



Semi-Supervised Support Vector Machine for Digital Twins Based Brain Image Fusion

Zhibo Wan¹, Youqiang Dong², Zengchen Yu¹, Haibin Lv³ and Zhihan Lv^{1*}

¹ College of Computer Science and Technology, Qingdao University, Qingdao, China, ² R&D Department, Qingdao Haily Measuring Technologies Co., Ltd., Qingdao, China, ³ North China Sea Offshore Engineering Survey Institute, Ministry Of Natural Resources North Sea Bureau, Qingdao, China

OPEN ACCESS

Edited by:

Yizhang Jiang,
Jiangnan University, China

Reviewed by:

Jing Xue,
Wuxi People's Hospital Affiliated
to Nanjing Medical University, China
Lu Wang,
Shandong Agricultural University,
China

*Correspondence:

Zhihan Lv
lvzhihan@gmail.com

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 05 May 2021

Accepted: 07 June 2021

Published: 09 July 2021

Citation:

Wan Z, Dong Y, Yu Z, Lv H and
Lv Z (2021) Semi-Supervised Support
Vector Machine for Digital Twins
Based Brain Image Fusion.
Front. Neurosci. 15:705323.
doi: 10.3389/fnins.2021.705323

The purpose is to explore the feature recognition, diagnosis, and forecasting performances of Semi-Supervised Support Vector Machines (S3VMs) for brain image fusion Digital Twins (DTs). Both unlabeled and labeled data are used regarding many unlabeled data in brain images, and semi supervised support vector machine (SVM) is proposed. Meantime, the AlexNet model is improved, and the brain images in real space are mapped to virtual space by using digital twins. Moreover, a diagnosis and prediction model of brain image fusion digital twins based on semi supervised SVM and improved AlexNet is constructed. Magnetic Resonance Imaging (MRI) data from the Brain Tumor Department of a Hospital are collected to test the performance of the constructed model through simulation experiments. Some state-of-art models are included for performance comparison: Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), AlexNet, and Multi-Layer Perceptron (MLP). Results demonstrate that the proposed model can provide a feature recognition and extraction accuracy of 92.52%, at least an improvement of 2.76% compared to other models. Its training lasts for about 100 s, and the test takes about 0.68 s. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) of the proposed model are 4.91 and 5.59%, respectively. Regarding the assessment indicators of brain image segmentation and fusion, the proposed model can provide a 79.55% Jaccard coefficient, a 90.43% Positive Predictive Value (PPV), a 73.09% Sensitivity, and a 75.58% Dice Similarity Coefficient (DSC), remarkably better than other models. Acceleration efficiency analysis suggests that the improved AlexNet model is suitable for processing massive brain image data with a higher speedup indicator. To sum up, the constructed model can provide high accuracy, good acceleration efficiency, and excellent segmentation and recognition performances while ensuring low errors, which can provide an experimental basis for brain image feature recognition and digital diagnosis.

Keywords: semi-supervised support vector machines, brain image, digital twins, image segmentation, improved AlexNet

INTRODUCTION

Now that science and technology advances quickly, Artificial Intelligence (AI), Big Data (BD), the Internet of Things (IoT), and Deep Learning (DL) have been accepted in all walks of life. Health is vital to the happy life of people. Medical intelligentization has improved health monitoring techniques notably. Brain tumors have severely harmed people's health.

Statistics suggest that the annual mortality rate of brain cancer in China is second only to gastric cancer and lung cancer. About 80,000 people die from brain cancer each year, accounting for about 22.5% of the world's total number of brain cancers (Zhou et al., 2019; Kang et al., 2021). In only a few decades of the 21st century, the number of people diagnosed with brain cancer continues to increase, especially in Asian countries (Zheng et al., 2020). Thus, brain health has become the focus of brain disease experts, researchers, and brain surgeons.

Medical imaging techniques play a vital role in clinical diagnosis with the successive introduction of medical imaging equipment and the continuous development of computer science and technology. Results of brain disease diagnosis depend on brain image acquisition and accurate interpretation. However, due to the large number of patient groups, the number of brain and other medical images keeps increasing dramatically (Tobon Vasquez et al., 2020). Manual image reading is apparently inefficient, and the workload for clinicians is unbearable. Therefore, using computer-aided diagnosis to extract features from medical images is of great significance in improving diagnostic accuracy and reducing clinicians' workload.

Image segmentation is one of the necessary steps of brain image fusion, so it is necessary to segment and extract the features of the image. There are many approaches to extract brain image features; currently trendy ones include Region Growing (RG), Level Set, Fuzzy Clustering, and Machine Learning (ML) (Aheleroff et al., 2020; Angin et al., 2020; Li et al., 2020; Pan et al., 2021). However, these approaches require manual image segmentation, consuming lots of human resources and materials. As a burgeoning AI algorithm, ML extracts data features and investigates data relationships through self-learning on brain images, which dramatically increases image segmentation accuracy and stability while improving efficiency. As Digital Twins (DTs) evolve, their applications in brain image segmentation have enhanced data generation speed and scale, leading to a remarkable increase in big data processing complexity and brain image analysis (Gudigar et al., 2019). Digital twins can map complex brain images from real space to virtual space, and it is of great significance to extract features, segment, and fuse images.

To sum up, AI applications in medical fields, such as brain image processing, can help clinicians choose surgical treatment or radiotherapy and chemotherapy treatment options for patients while avoiding the steps of direct surgical resection of all tumors, which greatly reduces the risk of treatment. The innovative points are (1) the Semi-Supervised Support Vector Machines (S3VMs) proposed regarding the massive volume of unlabeled brain image data, (2) the improved AlexNet, and (3) the brain image fusion DTs diagnosis and forecasting model constructed based on

S3VMs and improved AlexNet, in an effort to provide theoretical support for brain image feature recognition and digital diagnosis.

RECENT WORKS

ML Applications

Now that complex data grow explosively, big data mining and analysis functions, especially DL, are demanded increasingly, which has been explored by many scholars. Yan et al. (2018) put forward Device ElectroCardioGram (DECG) and constructed a forecasting algorithm for the remaining service life of industrial equipment based on Deep Denoising Autoencoder (DDA) and regression calculation. Comparison with traditional factory information models validated the feasibility and effectiveness of the proposed algorithm (Yan et al., 2018). Wu et al. (2019) adopted deep learning algorithms in medical imaging to understand patients' physical condition, thereby improving the disease treatment. Ghosh et al. (2019) developed a DL approach for molecular excitation spectroscopy forecasting. They trained and evaluated Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), and Deep Tensor Neural Network (DTNN) to analyze the electronic state density of organic molecules. Eventually, they discovered that real-time spectroscopic forecasting could be performed on the structure of tiny organic molecules, and potential application molecules could be determined (Ghosh et al., 2019). Shao and Chou (2020) developed a CNN-based forecasting factor for the subcellular localization of viral proteins, called Ploc-Deep-mVirus, for the Coronavirus Disease 2019 (COVID-19) currently spreading globally. They found that this forecasting factor was particularly suitable for processing multi-site systems, and its forecasting performance was remarkably better than other advanced forecasting indicators (Shao and Chou, 2020). Ahmed et al. (2020) applied DL approaches to the clinical or behavioral recognition of Autism Spectrum Disorder (ASD). They designed an image generator, which could produce a single-volume brain image from the entire brain image by separately considering the voxel time point of each subject. Eventually, performances of four different DL approaches, as well as their corresponding integrated classifiers and the algorithms, were assessed. Results demonstrated a better effect on a large-scale multi-site brain imaging dataset (bedict) (Ahmed et al., 2020).

