v4i5.NOV162149
35. Elchuri S, Oberley TD, Qi W, Eisenstein RS, Roberts LJ, Van Remmen H, et al. CuZnSOD deficiency leads to persistent and widespread oxidative damage and hepatocarcinogenesis later in life. *Oncogene.* (2005) 24:367–80. doi: 10.1038/sj.onc.1208207
36. Dworzański J, Strycharz-Dudziak M, Kliszczewska E, Kiełczykowska M, Dworzańska A, Drop B, et al. Glutathione peroxidase (GPx) and superoxide dismutase (SOD) activity in patients with diabetes mellitus type 2 infected with Epstein-Barr virus. *PLoS ONE.* (2020) 15:e0230374. doi: 10.1371/journal.pone.0230374
37. Zedan H, Abdel-Motaleb AA, Kassem NMA, Hafeez HAA, Hussein MRA. Low glutathione peroxidase activity levels in patients with vitiligo. *J Cutan Med Surg.* (2015) 19:144–8. doi: 10.2310/7750.2014.14076

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 Public Health

BRIEF RESEARCH REPORT
published: 29 October 2020
doi: 10.3389/fpubh.2020.580327

Global Forecasting Confirmed and Fatal Cases of COVID-19 Outbreak Using Autoregressive Integrated Moving Average Model

Debabrata Dansana[1], *Raghvendra Kumar*[1], *Janmejoy Das Adhikari*[1], *Mans Mohapatra*[1], *Rohit Sharma*[2]*, *Ishaani Priyadarshini*[3] and *Dac-Nhuong Le*[4,5]*

[1] Department of Computer Science and Engineering, GIET University, Gunupur, India, [2] Department of Electronics & Communication Engineering, SRM Institute of Science and Technology, Ghaziabad, India, [3] Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, United States, [4] Institute of Research and Development, Duy Tan University, Da Nang, Vietnam, [5] Faculty of Information Technology, Duy Tan University, Da Nang, Vietnam

The world health organization (WHO) formally proclaimed the novel coronavirus, called COVID-19, a worldwide pandemic on March 11 2020. In December 2019, COVID-19 was first identified in Wuhan city, China, and now coronavirus has spread across various nations infecting more than 198 countries. As the cities around China started getting contaminated, the number of cases increased exponentially. As of March 18 2020, the number of confirmed cases worldwide was more than 250,000, and Asia alone had more than 81,000 cases. The proposed model uses time series analysis to forecast the outbreak of COVID-19 around the world in the upcoming days by using an autoregressive integrated moving average (ARIMA). We analyze data from February 1 2020 to April 1 2020. The result shows that 120,000 confirmed fatal cases are forecasted using ARIMA by April 1 2020. Moreover, we have also evaluated the total confirmed cases, the total fatal cases, autocorrelation function, and white noise time-series for both confirmed cases and fatalities in the COVID-19 outbreak.

Keywords: COVID-19, ARIMA, forecasting, global pandemic, time series analysis

OPEN ACCESS

Edited by:
Deepak Gupta,
Maharaja Agrasen Institute of Technology, India

Reviewed by:
Korhan Cengiz,
Trakya University, Turkey
Noor Zaman,
Taylor's University, Malaysia

***Correspondence:**
Rohit Sharma
rohitapece@gmail.com
Dac-Nhuong Le
ledacnhuong@duytan.edu.vn

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Public Health

Received: *05 July 2020*
Accepted: *31 August 2020*
Published: *29 October 2020*

Citation:
Dansana D, Kumar R, Das Adhikari J, Mohapatra M, Sharma R, Priyadarshini I and Le D-N (2020) Global Forecasting Confirmed and Fatal Cases of COVID-19 Outbreak Using Autoregressive Integrated Moving Average Model. Front. Public Health 8:580327. doi: 10.3389/fpubh.2020.580327

INTRODUCTION

The first case of the virus came to light in Wuhan city of China in November 2019. The population of Wuhan city is nearly 11 million, and it connects to many major cities in China. The number of cases changed to dozens and then hundreds by the end of December. Medical experts first suspected that it was viral pneumonia, which could not be cured with conventional medicines. Ever since the virus first started to infect people, it has continued to spread and affect thousands of people (1). Further, every patient infected with this virus was infecting two or three people ahead of them. Until December 30 2019, no information was released from China regarding the deadly virus. Finally, in December 2020, officials from the public health department of China informed the World Health Organization about the medical issue that affected people in Hubei Province, China. The infection was described as a pneumonia-like ailment in humans and caused by a coronavirus, an extreme group of pathogens. Coronaviruses are known to spread among people, mice, winged creatures, bats, domesticated animals, and other wild creatures (2–4). In December 2019, the WHO was alerted by China to certain occurrences of a respiratory infection

associated with specific people who had visited the seafood market in Wuhan city (5). Wuhan experienced the spread of a coronavirus, called Coronavirus Disease-19 (COVID-19). In (6), the author presumed that COVID-19 likely started in bats, since it is progressively like two bat-determined coronavirus strains. However, the origin of the COVID-19 has not yet been confirmed at this point, and it requires more investigation. In 2003 and 2012, the Middle East Respiratory Syndrome (MERS) coronavirus and Severe Acute Respiratory Syndrome (SARS) coronavirus were found to be zoonotic such that they may be transmitted among animals and humans (7). COVID-19 is the third profoundly pathogenic human coronavirus that has been identified over the most recent two decades. The individual-to-individual transmission has been depicted both in emergency clinics and family settings (8). Therefore, it is necessary to forestall any further spread in the general society and in human services settings. COVID-19 transmission through tainted dry surfaces makes it even easier to transmit. Hence self-immunization of the mucous layers of the nose, eyes, or mouth has been proposed (9–11). Biocidal products like hydrogen peroxide, alcohols, sodium hypochlorite, and benzalkonium chloride are being utilized worldwide for sanitization purposes, especially in social settings (12).

As of March 25 2020, 18,295 individuals had died from COVID-19 infection, while 107,089 patients recovered. As per the WHO, there were more than 411,242 confirmed cases worldwide, with the majority of revealed cases in Wuhan city. This led to Wuhan placing a citywide lockdown on January 23 2020, in which no individuals were permitted to enter or leave. The officials temporarily suspended all accessible transportation, including trains, metro, air terminals, and public vehicles to avoid the spread of COVID-19. Also, a few urban areas in Hubei territory were put under lockdown. One of the challenges posed by COVID-19 is its quarantine period, which is as long as 2–14 days (13), and during this period, it can spread to others. Besides, in (14), it is mentioned that the alone time may range from 0 to 24 days depending upon the situation of the patient. The spread of such sickness is unbelievably dangerous. It requires continuously extraordinary blueprints and plans, which have been executed in different Chinese urban districts, particularly in the Hubei area. Hence it is indispensable to explore the number of confirmed cases at this time to start the vital assertion plans. The main contribution of this research work is the use of an ARIMA model (15), which is capable of forecasting the global pandemic COVID-19 using the dataset, as shown in **Table 1**. The main contributions are per the following:

- We used a proficient forecasting model to find the confirmed cases of COVID-19 dependent on recently confirmed cases.
- An ARIMA model was used to forecast the exact confirmed fatalities of the coronavirus outbreak from February 1 2020 to April 1 2020.
- We evaluated total confirmed cases, total fatalities, confirmed cases concerning fatal cases, Q-Q plot of confirmed and fatal cases, white noise confirmed cases vs. fatal cases, and lastly the autocorrelation function between confirmed and fatalities cases.

TABLE 1 | Dataset used in this study.

Province state	Country region	Date (M/D/Y)	Confirmed cases	Fatalities
Hubei	China	1/22/2020	444	17
Hubei	China	1/23/2020	444	17
Hubei	China	1/24/2020	549	24
Hubei	China	1/25/2020	761	40
Hubei	China	1/26/2020	1,058	52
Hubei	China	1/27/2020	1,423	76
Hubei	China	1/28/2020	3,554	125
Hubei	China	1/29/2020	3,554	125
Hubei	China	1/30/2020	4,903	162
Hubei	China	1/31/2020	5,806	204
Hubei	China	2/1/2020	7,153	249
Hubei	China	2/2/2020	11,177	350
Hubei	China	2/3/2020	13,522	414
Hubei	China	2/4/2020	16,678	479
Hubei	China	2/5/2020	19,665	549

The rest of this research has been organized as follows: section Literature Survey provides a survey of the previous work. Dataset description and ARIMA model are discussed in section Material and Methods. The results and their analysis are illustrated in section Result and Discussion. Finally, we conclude the paper in section Conclusion.

LITERATURE SURVEY

Existing work has been conducted in the past to evaluate estimation problems, like an adaptive neuro-fuzzy inference system (ANFIS) (16), which is applied extensively in the time course of action desire and envisioning issues, and it indicated that there was incredible execution in the present application. It offers adaptability for handling non-linearity in time series data, by combining an artificial neural network (ANN) and a fuzzy approach. ARIMA models applied to historical hemorrhagic fever with renal syndrome (HFRS) occurrence information are a significant device for HFRS observation in China. Chinese HFRS information from 1975 to 2008 was taken into account for fitting the ARIMA model. Akaike information criterion (AIC) and the Ljung-Box test have been relied on for assessing the developed models. Along these lines, the fitted ARIMA model was applied to get the suited HFRS frequency from 1978 to 2008 and appeared differently concerning the corresponding observed values (17). This paper highlights the significance of embracing dynamic modeling approaches, proposes difficulties for performing model determination across long time periods, and relates comprehensively to the predictability of complex adaptive systems. (18) introduced an ensemble model for sequential forecasting using a frequent computational bootstrap approach to evaluate the Ebola outbreak

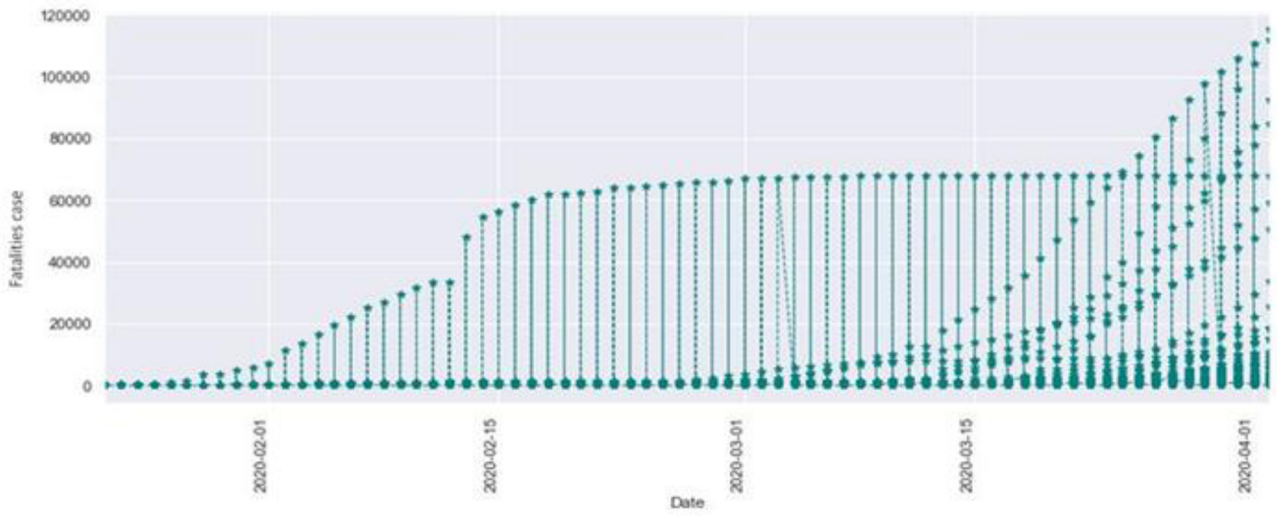

FIGURE 1 | Total confirmed cases of COVID-19.

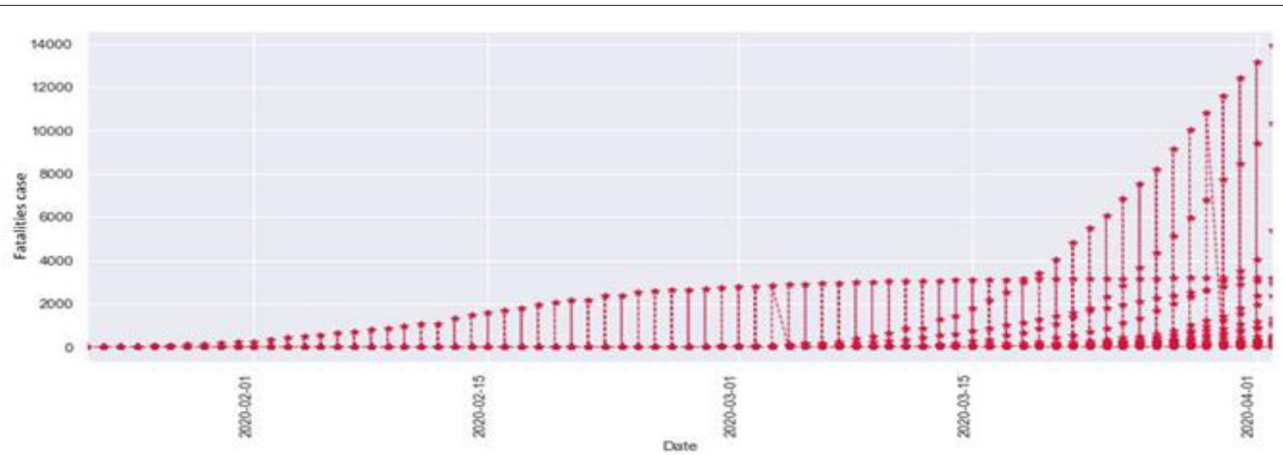

FIGURE 2 | Total fatalities of COVID-19.

and generated short-term forecasts of the epidemic outbreak by combining two models, the generalized-growth model (GGM) and the generalized-logistic model (GLM) (19). The seasonal autoregressive-integrated moving average (SARIMA) model is used to forecast monthly cases of hand, foot, and mouth disease (HFMD) in China (20). A short-term forecast of incidence in China has been done by applying ARIMA and exponential smoothing (ETS) that analyzed data from the Chinese Center for Disease Control and Prevention between 2005 and 2006 (21).

Ture and Kurt (22) proposed a comparative study among different types of time series methods to forecast Hepatitis A virus (HAV) infection. The methods considered were the ANN algorithm, radial basis function (RBF), time-delay neural networks (TDNN), and the ARIMA model, where the ANN algorithm was found to be more accurate than the others. In paper (15), the authors proposed to apply a susceptible–infectious–recovered–susceptible (SIRS) mathematical model estimating model dependence on gathering alteration Kalman channels for occasional flare-ups of flu. They assessed the proposed model utilizing the flu season information of New York City for a long period (2003–2008). Massad et al. (23), proposed a numerical model to break down and gauge the disease of the SARS epidemic to survey the viability of these techniques. Here the author worked to determine 13 years of time series data. In another work, Shaman et al. (24) formulated three scenarios based on a hypothesis about under-reporting of EVD cases and the EVD case fatality ratio using a standard life table technique to calculate the life expectancy

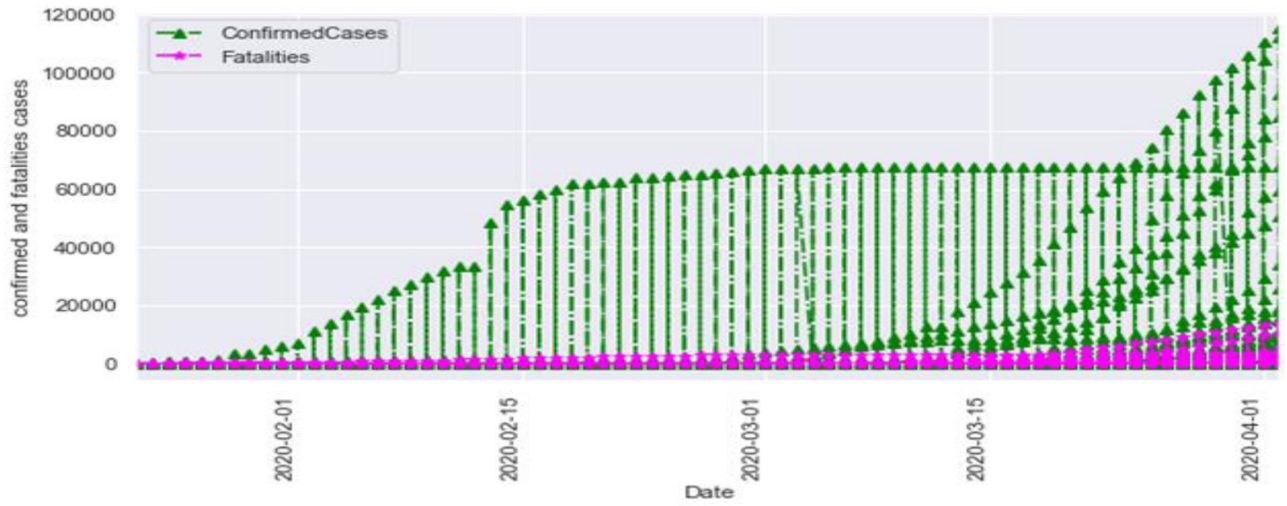

FIGURE 3 | The confirmed cases vs. fatalities.

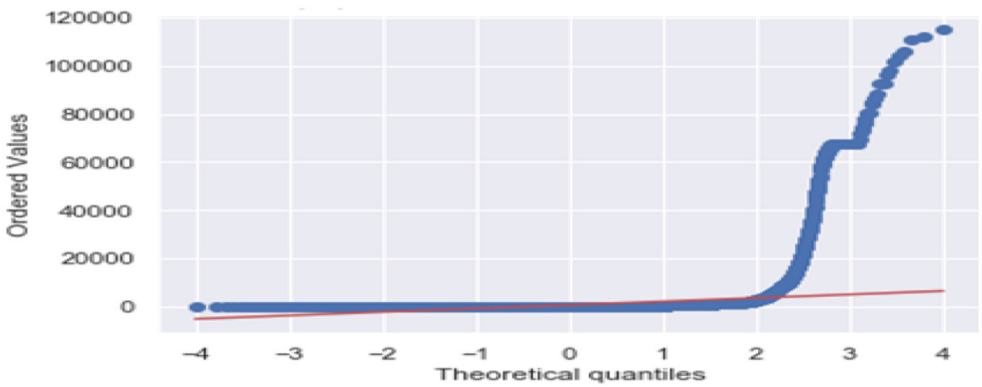

FIGURE 4 | The Q-Q plot of confirmed cases.

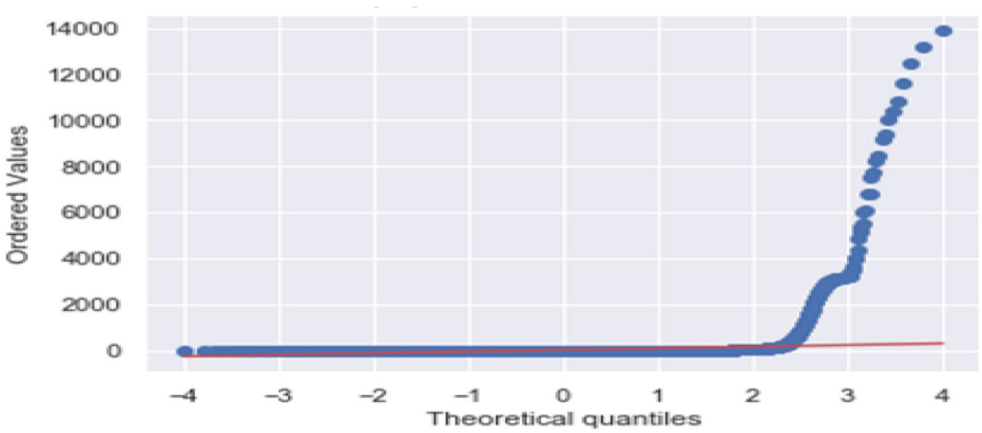

FIGURE 5 | The Q-Q plot of fatalities.

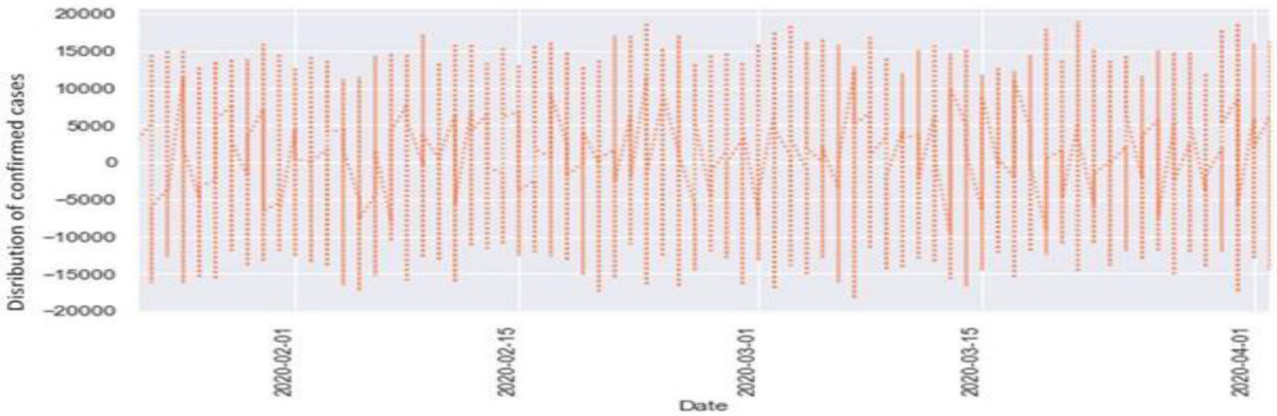

FIGURE 6 | White noise time-series of confirmed cases.

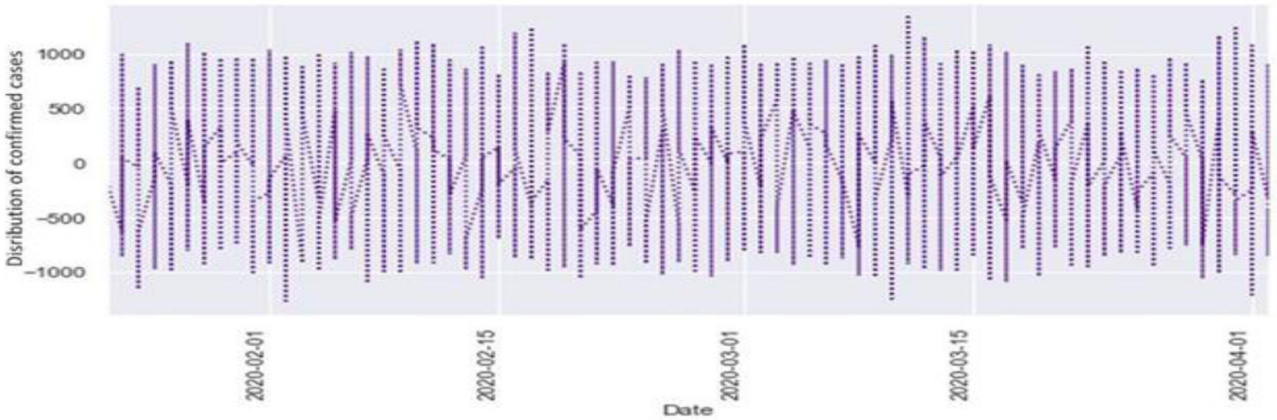

FIGURE 7 | White noise time-series of fatal case.

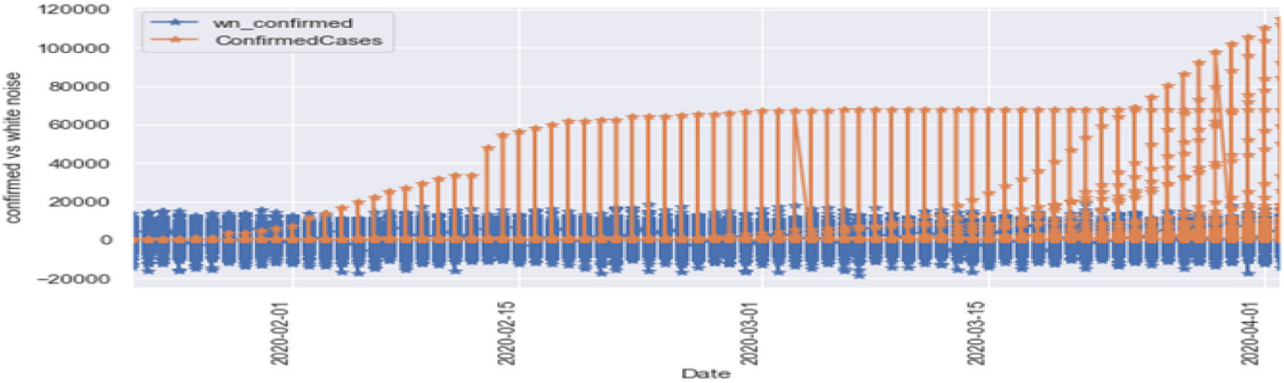

FIGURE 8 | White noise confirmed cases vs. confirmed cases.