Research Progress of Image Fusion DTs

Qi and Tao (2018) applied DTs to image information physical integration of the manufacturing industry. They compared the similarities and differences between BD and DTs from the overall and data perspectives. Consequently, they discussed using DTs to promote smart manufacturing (Qi and Tao, 2018). As embedded sensors, BD processing, and cloud computing got advanced, Barricelli et al. (2019) applied DTs to medical image diagnosis and recognition, and consequently, validated its effectiveness and feasibility. Liu et al. (2019) put forward a Cloud DTs Healthcare (CloudDTH) architecture to monitor, diagnose, and forecast image data collected by wearable

medical devices to forecast individual health status, thereby achieving individual health management, especially the health management of seniors. Li et al. (2021) proposed a novel semantic-enhanced DTs system for accurate and fine-grained real-time monitoring of Robot-Environment Interaction (REI). This system adopted the robot's external perception and ontology perception as the mechanism of interaction state monitoring and semantic reasoning. Meanwhile, they proposed a multi-feature fusion visual relationship detector and introduced a lightweight spatial relationship dataset to measure the visual relationship between entity pairs. They evaluated the real-time monitoring ability and semantic reasoning ability of the system and verified the effectiveness and feasibility of the scheme (Li et al., 2021).

The above works reveal that most ML algorithms applied in the medical field collect image data from historical cases. In contrast, DTs are mappings from virtual space to real cases. There are very few combinations of DTs and ML in the medical field. Hence, ML approaches and DTs are introduced for brain image fusion in the present work, which is of great significance to the accurate identification and diagnosis of brain images.

ML-Based Brain Image Fusion DTs Diagnosis and Forecasting

At present, although many achievements have been made in the fusion of medical images such as brain tumors, the results are still unsatisfactory, and the application effect in clinical medicine is extremely limited. As one of the effective methods of brain image fusion, image segmentation technology has great practical value and research significance in the diagnosis, prediction, and feature extraction of brain tumor and other medical images.

Demand Analysis of Brain Image Fusion DTs Diagnosis and Forecasting

With the aging of the population, the number of patients is increasing, and the number of medical images is growing. It has become a new direction to diagnose and treat diseases by using computer technology such as digital twins to assist the recognition and prediction of brain images. Sun et al. (2020) at present, common medical imaging techniques include X-ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Ultrasound Imaging (UI), Positron Emission Tomography (PET), and Molecular Imaging (MI) (Wang et al., 2020). MRI is very suitable for brain tumor research and diagnosis because of its non-invasiveness, high soft-tissue contrast, high resolution, and no direction limitations compared with other imaging techniques (Liu Y. et al., 2018).

Regarding brain image processing, excellent brain image segmentation results are of great significance for further target feature extraction and target recognition and analysis. Brain image segmentation differs from ordinary image segmentation in image complexity and diversity (Wang et al., 2018). Brain image segmentation is a crucial branch of medical image segmentation. Key difficulties are as follows: (1) the gray scale distribution between different tissues is not uniform, and the distribution of different tissues overlaps with each other. (2) Brain tissues have

complicated structures. Generally, brain MRI images include the brain cortex, gray matter, white matter, cerebrospinal fluid, and other tissues; each has a complicated shape and structure. (3) Boundaries of brain tumors are blurry, and the density distribution is uneven. Thus, in the analysis of brain image processing, the brain image and structure in real space can be mapped to virtual space by digital twins. After the virtual brain image feature extraction and image fusion processing, the treatment scheme for patients with brain diseases can be formulated, and the treatment purpose for patients can be achieved finally. Present brain image segmentation approaches include supervised learning, unsupervised learning, and DL, as illustrated in **Figure 1**.

To improve the effect of brain image segmentation diagnosis, both unlabeled and labeled data are adopted to put forward an S3VMs regarding the massive volume of unlabeled brain image data based on the above approaches of brain image segmentation diagnosis. S3VMs combine supervised learning and unsupervised learning, which can utilize labeled data and unlabeled data simultaneously to improve the model's generalization ability (Liu L. et al., 2018; Zavrak and İskefiyeli, 2020). Combining semi-supervised learning and Support Vector Machines (SVMs) based on labeled samples obtained by time-domain simulation and the unlabeled samples from the actual system to assess brain image data can extract actual data distribution comprehensively and improve the weak generalization ability of present supervised feature extraction and assessment models, in an effort to enhance model's adaptability to the actual system. This idea can solve the shortcomings of the current supervised ML feature extraction models and the starting point of the present work.

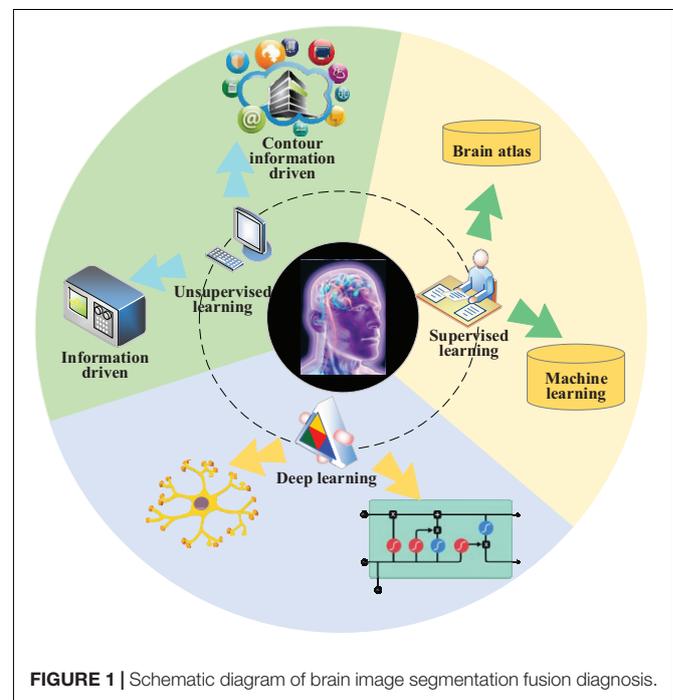


FIGURE 1 | Schematic diagram of brain image segmentation fusion diagnosis.

Performance of S3VMs in Brain Image Feature Extraction

Semi-supervised learning combines supervised learning and unsupervised learning. It utilizes both labeled data and unlabeled data for learning and mine the data distribution hidden in unlabeled samples such as brain images, thereby guiding the training of the classifier and improving its performance.

A graph-based semi-supervised learning algorithm is adopted to express the relationships between the entire training data during brain image feature extraction. Graph nodes represent training samples, including labeled samples and unlabeled samples. There are edges between nodes; these edges are given weights to represent the similarity between data nodes and can be obtained using various distance metrics. Overall, the greater the weight, the greater the similarity. No connection between nodes means zero similarity between them. Similar vertices are given the same category labels as much as possible so that the graph distribution is as smooth as possible.

The mathematical model of standard SVM can be formalized as a convex quadratic programming problem; then, a linear classification machine can be obtained through hard interval maximization learning (Zhang et al., 2020). If the training data are non-linear and inseparable, the kernel function will be utilized to map the input space to the feature space, and the linear SVM is implicitly learned in the high-dimensional feature space by maximizing the soft interval. The linearly separable optimal classification hyperplane is demonstrated in **Figure 2**.

As shown in **Figure 2**, a linearly separable sample set $T = \{(x_1, y_1), \dots, (x_n, y_n)\}$ is given, where (x_i, y_i) refers to the sample point, $x_i \in R^n$, and $y_i \in \{+1, -1\}$, $i = 1, 2, \dots, n$. The red and the black circles represent two types of training samples, respectively. At this point, there is an infinite number of interfaces that can correctly separate the two types of samples. Hence, an optimal classification hyperplane is required, which should not only ensure the minimum empirical risk and correctly divide the two types of samples but also minimize the actual risk and meet the maximum classification interval. The core idea of linear separable SVM employs the maximum classification interval to find the optimal hyperplane, at which the solution is unique. In **Figure 2**, H denotes the optimal classification

hyperplane, H_1 and H_2 are the planes containing the closest sample points (support vectors) to the classification hyperplane. H_1 and H_2 are parallel to H , and the distance between them is called the classification interval. The classification hyperplane function of linear separable SVM is:

$$w \cdot x + b = 0 \tag{1}$$

In Eq. 1, w represents the normal vector of the hyperplane, and b denotes the bias. The following conditions must be met to classify samples accurately and ensure the maximum interval simultaneously:

$$y_i (w \cdot x + b) \geq 1 \tag{2}$$

The classification interval can be obtained as $2/\|w\|$. Solving the optimal hyperplane of the training sample set T can be transformed into a problem described by the following equation:

$$\begin{aligned} \min_{a,b} & \frac{\|w\|^2}{2} \\ \text{s.t.} & y_i (w \cdot x + b) \geq 1, i = 1, 2, \dots, n \end{aligned} \tag{3}$$

Equation (3) is a quadratic programming problem, which can be solved as a dual Lagrangian problem. Introducing the Lagrangian function can obtain:

$$\begin{aligned} L(w, b, \alpha) &= \frac{\|w\|^2}{2} - \sum_{i=1}^n \alpha_i [y_i (w \cdot x + b) - 1] \\ \text{s.t.} & \alpha_i \geq 1, i = 1, 2, \dots, n \end{aligned} \tag{4}$$

In Eq. 4, α_i refers to the Lagrangian multiplier. The partial derivative of the Lagrangian function can be obtained. According to the extreme conditions:

$$\begin{aligned} \frac{\partial L}{\partial b} = 0 &\rightarrow \sum_{i=1}^n y_i \alpha_i = 0 \\ \frac{\partial L}{\partial w} = 0 &\rightarrow w = \sum_{i=1}^n y_i \alpha_i x_i \end{aligned} \tag{5}$$

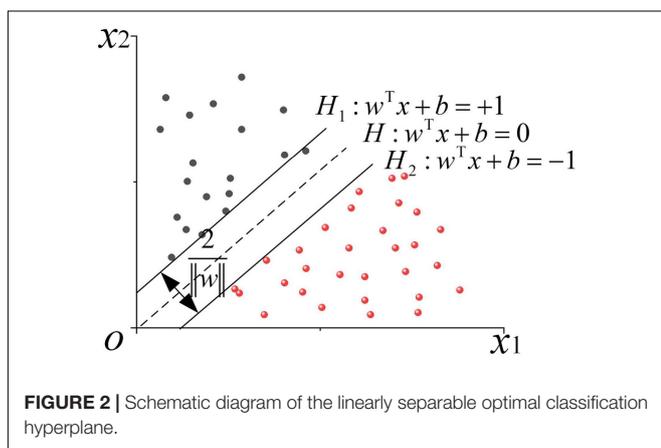
Substituting equation 5 into 4 can get the dual problem about the original problem:

$$\begin{aligned} \min_{\alpha} & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j (x_i \cdot x_j) - \sum_{i=1}^n \alpha_i \\ \text{s.t.} & \sum_{i=1}^n y_i \alpha_i = 0, \alpha_i y_i \geq 0, i = 1, 2, \dots, n \end{aligned} \tag{6}$$

Solving Eq. 6 can get the optimal solution $\alpha^* = (\alpha_1^*, \dots, \alpha_n^*)^T$. Calculating w^* and b^* can obtain the optimal classification function:

$$f(x) = \text{sgn} \left[\sum_{i=1}^n \alpha_i^* y_i (x_i \cdot x) + b^* \right] \tag{7}$$

According to Eq. 7, the optimal classification plane of SVM only depends on the sample point (x_i, y_i) corresponding to $\alpha_i^* > 0$ in the training data. SVM's optimal classification plane only depends on corresponding sample points in the training data, and these



samples just fall on the interval boundary. Usually, the sample point $x_i \in R^n$ at $\alpha_i^* > 0$ is called the support vector. The graph feature extraction of S3VMs is analyzed further.

A training set including labeled data and unlabeled data $\{(x_i, y_i)_{i=1}^l, (x_j)_{j=l+1}^{l+u}\}$ is given. Suppose that $g = (V, E)$ refers to a graph, V represents the node set, and E describes the edge set. The weight matrix W of g is symmetric, which can be described as:

$$W_{ij} = \begin{cases} w(e) = w_{ij} = 0 & \text{if } e = (x_i, x_j) \in E \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

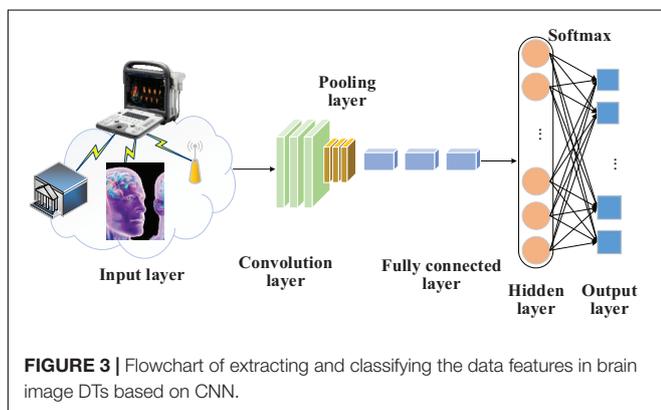
The weight of the edge $e = (x_i, x_j)$ $w(e) = w_{ij} = w_{ji}$ signifies the similarity between nodes x_i, x_j , which can be calculated using K Nearest Neighbor (KNN) graph. Each vertex in the KNN graph is only connected to its k nearest neighbors under a particular distance (such as Euclidean distance). Notably, a vertex x_i that is the k nearest neighbor of x_j does not necessarily indicate that x_j is the k nearest neighbor of x_i . While constructing the KNN graph, if x_i and x_j are k -nearest neighbors of each other, x_i and x_j will be connected by an edge, and the corresponding weights will be calculated using the Radial Basis Function (RBF) kernel:

$$W_{ij} = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (9)$$

In Eq. 9, σ refers to the kernel function that controls the weight reduction speed. When $x_i = x_j$, the weight $w_{ij} = 1$; when $\|x_i - x_j\|$ approaches infinite, the weight $w_{ij} = 0$.

To extract and forecast the obtained brain image features, CNN is introduced into the brain image feature extraction and diagnosis system. The proposed model can minimize the preprocessing workload and directly extract the most expressive features from the original data input without manual feature designation (Ogino et al., 2021). **Figure 3** illustrates how CNN extracts and classifies the features of the graph data in brain image DTs.

Convolutional Neural Network pooling layer performs down-sampling operations on the input feature maps in length and width dimensions. It can reduce the model parameters by down-sampling the input feature data. According to the number of parameters, the data complexity of different brain images can be



reduced, diminishing the over-fitting degree and the probability of local minimum. Moreover, the pooling layer can make the model more robust to translation and distortion in the image.