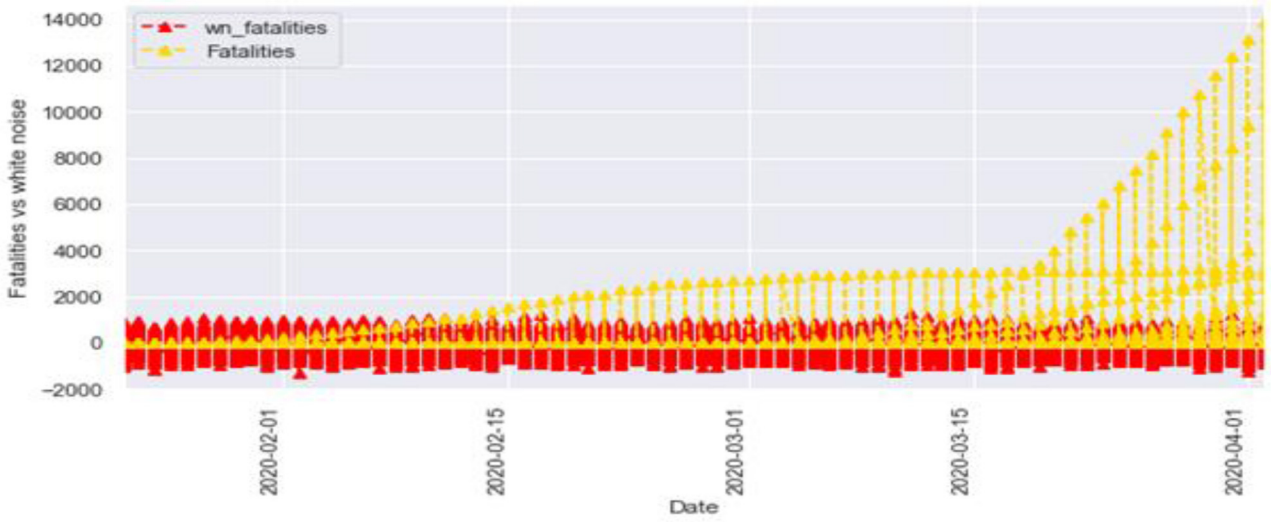

FIGURE 9 | White noise fatal case vs. fatal cases.

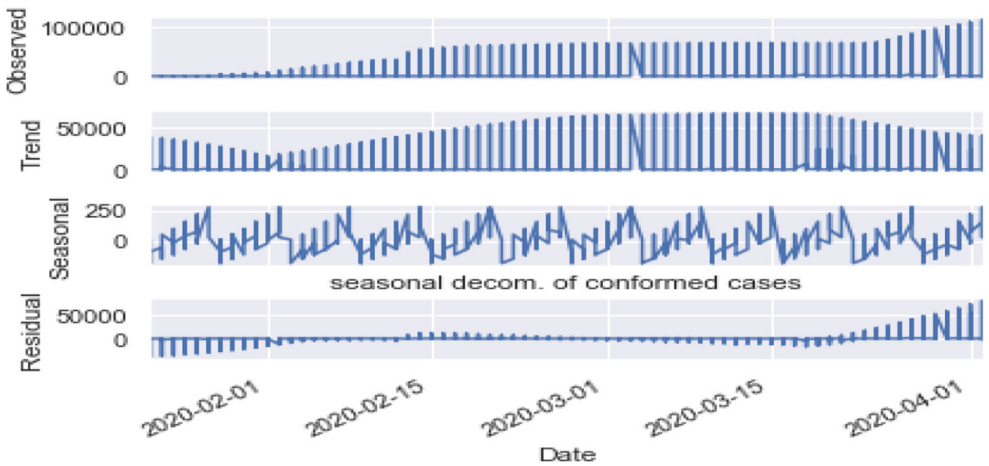

FIGURE 10 | Seasonal decomposition of confirmed cases.

of Ebola virus disease (EVD) patients in a couple of African countries.

On the basis of the existing studies, in this paper study, the ARIMA model was used for time series analysis to either get comprehensive information or to anticipate future qualities. This model is applicable *in situ*ations where information may be non-fixed. Non-fixed practices can be patterns, a cycle, random walks, or mixes of the three. Non-fixed information focuses are unusual and can not be displayed or estimated. An investigation utilizing non-fixed time arrangement information focuses may not be fitting as it might show the connection between two factors where one does not exist. To get predictable, reliable outcomes, the non-fixed information should be changed into fixed information. The non-fixed procedure and the fixed procedure around a consistent long haul has a steady difference autonomous of time.

MATERIALS AND METHODS

Dataset

The dataset considered for the study has been collected from relevant sources (https://www.kaggle.com/c/covid19-global-forecasting-week-2/data). It contains the day to day confirmed cases from all over the world between January 22 2020 and March 31 2020. An overview of the dataset has been shown in **Table 1**. The dataset consists of a total of 22,032 columns and 7 rows. The COVID-19 dataset also includes 5 attributes, i.e., id, prov_state, country_region, confirmed case, and fatal case. The data is in the form of time-series data points.

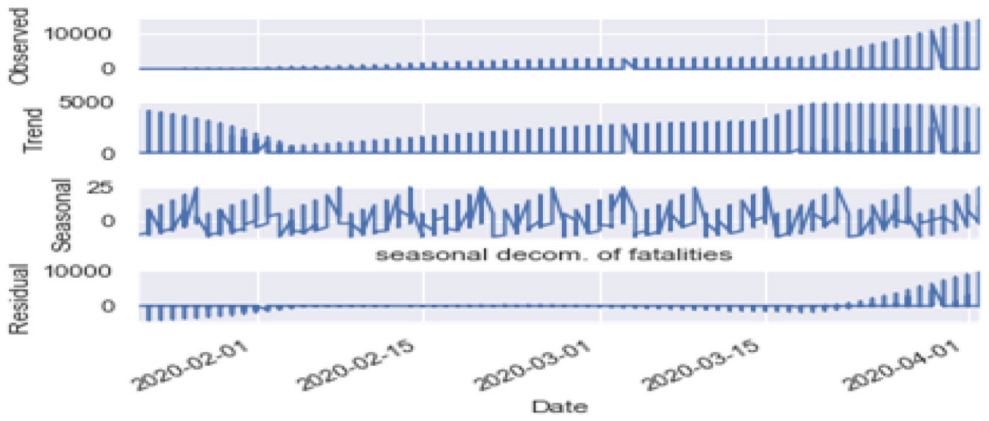

FIGURE 11 | Seasonal decomposition of fatal cases.

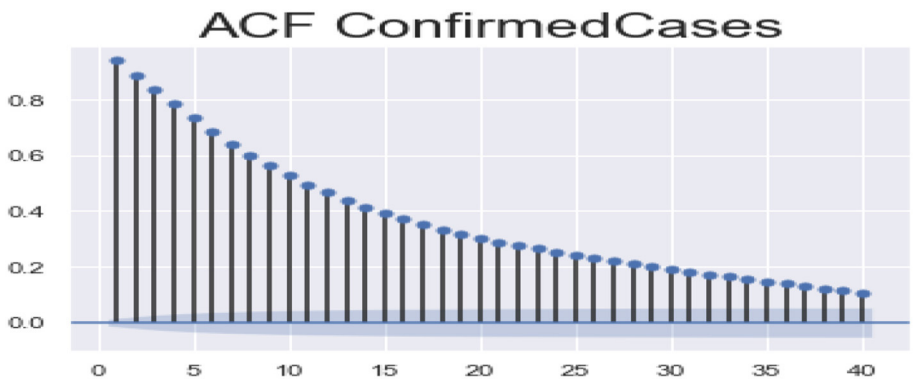

FIGURE 12 | Autocorrelation function confirmed cases.

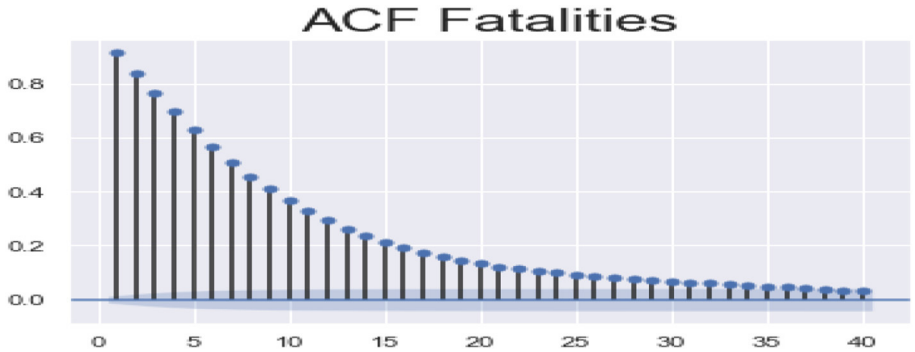

FIGURE 13 | Autocorrelation function fatalities.

Time-series is a sequence of information that describes the time period of each value. Generally, time-series data used for analysis and forecasting the future is based on historical data. Time-series data determines the stability of a situation over time and efficiency portfolios. Time-series datasets are time-dependent because values for every period are affected by outside factors and the values of the past period. During the dataset loading operation, we considered the date as our

FIGURE 14 | True vs. predicted values.

TABLE 2 | Comparative analysis.

S/No.	References	Studied countries	Sources of data or repository	Parameters	Methods/model used	Discussion
1	(1)	China	https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/	Fitness value	FPASSA-ANFIS	Predict the number of confirmed cases within 10 days based on previously confirmed cases
2	(25)	Italy	Italy government data	Transmission rate, recovery rate, and morality rate	ESIR	Forecasting total number COVID-19 cases
3	(26)	India (Maharashtra, Gujarat, Delhi)	https://www.mohfw.gov.in/.	Statistical parameter and metrics ETs and GEP	Genetic programming (GP)	Predict confirmed and death case
4	(27)	Brazil	World health organization	Total population of Brazil	Number of susceptible, exposed, infectious, recovered (SEIR)	Policy-making for avoiding outbreak in metropolitan cities
5	Our work	Global forecasting	https://www.kaggle.com/c/covid19-global-forecasting-week-2/data	Fraction of population	Autoregressive integrated moving average (ARIMA)	Forecast confirmed and fatalities case of COVID-19 across the globe

index column. Therefore, the date column is no longer a feature for us. This is because time-series data perform tasks related to the date. That is why it is the most used parameter in our methodology (**Table 1**).

Autoregressive Integrated Moving Average (ARIMA)

ARIMA is a famous and adaptable class of forecasting models that uses recorded data to make estimations. This model is an essential forecasting technique that can serve as a starting point for progressively complex models (15). It works effectively when the information displays a steady or predictable example after some time with a base measure of anomalies. The ARIMA approach endeavors to portray developments in a stationary time series as an element of what is designated as "autoregressive and moving normal" parameters. These are alluded to as autoregressive parameters and moving average (MA) parameters. We accept time is a discrete variable, Z_t shows the observation at time t and t demonstrates the zero-mean random noise term at time t. The MA(n) (moving average) model uses this procedure:

$$Zt = \sum_{i=1}^{n} \gamma_i \epsilon_{t-1} + \epsilon_t \quad (1)$$

where $\boldsymbol{\gamma_i}$ denotes coefficient, similar to MA(n) models, autoregression model, denoted by AR(m),

$$Zt = \sum_{i=1}^{m} \delta_i Z_{t-1} + \epsilon_t \quad (2)$$

Zt is a noisy linear combination of the previously taken *m* observations. An increasingly advanced model is the ARIMA (*m, n*), mix of AR(*m*), and MA(*n*) with a reduced structure and gives an adaptable demonstrating system. This model expects that Zt is created through the formula:

$$Z_{\mathrm{t}} = \sum_{i=1}^{n} \gamma_{\mathrm{i}} \epsilon_{\mathrm{t}-1} + \sum_{i=1}^{m} \delta_{\mathrm{i}} Z_{\mathrm{t}-1} + \epsilon_{\mathrm{t}}, \tag{3}$$

where *t* is the zero-mean noise term. On the off chance that we are adding imperative to the AR(*m*) part, it ensures a stationary process. A fixed and invertible ARIMA (*m, n*) model may be depicted either as an infinite AR model (AR(∞)) or an infinite MA model (MA(∞)). For the ARIMA model, one can compute the first-order differences of Zt by ∇Zt= Z_{t}-$Z_{\mathrm{t}-1}$ and second-order differences of Zt by ∇2Zt= ∇Zt–∇Zt-1 such that the sequence of ∇dZt satisfies an ARIMA (*m, n*). We state that the sequence of Zt satisfies the ARIMA (*m, d, n*).

$$\nabla^{d} Z_{t} = \sum_{i=1}^{n} \gamma_{\mathrm{i}} \epsilon_{\mathrm{t}-1} + \sum_{i=1}^{m} \delta i \nabla^{d} Z t_{-1} + \epsilon_{\mathrm{t}}, \tag{4}$$

which are specified by three order parameters terms *m*, *d*, *n* with specific weights vector $\delta \in \mathrm{R}$ *m* and $\gamma \in \mathrm{R}n$. Forecasting with ARIMA *(m, d, n)* is an inversion of the differential equation. Assuming the time-series sequence *Zt* fulfills ARIMA *(m, d, n)*, one can predict the *d*-th order differential of observation at time $t + 1$ as $\nabla^{d} Z_{t+1}^{\sim}$ and then predict the observation at time $t + 1$ as $Z_{t}^{\sim}$:

$$Z_{t}^{\sim} = \nabla^{d} Z_{t}^{\sim} + \sum_{i=0}^{d-1} . \nabla^{i} Z_{t-1} \tag{5}$$

RESULT AND DISCUSSION

For this study, data were analyzed using a python library called matplotlib. It is a popular package for plotting 2D data. This library has been used to derive the line charts of the dataset. We analyzed the COVID-19 data and performed data visualization, which gave a complete idea of the brief summarization of our dataset. For visualization, we used python modules like pandas, matplotlib, and seaborn. The study provided us with the summarized data using the described methods. This function prints the total distribution of the dataset, i.e., 50% of dataset, 75% of dataset, etc. We used further visualization techniques to get a better insight into our data. Using various parameters, we have analyzed our data and described the total confirmed case of COVID-19 starting from January 1 2020 to April 1 2020. We observed that as the time period increased the number of confirmed cases also increased. In **Figure 1**, the x-axis indicates the period as months, and the y-axis indicates the number of fatalities.

Figure 2 describes the total fatalities of COVID-19 starting from January. As the time period increased, the number of confirmed cases also increased. We can observe that as the time period increased at the same time the number of fatalities also increased. Here the x-axis indicates the time period in months, and the y-axis indicates the numbers of fatalities.

Figure 3 is the comparison of increasing trends of confirmed and fatal cases over the same period of time. As we can see the legends are mentioned above in the diagram. From this graph, we observed that the number of confirmed cases were more than the number of fatalities. While the number of confirmed cases increased gradually, it was not the same for fatalities. **Figures 4**, **5** are referred to as Q-Q plots in statistics. These plots are a graphical technique for determining if two data sets come from populations with a standard distribution. For such modules in python, we used scipy.stats and pylab. The above Q-Q plots of confirmed cases and fatalities describe the theoretical quintiles of both cases. It means that based on the numbers and statistics, the theoretical increase in the graph should follow the red line. The quantiles-quantiles (Q-Q) plots are only used to draw the theoretical quintiles.

Figures 6, **7** represent the white noise time-series data of the above COVID-19 data. White noise is a sequence of independent and identically distributed random variables with finite mean and variance. **Figure 6** is the white noise figure of the confirmed case and **Figure 8** is the white noise figure of the fatal case. It is worth mentioning that our dataset is stationary in nature because most values are around the mean figure in the white noise data. White noise describes the particular behavior of the time-series data.

Figure 8 is the comparison of the confirmed cases and white noise data. From the graph, we can see that the initial values are mostly around white noise, which means our dataset is distributed well.

Figure 9 represents the comparison of the fatalities and white noise data. From the graph, we can observe that most of the initial values are around the white noise, which means our dataset is distributed well.

Seasonality: **Figures 10**, **11** are the seasonality analysis of our data. A repeating pattern within a given time period is known as seasonality, although the term is applied more generally to repeating patterns within any fixed period. It means that we decomposed the time-series data and split them into trend, seasonal, residual, and observation. Seasonal decomposition can be performed in two ways, i.e., multiplicative. Here the term trend refers to a general systematic linear or (most often) non-linear component that changes over time and does not repeat, i.e., distribution throughout data. Seasonal refers to the cyclical effects of the dataset. Residual means the error of prediction. In a time series, it depicts what is left over after fitting a model.

ACF (autocorrelation function): **Figures 12**, **13** depict the ACF of both confirmed cases and fatalities over time. Autocorrelation is the correlation between a sequence and itself. Statistically, it can be referred to as the correlation among the members of a variable. But in general, when the values of the observation are somehow related to each other, the corresponding stage is referred to as autocorrelation.

For model building, we have used the ARIMA model. By applying the ARIMA model, we forecast the future trend of confirmed cases and fatalities which is shown in **Figure 14**.

A comparative analysis of COVID-19 has been discussed in **Table 2**. In this study the ARIMA model focused on different global forecasting of COVID-19 total confirmed cases and total fatal cases in the earlier stage. Here in the current research work, population has been taken as a parameter.

CONCLUSION

In response to the COVID-19 pandemic, we applied time series analysis to find different measurements like the stationary, trend, and the pattern of the dataset. Various visualization techniques have been applied to the dataset for studying the outbreak related to COVID-19. We have relied on seaborn and matplotlib modules for the same. The graphs appropriately describe the trend and pattern of the COVID-19 pandemic outbreak. The time-series model ARIMA has been used to forecast the cases of COVID-19 in the future and has successfully calculated the total confirmed cases and fatalities over the studied dates and Q-Q plots of confirmed cases and fatalities. We have also estimated the total confirmed cases and fatalities over the date-Q plots. The dataset is stationary in nature; we have presented the ACF of both confirmed cases and fatalities over time and forecasted the future trend for the same. This study provides an advanced level of work, which may be useful in analyzing as well as fighting the pandemic. In future work, we can apply advanced algorithms and techniques for preparing the model which will improve and forecast more precisely.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

AUTHOR CONTRIBUTIONS

DD and RK: conceptualization and writing—original draft preparation. DD, JD, and MM: methodology. RS and RK: software. D-NL: validation and visualization. RK and RS: writing—review and editing, formal analysis, and supervision. RK: investigation. D-NL and JD: data curation. All authors: contributed to the article and approved the submitted version.

REFERENCES

1. Priyadarshini I, Mohanty P, Kumar R, Son LH, Chau HTM, Nhu VH, et al. Analysis of outbreak and global impacts of the COVID-19. *Healthcare*. (2020) 8:148. doi: 10.3390/healthcare8020148
2. Chen Y, Liu Q, Guo D. Emerging coronaviruses: genome structure, replication, and pathogenesis. *J Med Virol*. (2020) 92:418–23. doi: 10.1002/jmv.25681
3. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature*. (2013) 503:535–8. doi: 10.1038/nature12711
4. Wang LF, Shi Z, Zhang S, Field H, Daszak P, Eaton BT. Review of bats and SARS. *Emerg. Infect Dis*. (2006) 12:1834. doi: 10.3201/eid1212.060401
5. Organization, W. H. Novel Coronavirus (2019-nCoV) (2020). Available online at: https://www.who.int/ (accessed January 27, 2020).
6. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. (2020) 395:565–74. doi: 10.1016/S0140-6736(20)30251-8
7. Cauchemez S, Van Kerkhove M, Riley S, Donnelly C, Fraser C, Ferguson N. Transmission scenarios for Middle East Respiratory Syndrome Coronavirus (MERS-CoV) and how to tell them apart. *Euro Surveill*. (2013) 18:20503.
8. Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet*. (2020) 395:30154–9. doi: 10.1016/S0140-6736(20)30154-9
9. Otter JA, Donskey C, Yezli S, Douthwaite S, Goldenberg SD, Weber DJ. Transmission of SARS and MERS coronaviruses and influenza virus in healthcare settings: the possible role of dry surface contamination. *J Hosp Infect*. (2016) 92:235e50. doi: 10.1016/j.jhin.2015.08.027
10. Dowell SF, Simmerman JM, Erdman DD, Wu JS, Chaovavanich A, Javadi M. Severe acute respiratory syndrome coronavirus on hospital surfaces. *Clin Infect Dis*. (2004) 39:652e7. doi: 10.1086/422652
11. Geller C, Varbanov M, Duval RE. Human coronaviruses: insights into environmental resistance and its influence on the development of new antiseptic strategies. *Viruses*. (2012) 4:3044e68. doi: 10.3390/v4113044
12. Kampf G. *Antiseptic Stewardship: Biocide Resistance and Clinical Implications*. Cham: Springer International Publishing (2018). p. 78–84. doi: 10.1007/978-3-319-98785-9
13. Cheng ZJ, Shan J. Novel Coronavirus: Where we are and What We Know. *Infection*. (2019) 48:155–63. doi: 10.1007/s15010-020-01401-y
14. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of 2019 novel coronavirus infection in China. *medRxiv*. (2020) 382:1708–20. doi: 10.1101/2020.02.06.20020974
15. Shaman J, Karspeck A. Forecasting seasonal outbreaks of influenza. *Proc Natl Acad Sci USA*. (2012) 109:20425–30. doi: 10.1073/pnas.1208772109
16. Jang JS. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern*. (1993) 23:665–85. doi: 10.1109/21.256541
17. Liu Q, Liu X, Jiang B, Yang W. Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC Infect Dis*. (2011) 11:218. doi: 10.1186/1471-2334-11-218
18. Scarpino SV, Petri G. On the predictability of infectious disease outbreaks. *Nat Commun*. (2019) 10:47–54. doi: 10.1038/s41467-019-08616-0
19. Chowell G, Luo R, Sun K, Roosa K, Tariq A, Viboud C. Real-time forecasting of epidemic trajectories using computational dynamic ensembles. *Epidemics*. (2020) 30:100379. doi: 10.1016/j.epidem.2019.100379
20. Tian CW, Wang H, Luo XM. Time-series modelling and forecasting of hand, foot and mouth disease cases in China from 2008 to 2018. *Epidemiol Infect*. (2019) 147:1–3. doi: 10.1017/S095026881800362X
21. Zeng Q, Li D, Huang G, Xia J, Wang X, Zhang Y, et al. Time series analysis of temporal trends in the pertussis incidence in Mainland China from 2005 to 2016. *Sci Reports*. (2016) 6:32367. doi: 10.1038/srep32367
22. Ture M, Kurt I. Comparison of four different time series methods to forecast hepatitis A virus infection. *Expert Systems Applications*. (2006) 31:41–6. doi: 10.1016/j.eswa.2005.09.002
23. Massad E, Burattini MN, Lopez LF, Coutinho FA. Forecasting versus projection models in epidemiology: the case of the SARS epidemics. *Med. Hypotheses*. (2005) 65:17–22. doi: 10.1016/j.mehy.2004.09.029
24. Shaman J, Yang W, Kandula S. Inference and forecast of the current West African Ebola outbreak in Guinea, Sierra Leone and Liberia. *PLoS Curr*. (2014) 6:1–17 doi: 10.1371/currents.outbreaks.3408774290b1a0f2dd7cae877c8b8ff6

25. Jia W, Han K, Song Y, Cao W, Wang S, Yang S, et al. Extended SIR prediction of the epidemics trend of COVID-19 in Italy and compared with Hunan, China. *medRxiv*. (2020) 7:1–7. doi: 10.1101/2020.03.18.20038570
26. Salgotra R, Gandomi M, Gandomi AH. Time series analysis and forecast of the COVID-19 pandemic in India using genetic programming. *Chaos Solitons Fractals*. (2020) 138:109945. doi: 10.1016/j.chaos.2020.109945
27. Rocha Filho TM, dos Santos FSG, Gomes VB, Rocha TA, Croda JH, Ramalho WM, et al. Expected impact' of COVID-19 outbreak in a major metropolitan area in Brazil. *medRxiv*. (2020) 14:20035873. doi: 10.1101/2020.03.14.20035873