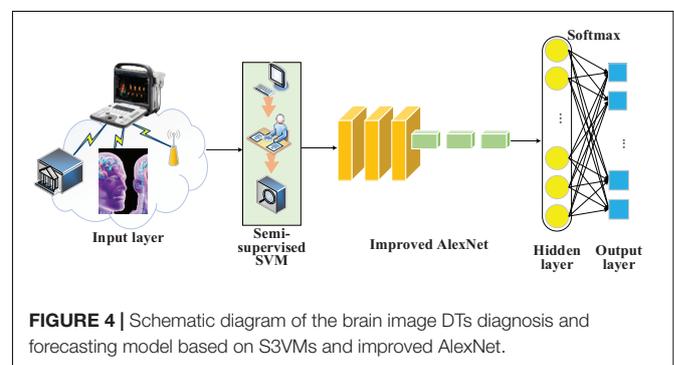
Data in CNN go through multiple convolutional layers and pooling layers; then, they are connected via one or more fully connected layers. All neurons in the current layer are connected to those in the last layer. Usually, this layer depends on two 1D network layers. Local information of the convolutional layer or the pooling layer is grouped together. The activation function used by all neurons is often the Rectified Linear Unit (ReLU).

AlexNet (Tanveer et al., 2019), a deep CNN model, is selected considering its multiple network layers and stronger learning ability to reduce the calculation amount and enhance CNN's generalization performance. Furthermore, the functional layer of AlexNet's convolutional layer is improved. The operation of "local normalization before pooling" is advanced to "pooling before local normalization." This improvement brings two benefits. First, the generalization ability of AlexNet can be enhanced while the over-fitting can be weakened, which greatly shortens the training time. Second, overlapping pooling before local normalization can reserve more data and weaken redundant information during pooling and accelerate the convergence rate of the brain image DTs diagnosis and forecasting model during training, highlighting its superiority over other pooling approaches.

Brain Image Fusion DTs Diagnosis and Forecasting Model Based on S3VMs and Improved AlexNet

Digital Twins (Sun et al., 2020) can forecast and analyze data of the physical space in the corresponding virtual space regarding the smart city development in physical spaces. To extract the data features from brain images, AlexNet, a deep CNN model, is selected considering its multiple network layers and stronger learning ability. Furthermore, the functional layer of AlexNet's convolutional layer is improved. A brain image DTs diagnosis and forecasting model is designed based on S3VMs and improved AlexNet while considering the model's security performance, as demonstrated in **Figure 4**.

In this diagnosis and forecasting model, different penalty coefficients are provided considering the different contribution degrees of the labeled and unlabeled data samples in the brain



images to the hinge loss. The objective function can be written as the following equation:

$$\min_{w,b,s,y_u} \frac{1}{2} w^T w + \lambda_1 \sum_{i=1}^n u_{ii} \left(1 - (w^T x^i + b) y_i \right) + \lambda_2 \text{tr} (w^T X L_s X^T w) + \lambda_3 \|s\|_2^2 + \lambda_4 \text{tr} (F^T L_s F) \quad (10)$$

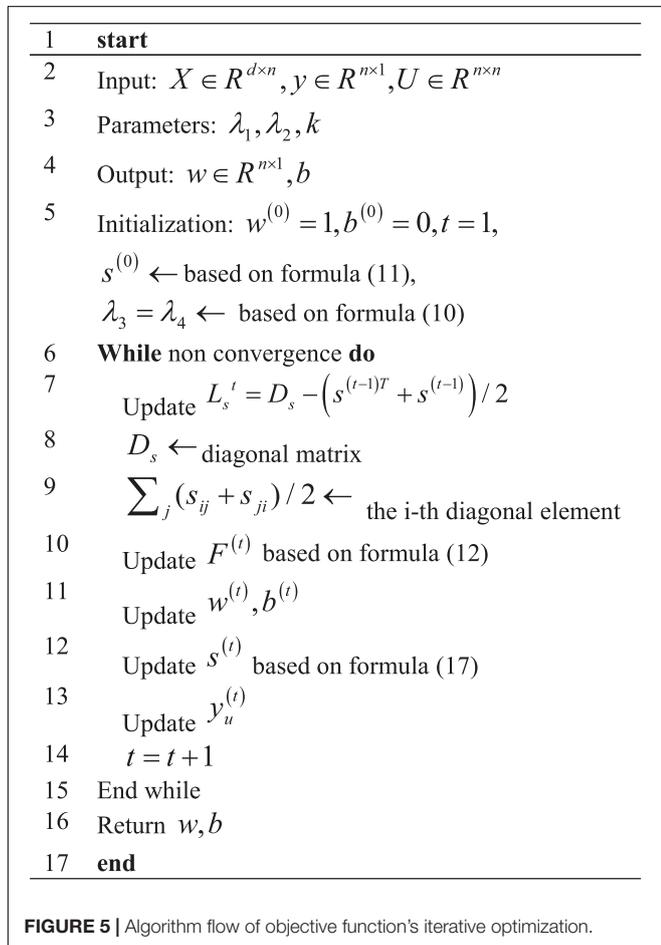
s.t. $\forall i, s_i \geq 0, F \in R^{n \times c}, F^T F = I$

In Eq. 9, s signifies a similarity matrix between samples, $L_s = D_s - \frac{s^T + s}{2}$, and $D_s \in R^{n \times n}$ refers to a diagonal element, the diagonal matrix of $\sum_j (s_{ij} + s_{ji})/2 \leftarrow y_u^{(t)} t = t + 1$. The above objective function can be solved by iterative optimization. The optimization process is described in **Figure 5**.

In this algorithm flow, $\lambda_3 = \lambda_4$ and s are initialized:

$$\lambda_3 = \lambda_4 = \frac{1}{n} \sum_{i=2}^n \left(\frac{k}{2} d_{i,k+1}^x - \frac{1}{2} \sum_{j=1}^k d_{ij}^x \right) \quad (11)$$

$$s_{ij} = \begin{cases} 0, & x^j \notin N_k(x^i) \\ \frac{d_{i,k+1}^x - d_{ij}^x}{k d_{i,k+1}^x - \sum_{j=1}^k d_{ij}^x}, & x^j \in N_k(x^i) \end{cases} \quad (12)$$



Here, $d_{ij}^x = \|x^i - x^j\|_2^2$. The smallest top k d_{ij}^x are found by sorting in increasing order, and $N_k(x^i)$ refers to the k samples nearest to x^i . $F^{(t)}$ is updated as the following equation:

$$\min_F \text{tr} (F^T L_s F) \quad (13)$$

s.t. $F \in R^{n \times c}, F^T F = I$

Eigenvectors corresponding to the smallest c eigenvalues of the Laplacian matrix L_s can form F .

Solving $s^{(t)}$ only needs to solve Eq. 13:

$$\min_{s_i=1, s_i \geq 0} \sum_{i=1}^n \left(\lambda_2 \|w^T x^i - w^T x^j\|_2^2 s_{ij} + 2\lambda_3 s_{ij}^2 \right) + \sum_{i=1}^n \lambda_4 \|f_i - f_j\|_2^2 s_{ij} \quad (14)$$

Features of brain images are extracted. The t -th feature map $y_t^l(i, j)$ of the l -th convolutional layer is sampled using overlapping pooling:

$$a_t^l(i, j) = \max\{y_t^l(i, j), i_s \leq i \leq i_s + w_c - 1, \} \quad (15)$$

$j_s \leq j \leq j_s + w_c - 1$

In Eq. 14, s is the pooling movement step size, w_c refers to the width of the pooling area, and $w_c > s$.

A local normalization layer is added after the first and second pooling layers of AlexNet to standardize the feature map $c_t^l(i, j)$:

$$c_t^l(i, j) = a_t^l(i, j) / \left(k + \alpha \sum_{\max(0, t-m/2)}^{\min(N-1, t+m/2)} (a_t^l(i, j))^2 \right)^\beta \quad (16)$$

In Eq. 15, k, α, β, m are all hyperparameters valuing 2, 0.78, 10^{-4} , and 7, respectively, and N is the total number of convolution kernels in the l -th convolutional layer. To prevent "gradient dispersion" (Shen et al., 2019), the activation function takes Rectified Linear Unit (ReLU) to activate the convolution output $S_t^l(i, j)$:

$$y_t^l(i, j) = f(S_t^l(i, j)) = \max\{0, S_t^l(i, j)\} \quad (17)$$

In Eq. 16, $f(\cdot)$ represents ReLU. To prevent over-fitting in the fully connected layer, the dropout parameter is set to 0.5. All the feature maps when l values 5 in Eq. 12 are reconstructed into a high-dimensional single-layered neuron structure C^5 ; thus, the input Z_i^6 of the i -th neuron in the sixth fully connected layer is:

$$Z_i^6 = W_i^6 C^5 + b_i^6 \quad (18)$$

In Eq. 17, W_i^6 and b_i^6 are the weight and bias of the i -th neuron in the sixth fully connected layer, respectively.

While improving the generalization ability, the neurons C^l of the sixth and seventh fully connected layers are discarded and output, $r_j^l \sim \text{bernoulli}(dp)$, $\tilde{C}^l = r^l C^l$; in this regard, the i -th neuron's input in the seventh and eighth fully connected layers Z_i^{l+1} is $W_i^{l+1} \tilde{C}^l + b_i^{l+1}$, where the i -th neuron's input in the

sixth and seventh fully connected layers C_i^l is $f(Z_i^l)$, namely $\max\{0, Z_i^l\}$. Finally, the input q^i of the i -th neuron in the eighth fully connected layer can be obtained:

$$q^i = \text{soft max}(Z_i^8) = \frac{e^{Z_i^8}}{\sum_{j=1}^{12} e^{Z_j^8}} \quad (19)$$

The cross-entropy loss function suitable for classification is taken as the model's error function, and the equation is:

$$\text{Loss} = \sum_{i=1}^K y_i \cdot \log(p_i) \quad (20)$$

$$p_i = \frac{\exp(\tilde{y}_i)}{\sum_{j=1}^K \exp(\tilde{y}_j)} \quad (21)$$

In Eqs. 19 and 20, K denotes the number of categories, y_i describes the true category distribution of the sample, \tilde{y}_i describes the network output, and p_i represents the classification result after the SoftMax classifier. SoftMax's input is an N -dimensional real number vector, denoted as x . Its equation is:

$$\zeta(x)_i = \frac{e^{x_i}}{\sum_{n=1}^N e^{x_n}}, i = 1, 2, \dots, N \quad (22)$$

Essentially, SoftMax can map an N -dimensional arbitrary real number vector to an N -dimensional vector whose values all fall in the range of $(0,1)$, thereby normalizing the vector. To reduce the computational complexity, the output data volume is reduced to 2^8 through μ companding conversion, that is, $\mu = 255$, thereby improving the model's forecasting efficiency.

$$f(x_t) = \text{sign}(x_t) \frac{\ln(1 + \mu |x_t|)}{\ln(1 + \mu)}, |x_t| < 1 \quad (23)$$

The proposed model is trained through learning rate updating using the polynomial decay approach (Poly) (Zhao et al., 2020). The equation is:

$$\text{init_lr} \times \left(1 - \frac{\text{epoch}}{\text{max_epoch}}\right)^{\text{power}} \quad (24)$$

In Eq. 23, the initial learning rate init_lr is 0.0005 (or $5e^{-4}$), and power is set to 0.9.

The Weighted Cross-Entropy (WCE) is accepted as a cost function to optimize model training. Suppose that $z_k(x, \theta)$ describes the unnormalized logarithmic probability of pixel x in the k -th category under the given network parameter θ . In that case, the SoftMax function $p_k(x, \theta)$ is defined as:

$$p_k(x, \theta) = \frac{\exp\{z_k(x, \theta)\}}{\sum_{k'}^K \exp\{z_{k'}(x, \theta)\}} \quad (25)$$

In Eq. 24, K represents the total number of image categories. During forecasting, once Eq. 24 reaches the maximum, pixel x will be labeled as the k -th category, namely $k^* = \arg \max\{P_k(x, \theta)\}$. A semantic segmentation task needs to

sum the pixel data loss in each input mini-batch. Here, N denotes the total number of pixels in the training batch of image data, y_i refers to the real semantic annotation of the pixel x_i , and $p_k(x_i, \theta)$ describes the forecasted probability of pixel x_i belonging to the k -th semantic category, that is, the log-normalized probability, abbreviated as p_{ik} . Hence, the training process aims to find the optimal network parameter θ^* by minimizing the WCE loss function $\ell(x, \theta)$, denoted as $\theta^* = \min_{\theta} \ell(x, \theta)$. Training samples with unbalanced categories in brain images usually make the network notice some easily distinguishable categories, resulting in poor recognition on some more difficult samples. In this regard, the Online Hard Example Mining (OHEM) strategy (Ghafoori et al., 2018) is adopted to optimize the network training process. The improved loss function is:

$$\ell(x, \theta) = - \frac{1}{\sum_{i=1}^N \sum_{k=1}^K \delta(y_i = k, p_{ik} < \eta)} \sum_{i=1}^N \sum_{k=1}^K \delta(y_i = k, p_{ik} < \eta) \log p_{ik} \quad (26)$$

In Eq. 25, $\eta \in (0, 1]$ refers to the predefined threshold, and $\delta(\cdot)$ describes the symbolic function, which will take 1 if the condition is met and 0 otherwise. The weighted loss function for brain image fusion is defined as:

$$\ell(x, \theta) = - \sum_{i=1}^N \sum_{k=1}^K w_{ik} q_{ik} \log p_{ik} \quad (27)$$

In Eq. 26, $q_{ik} = q(y_i = k|x_i)$ denotes the true label distribution of the k -th category of the pixel x_i , w_{ik} refers to the weighting coefficient. The following strategy is employed during training:

$$w_{ik} = \frac{1}{\ln(c + p_{ik})} \quad (28)$$

In Eq. 27, c is an additional hyperparameter, set to 1.10 based on experience during simulation experiments.

Simulation Experiment

In the present work, MATLAB is adopted for simulation to validate the performance of the proposed model. Brain image data in experiments come from MRI records of the Brain Tumor Department of a Hospital. It contains 20 groups of clinical data, and each group of data includes four kinds of multi sequence MRI images: T1, T2, T1c, and FLAIR. The selected image data are processed by registration, skull peeling, contrast enhancement and other operations. Each case of MRI image contains the standard segmentation results given by artificial experts. Data collected are divided into a training dataset and a test dataset in 7:3; the ratio of each data type in the two datasets shall be consistent. Hyperparameters of AlexNet are set as follows: 120 iterations, 2,000 seconds of simulation, and 128 Batch Size. Some state-of-art models are included for performance comparison, including LSTM (Nguyen et al., 2021), CNN, RNN (Minerva et al., 2021), AlexNet, and MLP (Dong et al., 2019). Experimental environment configuration includes software and hardware. As

for software, the operating system is Linux 64bit, the Python version is 3.6.1, and the development platform is PyCharm. As for hardware, the Central Processing Unit (CPU) is Intel Core i7-7700@4.2 GHz 8 Cores, the internal memory is Kingston DDR4 2400 MHz 16G, and the Graphics Processing Unit (GPU) is NVIDIA GeForce 1060 8G.