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 Computer Science

ORIGINAL RESEARCH
published: 20 November 2020
doi: 10.3389/fcomp.2020.551481

Latent Class and Transition Analysis of Alzheimer's Disease Data

Hany Alashwal[1], Thierno M. O. Diallo[2,3], Richard Tindle[4], Ahmed A. Moustafa[5,6,7] and for the Alzheimer's Disease Neuroimaging Initiative[†]*

[1] College of Information Technology, United Arab Emirates University, Al-Ain, United Arab Emiratest, [2] School of Social Science, Western Sydney University, Sydney, NSW, Australia, [3] Statistiques & M.N., Canada, [4] School of Psychology, Charles Stuart University, Bathurst, NSW, Australia, [5] MARCS Institute for Brain and Behaviour, Western Sydney University, Sydney, NSW, Australia, [6] School of Psychology, Western Sydney University, Sydney, NSW, Australia, [7] Department of Human Anatomy and Physiology, The Faculty of Health Sciences, University of Johannesburg, Johannesburg, South Africa

OPEN ACCESS

Edited by:
Deepak Gupta,
Maharaja Agrasen Institute of Technology, India

Reviewed by:
Angelo D'Ambrosio,
University of Turin, Italy
Shankar Kathiresan,
Alagappa University, India

***Correspondence:**
Hany Alashwal
halashwal@uaeu.ac.ae

[†] *Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how\protect_to\protect_apply/ADNI\protect_Acknowledgment\protect_List.pdf*

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Computer Science

Received: 13 April 2020
Accepted: 27 October 2020
Published: 20 November 2020

Citation:
Alashwal H, Diallo TMO, Tindle R, Moustafa AA and for the Alzheimer's Disease Neuroimaging Initiative (2020) Latent Class and Transition Analysis of Alzheimer's Disease Data. Front. Comput. Sci. 2:551481. doi: 10.3389/fcomp.2020.551481

This study uses independent latent class analysis (LCA) and latent transition analysis (LTA) to explore accurate diagnosis and disease status change of a big Alzheimer's disease Neuroimaging Initiative (ADNI) data of 2,132 individuals over a 3-year period. The data includes clinical and neural measures of controls (CN), individuals with subjective memory complains (SMC), early-onset mild cognitive impairment (EMCI), late-onset mild cognitive impairment (LMCI), and Alzheimer's disease (AD). LCA at each time point yielded 3 classes: Class 1 is mostly composed of individuals from CN, SMC, and EMCI groups; Class 2 represents individuals from LMCI and AD groups with improved scores on memory, clinical, and neural measures; in contrast, Class 3 represents LMCI and from AD individuals with deteriorated scores on memory, clinical, and neural measures. However, 63 individuals from Class 1 were diagnosed as AD patients. This could be misdiagnosis, as their conditional probability of belonging to Class 1 (0.65) was higher than that of Class 2 (0.27) and Class 3 (0.08). LTA results showed that individuals had a higher probability of staying in the same class over time with probability >0.90 for Class 1 and 3 and probability >0.85 for Class 2. Individuals from Class 2, however, transitioned to Class 1 from time 2 to time 3 with a probability of 0.10. Other transition probabilities were not significant. Lastly, further analysis showed that individuals in Class 2 who moved to Class 1 have different memory, clinical, and neural measures to other individuals in the same class. We acknowledge that the proposed framework is sophisticated and time-consuming. However, given the severe neurodegenerative nature of AD, we argue that clinicians should prioritize an accurate diagnosis. Our findings show that LCA can provide a more accurate prediction for classifying and identifying the progression of AD compared to traditional clinical cut-off measures on neuropsychological assessments.

Keywords: Alzheimer's disease, latent class analysis, latent transition analysis, neural markers, misdiagnosis

INTRODUCTION

The World Health Organization has identified Alzheimer's disease (AD) as a public health priority, with ∼30–35 million cases worldwide (World Health Organization, 2012). Alzheimer's disease is a chronic neurodegenerative syndrome which causes severe progressive deterioration in cognitive impairment (Alzheimer Association, 2019). Impairments include detriments in memory, learning

ability, language, judgment, decision making, and disordered thinking (Alzheimer Association, 2019). Patients are diagnosed with AD after being assessed on multiple neuropsychological assessments, including memory, language functioning, personality, and behavioral changes. Assessments of specific biomarkers of AD are also being used to identify structural changes within specific brain regions as well as measure levels of Amyloid-β, tau, and phospho-tau (Alzheimer Association, 2019). Typically, the assessment of AD is based on clinical cut-off points for neuropsychological assessments and biomarkers. This technique allows a medical professional to identify those who have symptoms of AD. While clinical cut-offs are important for categorizing individuals with and without AD, it does not always contribute to our understanding of the progression of AD or identify individuals at risk of developing AD. Understanding the progression of AD is important to developing preventative interventions and earlier detection.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) study has collected longitudinal data from more than 50 sites in North America on "… elderly individuals with normal cognition, mild cognitive impairment (MCI)" (Jack et al., 2008). In total, the ADNI project has collected data using 292 measurements (see http://adni.loni.usc.edu/data-samples/adni-data-inventory for a full list of items). These measurements include diagnostic assessments; neuropsychological assessments; bio-specimens; genetics; imaging—including different MRI and PET imaging techniques; demographic and medical history; and a participation record. Data were collected at 0, 6, 12, 24, and 36 months for participants in the normal cognition and mild cognitive impairment groups (Jack et al., 2008). However, the AD group's data were only recorded at 0, 6, 12, and 24 months (Jack et al., 2008). The main aims of the ADNI project are to improve early detection and track disease progression using biomarkers and advance early intervention, prevention, and treatment.

AD Diagnosis

The assessment and diagnosis of AD have primarily relied on cut-off scores on neuropsychological assessments. For example, the Clinical Dementia Rating Scale (Morris, 1997) can be used to categorize individuals into differing levels of severity ranging from normal cognitive functioning, questionable cognitive impairment, questionable impairment, very mild dementia, mild dementia, moderate dementia, and severe dementia (O'Bryant et al., 2008). An advantage of utilizing categories for the Clinical Dementia Rating Scale is that it reliably identifies individuals with mild cognitive impairment (Duara et al., 2013). This allows clinicians to use the results to identify patients who are suffering from differing degrees of dementia severity. However, utilizing the scale with cut-offs does not allow health professionals to track the progression of the disease or identify at-risk patients before the presentation of symptoms.

Other neuropsychological assessments such as the Functional Activities Questionnaire (Pfeffer et al., 1982), the Alzheimer's Disease Assessment Scale (Mohs and Cohen, 1988), Clinical Dementia Rating Scale (Morris, 1997), Everyday Cognition Scale (Marshall et al., 2014), Montreal Cognitive Assessment (Nasreddine et al., 2005), the Mini-Mental State Exam (Folstein et al., 1975), and the Cognitive Change Index (Saykin et al., 2006; Rattanabannakit et al., 2016) have also been used to categorize cognitive impairment and AD. For instance, within the ADNI, participants are classified with AD if they obtain a score between 20 and 26 on the mini-mental state examination; a score between 0.5 and 1.0 on the global clinical dementia rating, a score between 1.0 and 9.0 for the summed box-score for the clinical dementia rating (Shaw et al., 2009). These standardized assessments are useful for diagnosing probable AD, with most yielding good sensitivity, specificity, and classification scores. That is, they reliably distinguish between individuals with mild cognitive impairment and AD—making them good diagnostic tools. However, utilizing these techniques is only useful for determining probable AD. The use of cognitive assessments only allows for a measure of current cognitive function and does not indicate if an individual may progress from mild cognitive impairment to severe cognitive impairment or AD. In the absence of objective diagnostic assessments for AD, a positive diagnosis is currently only determined through an autopsy (Perrin et al., 2009; Shaw et al., 2009). However, recent advances in imaging techniques (i.e., MRI and PET) and acquiring cerebral spinal fluid have allowed researchers to identify potential biomarkers of AD and what structural changes occur within specific brain regions (e.g., hippocampus).

Shaw et al. (2009) collected cerebral spinal fluid from elderly individuals with normal cognitive functioning, mild cognitive impairment, and mild AD (classification were determined using the mini-mental state examination and the Alzheimer's Disease Assessment Scale). The levels of Amyloid-β 1 to 42 peptide ($A\beta_{1-42}$), total tau (t-tau), and tau phosphorylated (p-tau) were assessed to determine potential biomarkers of AD. To gain more accurate cut-off points, models of the levels of $A\beta_{1-42}$, t-tau, and p-tau were determined from cerebral spinal fluid samples from autopsy-confirmed AD cases. The results indicated that $A\beta_{1-42}$ showed excellent sensitivity (96.4%) and specificity (76.9%) with a clinical cut-off of 192 pg/ml; t-Tau showed acceptable sensitivity (69.6%) and excellent specificity (92.3%) with a clinical cut-off of 93 pg/ml, and p-tau showed acceptable sensitivity (67.9%) and specificity (73.1%) with a clinical cut-off of 23 pg/ml. Further, the interaction between decreasing levels of Aβ and increasing levels of p-tau have recently been implicated with neuronal death, atrophy, and cognitive changes (Gomar et al., 2016; Veitch et al., 2019). These results suggest that $A\beta_{1-42}$ and p-tau are the most sensitive measures and best predictors of early diagnoses of AD.

The diagnosis of probable AD can also be assessed by measuring specific biomarkers (i.e., Aβ, t-tau, &p-tau). However, similar to the use of neuropsychological assessments, diagnosis relies on patients exceeding a clinical threshold for the levels of each biomarker. While biomarkers of AD appear to reliably distinguish between those diagnosed with (i.e., sensitivity) and without AD (i.e., specificity), some of the measures are still below the recommended threshold of 85% for sensitivity and specificity (Ronald and National Institute on Aging Working Group, 1998; Frank et al., 2003; Shaw et al., 2009). Again, the use of clinical cut-offs only provides clinicians with a measure to differentiate between mild cognitive impairment and probable AD based on particular biomarkers. Therefore, the use of cut-off

scores is essential for diagnosis but does not identify at-risk patients or to accurately track the progression of AD from mild cognitive impairment to pre-clinical AD, probable AD, and a final diagnosis of AD.

To promote the early detection of AD and to possibly identify at-risk individuals, research should not solely rely on clinical cut-off points, which are only useful once an individual presents with neuropsychological symptoms or biomarkers associated with probable AD. Secondly, there are criticisms of using cut-off points on continuous neuropsychological assessments because patients on either side of the cut-off are likely similar (Berlin et al., 2014; Petersen et al., 2019).

Latent Class Analysis

Instead, Latent Class Analysis (LCA) can be used to identify homogeneous subgroups of individuals who are externally heterogeneous to other sub-groups (Berlin et al., 2014; Eppig et al., 2017; Mooney et al., 2018; Petersen et al., 2019; Villeneuve et al., 2019; Zammit et al., 2019a). Latent class analysis can be used to identify homogenous subgroups of AD based on psychological assessments (e.g., Scheltens et al., 2016; Eppig et al., 2017; Zammit et al., 2019b). For example, Scheltens et al. (2016) identified eight cognitive subtypes of AD within their sample of probable AD patients ($N = 938$). The cognitive subtypes included patients with mild-memory impairment, moderate memory impairment, mild-visuospatial-language impairment; moderate-visuospatial impairment, mild-executive functioning impairment, moderate diffuse (cognitive impairment), and severe-diffuse (cognitive impairment). The authors suggest that the identification of cognitive subtypes highlights that AD is a complex disease and rather than classifying individuals with AD, we should consider differential diagnoses.

Zammit et al. (2019b) also used LCA to identify cognitive subtypes of AD within participants from the Rush Memory and Aging Project. Participants included in their study had no dementia at baseline; displayed signs of dementia at follow-up; were deceased at the time of the study, and had neuropathological data available. Neuropathological data were obtained from autopsies. Based on the neuropsychological outcomes at baseline (i.e., Episodic-, Semantic-, working-, and logical-memory; perceptual-and line orientation; and Perceptual Speed–Symbol Digits Modalities Test) latent class analysis was used to categories participants into 5 classes within two categories (i.e., impaired cognition and intact cognition). The impaired cognition classes included participants with *mixed-domains impairment*, *memory-specific impairment*, and *frontal impairment*. The intact cognition classes included participants with *average* cognition and *superior* cognition.

The aim of Zammit's (2019b) study was to identify if neuropathological evaluations at autopsy (i.e., Aβ, tau, hippocampal sclerosis, DNA-binding protein 43, Lewy bodies, cerebral amyloid angiopathy, atherosclerosis, and arteriolosclerosis) were predicted by the five classes of cognitive impairment and intact cognition at baseline and if the neuropathological measures differ between each class. Their results showed that baseline measurements on neuropsychological assessments were predictive of neuropathology measured at autopsy, suggesting that neuropsychological assessments are reliable for the assessment and prognosis of cognitive impairments associated with AD.

One of the main findings of Zammit et al. (2019b) study was that the biomarkers Aβ and Tau are strongly predictive of AD and can possibly be used as an early detector. Indeed, abnormal levels of Aβ and Tau were strongly associated with participants within the mixed-domains class, the memory-specific class, and the frontal impairment classes. With fewer abnormalities in the average cognition class and the superior cognition class. That is, abnormal Aβ and Tau were associated with impaired cognition but not intact cognition. One of the limitations of their study was that it did not account for individuals who might change classes from baseline to follow–up. For example, participants could progress from average intact cognition to memory-specific impairment. As such, the results are only capturing the class an individual belongs to at a single point in time.

Zammit et al. (2020) extended their previous work by using latent transition analysis to identify participants within the Rush Memory and Aging Project who transitioned from non-impairment to cognitive impairment. A second aim was to compare the classification of individuals within the LTA to the clinical criteria of MCI. The results showed that across three measurements (within 12 months) cognition remained relatively stable. That is, participants did not regularly change between the five classes of impairment; identified as mixed domains impairment, memory-specific impairment, frontal Impairment, average cognition, and superior cognition. However, of the 1,924 participants, 98 individuals did change membership class from time 1 to time 2 ($n = 62$) and from time 2 to time 3 ($n = 37$). A majority of the transitions were associated with a decline in cognitive impairment at both time points. These results identified that participants who changed classes had an 86% higher risk of developing AD than those who did not change status. Further, their study identified 541 participants with cognitive impairment at time 2, 10.5% of these participants progressed to developing dementia at time 3. While a majority of older adults cognition remains stable, those who are experiencing some level of cognitive impairment have an elevated risk of progressing to developing dementia. The authors provide evidence that using LTA is a robust tool to identify individuals at risk of cognitive decline, identifying risk factors for interventions to target.

Zammit's (2020) study was not without limitations. Specifically, their LTA only used neuropsychological measures of episodic memory, semantic memory, working memory, and perceptual speed and orientation. With evidence suggesting that neurological biomarkers are significant and sensitive predictors of early diagnoses of AD (Shaw et al., 2009; Gomar et al., 2016; Veitch et al., 2019), it is important to identify if biological markers of AD can predict cognitive impairment transitions. However, their paper does highlight that LCA and LTA are at the forefront of research aiming to improve diagnostic methods and to identify individuals at risk of progressing toward AD. Zammit et al. (2020) also note the need to validate these methods through replication of their findings, and efforts to identify homogeneous classes of cognitive impairment using other neuropsychological measures of AD. As mentioned earlier, recent studies using

LTA have not included biomarkers of AD. Our study adds a novel contribution to this emerging area by identifying if neuropsychological measures and neurological biomarkers of AD are indicators of individuals transitioning from healthy individuals to individuals with mild cognitive impairment and AD.

The Current Study

To our knowledge, Latent Transition Analyses (LTA) has not been used to identify the neuropsychological and biomarkers associated with the progression of AD in terms of patients transitioning from one AD class to another. In the present study rather than using a set cut off point to diagnose individuals as Alzheimer's patients, Latent Class Analysis (LCA) was used to identify individuals that are more likely to develop dementia. In addition, the focus of the analysis was on the development of the individuals over time, that is, how an individual changes class membership over time. In total, the following three research aims were addressed in this study: (1) determine and describe the number of classes that best characterize individuals with respect to clinical measures and neurological biomarkers; (2) compare the classification results obtained from the LCA and the cut-off methods, to identify the misdiagnosed individuals and characterize these patients; and (3) explore the developmental course of individuals with respect to clinical and neural measures.

Below, we first describe the ADNI dataset, which we have utilized in the current study. Second, we provide details on our latent class analysis and latent transition analysis. Following that, we present the results from both latent class analysis and latent transition analysis, respectively. Finally, we discuss our results in terms of importance of our findings and clinical implications.

METHOD

ADNI Dataset

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD).

The ADNI dataset includes 2,132 participants: 512 controls, and 353 with EMCI, 621 with LMCI, and 279 with SMC, and 367 AD patients. All participants were tested at 3 different times annually. In all participants, ADNI dataset includes the following measures for all participants: APOE4 = Apolipoprotein E4 gene; FDG = Fluorodeoxyglucose CDRSB = Clinical Dementia Rating Sum of Boxes; ADAS11 = Alzheimer's Disease Assessment Scale (Cognitive Subscale), 11 item version; MMSE = Mini-Mental State Examination; RAVimD = Rey Auditory Verbal Learning Test (Immediate word recall score); MOCA = The Montreal Cognitive Assessment; EcPtMm = Everyday Cognition-Participant Self Report (8 memory items); EcPtLg = Everyday Cognition-Participant Self Report (9 language items); EcSPM = Everyday Cognition- Participant Study Partner Report (8 Memory items); EcSPLg = Everyday Cognition- Participant Study Partner Report (9 Language items); Hipc = Hippocampus volume; Entor = entorhinal cortex volume; Fusif = fusiform gyrus volume.

Statistical Analysis

A series of Latent Class Analysis (LCA), multivariate analysis of variance, and Latent Transition Analyses (LTA) were conducted. LTA is a longitudinal extension of LCA that explores changes in class membership over time by capturing individual movements in forward and backward directions across time points. This statistical method is based on Markov chain models (Kaplan, 2008) and uses an LCA model as a measurement model.

Latent Class Analysis

Latent Class Analysis (Lazarsfeld and Henry, 1968; Clogg, 1981) was employed to empirically identify the number of classes that best characterize individuals with respect to clinical and neural measures. LCA is a mixture model that classifies participants into optimal classes on the basis of shared characteristics that distinguish members of one class from another. Furthermore, unlike traditional cluster analysis, which is based on heuristic or distance procedures (Moustafa et al., 2018; Alashwal et al., 2019), this approach is a model-based statistical method that allows the LCA solution to be replicated with an independent sample (e.g., Nylund et al., 2007).

A commonly-used strategy to determine the optimal number of classes in LCA is to estimate a series of models by progressively increasing the number of classes and comparing the models through fit statistics and tests of significance and the quality of classification across models, as well as the usefulness and the interpretability of the latent classes (e.g., Muthén and Muthén, 2000; Vermunt and Magidson, 2002). To determine the optimal number of classes for the sample, each model was evaluated using three information criteria (IC), namely, the Akaike Information Criterion (AIC; Akaike, 1987), the Bayesian Information Criterion (BIC; Schwartz, 1978), sample size adjusted BIC (SBIC; Sclove, 1987), and the Lo-Mendell-Rubin likelihood ratio test (LMR; Lo et al., 2001). For AIC, BIC, and SBIC, a lower value indicates a better model. For the LRT, a significant *p-value* for a model with k classes followed by a non-significant *p-value* for a model with $k + 1$ classes indicates that the k class model is the best fitting model. The indices BIC, SBIC, and LMR have been shown to identify the appropriate number of groups within finite mixture models (e.g., Diallo et al., 2016a,b; Diallo et al., 2017). Furthermore, the entropy criterion was used to examine the quality of classification across models. The normalized entropy values ranged from 0 to 1 with values >0.80 representing a clear assignment of individuals to latent classes. Finally, class size was also considered when determining the optimal number of latent classes. Small classes (i.e., those that contain $<5\%$ of the sample) were considered spurious classes, as they are often associated with class over-extraction (Hipp and Bauer, 2006).

Latent Transition Analysis

LCA can be extended to accommodate longitudinal data through LTA. LTA is a type of Markov model that studies how individuals change membership in latent classes over time. LTA links LCA variables at different time points to each other using autoregressive models. A series of multinomial logistic regression, where the latent class variable at time t is regressed on the latent class variable at time t-1, is commonly used to estimate transitions over time in latent class membership.

Analytical Steps

The statistical analyses involved three steps. In the first step, we identified the optimal number of classes for each time point separately. In the second step, individuals were assigned to their most likely latent class (modal class assignment) and the latent class variable at time 1 is compared to the clinical diagnostic variable and misclassified patients are studied using multivariate analysis of variance. The third step involved exploring the developmental course of the patients with respect to clinical and neural measures. That is, transition probabilities were used to explore changes that had taken place in the latent classes. For this analysis, measurement invariance was assumed to ensure that the classes have the same meaning over time. Specifically, measurement model parameters were set to be equal over time. Hence, conditional item probabilities, item means, and item variances for the LCA were constrained to be equal at the three time points.

For this study, all models were estimated using a Full Information Maximum Likelihood (FIML) procedure available in Mplus 8.3 (Muthén and Muthén, 2019). FIML utilizes all available information during the estimation process and provides consistent and efficient population parameters (Enders, 2010). Furthermore, all LCA models with continuous indicators were estimated with residual variances of the outcomes constrained to be equal across classes and under local independence within classes assumption (i.e., indicators' residual covariances within classes were constrained to zero). All models with two classes or more were estimated using 500 sets of random starting values, 50 iterations for each of these sets, and the 20 best sets of random starting values associated with the highest likelihood values were retained for the final optimization stage.

RESULTS

The first aim of this study was to determine the number of classes that best characterize patients with respect to clinical and neural measures. **Table 1** provides an overview of patients' characteristics with respect to clinical and neural measures. As these data showed, there was substantial variability among the patients on their clinical and neural measures. This variability supports the value of using mixture methods to assess whether the patients can be grouped into different classes based on their clinical and neural measures.

Latent Class Analysis Results

Latent class models containing 1–7 classes at each time point were fitted to the data. The model fit statistics are available in **Table 2**. All LCA models converged at Time 1. The log-likelihood increased while no minimum was found for the ICs as their values decreased across the range of models considered. The LMR pointed to the three-class solution since the test of the two-class model against the three-class model has a p-value of 0.003, suggesting rejection, whereas the test of the three-class against the four-class has a p-value of 0.24. Further, an examination of the LCA models indicated that the four- and five-class models each included small classes that seemed to have splintered off from larger classes in the three-class model. Therefore, a three-class model was selected at time 1 based on the fit statistics (Muthen, 2004). The three-class model resulted in a log-likelihood value of −11010.52 with 60 parameters, an AIC of 22141.04, a BIC of 22481.49, a SBIC of 22290.86, and a high entropy value of 0.89. Moreover, the three-class solution satisfied the minimum class size required to be useful (each comprised at least 5% of the sample) and meaningful.

All models with fewer than seven classes converged at time 2. Consequently, only model results for classes between one and six were considered for further analysis. As in time 1, log-likelihood values increased, no minimum was found for the ICs but no solution was favored by the LRT. However, similar to the results at time 1, the results showed that the four- and five-class models each included small classes that seemed to have splintered off from larger classes in the three-class model. Hence, based on the interpretability and the usefulness of the classes, the three-class solution was also selected as the optimal number of classes at time 2. Fit indices for the three-class solution at time 2 were as follows: Log-likelihood = −6927.78, number of parameters = 60, AIC = 13975.57, BIC = 14303.75, SBIC = 14113.13, and entropy = 0.82.

Finally, all models with fewer than six classes converged at time 3, whereas models with six classes and more did not converge. Hence, only model results for classes between one and five were considered for further analysis. Similar to time 1, log-likelihood values increased, no minimum was found for the ICs, whereas the LRT selected the three-class solution. Based on the interpretability and the usefulness of the classes, the three-class solution was also selected as the optimal number of classes at time 3. Fit indices for the three-class solution at time 3 were as follows: Log-likelihood = −6984.23, number of parameters = 60, AIC = 14088.45, BIC = 14415.74, SBIC = 14225.12, and entropy = 0.81.