The similarity between the experimental segmentation result and the expert segmentation result can measure the model's segmentation quality (Dai et al., 2020). Common similarity metrics are the Jaccard coefficient and Dice Similarity Coefficient (DSC). DSC indicates the similarity degree between the segmentation result of the model and the label labeled by the expert. The higher the DSC, the higher the segmentation similarity. Positive Predictive Value (PPV) represents the proportion of correct brain tumor points in the experimental segmentation to the tumor points in the segmentation result. Sensitivity describes the proportion of the correct tumor points in the experimental segmentation to the true value of the tumor points. The above indicators are adopted to assess the proposed brain image segmentation model. The Jaccard coefficient, DSC, PPV, and Sensitivity are defined as follows:

$$Jaccard(T, P) = \frac{|T \cap P|}{|T \cup P|} \quad (29)$$

$$DSC(T, P) = \frac{2|T \cap P|}{|T| + |P|} \quad (30)$$

$$PPV = \frac{|T \cap P|}{P} \quad (31)$$

$$Sensitivity = \frac{|T \cap P|}{T} \quad (32)$$

In Eqs. 28–31), T represents the expert segmentation result, and P denotes the segmentation result of the proposed model based on S3VMs and the improved AlexNet.

RESULTS AND DISCUSSION

Forecasting Performance Analysis

To explore the forecasting performance of the proposed model, LSTM, CNN, RNN, AlexNet, and MLP are included for comparison. Values of Accuracy, Precision, Recall, and F1 of these models are compared through simulation experiments, and the results are illustrated in **Figure 6**. Furthermore, time durations required for training and test and errors are compared, as in **Figures 7, 8**.

As shown in **Figure 6**, the proposed model can provide a recognition accuracy of 92.52%, at least an improvement of 2.76% than other models. The Precision, Recall, and F1 of the proposed model are the highest, at least 3.35% higher than other models. Hence, the proposed brain image DTs diagnosis and forecasting model based on S3VMs and improved AlexNet can provide excellent recognition and forecasting accuracy.

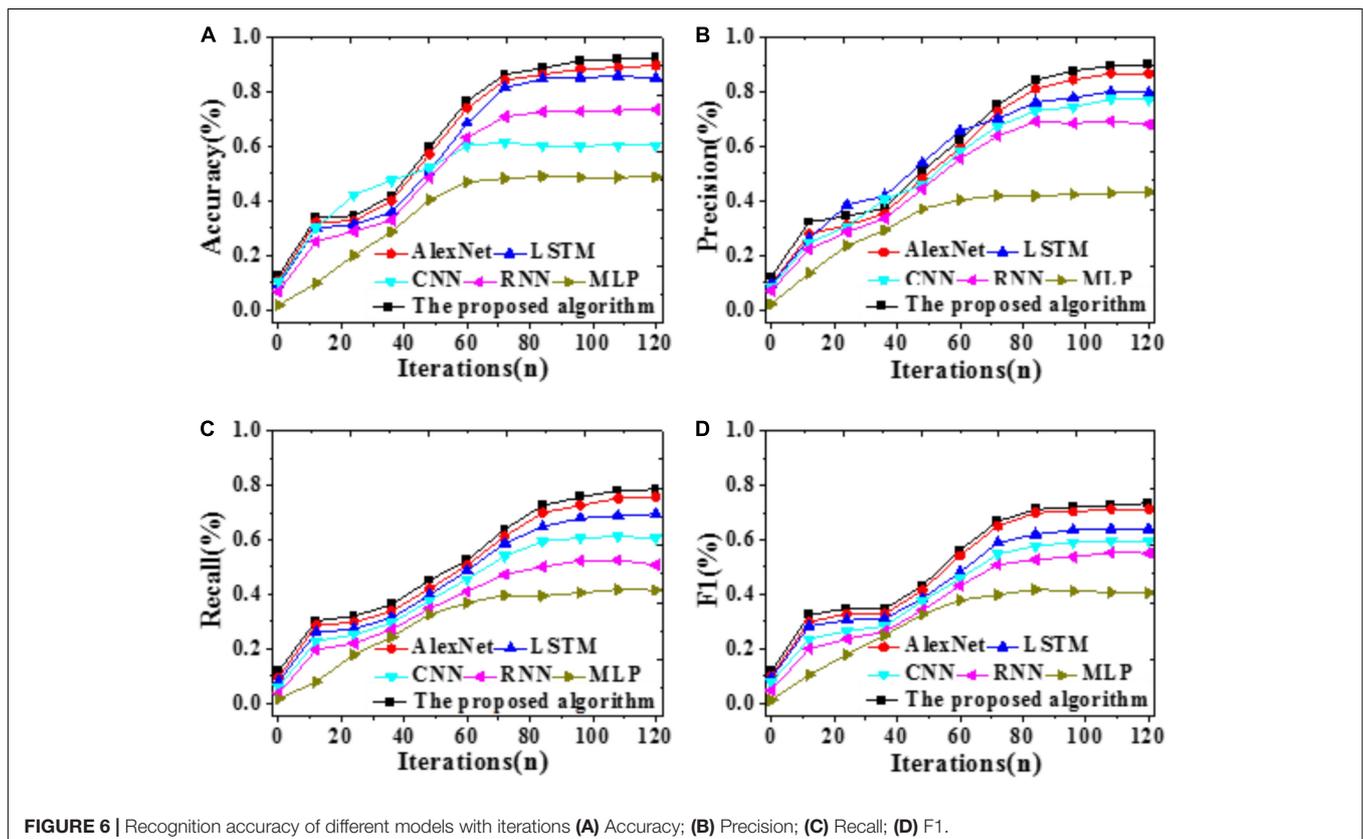
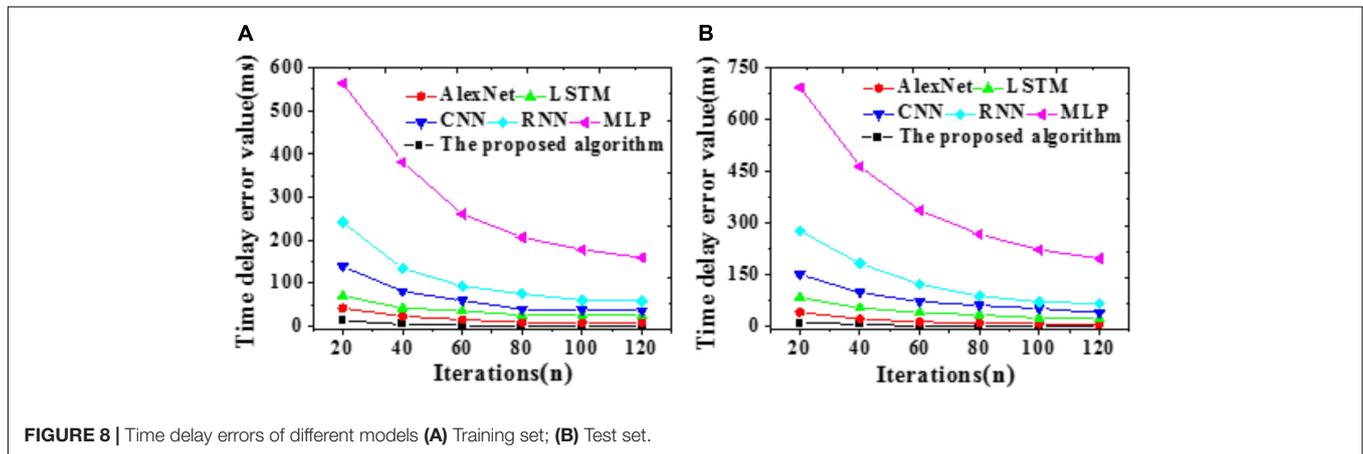
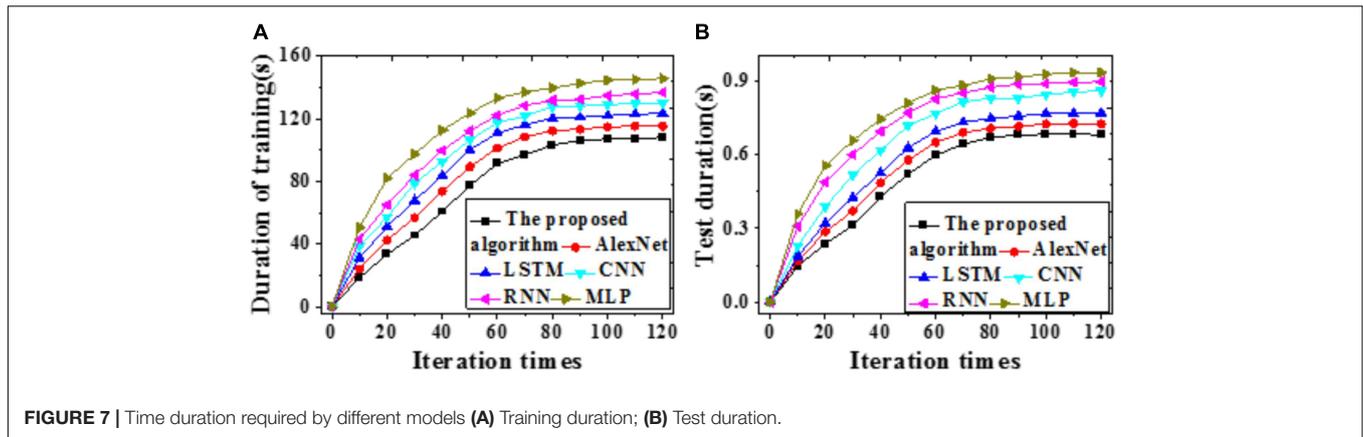


FIGURE 6 | Recognition accuracy of different models with iterations (A) Accuracy; (B) Precision; (C) Recall; (D) F1.



According to **Figures 7, 8**, the time required by all models first increases then stabilizes with iterations. Compared with LSTM, AlexNet, CNN, RNN, and MLP, the proposed model requires remarkably less time for training and test. A possible reason is that the improved AlexNet provides enhanced generalization ability while accelerating the convergence velocity of the model training process. Therefore, the proposed model can achieve higher forecasting accuracy more quickly.

In both the training and test sets, errors reduce gradually with iterations. MLP provides the longest time delay, 564.15 ms and 693.06 ms, respectively. In contrast, the proposed model provides a time delay approaching zero, the smallest among all tested

models. Comparisons of RMSE and MAE values are summarized in **Tables 1, 2** below:

According to **Tables 1, 2**, the RMSE and MAE of the proposed model are 4.91 and 5.59%, respectively, significantly lower than other models. Hence, it can provide stronger robustness and better accuracy for brain image diagnosis and forecasting.

Comparing Assessment Indicators of Brain Image Segmentation and Fusion

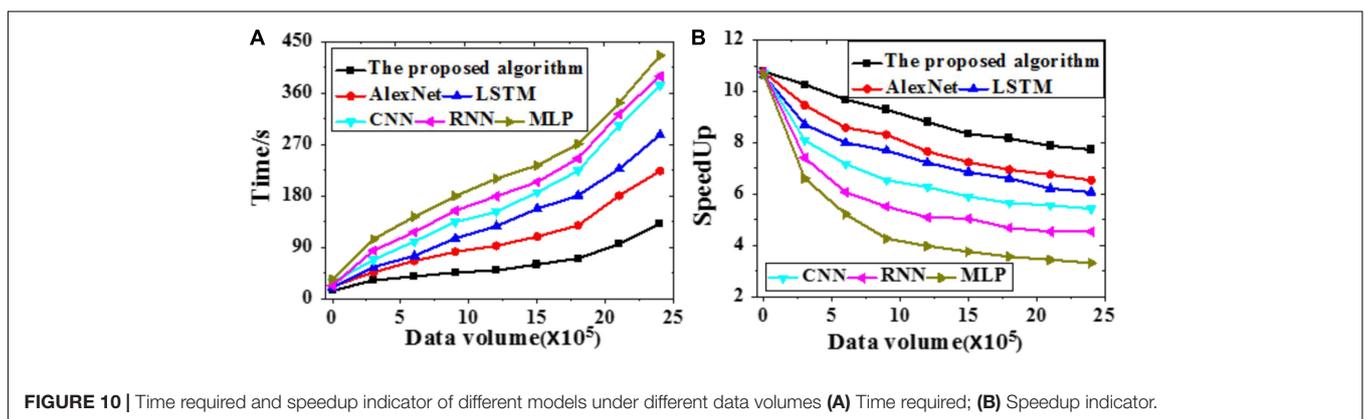
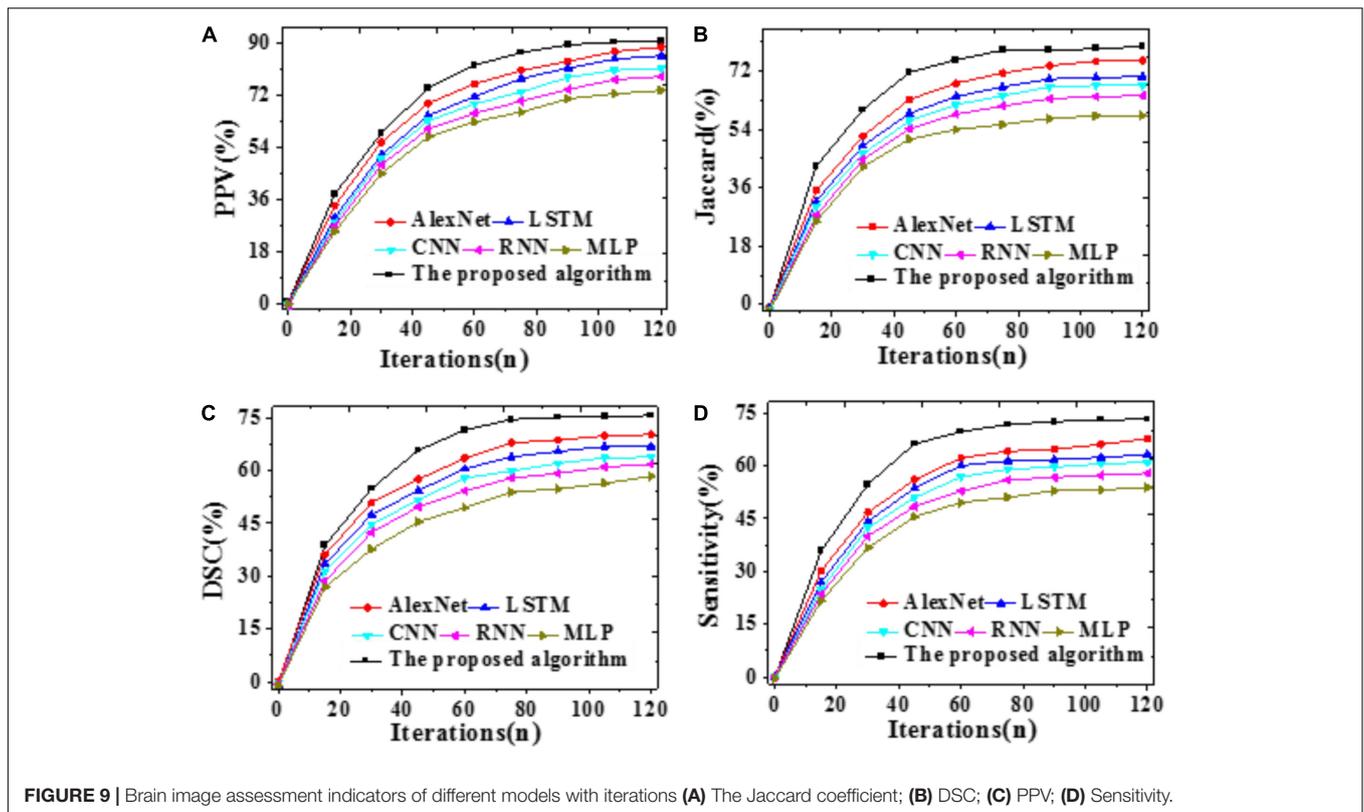
The Jaccard coefficients, DSCs, PPVs, and Sensitivity of the proposed model, LSTM, CNN, RNN, AlexNet, and MLP are compared. The results are demonstrated in **Figure 9**.