Explanation of Latent Class Solutions

Here, we describe the latent class solutions at the three time points. Across the three time points, Class 1 showed a pattern of low means on CDRSB, ADAS11, EcPtMm, EcPtLg, EcSPM, EcSPLg, a pattern of high means on FDG, MMSE, RAVimD, MOCA, Hipc, Entor, Fusif, and selected category zero of with item probability >0.65. Class 1 is composed of 63% of the sample at time 1, 59% at time 2, and 65% at time three. In contrast, Class 3 showed a pattern of low means on FDG, MMSE, RAVimD, MOCA, Hipc, Entor, Fusif, a pattern of high means on CDRSB, ADAS11, EcPtLg, EcSPM. We, therefore, interpreted this class as the AD class. Class 3 is composed of 7% of the sample at time 1, 10% at time 2, and 6% at time three. Class 2, however, showed scores that overall were between Class 1 and Class 3. Class 2 was composed of 30% of the sample at time 1, 31% at

TABLE 1 | Descriptive statistics of the clinical and neural measures.

Variables	Time 1				Time 2				Time 3			
	%	*N*	M	SD	%	*N*	M	SD	%	*N*	M	SD
APOE4		1,726				1,210				1,189		
Zero	53.10				55.02				50.10			
One	37.00				36.26				38.01			
Two	9.90				8.72				11.89			
FDG		735	1.23	0.15		474	1.19	0.16		464	1.18	0.16
CDRSB		1,990	1.69	2.39		1,342	2.26	2.62		1,315	2.12	2.54
ADAS11		1,962	11.18	7.60		1,335	12.21	8.47		1,323	11.51	8.35
MMSE		1,978	27.07	3.32		1,326	26.30	3.90		1,325	26.49	3.91
RAVimD		1,956	36.33	13.47		1,340	32.41	13.56		1,316	34.01	13.23
MOCA		1,121	23.70	4.86		1,329	22.56	5.19		586	23.53	4.18
EcPtMm		1,144	1.98	0.72		537	2.12	0.73		589	2.06	0.73
EcPtLg		1,137	1.67	0.62		545	1.73	0.65		587	1.72	0.63
EcSPM		1,135	1.98	0.97		557	2.56	1.03		590	2.08	0.97
EcSPLg		1,138	1.60	0.80		557	1.82	0.91		588	1.67	0.79
Hipc		1,222	6676.92	1211.87		1,177	6609.89	1243.53		1,107	6625.67	1269.12
Entor		1,170	3437.39	810.34		1,105	3397.00	815.24		1,032	3414.71	838.77
Fusif		1,170	16942.23	2792.54		1,105	16886.15	2780.76		1,032	16877.08	2822.98

APOE4, Apolipoprotein E4 gene; FDG, Fluorodeoxyglucose; CDRSB, Clinical Dementia Rating Sum of Boxes; ADAS11, Alzheimer's Disease Assessment Scale (Cognitive Subscale), 11 item version; MMSE, Mini-Mental State Examination; RAVimD, Rey Auditory Verbal Learning Test (Immediate word recall score); MOCA, The Montreal Cognitive Assessment; EcPtMm, Everyday Cognition-Participant Self Report (8 memory items); EcPtLg, Everyday Cognition-Participant Self Report (9 language items); EcSPM, Everyday Cognition- Participant Study Partner Report (8 Memory items); EcSPLg, Everyday Cognition- Participant Study Partner Report (9 Language items); Hipc, Hippocampus volume; Entor, entorhinal cortex volume; Fusif, fusiform gyrus volume. N stands for number, and M is for mean.

time 2, and 29% at time three. The latent class estimated for the three-class solution for the three time points are shown in **Figures 1**–**3**. Finally, cross-tabulation analysis between the LCA solution and the diagnostic variable at time 1 (**Table 3**) showed that 37.5% of the patients from Class 1 were CN, 20.10% were SMC, 19.60% were EMCI, 18.10% were LMCI, and 4.6% were AD. Similar figures were 0.5, 0.9, 12.90, 53, and 32.60%; and 0.7, 0, 3.5, 28, and 67.80%, for Class 2 and Class 3, respectively. Class 1 can be seen as composed by individuals from CN, SMC, and EMCI groups. Class 2 represents people from LMCI and from AD groups with improved scores on memory, clinical, and neural measures. In contrast, Class 3 represents people from LMCI and from AD groups with deteriorated scores on memory, clinical, and neural measures. However, 63 individuals from Class 1 were classified as AD individuals with the diagnosis condition.

Multivariate Analysis of Variance

Multivariate analysis of variance was used to compare the group of 63 individuals with two groups: Class 1 from the LCA without the 63 individuals and the AD individuals from the diagnoses variable without the 63 individuals. The results of the multivariate analysis of variance with the 13 clinical and neural measures as dependent variables resulted in multivariate F statistics of $F_{(13,1,636)} = 163.85$, $p < 0.001$, partial $\eta^2 = 0.57$. Detailed analyses revealed that group differences were significant ($0.001 < p < 0.01$). The mean of the 63 individuals on the clinical and neural measures were between the mean of Class 1 without the 63 individuals and those from the AD individuals from the diagnoses variable without the 63 individuals (**Table 4**). The LCA classified the 63 individuals within Class 1 as their condition probability of belonging to this class was higher than those of belonging to the other two classes (with conditional probabilities of 0.65, 0.27, and 0.08 for Class 1, 2, 3, respectively). LTA will be used to study the development of these individuals over time.

Latent Transition Analysis Results

The LTA was conducted under the measurement invariance assumption. Consequently, measurement model parameters were set to be equal over time. This ensured that the classes have the same meaning over time. Transition probabilities for the whole sample are presented in **Table 5** and provide information on patient's status at time 2 given their latent status at time 1, and patient's status at time 3 given their latent status at time 2.The results showed most individuals stayed in the same class from time 1 to time 2 but some changes in class membership for some individuals were seen from time 2 to time 3. Individuals who were in the Class 1 at time 1 had a 0.99 probability of remaining there at time 2 (0.88 for Class 2 and 0.98 for Class 3, respectively). The probability (0.04) that individuals would move from the Class 2 to Class 1 by time 2 was not statistically significantly. Similarly, the probability (0.07) that individuals would move from the Class 2 to the Class 1 by time 2 was not significantly different from zero. The probability (0.02) that individuals would move from the Class 3 to the Class 2 by time 2 was not statistically significant either. There was, however, a 0.10 probability that individuals would transition from Class 2 to Class

TABLE 2 | Fit statistics for model specifications at time point 1, time point 2, and time point 3.

Model	Loglikelihood	#Free parameters	AIC	BIC	SBIC	*p* LMR	Entropy
Time point 1							
Two-Class	−12514.64	44	25117.28	25366.95	25227.15	<0.001	0.88
Three-Class	−11010.52	60	22141.04	22481.49	22290.86	0.003	0.89
Four-Class	−10133.10	76	20418.21	20849.44	20607.98	0.24	0.89
Five-Class	−9766.25	92	19716.50	20238.52	19946.22	0.189	0.87
Six-Class	−9492.06	108	19200.12	19812.93	19469.80	0.131	0.83
Seven-Class	−9279.09	124	18806.18	19509.77	19115.81	0.62	0.80
Time point 2							
Two-Class	−7936.22	44	15960.44	16201.10	16061.32	<0.001	0.81
Three-Class	−6927.78	60	13975.57	14303.75	14113.13	<0.001	0.82
Four-Class	−6513.27	76	13178.53	13594.23	13352.78	0.006	0.83
Five-Class	−6256.23	92	12696.45	13199.66	12907.38	0.004	0.75
Six-Class	−6087.00	108	12390.00	12980.73	12637.62	0.031	0.77
Time point 3							
Two-Class	−7936.56	44	15961.11	16201.12	16061.33	<0.001	0.79
Three-Class	−6984.23	60	14088.45	14415.74	14225.12	0.004	0.81
Four-Class	−6549.12	76	13250.23	13664.79	13423.35	0.514	0.78
Five-Class	−6374.21	92	12932.43	13434.26	13141.99	0.353	0.71

#, number; AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; SBIC, Sample Size Adjusted BIC; p LMR, p- values for the Lo-Mendell-Rubin Likelihood ratio test for k vs. k+1Classs. LCA models converged at Time 1. The log-likelihood increased while no minimum was found for the ICs as their values decreased across the range of models considered.

1 from time 2 to time 3. Multivariate analysis of variance showed that individuals who moved from Class 2 to Class 1 from time 2 to time 3 had significantly higher means than other individuals of Class 2 on FDG, MMSE, RAVimD, MOCA, EcPtLg, Hipc, Entor, and Fusif. But these groups of individuals had significantly lower on CDRSB, ADAS11, EcPtMm, EcSPM. No significant mean difference was found between the two groups on EcSPLang. Finally, the small number of misclassified individuals prevented us from computing the transition probabilities since LTA rely on larger sample sizes to be trustworthy. However, descriptive data show that at time 2, 42 (66.7%) of the 63 individuals that were classified in Class 1 from the LCA, but AD individuals by the diagnosis variable, stayed in the Class 1 at time 2, 18 (28.60%) moved to the Class 2, and 3 (4.80%) to the AD class. Similar figures were also found at time 3.

DISCUSSION

The aim of this study was to use LCA to identify and describe the number of classes that best characterize CN, SMC, EMCI, LMCI, and AD individuals with respect to clinical and neural measures. Our second aim was to compare the classification results obtained from the LCA to more traditional cut-off methods for classifying individuals with dementia. This can help us identify and characterize potentially misdiagnosed individuals. Finally, we used LTA to investigate changes in class membership over time. Our results showed that while there was substantial variability among individuals on their clinical and neural measures, the use of LCA with mixture methods to assess grouping individuals into optimal classes yields meaningful results. We confirm that using LCA, observing model fit indices, and entropy criterion were effective for selecting the optimal number of classes of individuals. Our results identified three classes of individuals with the following characterization: 37.5% of the individuals from Class 1 were CN, 20.10% with SMC, 19.60% with EMCI, 18.10% with LMCI, and 4.6% (63 individuals) with AD. Similar figures were 0.5, 0.9, 12.90, 53, and 32.60%; and 0.7, 0, 3.5, 28, and 67.80%, for Class 2 and Class 3, respectively. Further, our results showed that the LCA identified 63 individuals that were potentially misdiagnosed with AD. Indeed, based on the clinical and neural measures, it was deemed more probable that the misdiagnosed individuals be classified within the Class 1 instead of Class 2 or 3. Further, LTA did not show any significant change in class over time. Most individuals remained within their initial class (i.e., determined at baseline) and did not show a transition from Class 1 to Class 2, or from Class 2 to Class 3 between any time points. These results indicate that classifying individuals based on their cognitive and pathological parameters into different categories is an essential step toward understanding dementia and AD. To our knowledge, this is the first study to successfully use neuropsychological assessments and biomarkers of AD (e.g., Fluorodeoxyglucose, entorhinal cortex volume, and fusiform gyrus volume) to classify and predict individuals likely to transition from MCI to AD.

Similar to previous results, we identified multiple classes of cognitive impairment (Scheltens et al., 2016; Zammit et al., 2019b, 2020). For example, the results from our LCA identified 3 classes of individuals at each time point: Class 1, which is more healthy than the other classes representing 63, 59, and 65% of the sample at, respectively, time 1, 2, and time 3; Class

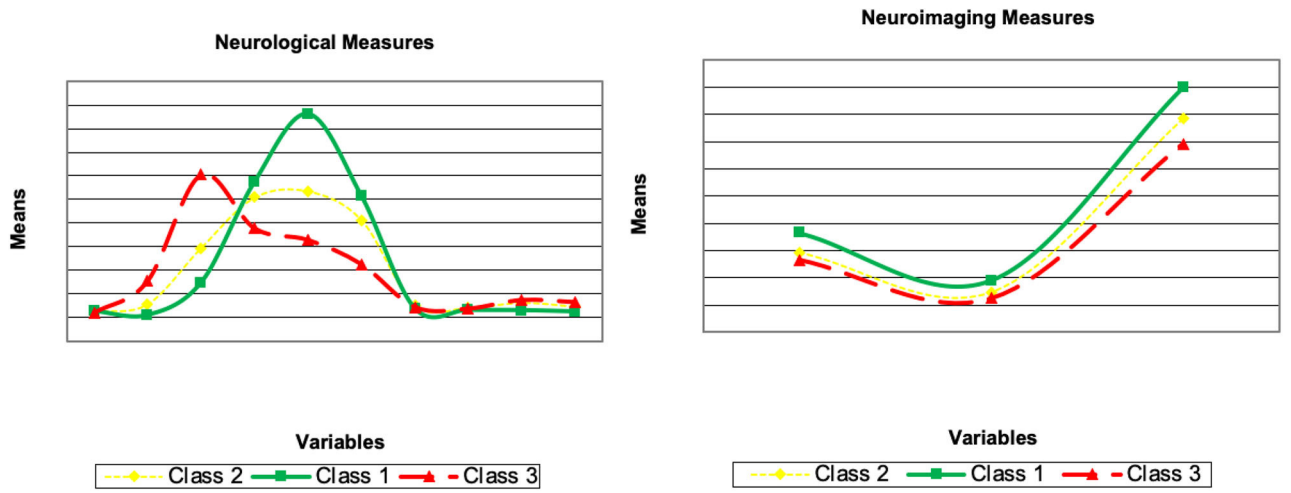

FIGURE 1 | Three-class solutions for neurological and neuroimaging measures at time 1.

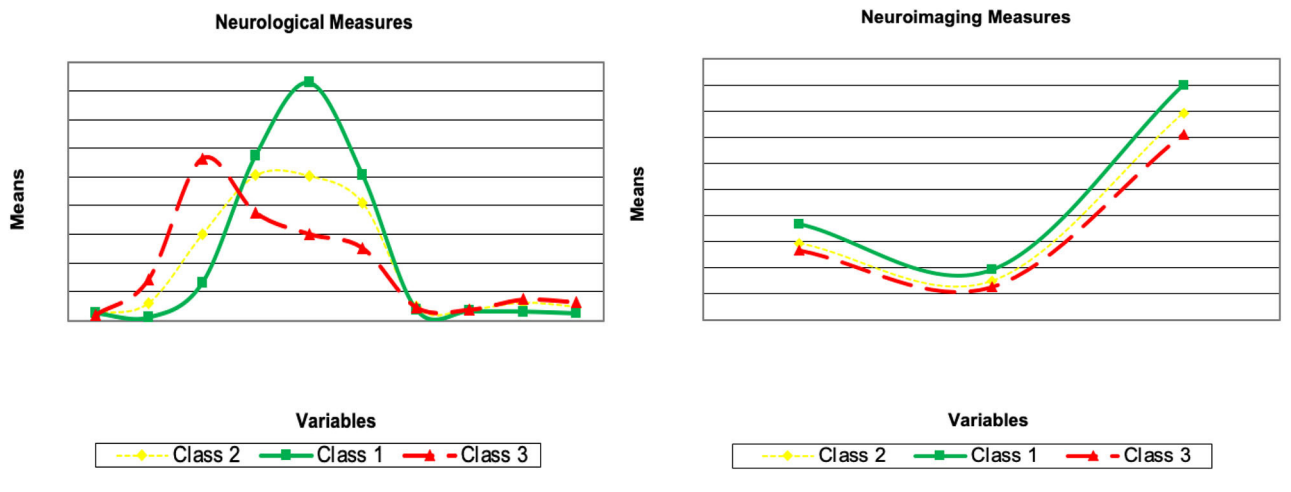

FIGURE 2 | Three-class solutions for neurological and neuroimaging measures at time 2.

2, which lies in between Class 1 and 3, representing ∼30% of the sample across the three time points; and Class 3, which include the least healthy individuals, representing 7, 10, and 6% of the sample at time 1, 2, and time 3, respectively. In other words, our LCA results reveal 3 classes that most likely match healthy individuals, individuals with MCI, and individual with AD, respectively. In comparison, Zammit et al. (2019b) classified participants into 5 classes within two categories (i.e., impaired and intact cognition) and Scheltens et al. (2016) classified participants into eight cognitive subtypes of AD. While there are differences in the number of classes identified with our results compared to others (Scheltens et al., 2016; Zammit et al., 2019b), it is important to note that the previous studies only utilized participants diagnosed with probable AD using neuropsychological assessments. Our research extends these findings by indicating that there are distinct classes of individuals who can be categorized as being healthy, experiencing MCI, and probable AD by using neuropsychological assessments and neurological biomarkers of AD. This has important clinical implications as individuals can be classified as experiencing different kinds of cognitive impairment early on (i.e., at baseline) and this categorization does not change significantly across time. Consistent with recent findings (i.e., Zammit et al., 2020)., we also showed that using LCA to classify individuals with MCI or AD remains relatively stable over time (as indicated by LTA) and that LCA might better categorize and reduce the risk of misdiagnosis.

Our study has replicated previous findings that LCA and LTA can be used to identify homogeneous classes of cognitive impairment (Zammit et al., 2020). However, we have uniquely identified that neuropsychological measures of AD and the associated neurological biomarkers are indicators of an individual's class membership and can predict their likelihood to transition between the healthy class (Class 1),

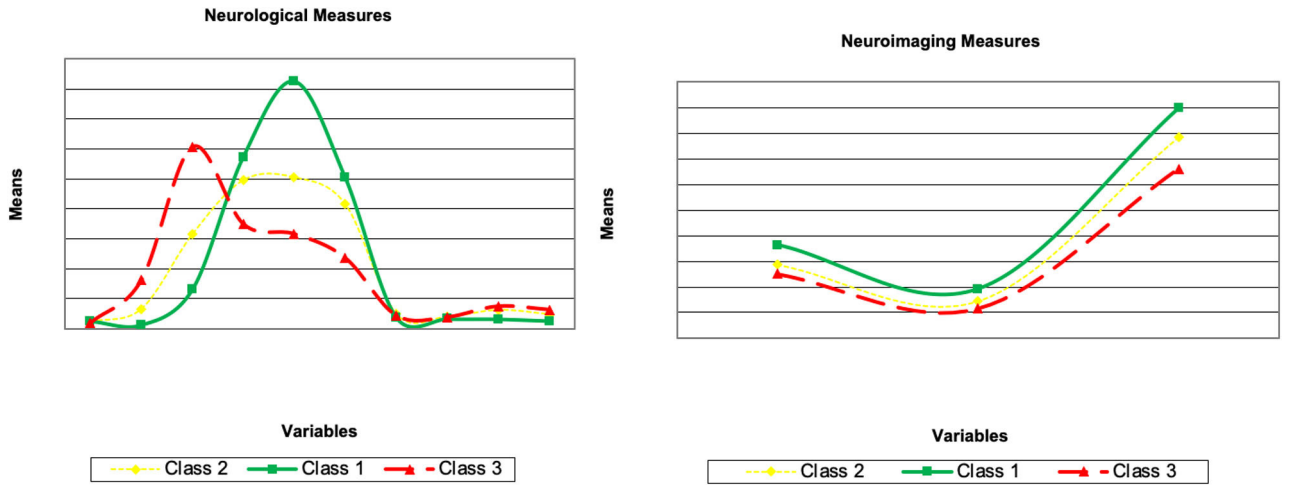

FIGURE 3 | Three-class solutions for neurological and neuroimaging measures at time 3.

TABLE 3 | Cross tabulation of diagnostic variable and latent class variable at Time 1.

Diagnostic variable at time 1		Latent class solution at Time 1 Class 1	Class 2	Class 3	
CN	Frequency	508	3	1	512
	Row %	99.2%	0.6%	0.2%	
	Column %	37.5%	0.5%	0.7%	
SMC	Frequency	273	6	0	279
	Row %	97.8%	2.2%	0.0%	
	Column %	20.1%	0.9%	0.0%	
EMCI	Frequency	266	82	5	353
	Row %	75.4%	23.2%	1.4%	
	Column %	19.6%	12.9%	3.5%	
LMCI	Frequency	245	336	40	621
	Row %	39.5%	54.1%	6.4%	
	Column %	18.1%	53.0%	28.0%	
AD	Frequency	**63**	207	97	367
	Row %	17.2%	56.4%	26.4%	
	Column %	4.6%	32.6%	67.8%	
		1,355	634	143	

AD, Alzheimer's disease; CN, controls; EMCI, early-stage mild cognitive impairment; LMCI, late-stage mild cognitive impairment; SMC, subjective memory complains.

the MCI class (Class 2), and the AD class (Class 3). As mentioned earlier, recent studies using LTA have only utilize neuropsychological assessments of AD or MCI (e.g., Scheltens et al., 2016; Eppig et al., 2017; Zammit et al., 2019b, 2020). Our study adds a novel contribution to this emerging area by identifying that neurological biomarkers of AD can also be used to correctly classify individuals with MCI and AD and identify those at risk of transitioning from healthy cognitive function, to mild-cognitive impairment, and finally to AD.

A comparison of the classification results obtained from the LCA with the cutoff methods at Time 1 (baseline) revealed a group of 63 misclassified individuals. This group of individuals were classified as healthy individuals by the LCA, but were classified as AD by using clinical cut-off scores in neuropsychological assessments. The multivariate analysis of variance revealed that the misclassified individuals' scores on the clinical and neuropsychological assessments and neurological biomarkers were bounded between the mean of the Healthy individuals (i.e., Class 1) from the LCA without the 63 individuals and those from the AD individuals (i.e., Class 3) from the cut-off method without the 63 individuals. However, further analysis showed that it was more probable that the misclassified participants belonged to the healthy class (i.e., Class 1) rather than the MCI class (Class 2). As LCA takes into account several clinical and neural variables, longitudinal data, as well as also considers different groups of participants, it is likely to be more accurate than standard clinical cut-off methods, which often relies on one measure and does not compare data across different groups of participants.

There is a large discrepancy between the two methods of classification (i.e., clinical assessment vs. statistical), and perhaps reinforces the criticisms of using cut-off points on continuous neuropsychological assessments. Our results show that participants either side of the cut-off are similar (Berlin et al., 2014; Petersen et al., 2019). That is, based on clinical cut-off scores, some healthy participants were identified as similar to AD individuals, which resulted in misclassification. By using LCA, we have identified homogeneous sub-groups of individuals who are externally heterogeneous to other sub-groups (Berlin et al., 2014; Eppig et al., 2017; Mooney et al., 2018; Petersen et al., 2019; Villeneuve et al., 2019; Zammit et al., 2019a). For example, the healthy class (i.e., Class 1) is externally heterogeneous compared to the MCI class (Class 2) and the AD class (i.e.,

TABLE 4 | Analysis of variance results with mean and standard deviations.

	The 63 patients (MD)	Diagnostics minus the 63 (DG)	Healthy minus the 63 (HT)	*d*	Contrasts
Measures					
FDG *M (SD)*	1.26 (0.13)	1.06 (0.15)	1.29 (0.12)	0.29***	MD = HT > DG
CDRSB*M (SD)*	1.69 (2.05)	4.87 (2.45)	0.52 (0.93)	0.59***	DG > MD > HT
ADAS11*M (SD)*	1.14 (0.61)	2.16 (0.90)	0.75 (0.39)	0.52***	DG > MD > HT
MMSE*M (SD)*	2.70 (0.28)	2.24 (0.34)	2.87 (0.15)	0.52***	HT > MD > DG
RAVimD*M (SD)*	3.58 (1.26)	2.17 (0.98)	4.30 (1.13)	0.36***	HT > MD > DG
MOCA*M (SD)*	2.30 (0.43)	1.70 (0.47)	2.54 (0.30)	0.48***	HT > MD > DG
EcPtMm*M (SD)*	2.09 (0.74)	2.27 (0.70)	1.86 (0.66)	0.05***	DG > MD > HT
EcPtLg*M (SD)*	1.77 (0.66)	1.84 (0.66)	1.60 (0.57)	0.02***	DG > MD > HT
EcSPM*M (SD)*	2.08 (0.83)	2.99 (0.76)	1.59 (0.61)	0.41***	DG > MD > HT
EcSPLg*M (SD)*	1.70 (0.69)	2.36 (0.79)	1.31 (0.46)	0.36***	DG > MD > HT
Hipc*M (SD)*	0.70 (0.12)	0.56 (0.11)	0.72 (0.10)	0.28***	MD = HT > DG
Entor*M (SD)*	0.34 (0.07)	0.27 (0.07)	0.37 (0.07)	0.26***	HT > MD > DG
Fusif *M (SD)*	1.71 (0.25)	1.49 (0.28)	1.78 (0.26)	0.16***	HT > MD > DG

****p < 0.001. All variables are defined in the text and in the captions of prior tables.*

TABLE 5 | Transition probabilities from Time 1 to Time 2 and from Time 2 to Time 3.

		T2					T3		
		Class 1	**Class 2**	**Class 3**			**Class 1**	**Class 2**	**Class 3**
T1	Class 1	0.99a*	0.01	0.00	**T2**	Class 1	0.93*	0.07	0.00
	Class 2	0.04[b]	0.88*	0.07		Class 2	0.10*	0.85*	0.05
	Class 3	0.00	0.02	0.98*		Class 3	0.00	0.09	0.91*

[a]*Read as 99% of patients who were in class 1 at time 1 were predicted to remain in class 1 at time 2.* [b]*4% of patients who were in class 2 at time 1 were predicted to transition to class 1 at time 2.*

**probabilities significantly different from 0 at p < 0.05.*

Class 3). Conversely, the AD class (i.e., Class 3) is relatively more homogenous (i.e., after identifying misdiagnosed individuals) and externally heterogeneous compared to the healthy class and MCI class.

The LTA results were used to examine the individuals' transition probabilities with respect to clinical assessments, neurological measures, and neurological biomarkers to explore changes in the latent classes over time. The LTA results showed that individuals in the Healthy (i.e., Class 1) and AD (i.e., Class 3) classes were stable (with a probability >0.90 of staying in the same class over time). The results show that the transition from Healthy to AD classes was non-existent. This is consistent with the nature of the AD as a progressive disease. Individuals at early stages do not exhibit symptoms of AD. However, there is an insignificant probability for the individuals in the healthy class to transition to the MCI class (0.01 from time 1 to time 2 and 0.07 from time 2 to time 3). Furthermore, the results confirm the nature of AD as a neurodegenerative disease (Alzheimer Association, 2019). For example, at no time-point did individuals who are in the AD class show cognitive improvement by transitioning to the Healthy class or MCI class.

In contrast, individuals from the MCI class have non-zero probabilities moving to other classes over time. However, only the probability of transitioning to the Healthy class from time 2 to time 3 was significant (with a probability of 0.10). That is, some individuals classified within the MCI class (Class 2) showed cognitive improvement from time 2 to time 3. This further emphasizes the differences between MCI and AD. While AD is a regressive disease that does not allow cognitive improvements, MCI is not necessarily degenerative. As such, we should be cautious about suggesting individuals with MCI are on a progression toward AD. As we only observed movement from the MCI class into the healthy class, it is less probable that individuals with MCI will progress into AD. Therefore, based on our findings, any transition from the MCI class is likely to resemble cognitive improvement rather than decline. It is worth noting that we did not identify the characteristics that predict the movement of individuals from the MCI class. Further research is needed to investigate possible factors that may contribute to this movement, either to the Healthy or AD class.

LIMITATIONS

It is important to note that one main advantage of traditional rule-based diagnostic methods (as often used by most clinicians and doctors) is easy utilization in the everyday clinical setting. However, our methods used here are more complex and

require applying analytical and statistical method to be able to reach a more robust diagnosis. Accordingly, because of its complexity (e.g., conducting the analysis and interpretations of results), it is expected that latent class analysis methods may not be widely used. However, we also agree that it is exactly LCA complexity over the traditional discrete diagnostic methods (e.g., surveys) that allow it to be a better predictor of class membership (e.g., an individual is healthy, has mild cognitive impairment, or has Alzheimer's disease). This is due to the fact that rule-based diagnostic methods are inherently additive, as they rely on discrete methods. However, LCA is a multivariate approach that attempts to find complex joint probability distributions that create a richer risk profile which is difficult to define using discrete rule-based decision tools and diagnostic methods.

CONCLUSION

In conclusion, this study demonstrated that latent class analysis can be used to classify participants within the ADNI project into three distinct classes: Healthy, MCI, or AD. We argue that LCA is a more suitable method for classifying individuals with SMC, MCI, and AD rather than using clinical cut-off measures. This is due to LCA's ability to create internally homogenous and externally heterogeneous sub-groups. This technique might help reduce the number of misclassifications of individuals incorrectly diagnosed with probable AD—as demonstrated by the misclassified individuals in our study. By using latent transition analysis, we showed that individuals classified as healthy or with AD had a high probability of staying in the same class over time. However, it was more probable for individuals to transition from the MCI class to the healthy class. Our results emphasize that AD is a neurodegenerative syndrome, with individuals within the AD class showing no evidence of cognitive improvement over time. However, individuals with MCI can show improvement over time. Therefore, we argue that LCA can be used to differentiate between individuals with AD and that this diagnosis remains stable across time and produces fewer misdiagnoses than using clinical cut-offs. Robust methods should be used to accurately diagnose patients and to identify individuals at a highest risk of developing AD. While using cut-off scores using traditional discrete diagnostic methods are quicker, our study has shown that LCA can provide a more accurate prediction for classifying individuals with SMC, MCI, and AD. While the time requirement to conduct LCA is burdensome, ensuring an accurate diagnosis for patients should be a prioritized. Especially given the severity and neurodegenerative nature of AD (Alzheimer Association, 2019). Using LCA and LTA can provide more accurate diagnoses and improve the outcomes for patients. Clinicians should consider alternative diagnostic methods for AD instead of relying solely on the clinical cut-off measures on neuropsychological assessments.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: http://adni.loni.usc.edu.

AUTHOR CONTRIBUTIONS

All authors contributed to data analysis, writing, and editing the manuscript.

FUNDING

This work received financial support from the United Arab Emirates University (grant no. CIT 31T129).

ACKNOWLEDGMENTS

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI was funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

REFERENCES

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika* 52, 317–332. doi: 10.1007/BF02294359

Alashwal, H., El Halaby, M., Crouse, J. J., Abdalla, A., and Moustafa, A. A. (2019). The application of unsupervised clustering methods to Alzheimer's disease. *Front. Comput. Neurosci.* 13:31. doi: 10.3389/fncom.2019.00031

Alzheimer Association (2019). *Early Signs and Symptoms of Alzheimer's. Alzheimer's and Dementia*, 1–88. Retrieved from https://www.alz.org/alzheimers-dementia/10_signs

Berlin, K. S., Williams, N. A., and Parra, G. R. (2014). An introduction to latent variable mixture modeling (Part 1): overview and cross-sectional latent class and latent profile analyses. *J. Pediatr. Psychol.* 39, 174–187. doi: 10.1093/jpepsy/jst084

Clogg, C. C. (1981). "New developments in latent structure analysis," in *Factor Analysis and Measurement in Sociological Research*. eds D. J. Jackson and E. F. Borgatta (London: Sage), 77–108.

Diallo, T. M. O., Morin, A. J. S., and Lu, H. (2016a). Impact of misspecifications of the latent variance–covariance and residual matrices on the class enumeration accuracy of growth mixture models. *Struct. Equat. Model.* 23, 507–531. doi: 10.1080/10705511.2016.1169188

Diallo, T. M. O., Morin, A. J. S., and Lu, H. (2016b). Performance of growth mixture models in the presence of time-varying covariates. *Behav. Res. Methods* 49, 1951–1965. doi: 10.3758/s13428-016-0823-0

Diallo, T. M. O., Morin, A. J. S., and Lu, H. (2017). The impact of total and partial inclusion or exclusion of active and inactive time invariant covariates in growth mixture models. *Psychol. Methods* 22, 166–190. doi: 10.1037/met0000084

Duara, R., Loewenstein, D. A., Shen, Q., Barker, W., Varon, D., Greig, M. T., et al. (2013). The utility of age-specific cut-offs for visual rating of medial temporal atrophy in classifying Alzheimer's disease, MCI and cognitively normal elderly subjects. *Front. Aging Neurosci.* 5:47. doi: 10.3389/fnagi.2013.00047

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York, NY: Guilford Press.

Eppig, J. S., Edmonds, E. C., Campbell, L., Sanderson-Cimino, M., Delano-Wood, L., and Bondi, M. W. (2017). Statistically derived subtypes and associations with cerebrospinal fluid and genetic biomarkers in mild cognitive impairment: a latent profile analysis. *J. Int. Neuropsychol. Soc.* 23, 564–576. doi: 10.1017/S135561771700039X

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Frank, R. A., Galasko, D., Hampel, H., Hardy, J., de Leon, M. J., Mehta, P. D., et al. (2003). Biological markers for therapeutic trials in Alzheimer's disease—proceedings of the biological measures working group: NIA Initiative on Neuro-imaging in Alzheimer's disease. *Neurobiol. Aging* 24, 521–536. doi: 10.1016/S0197-4580(03)00002-2

Gomar, J. J., Conejero-Goldberg, C., Huey, E. D., Davies, P., Goldberg, T. E., and Alzheimer's Disease Neuroimaging Initiative. (2016). Lack of neural compensatory mechanisms of BDNF val66met met carriers and APOE E4 carriers in healthy aging, mild cognitive impairment, and Alzheimer's disease. *Neurobiol. Aging* 39, 165–173. doi: 10.1016/j.neurobiolaging.2015.12.004

Hipp, J. R., and Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychol. Methods* 11, 36–53. doi: 10.1037/1082-989X.11.1.36

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

Kaplan, D. (2008). An overview of markov chain methods for study of stage-sequential developmental processes. *Dev. Psychol.* 44, 457–467. doi: 10.1037/0012–1649.44.2.457

Lazarsfeld, P. F., and Henry, N. W. (1968). *Latent Structure Analysis*. Boston, MA: Houghton Mill.

Lo, Y., Mendell, N., and Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika* 88, 767–778. doi: 10.1093/biomet/88.3.767

Marshall, G.A,. Zoller, A. S., Kelly, K. E., Amariglio, R. E., Locascio, J. J., Johnson, K. A., et al. (2014). Everyday cognition scale items that best discriminate between and predict progression from clinically normal to mild cognitive impairment. *Curr. Alzheimer Res.* 11, 853–861. doi: 10.2174/1567205011666141001120903

Mohs, R. C., and Cohen, L. (1988). Alzheimer's disease assessment scale (ADAS). *Psychopharmacol. Bull.* 24, 627–628.

Mooney, S. J., Joshi, S., Cerdá, M., Kennedy, G. J., Beard, J. R., and Rundle, A. G. (2018). Longitudinal patterns of physical activity among older adults: a latent transition analysis. *Am. J. Epidemiol.* 187, 1549–1558. doi: 10.1093/aje/kwy027

Morris, J. C. (1997). Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *Int. Psychogeriatr.* 9, 173–176. doi: 10.1017/S1041610297004870

Moustafa, A. A., Diallo, T., Amoroso, N., Zaki, N., Hassan, M., and Alashwal, H. (2018). Applying big data methods to understanding human behavior and health. *Front. Comput. Neurosci.* 12:84. doi: 10.3389/fncom.2018.00084

Muthen, B. (2004). "Latent variable analysis: growth mixture modeling and related techniques for longitudinal data," in *Handbook of Quantitative Methodology for the Social Sciences,* ed D. Kaplan. (Thousand Oaks, CA: Sage), 345–368.

Muthén, B. O., and Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol. Clin. Exp. Res.* 24, 882–891. doi: 10.1111/j.1530–0277.2000.tb02070.x

Muthén, L. K., and Muthén, B. O. (2019). *Mplus User's Guide*. New York, NY: Author.

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

Nylund, K. L., Asparouhov, T., and Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a monte carlo simulation study. *Struc. Equat. Model.* 14, 535–569. doi: 10.1080/10705510701575396

O'Bryant, S. E., Waring, S. C., Cullum, C. M., Hall, J., Lacritz, L., Massman, P. J., et al. (2008). Staging dementia using clinical dementia rating scale sum of boxes scores: a texas Alzheimer's research consortium study. *Arch. Neurol.* 65, 1091–1095. doi: 10.1001/archneur.65.8.1091

Perrin, R. J., Fagan, A. M., and Holtzman, D. M. (2009). Multimodal techniques for diagnosis and prognosis of Alzheimer's disease. *Nature* 461, 916–922. doi: 10.1038/nature08538

Petersen, K. J., Qualter, P., and Humphrey, N. (2019). The application of latent class analysis for investigating population child mental health: a systematic review. *Front. Psychol.* 10:1214. doi: 10.3389/fpsyg.2019.01214

Pfeffer, R. I., Kurosaki, T. T., Harrah C. H. Jr, Chance, J. M., and Filos, S. (1982). Measurement of functional activities in older adults in the community. *J. Gerontol.* 37, 323–329. doi: 10.1093/geronj/37.3.323

Rattanabannakit, C., Risacher, S. L., Gao, S., Lane, K. A., Brown, S. A., McDonald, B. C., et al. (2016). The cognitive change index as a measure of self and informant perception of cognitive decline: relation to neuropsychological tests. *J. Alzheimers Dis.* 51, 1145–1155. doi: 10.3233/JAD-150729

Ronald, T., and National Institute on Aging Working Group (1998). Consensus report of the working group on:"Molecular and biochemical markers of Alzheimer's disease". *Neurobiol. Aging* 19, 109–116. doi: 10.1016/S0197-4580(98)00022-0

Saykin, A. J., Wishart, H. A., Rabin, L. A., Santulli, R. B., Flashman, L. A., West, J. D., et al. (2006). Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. *Neurology* 67, 834–842. doi: 10.1212/01.wnl.0000234032.77541.a2

Scheltens, N. M. E., Galindo-Garre, F., Pijnenburg, Y. A. L., Van Der Vlies, A. E., Smits, L. L., Koene, T., et al. (2016). The identification of cognitive subtypes in Alzheimer's disease dementia using latent class analysis. *J. Neurol. Neurosurg. Psychiatry* 87, 235–243. doi: 10.1136/jnnp-2014-309582

Schwartz, G. (1978). Estimating the dimensions of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136

Sclove, L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika* 52, 333–343. doi: 10.1007/BF02294360

Shaw, L. M., Vanderstichele, H., Knapik-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., et al. (2009). Cerebrospinal fluid biomarker signature in alzheimer's disease neuroimaging initiative subjects. *Ann. Neurol.* 65, 403–413. doi: 10.1002/ana.21610

Veitch, D. P., Weiner, M. W., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., et al. (2019). Understanding disease progression and improving Alzheimer's disease clinical trials: recent highlights from the Alzheimer's disease neuroimaging initiative. *Alzheimers Dement.* 15, 106–152. doi: 10.1016/j.jalz.2018.08.005

Vermunt, J. K., and Magidson, J. (2002). "Latent class cluster analysis," in *Applied Latent Class Analysis,* eds J. A. Hagenaars and A. L. McCutcheon (Cambridge: Cambridge University Press), 89–106. doi: 10.1017/CBO9780511499531.004

Villeneuve, S. C., Houot, M., Cacciamani, F., Verrijp, M., Dubois, B., Sikkes, S., et al. (2019). Latent class analysis identifies functional decline with Amsterdam IADL in preclinical Alzheimer's disease. *Alzheimers Dement.* 5, 553–562. doi: 10.1016/j.trci.2019.08.009

World Health Organization (2012). *Dementia: A Public Health Priority*. Dementia.

Zammit, A. R., Bennett, D. A., Hall, C. B., Lipton, R. B., Katz, M. J., and Muniz-Terrera, G. (2020). A Latent transition analysis model to assess change in cognitive states over three occasions: results from the rush memory and aging project. *J. Alzheimers Dis.* 73, 1063–1073. doi: 10.3233/JAD-190778

Zammit, A. R., Hall, C. B., Bennett, D. A., Ezzati, A., Katz, M. J., Muniz-terrera, G., et al. (2019a). Neuropsychological latent classes at enrollment and postmortem neuropathology. *Alzheimers Dement.* 15, 1195–1207. doi: 10.1016/j.jalz.2019.05.012

Zammit, A. R., Terrera, G. M., Hall, C. B., Katz, M. J., Bennett, D. A., Ezzati, A., et al. (2019b). A latent transition analysis model to assess change in cognitive states and predict incident dementia: results from the rush memory and aging project. *Alzheimers Dement.* 15:P445. doi: 10.1016/j.jalz.2019.06.1067

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 Public Health

ORIGINAL RESEARCH
published: 20 November 2020
doi: 10.3389/fpubh.2020.599550

Analysis of COVID-19 Infections on a CT Image Using DeepSense Model

Adil Khadidos[1], *Alaa O. Khadidos*[2], *Srihari Kannan*[3]*, *Yuvaraj Natarajan*[4], *Sachi Nandan Mohanty*[5] *and Georgios Tsaramirsis*[6]

[1] Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia, [2] Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia, [3] Department of Computer Science and Engineering, SNS College of Engineering, Coimbatore, India, [4] Research and Development, Information Communication Technology Academy, Chennai, India, [5] Department of Computer Science and Engineering, Institute of Chartered Financial Analysts of India Foundation of Higher Education, Hyderabad, India, [6] Higher Colleges of Technology, Women's College, Abu Dhabi, United Arab Emirates

In this paper, a data mining model on a hybrid deep learning framework is designed to diagnose the medical conditions of patients infected with the coronavirus disease 2019 (COVID-19) virus. The hybrid deep learning model is designed as a combination of convolutional neural network (CNN) and recurrent neural network (RNN) and named as DeepSense method. It is designed as a series of layers to extract and classify the related features of COVID-19 infections from the lungs. The computerized tomography image is used as an input data, and hence, the classifier is designed to ease the process of classification on learning the multidimensional input data using the Expert Hidden layers. The validation of the model is conducted against the medical image datasets to predict the infections using deep learning classifiers. The results show that the DeepSense classifier offers accuracy in an improved manner than the conventional deep and machine learning classifiers. The proposed method is validated against three different datasets, where the training data are compared with 70%, 80%, and 90% training data. It specifically provides the quality of the diagnostic method adopted for the prediction of COVID-19 infections in a patient.

Keywords: DeepSense, artificial intelligence, convolutional neural network, CT images, prediction, COVID-19

OPEN ACCESS

Edited by:
Deepak Gupta,
Maharaja Agrasen Institute of Technology, India

Reviewed by:
Abdulsattar Abdullah Hamad,
University of Tikrit, Iraq
George Halikias,
City University of London, United Kingdom
Mohammad Yamin,
King Abdulaziz University, Saudi Arabia

***Correspondence:**
Srihari Kannan
harionto@gmail.com

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Public Health

Received: *27 August 2020*
Accepted: *16 October 2020*
Published: *20 November 2020*

Citation:
Khadidos A, Khadidos AO, Kannan S, Natarajan Y, Mohanty SN and Tsaramirsis G (2020) Analysis of COVID-19 Infections on a CT Image Using DeepSense Model. Front. Public Health 8:599550. doi: 10.3389/fpubh.2020.599550

INTRODUCTION

The novel coronavirus disease 2019 (COVID-19) is a pandemic outbreak (1). COVID-19 patients are classified essentially based on computerized tomography (CT) lung images, and it is used widely for testing. The healthcare institutions fitted with CT scans help in the process of image acquisition and classification of CT images at a faster rate. However, the need for an expert medical practitioner is hence required for the verification of the final results, which increases the time of computation (2). On the other hand, the supervised learning models (3–10) can be utilized for classifying the patients from the CT images.

Infections based on CT images are not classified using very little unattended methods (11–24). We have developed a model that mainly includes supervised and unsupervised learning models in order to improve the classification process. The aim is to classify the infected patients automatically based on their CT images.

In this paper, a DeepSense algorithm is utilized to diagnose COVID-19 infections among the medical community. The deep learning method is designed as a combination of convolutional neural network (CNN) and recurrent neural network (RNN) that reduces the classifier burden on optimal classification of the multidimensional data features.

The main contribution of the work includes the following:

(a) The authors develop a combined CNN and RNN to classify the medical image datasets.
(b) The experimental results are conducted to measure the correctness in terms of its accuracy, precision, and recall values against artificial neural network (ANN), feedforward neural network (FFNN), back propagation neural network (BPNN), deep neural network (DNN), and RNN.

The outline of the paper is presented as follows: The *Methods* section provides the details of the ensemble classifier. The *DeepSense Model* section evaluates the entire work. The *Results and Discussions* section concludes the work with future enhancement.

METHODS

The deep learning model namely DeepSense algorithm is a combination of CNN and RNN designed to improve the performance of the classification accuracy. DeepSense learning is regarded as a module for accurate predictions of lung infections caused by the COVID-19 virus. **Figure 1** shows the architecture of the proposed classification model using the DeepSense algorithm.

DEEPSENSE MODEL

Figure 2 shows the DeepSense DNN (25) model that has three components, including convolutional, recurrent, and output

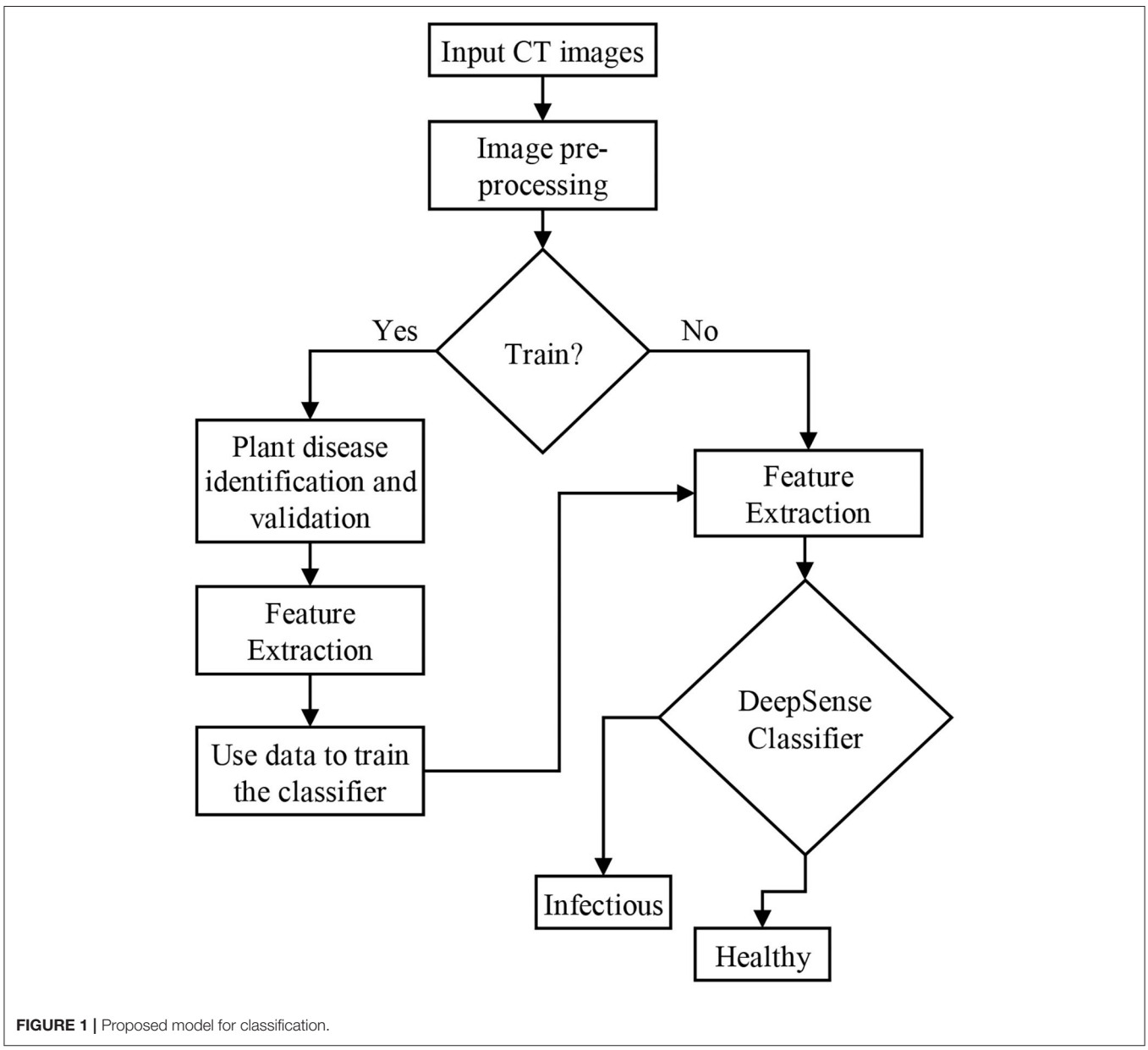

FIGURE 1 | Proposed model for classification.

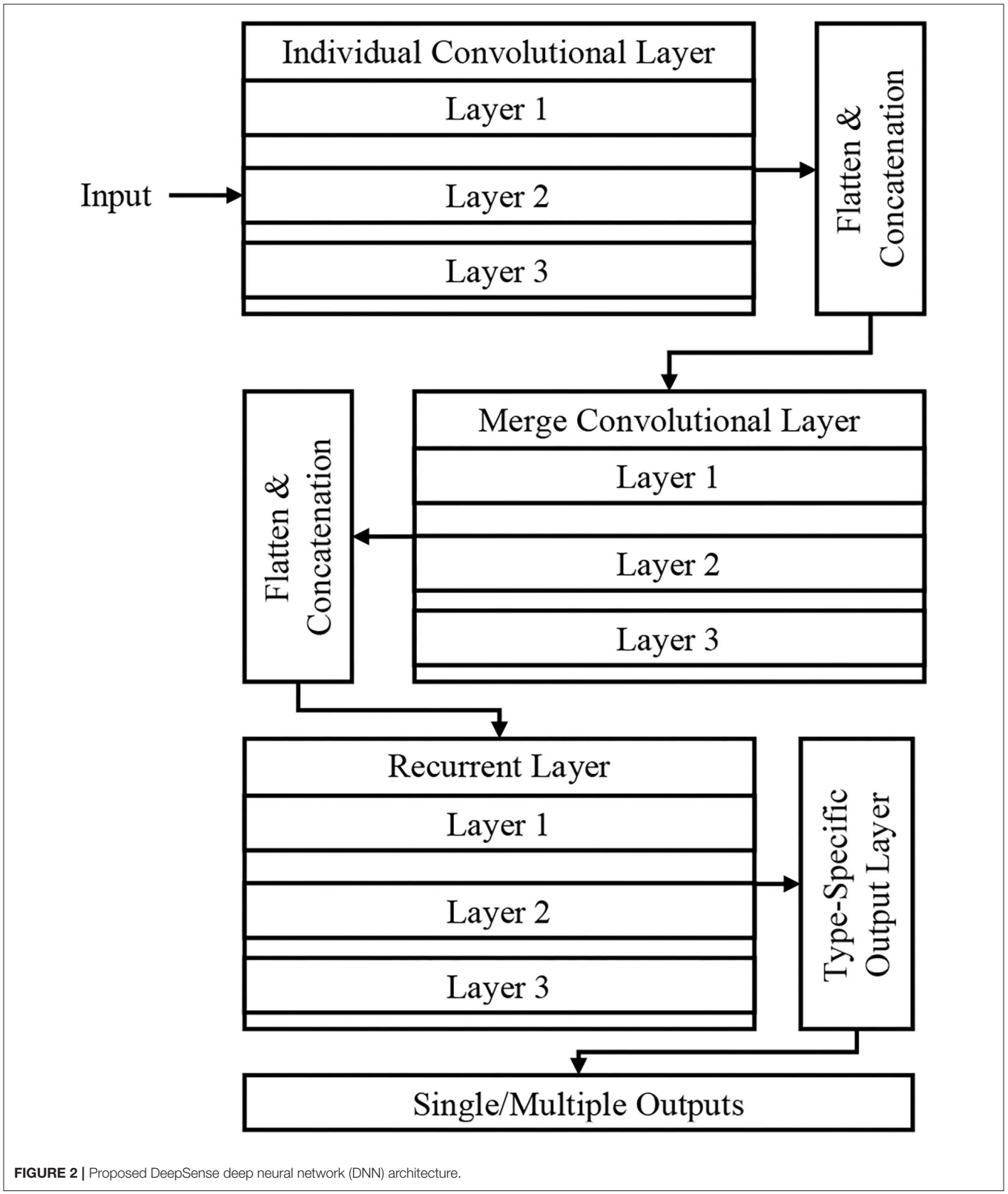

FIGURE 2 | Proposed DeepSense deep neural network (DNN) architecture.

layer that are stacked upon one another. The convolutional and recurrent layers are regarded as the significant building blocks (**Figure 1**), and the output layer is considered as a specific layer that classifies the images. The DeepSense DNN model is designed for the classification of input CT images for COVID-19-related infections.

The DeepSense network avoids gradient exploding and improves the rate of convergence using residual learning, adjustable learning rate, and gradient clipping that helps in optimizing the process of training.

The features are extracted through the DeepSense model that increases the reconstruction accuracy and reduces the time of training. Such optimization helps in obtaining the rich text information, and it has better ability for classification.

Convolutional Layers

The convolutional layers have three different parts that include individual convolutional subnets for input from CT device $X(k)$, where k is the number of CT device. The other subnets include a merged convolutional subnet for K convolutional subnets' outputs.

For a time interval t, the matrix $X(k)$ is used as an input to the DNN architecture that extracts the relationship of $X(k,t)$, which includes the relationships lying inside the frequency domain. The sensor measurement interactions include entire dimension, where the frequency domain usually has several local patterns. These interactions are studied using 2D filters and produces the output $X(k,1,t)$ based on the local patterns and dimensions in frequency domain. The high-level relationships are learned hierarchically using the application of a 1D filter. The matrix is then flattened into a vector, and they are concatenated to produce the input for RNN layers. The activation function in the convolutional layer is a rectified linear unit (ReLU) function, and batch normalization eliminates the internal covariate shift.

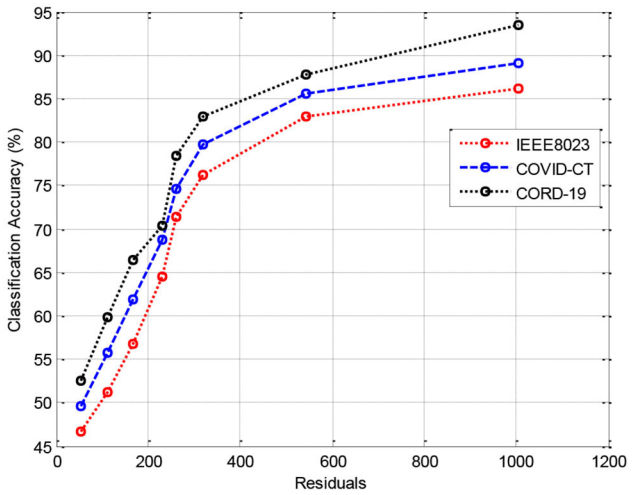

FIGURE 3 | Results of classification accuracy during training with 70% training data.

Recurrent Layers

The RNN architecture learns the needed features having long-term dependencies (long paths). The study uses Gated Recurrent Unit (GRU) on long and short path selection to reduce well the network complexity. A set of three layers stacked in GRU is used in this paper that uses time flow that runs the stacked GRU incrementally for faster input data processing. The recurrent layer outputs vector series $\{x(r,t)\}$ where $t = 1,2,\cdots,T$ for the process of classification at the output layer.

Output Layer

For the purpose of classification, $\{x(r,t)\}$ is selected as the feature vector, and this layer converts the vector of variable length into fixed length. The final feature is generated by averaging the features over a specific time interval based on long or short paths, $x(r)\ x\,(r) = \frac{\sum_{t=1}^{T} x(t,r)}{T}$. Finally the probability of predicted category is generated by feeding the averaging features into the softmax layer.

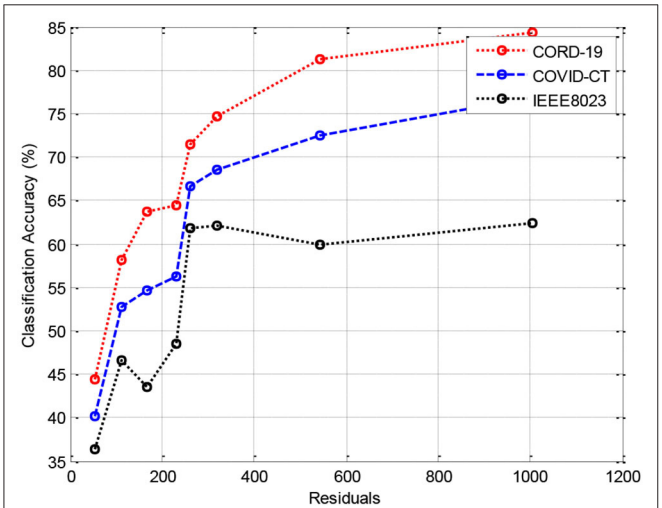

FIGURE 4 | Results of classification accuracy during training with 80% training data.

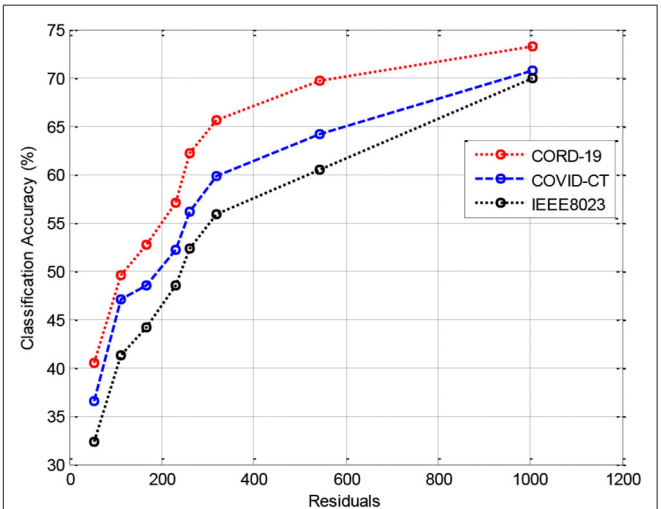

FIGURE 5 | Results of classification accuracy during training with 90% training data.

Type-Specific Layer

For the customization of the DeepSense layer to operate the process of classification, we specifically use the following process:

Step 1: Identify the input image
Step 2: Preprocessing input image for temporal and spectral noise
Step 3: Extract the features related to COVID-19 infections

TABLE 1 | Results of statistical parameters for IEEE8023 with 70% training data on 1,000 images.

Statistical parameters	ANN	FFNN	BPNN	DNN	RNN	DeepSense
Accuracy	55.67145	55.97152	58.06198	58.32304	59.68335	80.475
F-measure	38.39159	40.49205	51.72857	51.8886	54.26013	83.65671
G-mean	72.54022	72.77127	74.27161	74.31162	74.72171	85.57814
MAPE	28.32533	25.38368	23.98336	21.40179	20.82166	16.1186
Sensitivity	61.74481	65.25659	73.16136	85.54813	86.20828	96.25452
Specificity	74.18159	74.37163	77.88342	77.90342	79.27473	80.11492

ANN, artificial neural network; BPNN, back propagation neural network; DNN, deep neural network; FFNN, feedforward neural network; MAPE, mean absolute percentage error; RNN, recurrent neural network.

TABLE 2 | Results of statistical parameters for COVID-CT with 70% training data on 1,000 images.

Statistical parameters	ANN	FFNN	BPNN	DNN	RNN	DeepSense
Accuracy	56.26158	58.87317	61.23469	62.605	65.84672	84.68794
F-measure	66.74693	66.79694	67.68814	68.80839	73.84151	79.40476
G-mean	43.50373	56.43162	59.4633	44.705	76.09302	85.98823
MAPE	19.37033	16.69873	16.60871	11.75563	10.42533	9.275074
Sensitivity	76.23305	78.91465	79.00467	83.84675	85.18805	86.33831
Specificity	73.39141	76.27306	77.21327	80.37497	82.37642	84.46789

ANN, artificial neural network; BPNN, back propagation neural network; COVID-19, coronavirus disease 2019; DNN, deep neural network; FFNN, feedforward neural network; MAPE, mean absolute percentage error; RNN, recurrent neural network.

TABLE 3 | Results of statistical parameters for CORD-19 with 70% training data on 1,000 images.

Statistical parameters	ANN	FFNN	BPNN	DNN	RNN	DeepSense
Accuracy	59.07321	65.85673	68.8784	74.09157	77.93343	82.41643
F-measure	69.68858	69.93964	70.10968	70.28972	74.85174	80.39498
G-mean	69.98965	70.2197	71.94009	74.00155	76.4631	79.21471
MAPE	68.11823	64.47642	57.75191	39.63186	36.77022	34.91881
Sensitivity	77.52334	71.14991	71.87007	73.64147	73.84151	80.69505
Specificity	70.39974	72.28016	75.35185	80.58502	81.89631	82.30641

ANN, artificial neural network; BPNN, back propagation neural network; CORD-19, COVID-19 Open Research Dataset Challenge; COVID-19, coronavirus disease 2019; DNN, deep neural network; FFNN, feedforward neural network; MAPE, mean absolute percentage error; RNN, recurrent neural network.

TABLE 4 | Results of statistical parameters for IEEE8023 with 80% training data on 1,000 images.

Statistical parameters	ANN	FFNN	BPNN	DNN	RNN	DeepSense
Accuracy	96.46457	97.18473	97.21474	97.29476	97.30476	97.43479
F-measure	52.36871	69.72859	70.04966	72.93131	76.16303	79.36475
G-mean	81.88631	82.7365	84.35786	85.91821	90.96134	92.48168
MAPE	26.90502	25.51371	22.74209	20.07049	10.60537	90.12115
Sensitivity	68.69836	70.08967	72.86129	75.54189	84.99801	88.59981
Specificity	96.53459	97.32476	97.52481	97.60483	97.62483	97.68484

ANN, artificial neural network; BPNN, back propagation neural network; DNN, deep neural network; FFNN, feedforward neural network; MAPE, mean absolute percentage error; RNN, recurrent neural network.

Step 4: Apply DeepSense classifier for optimal classifier.

RESULTS AND DISCUSSIONS

This section provides the results of comparison between the machine/deep learning classifiers for predicting COVID-19 infections using IEEE8023 (26), COVID-CT-Dataset (27), and COVID-19 Open Research Dataset Challenge (CORD-19) (28) datasets.

IEEE8023 has the image collection from various sources including COVID-19 or viral and bacterial pneumonias in the form of CT images. COVID-CT has 349 COVID-19 CT images from 216 patients and 463 non-COVID-19 CTs. CORD-19 has collected the CT image resources from 52,000 scholarly articles.

The study is experimented using a 10-fold cross validation, which is tested with all these three base classifiers.

Experiment

The performance measures for evaluating the DeepSense classifier is estimated against various metrics: accuracy, geometric mean (G-mean), F-measure, precision, percentage error, specificity, and sensitivity.

Accuracy for optimal classification is given below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

where:

TP is defined as the true positive
TN is defined as the true negative
FP is defined as the false positive
FN is defined as the false negative

TABLE 5 | Results of statistical parameters for COVID-CT with 80% training data on 1,000 images.

Statistical parameters	ANN	FFNN	BPNN	DNN	RNN	DeepSense
Accuracy	97.73486	97.75486	97.77486	97.77486	97.78487	97.78487
F-measure	89.19995	90.63127	90.7813	91.29141	91.50146	92.1416
G-mean	93.27286	96.67462	97.22474	97.52481	97.66484	97.66484
MAPE	86.64838	27.01504	20.25053	9.275074	54.63022	21.0017
Sensitivity	88.94989	95.58337	96.68462	97.26475	97.54481	97.55482
Specificity	96.76464	96.78464	96.78464	96.78464	96.78464	97.42479

ANN, artificial neural network; BPNN, back propagation neural network; COVID-19, coronavirus disease 2019; DNN, deep neural network; FFNN, feedforward neural network; MAPE, mean absolute percentage error; RNN, recurrent neural network.

TABLE 6 | Results of statistical parameters for CORD-19 with 80% training data on 1,000 images.

Statistical parameters	ANN	FFNN	BPNN	DNN	RNN	DeepSense
Accuracy	93.97301	94.12305	94.20307	94.25308	94.43312	94.44312
F-measure	58.25303	60.04343	60.3835	60.84361	62.37495	62.51498
G-mean	79.1447	79.65481	80.18493	80.41498	81.52623	81.98633
MAPE	29.90669	29.17552	28.28533	27.93525	26.10384	25.28365
Sensitivity	65.69669	66.42685	67.31805	67.67813	69.49854	70.32973
Specificity	95.2533	95.41334	95.42334	95.47335	95.51336	95.56337

ANN, artificial neural network; BPNN, back propagation neural network; CORD-19, COVID-19 Open Research Dataset Challenge; COVID-19, coronavirus disease 2019; DNN, deep neural network; FFNN, feedforward neural network; MAPE, mean absolute percentage error; RNN, recurrent neural network.

TABLE 7 | Results of statistical parameters for IEEE8023 with 90% training data on 1,000 images.

Statistical parameters	ANN	FFNN	BPNN	DNN	RNN	DeepSense
Accuracy	95.91445	95.93445	95.94446	96.03448	96.05448	96.11449
F-measure	77.3833	77.51333	78.03345	79.11469	79.79484	80.08491
G-mean	79.43476	79.67482	79.94488	80.9451	81.26517	81.45622
MAPE	31.14697	30.78688	30.26677	28.65541	28.1553	27.83522
Sensitivity	64.45641	64.81649	65.33661	66.94797	67.44808	67.76815
Specificity	94.72318	94.7832	94.8232	96.05448	96.46457	96.81465

ANN, artificial neural network; BPNN, back propagation neural network; DNN, deep neural network; FFNN, feedforward neural network; MAPE, mean absolute percentage error; RNN, recurrent neural network.

TABLE 8 | Results of statistical parameters for COVID-CT with 90% training data on 1,000 images classifier.

Statistical parameters	ANN	FFNN	BPNN	DNN	RNN	DeepSense
Accuracy	97.37478	97.37478	97.45479	97.45479	97.4748	97.52481
F-measure	85.88821	86.00823	87.94967	87.97967	89.33998	89.34998
G-mean	94.03303	94.03303	94.40311	94.47313	94.8032	94.84321
MAPE	70.79983	70.61979	62.72503	61.39473	54.04008	53.35993
Sensitivity	90.53124	90.53124	91.34143	91.47146	92.21162	92.28164
Specificity	97.4648	97.4748	97.56482	97.56482	97.65484	97.65484

ANN, artificial neural network; BPNN, back propagation neural network; COVID-19, coronavirus disease 2019; DNN, deep neural network; FFNN, feedforward neural network; MAPE, mean absolute percentage error; RNN, recurrent neural network.

TABLE 9 | Results of statistical parameters for CORD-19 with 90% training data on 1,000 images.

Statistical parameters	ANN	FFNN	BPNN	DNN	RNN	DeepSense
Accuracy	97.44379	97.44379	97.52381	97.52381	97.54381	97.59382
F-measure	85.94922	86.06925	88.01168	88.04169	89.40299	89.41399
G-mean	94.09904	94.09904	94.47013	94.54014	94.87022	94.91022
MAPE	70.84984	70.6698	62.77004	61.43874	54.07909	53.39794
Sensitivity	90.59526	90.59526	91.40644	91.53647	92.27664	92.34765
Specificity	97.53381	97.54381	97.63383	97.63383	97.72385	97.72385

ANN, artificial neural network; BPNN, back propagation neural network; CORD-19, COVID-19 Open Research Dataset Challenge; COVID-19, coronavirus disease 2019; DNN, deep neural network; FFNN, feedforward neural network; MAPE, mean absolute percentage error; RNN, recurrent neural network.

F-measure of the DeepSense classifier is defined as follows:

$$F - measure = \frac{2TP}{2TP + FP + FN} \tag{2}$$

G-mean of the DeepSense classifier is defined as follows:

$$G - mean = \sqrt{\frac{TP}{TP + FN} \times \frac{TN}{TN + FP}} \tag{3}$$

Mean absolute percentage error (MAPE) of the DeepSense classifier is defined as follows:

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \tag{4}$$

Where

A_t is defined as the actual class
F_t is defined as the predicted class, and
n is defined as the fitted points

Sensitivity of the DeepSense is defined as:

$$Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

Specificity of the DeepSense is defined as:

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

Analysis

In this section, we provide the results of various meta-ensemble classifiers that include FFNN (29), ANN (25), DNN (30), BPNN (31), and RNN (32). The proposed method is validated against three different datasets, where the training data are compared with 70% (**Figure 3**), 80% (**Figure 4**), and 90% (**Figure 5**) training data.

Figure 3 shows the results of classification accuracy of CORD-19 datasets for all residuals are higher, and with increasing residuals, the accuracy increases. Same is the case for other training sets; however, with 80% datasets, the accuracy is fluctuating due to the extraction of on-optimal features from IEEE8023 datasets.

Tables 1, **4**, **7** provide the results of statistical parameters on predicting COVID-19 infections over 70, 80, and 90% training data over IEEE8023 datasets.

Tables 2, **5**, **8** provide the results of statistical parameters on predicting COVID-19 infections over 70, 80, and 90% training data over COVID-CT datasets.

Tables 3, **6**, **9** provide the results of statistical parameters on predicting COVID-19 infections over 70, 80, and 90% training data over CORD-19 datasets.

Evaluation Criteria

The simulation results show that the DeepSense classifier has higher classification accuracy than the existing meta-ensemble classifiers. In addition, the CORD-19 datasets offer optimum selection of features to increase classification accuracy by 90% training data over 80 or 90%. The other measurements are optimal for CORD-19 than the other selection tools.

Furthermore, MAPE is less than the other methods in the deep learning model.

The result shows that the CORD-19 datasets are more accurate than RNN and DNN. The results also show that the classification accuracy with IEEE8023 as a functional selection tool decreases at some point as the number of residues increases compared to COVID-CT and CORD-19. The class of infections is therefore accurately determined with the proposed classification.

CONCLUSIONS AND FUTURE WORK

In this paper, a DeepSense algorithm is designed for the classification of COVID-19 infections. The DeepSense algorithm helps in optimal classification of multidimensional features from CT images. The classifier combined with hybrid deep learning classifier, namely, CNN and RNN, helps in improving the prediction of events from a medical image. The extraction of optimal features from the feature extraction model helps the classifier to optimally detect whether the patient is infected or not. The experimental results show that the proposed method has higher accuracy than the other methods. In the future, the model can be designed with an ensemble data model to classify the highly rated multidimensional dataset.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

AK: visualization and investigation. AOK: data curation, software, and validation. SK: methodology, data curation, review and editing, and supervision. YN: conceptualization, methodology, writing original draft, software, and data curation. SM: software and validation. GT: writing—review and editing and supervision.

REFERENCES

1. Pathak Y, Shukla PK, Tiwari A, Stalin S, Singh S, Shukla PK. Deep transfer learning based classification model for COVID-19 disease. *Ing Rech BioMed*. (2020). doi: 10.1016/j.irbm.2020.05.003. [Epub ahead of print].
2. Singh D, Kumar V, Kaur M. Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks. *Eur J Clin Microbiol Infect Dis*. (2020) 39:1379–89. doi: 10.1007/s10096-020-03901-z
3. Gao XW, James-Reynolds C, Currie E. Analysis of tuberculosis severity levels from ct pulmonary images based on enhanced residual deep learning architecture. *Neurocomputing*. (2020) 392:233–44. doi: 10.1016/j.neucom.2018.12.086
4. Gerard SE, Patton TJ, Christensen GE, Bayouth JE, Reinhardt JM. FissureNet: a deep learning approach for pulmonary fissure detection in CT images. *IEEE Trans Med Imaging*. (2018) 38:156–66. doi: 10.1109/TMI.2018.2858202
5. Hagerty JR, Stanley RJ, Almubarak HA, Lama N, Kasmi R, Guo P, et al. Deep learning and handcrafted method fusion: higher diagnostic accuracy for melanoma dermoscopy images. *IEEE J Biomed Health Inform*. (2019) 23:1385–91. doi: 10.1109/JBHI.2019.2891049
6. Nardelli P, Jimenez-Carretero D, Bermejo-Pelaez D, Washko GR, Rahaghi FN, Ledesma-Carbayo MJ, et al. Pulmonary artery–vein classification in CT images using deep learning. *IEEE Trans Med Imaging*. (2018) 37:2428–40. doi: 10.1109/TMI.2018.2833385
7. Pannu HS, Singh D, Malhi AK. Improved particle swarm optimization based adaptive neuro-fuzzy inference system for benzene detection. *CLEAN Soil Air Water*. (2018) 46:1700162. doi: 10.1002/clen.201700162
8. Xia K, Yin H, Qian P, Jiang Y, Wang S. Liver semantic segmentation algorithm based on improved deep adversarial networks in combination of weighted loss function on abdominal CT images. *IEEE Access*. (2019) 7:96349–58. doi: 10.1109/ACCESS.2019.2929270
9. Xie Y, Xia Y, Zhang J, Song Y, Feng D, Fulham M, et al. Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT. *IEEE Trans Med Imaging*. (2018) 38:991–1004. doi: 10.1109/TMI.2018.2876510
10. Zreik M, Van Hamersvelt RW, Wolterink JM, Leiner T, Viergever MA, Išgum I. A recurrent CNN for automatic detection and classification of coronary artery plaque and stenosis in coronary CT angiography. *IEEE Trans Med Imaging*. (2018) 38:1588–98. doi: 10.1109/TMI.2018.2883807
11. Carrillo-Larco RM, Castillo-Cara M. Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: an unsupervised machine learning approach. *Wellcome Open Res*. (2020) 5:56. doi: 10.12688/wellcomeopenres.15819.2
12. Hemdan EED, Shouman MA, Karar ME. Covidx-net: a framework of deep learning classifiers to diagnose covid-19 in x-ray images. *arXiv*. (2020). p. 1–14.
13. Khmaissia F, Haghighi PS, Jayaprakash A, Wu Z, Papadopoulos S, Lai Y, et al. An Unsupervised machine learning approach to assess the ZIP code level impact of COVID-19 in NYC. *arXiv*. (2020). p. 1–16.
14. Pereira RM, Bertolini D, Teixeira LO, Silla Jr CN, Costa YM. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. *Comput Methods Programs Biomed*. (2020) 194:105532. doi: 10.1016/j.cmpb.2020.105532
15. Wang X, Deng X, Fu Q, Zhou Q. A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. *IEEE Trans. Med. Imag.* 39:26150–25. doi: 10.1109/TMI.2020.2995965
16. Gupta D, Rodrigues JJ, Sundaram S, Khanna A, Korotaev V, de Albuquerque VHC. Usability feature extraction using modified crow search algorithm: a novel approach. *Neural Comput Appl*. (2018) 32:10915–925. doi: 10.1007/s00521-018-3688-6
17. Gupta D, Ahlawat AK. Usability determination using multistage fuzzy system. *Procedia Comput Sci*. (2016) 78:263–70. doi: 10.1016/j.procs.2016.02.042
18. Khamparia A, Saini G, Gupta D, Khanna A, Tiwari S, de Albuquerque VHC. Seasonal crops disease prediction and classification using deep convolutional encoder network. *Circuits Syst Signal Process*. (2020) 39:818–36. doi: 10.1007/s00034-019-01041-0
19. Khamparia A, Singh A, Anand D, Gupta D, Khanna A, Kumar NA, et al. A novel deep learning-based multi-model ensemble method for the prediction of neuromuscular disorders. *Neural Comp Appl*. (2018) 32:11083–95. doi: 10.1007/s00521-018-3896-0
20. Gochhayat SP, Kaliyar P, Conti M, Tiwari P, Prasath VBS, Gupta D, et al. LISA: Lightweight context-aware IoT service architecture. *J Clean Prod*. (2019) 212:1345–56. doi: 10.1016/j.jclepro.2018.12.096
21. Khan FQ, Musa S, Tsaramirsis G, Buhari SM. SPL features quantification and selection based on multiple multi-level objectives. *Appl Sci*. (2019) 9:2212. doi: 10.3390/app9112212
22. Raj RJS, Shobana SJ, Pustokhina IV, Pustokhin DA, Gupta D, Shankar K. Optimal feature selection-based medical image classification using deep learning model in internet of medical things. *IEEE Access*. (2020) 8:58006–17. doi: 10.1109/ACCESS.2020.2981337

23. Tsaramirsis K, Tsaramirsis G, Khan FQ, Ahmad A, Khadidos AO, Khadidos A. More agility to semantic similarities algorithm implementations. *Int J Environ Res Public Health*. (2020) 17:267. doi: 10.3390/ijerph17010267
24. Pustokhina IV, Pustokhin DA, Gupta D, Khanna A, Shankar K, Nguyen GN. An effective training scheme for deep neural network in edge computing enabled Internet of medical things (IoMT) systems. *IEEE Access.* (2020) 8:107112–23. doi: 10.1109/ACCESS.2020.3000322
25. Saritas MM, Yasar A. Performance analysis of ANN and naive bayes classification algorithm for data classification. *Int J Intell Syst Appl Eng.* (2019) 7:88–91. doi: 10.18201/ijisae.2019252786
26. IEEE8023. Available online at: https://github.com/ieee8023/covid-chestxray-dataset (accessed March 3, 2020).
27. Zhao J, Zhang Y, He X, Xie P. COVID-CT-dataset: a CT scan dataset about COVID-19. *arXiv*. (2020). p. 1–15.
28. COVID-19 Open Research Dataset Challenge (CORD-19): Available online at: https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge (accessed August 17, 2020).
29. Alrikabi HA, Annajjar W, Alnasrallah AM, Mustafa ST, Rahim MSM. Using FFNN classifier with HOS-WPD method for epileptic seizure detection. In *2019 IEEE 9th International Conference on System Engineering and Technology (ICSET).* Shah Alam: IEEE (2019). p. 360–3. doi: 10.1109/ICSEngT.2019.8906408
30. Nurmaini S, UmiPartan R, Caesarendra W, Dewi T, NaufalRahmatullah M, Darmawahyuni A, et al. An automated ECG beat classification system using deep neural networks with an unsupervised feature extraction technique. *Appl Sci.* (2019) 9:2921. doi: 10.3390/app9142921
31. Dandu JR, Thiyagarajan AP, Murugan PR, Govindaraj V. Brain and pancreatic tumor segmentation using SRM and BPNN classification. *Health Technol.* (2020) 10:187–95. doi: 10.1007/s12553-018-00284-2
32. Liu H, Lang B, Liu M, Yan H. CNN and RNN based payload classification methods for attack detection. *Knowledge-Based Syst.* (2019) 163:332–s41. doi: 10.1016/j.knosys.2018.08.036

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 Public Health

ORIGINAL RESEARCH
published: 26 November 2021
doi: 10.3389/fpubh.2021.755644

Determinants of Intention to Use Artificial Intelligence-Based Diagnosis Support System Among Prospective Physicians

Anh Quynh Tran[1†], Long Hoang Nguyen[2†], Hao Si Anh Nguyen[3], Cuong Tat Nguyen[4,5], Linh Gia Vu[4,5], Melvyn Zhang[6], Thuc Minh Thi Vu[3], Son Hoang Nguyen[7], Bach Xuan Tran[1,8], Carl A. Latkin[8], Roger C. M. Ho[9,10] and Cyrus S. H. Ho[9]*

[1] Institute for Preventive Medicine and Public Health, Hanoi Medical University, Hanoi, Vietnam, [2] Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden, [3] Institute of Health Economics and Technology, Hanoi, Vietnam, [4] Institute for Global Health Innovations, Duy Tan University, Da Nang, Vietnam, [5] Faculty of Medicine, Duy Tan University, Da Nang, Vietnam, [6] National Addictions Management Service (NAMS), Institute of Mental Health, Singapore, Singapore, [7] Center of Excellence in Evidence-Based Medicine, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam, [8] Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States, [9] Department of Psychological Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore, [10] Institute for Health Innovation and Technology (iHealthtech), National University of Singapore, Singapore, Singapore

OPEN ACCESS

Edited by:
Shabana Urooj, Princess Nourah Bint Abdulrahman University, Saudi Arabia

Reviewed by:
Navaid Z. Rizvi, Gautam Buddha University, India
Rashmi Sinha, Forsyth Technical Community College, United States
Bharat Singh, Bharati Vidyapeeth's College of Engineering, India

***Correspondence:**
Linh Gia Vu
vugialinh@duytan.edu.vn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Digital Public Health, a section of the journal Frontiers in Public Health

Received: *09 August 2021*
Accepted: *19 October 2021*
Published: *26 November 2021*

Citation:
Tran AQ, Nguyen LH, Nguyen HSA, Nguyen CT, Vu LG, Zhang M, Vu TMT, Nguyen SH, Tran BX, Latkin CA, Ho RCM and Ho CSH (2021) Determinants of Intention to Use Artificial Intelligence-Based Diagnosis Support System Among Prospective Physicians. Front. Public Health 9:755644. doi: 10.3389/fpubh.2021.755644

Background: This study aimed to develop a theoretical model to explore the behavioral intentions of medical students to adopt an AI-based Diagnosis Support System.

Methods: This online cross-sectional survey used the unified theory of user acceptance of technology (UTAUT) to examine the intentions to use an AI-based Diagnosis Support System in 211 undergraduate medical students in Vietnam. Partial least squares (PLS) structural equational modeling was employed to assess the relationship between latent constructs.

Results: Effort expectancy ($\beta = 0.201$, $p < 0.05$) and social influence ($\beta = 0.574$, $p < 0.05$) were positively associated with initial trust, while no association was found between performance expectancy and initial trust ($p > 0.05$). Only social influence ($\beta = 0.527$, $p < 0.05$) was positively related to the behavioral intention.

Conclusions: This study highlights positive behavioral intentions in using an AI-based diagnosis support system among prospective Vietnamese physicians, as well as the effect of social influence on this choice. The development of AI-based competent curricula should be considered when reforming medical education in Vietnam.

Keywords: artificial intelligence, diagnosis, theoretical model, intention, medical students

INTRODUCTION

Artificial intelligence (AI) was first introduced some years ago, but in recent years, there has been increasing exploration of the utility and cost-saving of such technology (1, 2). AI brings about great potential in changing existing healthcare practice, from prevention, screening, diagnosis, treatment, and care (2, 3). AI could tap onto data from existing medical records; or even data from smartphones that individuals possess, and data from the applications that individuals use, including

their social media posts (3, 4). By using large datasets and employ advanced techniques such as machine learning and deep learning approaches, AI informs more precise predictions of behavioral patterns and understanding of existent medical conditions (3). These benefits would facilitate the clinical decision process, improve the efficacy and accuracy of diagnosis, and diminish physician's workload. Evidence on the utility of AI in healthcare has been widely recorded from dentistry (5), primary care (6), radiology (7), ophthalmology (8) or pathology (9). AI has been recommended for inclusion in routine workflow processes (7). It is thus evident from these studies that the use of AI has been explored in various domains, and it is a promising technology for healthcare.

However, although many reports show the promising role of AI, the usage of AI is still in the early stage. Recent studies indicated low rates of physicians who were familiar or had chances to adopt AI in their clinical practices, even in technologically advanced nations such as 5.9% in South Korea (10) or 23% in the United States (11). Many technological, social, organizational, and individual challenges to apply AI principles in healthcare facilities have been discussed thoroughly in literature (2, 12–15). Nonetheless, the most important factor was physician's attitudes and perceptions toward AI, which can decide whether they would want to integrate AI in their practice or not (13, 14, 16). In healthcare, when the clinical decision is closely related to the patient's lives, health professionals are more likely to be cautious to use new technology in treatment and care; thus, it is not easy for them to trust and use a new product to support their practice.

Health systems can actively involve in the roles of AI adopters and innovators. Therefore, given the rapid expansion of AI applications in healthcare, it is crucial for future health workforces to prepare their capacities, as well as positive perceptions and attitudes to participate in the development of these novel tools. Prior studies indicated some controversial results about the attitudes and intentions to use AI in healthcare practices among medical students. For example, a study in the United States revealed that although the majority of radiology students had a belief in the future role of AI, they felt less interest in the applications of AI in the radiology field (17). Another study in the United Kingdom showed that 49% of medical students were more likely to apply for a radiology career due to AI (18). Understanding determinants of their behavioral intention to use and adopt AI in healthcare delivery is thus necessary for developing medical education curriculum to facilitate AI competence.

In Vietnam, it was been reported in 2019 that more organizations (both healthcare and non-healthcare related) have started developing AI technologies, and utilizing such technologies (19). In 2019, the Vietnam Ministry of Health has issued Decision No. 4888/QD-BYT about the applications and development of smart health care during 2019–2025, which underlines the importance of digital health and strategies to integrate digital health, including AI, into routine health service delivery (20). To date, there remains limited evaluation of AI amongst Vietnamese healthcare services. From our knowledge, there has been only prior publication, that of Vuong et al. (21) that presents a framework seeking to evaluate the AI readiness of the Vietnamese healthcare sector. The authors reported that the implementation of AI in healthcare in Vietnam is limited by several factors, such as the lack of funding; the necessary information infrastructure; and most importantly, the lack of understanding and misunderstanding of AI. Whilst the previous article by Vuong et al. (21) has provided some insights into the challenges with AI implementation and utilization, the review focused on issues at a macro-level, and has not evaluated the perspectives of individual healthcare professionals. For there to be a high uptake rate of AI on the ground, there needs to be an understanding of existing attitudes, preferences, and perspectives of future physicians.

In healthcare, various theories have been used to understand comprehensive facilitating factors in the individual's adoption and acceptance of a novel technology. For instance, several theories included the theory of planned behavior, the theory of diffusion of innovations, the technology acceptance model, or the unified theory of user acceptance of technology (UTAUT). Of which, UTAUT has been recognized as one of the most common theories to examine the adoption behavior of one individual (22–25). UTAUT was developed based on other dominant behavioral theories. Venkatesh et al. showed a higher explanatory level of UTAUT compared to other theories in exploring the information technology adaptation, with 70% of the variance for behavioral intentions and 50% of the variance for actual use (26, 27). A previous study in Chinese physicians showed that initial trust and performance expectancy were significant predictors for the AI adoption intentions (28). This study aimed to use UTAUT to explore the behavioral intentions of medical students to adopt an AI-based Diagnosis Support System. Understanding medical student's attitudes and perspectives would help to resolve potential barriers in adoption at the ground level, and such a survey would also help guide AI policy formulation at different levels.

MATERIALS AND METHODS

In this section, we presented literature review and conceptual framework of this study. Moreover, study design, data collection method, and statistical analysis were described.

Literature Review and Conceptual Framework

UTAUT has been used widely in the literature to examine the behaviors of an individual in adopting the technology. UTAUT explains individual's behaviors *via* four constructs: (1) performance expectancy, (2) effort expectancy; (3) social influence, and (4) facilitating conditions (26). Because AI-based Diagnosis Support System has not been implemented in entire Vietnam, we supposed that there was very difficult for medical students to have a chance to use AI systems during their clerkship or when they studied in the medical university. Therefore, we used UTAUT to explore the behavioral intention, which was defined as the willingness of medical students to use this system in the future if they had an opportunity. The behavioral intention

was a significant predictor of actual use; thus, it is valid to determine the factors associated with the behavioral intention of AI use, which would partly reflect the AI practice in the future (13).

Firstly, three main constructs of the UTAUT model (i.e., performance expectancy, effort expectancy, and social influence) were included. The performance expectancy refers to "*the degree to which a person believes that using a particular system would enhance his or her job performance*," while the effort expectancy is defined as "*the degree of ease associated with the use of the system*," and the social influence refers to "*the degree to which an individual perceives that important others believe he or she should use the new system*" (26). All of them have been revealed to have positive associations with behavioral intentions in different studies regarding IT adoption (26). Performance expectancy is found to be related to effort expectancy because it is supposed that people were more likely to perceive that one technology is useful if they ease using this technology (29). In literature, previous studies showed that medical students believed that AI would help to enhance the performance of practices and AI would be integrated deeply in healthcare, from administrative works to clinical routine (30–33). Indeed, medical students are considered to have high AI literacy than current health professionals. A survey in the United States indicated that medical students were more likely to have basic knowledge about AI and prefer to use AI in patient care when comparing to their faculties (31). Another survey in the United Kingdom found that medical students who were taught about AI were more likely to adopt AI in their practices (18). Social influence may also affect the intention to use AI in healthcare. Prior research in both the general public and health professionals recommended that medical students should learn and practice AI during their studies (31, 34–36). Experts shared that future physicians should have a good understanding and can transforming AI from potential threats to become helpful assistants (37).

Via literature review, we also decided to develop the model with three additional constructs: task complexity, personal innovativeness in IT, and technology characteristics. Task complexity is the level of difficulty for completing an assigned task (38); hence, technology can have different roles in different tasks. Health professionals in their daily practices will face a variety of tasks, from simple to complex tasks. If they perceived that their tasks are difficult, they are more likely to accept the support from the AI system to increase their performance (i.e., performance expectancy). A study in Canada showed that medical students perceived the usefulness of AI in providing diagnosis, prognosis, building personalized medication, and performing robotic surgery, which indicated the promising roles of AI in addressing task complexity (33). Meanwhile, personal innovativeness in information technology (IT) means that one person is willing to try an innovation (particularly in IT) (39), while technology characteristics refer to the system, interface, etc. which allow users to use the technology for completing their tasks (40). Prior studies showed the potential relationships between these two constructs with effort expectancy (41, 42). Overall, we attempted to examine the association between task complexity and performance expectancy; and between personal innovativeness in IT and technology characteristics with effort expectancy.

Along with these three constructs, we added perceived substitution crisis and initial trust constructs aiming to examine the facilitating conditions to behavioral intentions. Perceived substitution crisis was served as a potential barrier for medical students to adopt technology in their future practice. Several obstacles such as the likelihood of being replaced by AI, being dependent on AI, being unemployed due to AI, and decreasing diagnosis capacity due to AI would greatly affect the benefits of physicians. Previous research found that 17% of German medical students agreed that AI could replace health professionals (30), and 49% of English medical students stated that they did not prefer the radiology field because of AI (18). Therefore, the perceived substitution crisis was suggested to be included when examining the intention to use AI among health professionals (2, 12–15).

For the initial trust, Mcknight et al. defined trust in the field of technology as "*beliefs about a technology's capability rather than its will or its motives*" (43). Trust is an important determinant of technology acceptance and adoption (44–47). Physicians are more likely to be cautious when adopting new technology in patients to prevent any potential harm; thus, trusting can help to reduce any suspicions and facilitate the use of the AI system among physicians. In a previous study, lack of trust in AI was the main contributor to the negative attitude among Chinese people toward the application of AI in healthcare (34). Another study in Canada found that medical students did not believe AI could deliver personalized and empathetic care (33). Thus, we hypothesized that trust would be positively associated with the behavioral intention to use AI systems in medical students. Given the matter that in Vietnam's medical education curriculum, none of course about AI was tough, we supposed that our medical students did not have any previous experience with AI and AI-based Diagnosis Support System. Thus, among different stages in trust formation, we concentrated on the initial stage, i.e., initial trust, which reflected how people trust in a technology that they have no experience.

Additionally, to identify the relationship between initial trust and behavioral intention, we developed a trust-based theoretical model to explore the trust of medical students in a novel technology as an AI-based Diagnosis Support System. We estimated the associations between performance expectancy, effort expectancy, and social influence with initial trust. Previous studies indicated that performance expectancy and effort expectancy were two forms of technology-specific expectations as discussed above, which are believed to result in trust formation (48). Social influence was also found to be an associated factor with trust in other settings. Prior research revealed that those without any experience with technology were more likely to be dependent on the opinions of their important people, which in turn formulated their trust (48–50). The final conceptual framework used in this study is illustrated in **Figure 1**.

Study Design and Data Collection

Data of this study was obtained from an Internet survey from December 2019 to February 2020. This online survey was

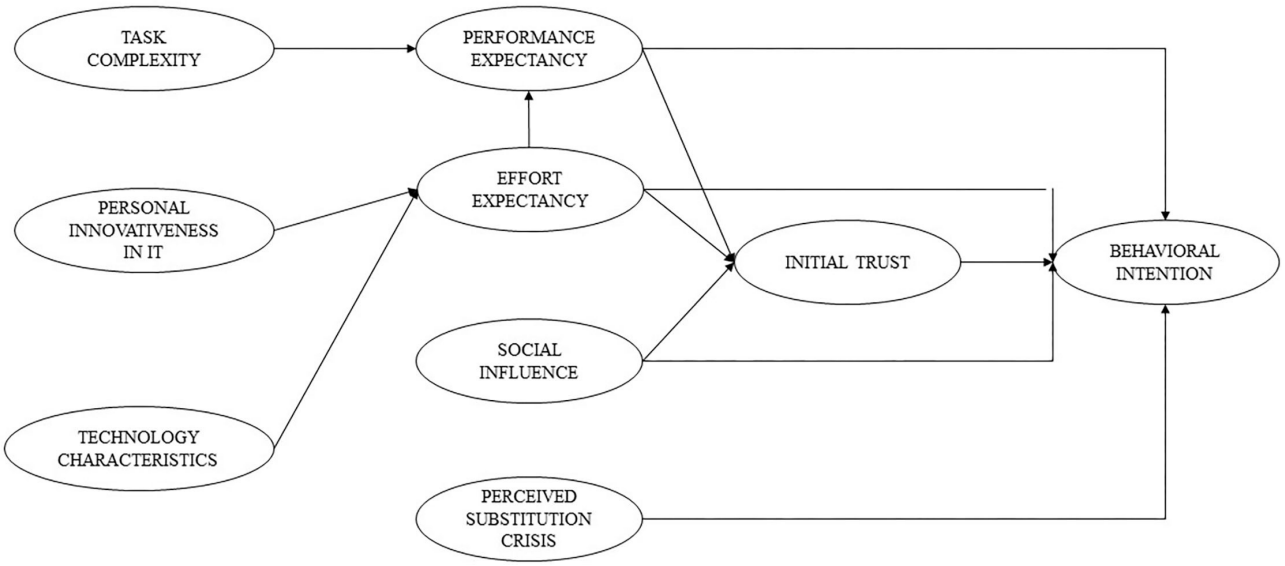

FIGURE 1 | A theoretical model to explore trusts and intentions to use AI-based diagnosis support system.

designed by using an online platform called Survey Monkey (https://www.surveymonkey.com/), which is a highly secure online platform. This survey was sent to medical students at medical university in Vietnam, with the inclusion criteria as follow: (1) aged 18 years or above; (2) currently studying undergraduate medical doctor programs in a medical university in Vietnam; (3) having a valid online account (such as email or social network sites) to help to recruit other medical students. We used the snowball sampling technique to recruit participants. First, we sent out the survey to a core group with twenty medical students who were from different medical universities. After they completed the survey, they were asked to invite other medical students in their networks to do the survey. The recruitment chain stopped when no one was invited or completed the survey within 7 days. A total of 223 medical students from different provinces (Hanoi, Ho Chi Minh city, and other provinces) were enrolled in the study. We obtained their electronic informed consent before doing the survey. After excluding invalid responses, data of 211 (completion rate 94.6%) medical students were used for analysis.

Variables and Questionnaire

In this study, we developed a structured questionnaire with two parts: the demographic characteristics section (including age, gender, living area, specialty, and location), and 26 items that reflected the 9 latent constructs for our theoretical models. These items were about performance expectancy (PE), effort expectancy (EE), social influence (SI), task complexity (TC), personal innovativeness in IT (PI), technology characteristics (TECH), perceived substitution crisis (PC), initial trust (IT), and behavioral intention. These items were selected based on a literature review (26, 28, 41, 42, 48). Participants were asked to respond using a 5-point Likert scale ranging from "strongly disagree" (1), "disagree" (2), "somewhat agree" (3), "agree" (4) to "strongly agree" (5). The proposed constructs and profiles are shown in the **Supplementary Material**.

Data Analysis

Stata software version 15.0 was used to analyze the data. Properties of measurement were evaluated. Internal consistency reliability was assessed by using Cronbach's alpha. Good internal consistency was defined as a Cronbach's alpha $\geq$0.7. Validity was examined, including convergent, discriminant, and construct validities. Convergent validity was assessed *via* two criteria: factor loading >0.70 and average variance extracted of each construct $\geq$0.5 (41). Regarding discriminant validity, we computed the variance inflation factor (VIF) to examine the multicollinearity of each construct. Construct with VIF value >10 indicated that it was not appropriate as a component of regression analysis. The square root of AVE per construct was also computed, and good discriminant validity was achieved when the square root of AVE of a construct was higher than its correlations with other constructs. Given that a sample size of 211 medical students might not be sufficient for the structural equation modeling (SEM) method, we employed partial least squares (PLS) SEM, which is a 2nd-generation SEM, to assess the relationship between latent constructs. We considered a statistical significance when the $p < 0.05$.

RESULTS

Table 1 depicts the demographic characteristics of our sample. The mean age of selected medical students was 20.6 years old (SD = 1.5). The majority of them were female at 73.5%, lived in urban areas (89.1%), and Ho Chi Minh city (59.7%). Most of the respondents belonged to the general physician program (57.8%).

TABLE 1 | Characteristics of respondents ($n = 211$).

Characteristics	
Age, years, Mean (SD)	20.6 (1.5)
Gender, *n* (%)	
Male	55 (26.5)
Female	155 (73.5)
Living area, *n* (%)	
Urban	188 (89.1)
Rural	23 (10.9)
Specialty	
General physician	122 (57.8)
Odonto-Stomatology	48 (22.7)
Traditional medicine	41 (19.4)
Location	
Hanoi	51 (24.2)
Ho Chi Minh city	126 (59.7)
Other provinces	34 (16.1)

TABLE 2 | Reliability and validity of the measure ($n = 211$).

Factor	No. of items	Factor loading	Mean	SD	Cronbach's alpha	AVE
PE	4	0.847–0.915	3.7	0.8	0.903	0.775
EE	2	0.945–0.953	3.3	0.9	0.89	0.901
SI	4	0.827–0.894	3.4	0.7	0.88	0.736
PI	4	0.771–0.869	3.4	0.7	0.854	0.696
TC	2	0.879–0.901	3.8	0.9	0.738	0.791
TECH	3	0.824–0.916	3.1	0.8	0.846	0.765
PC	4	0.710–0.862	3.1	0.8	0.825	0.646
IT	2	0.957–0.957	3	0.9	0.909	0.916
BI	1	–	3.4	0.9	–	1

PE, performance expectancy; EE, effort expectancy; SI, social influence; PI, perceived innovativeness in IT; IT, initial trust; TC, task complexity; TECH, technology characteristics; PC, perceived substitution crisis; BI, behavioral intention.

Table 2 showed that the initial trust construct had the lowest mean score at 3.0 (SD = 0.9), while TC had the highest mean score at 3.8 (SD = 0.9). Overall, the Cronbach's alpha of each construct ranged from 0.738 to 0.909, suggesting good reliability among constructs. All item loadings of these constructs were above 0.7, and all construct's AVE values were above 0.5, indicating good convergent validity.

In **Table 3**, regarding discriminant validity, the value of the square root of AVE per construct was higher than its correlation coefficient with other constructs. Moreover, the results of VIF analysis showed that all VIF values were below 10, suggesting no multicollinearity existed.

Figure 2 illustrates path coefficients and *p*-values of PLS analysis. Regarding the behavioral intention model, only social influence ($\beta = 0.527$, $p < 0.05$) was positively related to the behavioral intention. Meanwhile, other constructs such as performance expectancy, effort expectancy, initial trust, and perceived substitution crisis showed no associations with behavioral intentions to use AI. Overall, the model with five proposed constructs for behavioral intentions, including performance expectancy, effort expectancy, social influence, perceived substitution crisis, and initial trust, explained 47.6% ($R^2 = 0.476$) of the behavioral intention's variance.

TABLE 3 | Correlation of latent variables and square root of AVE of each construct ($n = 211$).

	PE	EE	SI	PI	IT	TC	TECH	PC	BI
PE	0.8803*								
EE	0.6936	0.9492*							
SI	0.6794	0.6656	0.8579*						
PI	0.7408	0.7391	0.7243	0.8343*					
IT	0.4937	0.5586	0.6834	0.5427	0.9571*				
TC	0.6002	0.5015	0.5763	0.6523	0.3213	0.8894*			
TECH	0.5527	0.6099	0.691	0.5801	0.7728	0.3925	0.8746*		
PC	0.3873	0.4568	0.523	0.463	0.3192	0.3935	0.4374	0.8037*	
BI	0.5458	0.5453	0.6856	0.5755	0.4904	0.4838	0.4686	0.3729	1.000*

PE, performance expectancy; EE, effort expectancy; SI, social influence; PI, perceived innovativeness in IT; IT, initial trust; TC, task complexity; TECH, technology characteristics; PC, perceived substitution crisis; BI, behavioral intention.
**Squared root of AVE.*

Figure 2 also shows that effort expectancy ($\beta = 0.201$, $p < 0.05$) and social influence ($\beta = 0.574$, $p < 0.05$) were positively associated with initial trust, while no association was found between performance expectancy and initial trust ($p > 0.05$). The model including performance expectancy, effort expectancy, and social influence explained 47.9% of the variance of initial trust ($R^2 = 0.479$).

DISCUSSION

Developing and adopting AI in healthcare are essential due to its great benefits in enhancing healthcare professional's performance and efficiency. Overall, the perceptions of our students about diagnosis-related capacities of AI, effort to use AI, and intention to use AI were positive. It is clear that the role of AI in healthcare delivery has been widely documented, where AI has shown its success in the interpretation of image and examination data, as well as clinical outcomes prediction and management (15, 51). Nonetheless, information about AI and its application in Vietnam has been disseminated in mainstream media but not in university settings. Our results indicated that undergraduate medical students in Vietnam had great confidence in the knowledge of their work characteristics, understanding how AI could assist them to promote diagnosis performance, and desire to use AI when available. However, there were still some gaps between their expectancy and preparation, including awareness of technology characteristics and capacities to use such technology. Equipping the medical students with the basics, as well as the correct understanding and attitudes about the application of AI in medicine, are crucial in the digitalization of the healthcare system. However, currently, the medical training program in Vietnam has not been systematically updated in this area. The AI content has been mainly shared through scientific

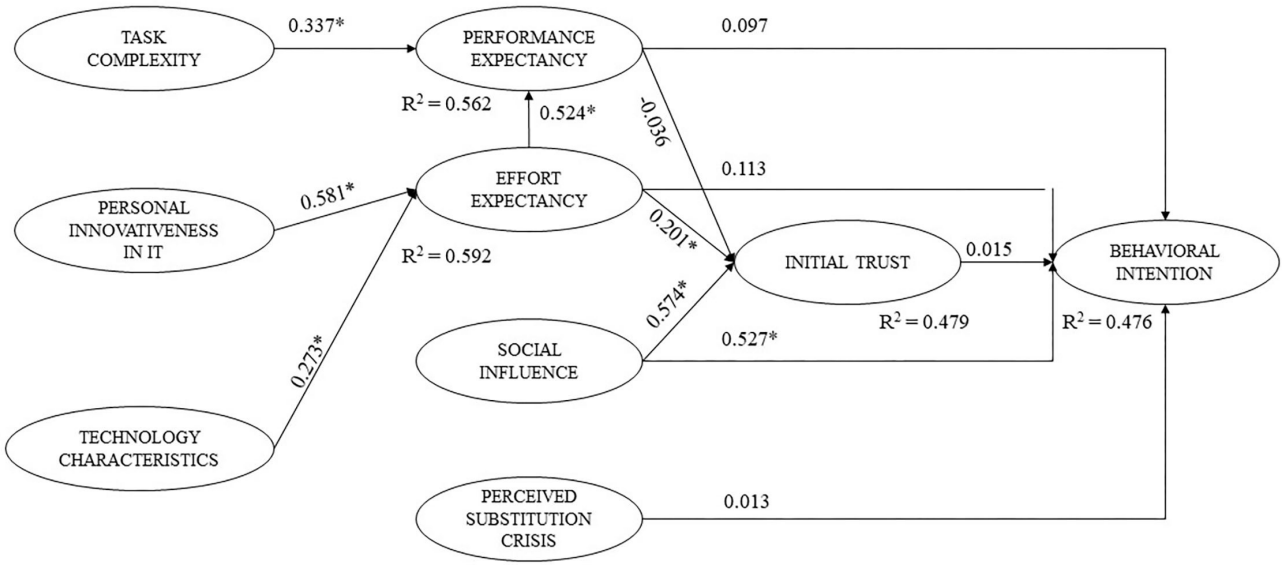

FIGURE 2 | Structural model and standardized path coefficients ($n = 211$). *$p < 0.05$.

seminars or short-term courses, without a specific program to develop the AI capabilities.

This lack of pre-paration might also lead to the findings that the majority of our sample somewhat agreed or agreed that AI would replace the position of physicians in healthcare. This result was congruent with findings among medical students worldwide, particularly those in the radiology field (13, 14, 18, 52). Several previous studies found contradict results where the medical students stated that AI could not have a role as an alternative for the physicians in the future (33, 53), particularly in some fields that need a "sense of caring" or "art of caring" such as psychological health or aging care (53–55). Many authors argued that AI should be treated as a virtual assistant rather than being a replacement for physicians in healthcare. However, prospective physicians should acquire fundamental knowledge about mathematics, data science, AI, as well as ethical and legal issues related to AI (56). They should understand the systemic bias behind AI algorithms due to the insufficient data, which might be a great reason for health equity issues when making a clinical decision (57, 58). Moreover, other humanistic aspects such as communication skills, empathy, decision-making, or leadership skills should also be required (53). Acquiring these capacities would enable physicians to take advantage of AI in integrating it into their routine clinical practices. Thus, it is needed to call actions to innovate the medical education programs in the digital area.

Our path analysis showed the dominance of social influence on the intention of using AI for future work among undergraduate medical students, instead of other factors such as performance expectancy or initial trust, which were found in the previous research (28). Although this result is unexpected compared to what we hypothesized, there were several reasons which can be used to explain this phenomenon. First, this study was conducted on undergraduate medical students, whose healthcare delivery experience, as well as perceptions about the diagnosis process, were constrained. Moreover, given that AI has not yet been scaled up in Vietnam and AI-related curriculums for medical students had not yet been developed, we supposed that the majority of our sample had no experience with an AI-based diagnosis support system. This limitation hinders the way medical students perceived their capacities in adopting AI, as well as results in the homogeneity in their competency and trust evaluation. Moreover, because of this lacking experience, it is understandable when undergraduate medical students tended to be heavily dependent on the experiences of senior physicians in their social networks and information they gathered in social media about AI. With the exchange and sharing of practical experiences from those who have used this AI system, students' trust and intention to use the AI system in the future would be improved.

The findings of this study suggested several implications. First, undergraduate medical students should actively find opportunities to update and involve in AI development and adoption to increase their necessary AI knowledge and capacities. Self-learning ability is important to acquire new knowledge in the context where AI curricula at medical schools have not been paid sufficiently. Second, our study suggested the importance of role model approaches for facilitating the use of AI in this group. Opportunities to gain hands-on experience in different teaching hospitals are critical. AI may be useful for diagnosing rare conditions, which are often only seen at large teaching hospitals. Finally, this study underlined the need to integrate AI curriculums in the current medical education, which helped medical students to prepare appropriate capacities in technology adoption. Further studies should be performed to measure the preference and effectiveness of different education

strategies to facilitate AI applications in healthcare among health professionals and medical students. Moreover, they should also assess whether training students with AI helps or hinders their diagnostic abilities.

Some limitations should be acknowledged in this study. First, since our study was conducted on medical students who had no experience with AI-based diagnostic support systems, we have not yet assessed whether they would use these systems or not in the future. A longitudinal follow-up study evaluating the rate of use of this system among medical students after graduation is essential to help refine the theoretical model. Second, our research was conducted online and had recruited medical students in entire Vietnam; however, this study may be limited to the group of medical students with Internet access, while other groups of medical students were not accessed. In addition, a small sample size might reduce the statistical power. Other studies on larger sample sizes need to be conducted, which help verify our results in other medical student groups. Third, in addition to constructs included in the theoretical model, the study has not assessed the mediating effects of other factors such as age, gender, and previous training in AI use during university studies, etc., which could affect the relationship among factors in the theoretical model.

CONCLUSIONS

This study highlights positive behavioral intentions in using an AI-based diagnosis support system among prospective Vietnamese physicians, as well as the effect of social influence on this choice. The development of AI-based competent curricula should be considered when reforming medical education in Vietnam.

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 Institutional Review Board of Youth Research Institute. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LN: conception, design, acquisition, interpretation of data, drafting the article, and final approval of the version to be published. HN: conception, design, drafting the article, and final approval of the version to be published. AT: design, analysis, interpretation of the data, drafting the article, and final approval of the version to be published. CN: acquisition, analysis, interpretation of data, drafting the article, and final approval of the version to be published. LV and MZ: conception and design of data, revising article, and final approval of the version to be published. CH: conception and design of data, drafting the article, and final approval of the version to be published. BT, CL, and RH: conception, design, interpretation of data, drafting the article, and final approval of the version to be published. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by NUS iHeathtech Other Operating Expenses (R-722-000-004-731) and NUS Department of Psychological Medicine Other Operating Expenses (R-177-000-003-001).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2021.755644/full#supplementary-material

REFERENCES

1. Polesie S, McKee PH, Gardner JM, Gillstedt M, Siarov J, Neittaanmäki N, et al. Attitudes toward artificial intelligence within dermatopathology: an international online survey. *Front Med.* (2020) 7:591952. doi: 10.3389/fmed.2020.591952
2. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. In: *Artificial Intelligence in Healthcare*. (2020). p. 25–60. doi: 10.1016/B978-0-12-818438-7.00002-2
3. Blease C, Locher C, Leon-Carlyle M, Doraiswamy M. Artificial intelligence and the future of psychiatry: qualitative findings from a global physician survey. *Digi Health.* (2020) 6:2055207620968355. doi: 10.1177/2055207620968355
4. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J-P. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. *Neuropsychopharmacology.* (2018) 43:1660–6. doi: 10.1038/s41386-018-0030-z
5. Shan T, Tay FR, Gu L. Application of artificial intelligence in dentistry. *J Dent Res.* (2020) 100:232–44. doi: 10.1177/0022034520969115
6. Mistry P. Artificial intelligence in primary care. *Br J Gen Pract.* (2019) 69:422–3. doi: 10.3399/bjgp19X705137
7. Chong LR, Tsai KT, Lee LL, Foo SG, Chang PC. Artificial intelligence predictive analytics in the management of outpatient MRI appointment no-shows. *AJR Am J Roentgenol.* (2020) 215:1155–62. doi: 10.2214/AJR.19.22594
8. Ting DSW, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol.* (2019) 103:167–75. doi: 10.1136/bjophthalmol-2018-313173
9. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol.* (2019) 20:e253–61. doi: 10.1016/S1470-2045(19)30154-8
10. Oh S, Kim JH, Choi S-W, Lee HJ, Hong J, Kwon SH. Physician confidence in artificial intelligence: an online mobile survey. *J Med Internet Res.* (2019) 21:e12422. doi: 10.2196/12422
11. Esmaeilzadeh P. Use of AI-based tools for healthcare purposes: a survey study from consumer's perspectives. *BMC Med Inform Decis Mak.* (2020) 20:170. doi: 10.1186/s12911-020-01191-1
12. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* (2019) 17:195. doi: 10.1186/s12916-019-1426-2
13. Petitgand C, Motulsky A, Denis JL, Régis C. Investigating the barriers to physician adoption of an artificial intelligence- based decision support system

in emergency care: an interpretative qualitative study. *Stud Health Technol Inform.* (2020) 270:1001–5.
14. Singh RP, Hom GL, Abramoff MD, Campbell JP, Chiang MF. Current challenges and barriers to real-world artificial intelligence adoption for the healthcare system, provider, and the patient. *Transl Vis Sci Technol.* (2020) 9:45. doi: 10.1167/tvst.9.2.45
15. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol.* (2017) 2:230–43. doi: 10.1136/svn-2017-000101
16. Sarwar S, Dent A, Faust K, Richer M, Djuric U, Van Ommeren R, et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. *NPJ digital medicine.* (2019) 2:28. doi: 10.1038/s41746-019-0106-0
17. Park CJ Yi PH, Siegel EL. Medical student perspectives on the impact of artificial intelligence on the practice of medicine. *Curr Probl Diagn Radiol.* (2020) 50:614–9. doi: 10.1067/j.cpradiol.2020.06.01
18. Sit C, Srinivasan R, Amlani A, Muthuswamy K, Azam A, Monzon L, et al. Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: a multicentre survey. *Insights Imaging.* (2020) 11:14. doi: 10.1186/s13244-019-0830-7
19. Insider SD-V. *Artificial Intelligence (AI) is developing rapidly in Vietnam*. (2021). Available online at: https://vietnaminsider.vn/artificial-intelligence-ai-is-developing-rapidly-in-vietnam/
20. Health VMo. *Decision No. 4888/QD-BYT Introducing the Scheme for Application and Development of Smart Healthcare Information Technology for the 2019 - 2025 Period*. Hanoi, Vietnam: Ministry of Health (2019).
21. Vuong QH, Ho MT, Vuong TT, La VP, Ho MT, Nghiem KP, et al. Artificial intelligence vs. natural stupidity: evaluating AI readiness for the Vietnamese medical information system. *J Clin Med.* (2019) 8:168. doi: 10.3390/jcm8020168
22. Kijsanayotin B, Pannarunothai S, Speedie SM. Factors influencing health information technology adoption in Thailand's community health centers: applying the UTAUT model. *Int J Med Inform.* (2009) 78:404–16. doi: 10.1016/j.ijmedinf.2008.12.005
23. AbuShanab E, Pearson JM. Internet banking in Jordan: the unified theory of acceptance and use of technology (UTAUT) perspective. *J Syst Inf Technol.* (2007) 9:78–97. doi: 10.1108/13287260710817700
24. Wang YS, Wu MC, Wang HY. Investigating the determinants and age and gender differences in the acceptance of mobile learning. *Br J Educ Technol.* (2009) 40:92–118. doi: 10.1111/j.1467-8535.2007.00809.x
25. Kim S, Lee K-H, Hwang H, Yoo S. Analysis of the factors influencing healthcare professional's adoption of mobile electronic medical record (EMR) using the unified theory of acceptance and use of technology (UTAUT) in a tertiary hospital. *BMC Med Inform Decis Mak.* (2015) 16:1–12. doi: 10.1186/s12911-016-0249-8
26. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of information technology: toward a unified view. *MIS Quarterly.* (2003) 27:425–78. doi: 10.2307/30036540
27. Venkatesh V, Thong JYL, Chan FKY, Hu PJ-H, Brown SA. Extending the two-stage information systems continuance model: incorporating UTAUT predictors and the role of context. *Inf Syst J.* (2011) 21:527–55. doi: 10.1111/j.1365-2575.2011.00373.x
28. Fan W, Liu J, Zhu S, Pardalos PM. Investigating the impacting factors for the healthcare professionals to adopt artificial intelligence-based medical diagnosis support system (AIMDSS). *Ann Oper Res.* (2020) 294:567–92. doi: 10.1007/s10479-018-2818-y
29. Cimperman M, Makovec Brenčič M, Trkman P. Analyzing older user's home telehealth services acceptance behavior-applying an Extended UTAUT model. *Int J Med Inform.* (2016) 90:22–31. doi: 10.1016/j.ijmedinf.2016.03.002
30. Pinto Dos Santos D, Giese D, Brodehl S, Chon SH, Staab W, Kleinert R, et al. Medical student's attitude towards artificial intelligence: a multicentre survey. *Eur Radiol.* (2019) 29:1640–6. doi: 10.1007/s00330-018-5601-1
31. Wood EA, Ange BL, Miller DD. Are we ready to integrate artificial intelligence literacy into medical school curriculum: students and faculty survey. *J Med Educ Curric Dev.* (2021) 8:23821205211024078. doi: 10.1177/23821205211024078
32. Cho SI, Han B, Hur K, Mun JH. Perceptions and attitudes of medical students regarding artificial intelligence in dermatology. *J Eur Acad Dermatol Venereol.* (2021) 35:e72–3. doi: 10.1111/jdv.16812
33. Mehta N, Harish V, Bilimoria K, Morgado F, Ginsburg S, Law M, et al. Knowledge of and attitudes on artificial intelligence in healthcare: a provincial survey study of medical students. *medRxiv*. (2021). doi: 10.1101/2021.01.14.21249830
34. Gao S, He L, Chen Y, Li D, Lai K. Public perception of artificial intelligence in medical care: content analysis of social media. *J Med Internet Res.* (2020) 22:e16649. doi: 10.2196/16649
35. Dumić-Cule I, Orešković T, Brkljačić B, KujundŽić Tiljak M, Orešković S. The importance of introducing artificial intelligence to the medical curriculum - assessing practitioner's perspectives. *Croat Med J.* (2020) 61:457–64. doi: 10.3325/cmj.2020.61.457
36. Yun D, Xiang Y, Liu Z, Lin D, Zhao L, Guo C, et al. Attitudes towards medical artificial intelligence talent cultivation: an online survey study. *Ann Transl Med.* (2020) 8:708. doi: 10.21037/atm.2019.12.149
37. Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ.* (2019) 7:e7702. doi: 10.7717/peerj.7702
38. Gallupe RB, DeSanctis G, Dickson GW. Computer-based support for group problem-finding: an experimental investigation. *MIS Quarterly.* (1988) 12:277–96. doi: 10.2307/248853
39. Agarwal R, Prasad JA. Conceptual and operational definition of personal innovativeness in the domain of information technology. *Inf Syst Res.* (1998) 9:204–15. doi: 10.1287/isre.9.2.204
40. Goodhue DL, Thompson RL. Task-technology fit and individual performance. *MIS Quarterly.* (1995) 19:213–36. doi: 10.2307/249689
41. Wu I-L, Li J-Y, Fu C-Y. The adoption of mobile healthcare by hospital's professionals: an integrative perspective. *Decis Support Syst.* (2011) 51:587–96. doi: 10.1016/j.dss.2011.03.003
42. Zhou T, Lu Y, Wang B. Integrating TTF and UTAUT to explain mobile banking user adoption. *Comput Human Behav.* (2010) 26:760–7. doi: 10.1016/j.chb.2010.01.013
43. McKnight DH. *Trust in Information Technology. The Blackwell Encyclopedia of Management*. Oxford: Blackwell (2005). p. 329–31.
44. Benbasat I, Wang W. Trust In and Adoption of Online Recommendation Agents. *J Assoc Inf Syst.* (2005) 6:4. doi: 10.17705/1jais.00065
45. Yan H, Pan K. Examining mobile payment user adoption from the perspective of trust transfer. *Int. J. Netw. Virtual Organ.* (2015) 15:136–51. doi: 10.1504/IJNVO.2015.070423
46. Chiu C-M, Hsu M-H, Lai H, Chang C-M. Re-examining the influence of trust on online repeat purchase intention: the moderating role of habit and its antecedents. *Decis Support Syst.* (2012) 53:835–45. doi: 10.1016/j.dss.2012.05.021
47. Bansal G, Zahedi FM, Gefen D. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decis Support Syst.* (2010) 49:138–50. doi: 10.1016/j.dss.2010.01.010
48. Li X, Hess TJ, Valacich J. Why do we trust new technology? A study of initial trust formation with organizational information systems. *J Strateg Inf Syst.* (2008) 17:39–71. doi: 10.1016/j.jsis.2008.01.001
49. Li X, Hess TJ, Valacich JS. Using attitude and social influence to develop an extended trust model for information systems. *SIGMIS Database.* (2006) 37:108–24. doi: 10.1145/1161345.1161359
50. Kelman HC. Compliance, identification, and internalization: three processes of attitude change. *J Conflict Resolut.* (1958) 2:51–60. doi: 10.1177/002200275800200106
51. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* (2019) 25:44–56. doi: 10.1038/s41591-018-0300-7
52. Gong B, Nugent JP, Guest W, Parker W, Chang PJ, Khosa F, et al. Influence of artificial intelligence on Canadian medical student's preference for radiology specialty: a national survey study. *Acad Radiol.* (2019) 26:566–77. doi: 10.1016/j.acra.2018.10.007

53. Johnston SC. Anticipating and training the physician of the future: the importance of caring in an age of artificial intelligence. *Acad Med.* (2018) 93:1105–6. doi: 10.1097/ACM.0000000000002175
54. Stokes F, Palmer A. Artificial intelligence and robotics in nursing: ethics of caring as a guide to dividing tasks between AI and humans. *Nurs Philos.* (2020) 21:e12306. doi: 10.1111/nup.12306
55. Kim JW, Jones KL, D'Angelo E. How to prepare prospective psychiatrists in the era of artificial intelligence. *Acad Psychiatry.* (2019) 43:337–9. doi: 10.1007/s40596-019-01025-x
56. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. *JMIR Med Educ.* (2019) 5:e16048. doi: 10.2196/16048
57. Gomolin A, Netchiporouk E, Gniadecki R, Litvinov IV. Artificial intelligence applications in dermatology: where do we stand? *Front Med (Lausanne).* (2020) 7:100. doi: 10.3389/fmed.2020.00100
58. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ global health.* (2018) 3:e000798. doi: 10.1136/bmjgh-2018-000798

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.

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