TABLE 1 | RMSE (%) changes of each model with iterations.

	Iterations		
	1.00	60.00	120.00
The proposed model	4.85	5.09	4.78
AlexNet	5.61	5.79	5.31
LSTM	6.43	6.11	6.01
CNN	7.28	6.51	6.45
RNN	7.99	6.87	7.04
MLP	8.42	7.33	7.48

TABLE 2 | MAE (%) changes of each model with iterations.

	Iterations		
	1	60	120
The proposed model	5.67	5.45	5.66
AlexNet	6.59	6.98	6.91
LSTM	8.92	8.26	8.44
CNN	9.93	9.11	8.92
RNN	10.53	9.59	9.48
MLP	11.22	10.19	10.01



The higher the indicator value, the more accurate the segmentation result. As shown in **Figure 9**, the proposed model can provide a 79.55% Jaccard coefficient, a 90.43% PPV, a 73.09% Sensitivity, and a 75.58% DSC, remarkably better than other models. Hence, the proposed model outperforms other state-of-art models for MRI brain tumor image segmentation.

Acceleration Efficiency Analysis

The acceleration efficiency of the proposed model is analyzed through simulation experiments. The time required and speedup indicator of different models under different data volumes are presented in **Figure 10**.

The proposed model is less sensitive to brain image data growth than other models. Hence, it is suitable to process

massive data. Simultaneously, the larger the data volume, the higher the speedup indicator. To sum up, the proposed model can provide better data recognition and segmentation results than other models.

A diagnosis and prediction model of brain image fusion digital twins based on semi supervised SVM and improved AlexNet is proposed to improve the accuracy of traditional CNN segmentation of MRI brain tumor images. The model mainly includes the preprocessing of brain image data based on semi supervised SVM. Then, the improved AlexNet network algorithm is used to extract and analyze the brain image features. Finally, through the comparative test and verification of simulation experiments, the model algorithm constructed can acquire more accurate segmentation results and prediction

accuracy, and can effectively solve a series of problems in the process of MRI brain tumor image fusion, such as under-segmentation and over-segmentation, and meet the needs of clinical pathological diagnosis.

CONCLUSION

Brain tumor images have complicated edge structures, artifacts, offset fields, and other defects that affect image segmentation. Extracting features from multi-sequence MRI brain tumor images becomes particularly vital. Both unlabeled and labeled data are utilized to construct the S3VMs regarding the massive number of unlabeled brain image data. Afterward, AlexNet is improved, and a brain image fusion DTs diagnosis and forecasting model is built based on S3VMs and improved AlexNet. Simulation experiments are performed to analyze the proposed model's performance. Results demonstrate a 92.52% feature recognition and extraction accuracy, a 79.55% Jaccard coefficient, a 90.43% PPV, a 73.09% Sensitivity, and a 75.58% DSC of the proposed model, providing an experimental basis for brain image feature recognition and digital diagnosis. There are some weaknesses in the present work. The constructed model does not grade brain tumors. In the future, the texture feature, shape feature, position information, and volume of the multi-sequence MRI image will be extracted based on brain tumor segmentation results before being classified by ML and DL algorithms, achieving

the key steps of precise brain tumor treatment and truly meeting the clinical needs, which will provide remarkable significance for the subsequent clinical diagnosis and treatment of brain tumors.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

ZW was responsible for the writing and implementation of the full text. YD was responsible for the implementation of the experiment. ZY was responsible for the collection of data and the writing of related research. HL was responsible for the implementation and data analysis of the experiment. ZL was responsible for the design of the experiment and the inspection of the results.

ACKNOWLEDGMENTS

We thank the editorial office of Frontiers in Neuroscience for providing us with important information on our sample of articles, and Sarah Mannion de Hernandez for proofreading.

REFERENCES

- Ahleroff, S., Polzer, J., Huang, H., Zhu, Z., Tomzik, D., Lu, Y., et al. (2020). *Smart Manufacturing Based on Digital Twin Technologies. Industry 4.0: Challenges, Trends, and Solutions in Management and Engineering*. (Boca Raton, FL: CRC Press), 77–82.
- Ahmed, M. R., Zhang, Y., Liu, Y., and Liao, H. (2020). Single volume image generator and deep learning-based ASD classification. *IEEE J. Biomed. Health Inform.* 24, 3044–3054. doi: 10.1109/jbhi.2020.2998603
- Angin, P., Anisi, M. H., Goksel, F., Gursoy, C., and Buyukgulcu, A. (2020). AgriLoRa: a digital twin framework for smart agriculture. *J. Wirel. Mob. Netw. Ubiquitous Comput. Dependable. Appl.* 11, 77–96.
- Barricelli, B. R., Casiraghi, E., and Fogli, D. (2019). A survey on digital twin: definitions, characteristics, applications, and design implications. *IEEE Access* 7, 167653–167671. doi: 10.1109/access.2019.2953499
- Dai, Y., Zhang, K., Maharjan, S., and Zhang, Y. (2020). Deep reinforcement learning for stochastic computation offloading in digital twin networks. *IEEE Trans. Ind. Inform.* 17, 4968–4977. doi: 10.1109/tii.2020.3016320
- Dong, R., She, C., Hardjawana, W., Li, Y., and Vucetic, B. (2019). Deep learning for hybrid 5G services in mobile edge computing systems: learn from a digital twin. *IEEE Trans. Wirel. Commun.* 18, 4692–4707. doi: 10.1109/twc.2019.2927312
- Ghafoori, Z., Erfani, S. M., Rajasegarar, S., Bezdek, J. C., Karunasekera, S., and Leckie, C. (2018). Efficient unsupervised parameter estimation for one-class support vector machines. *IEEE Trans. Neural Netw. Learn. Syst.* 29, 5057–5070. doi: 10.1109/tnnls.2017.2785792
- Ghosh, K., Stuke, A., Todorović, M., Jørgensen, P. B., Schmidt, M. N., Vehtari, A., et al. (2019). Deep learning spectroscopy: neural networks for molecular excitation spectra. *Adv. Sci.* 6:1801367. doi: 10.1002/adv.201801367
- Gudigar, A., Raghavendra, U., Ciaccio, E. J., Arunkumar, N., Abdulhay, E., and Acharya, U. R. (2019). Automated categorization of multi-class brain abnormalities using decomposition techniques with MRI images: a comparative study. *IEEE Access* 7, 28498–28509. doi: 10.1109/access.2019.2901055
- Kang, J. S., Chung, K., and Hong, E. J. (2021). Multimedia knowledge-based bridge health monitoring using digital twin. *Multimed. Tools Appl.* 80, 1–16.
- Li, X., Cao, J., Liu, Z., and Luo, X. (2020). Sustainable business model based on digital twin platform network: the inspiration from Haier's case study in China. *Sustainability* 12:936. doi: 10.3390/su12030936
- Li, X., He, B., Wang, Z., Zhou, Y., Li, G., and Jiang, R. (2021). Semantic-enhanced digital twin system for robot–environment interaction monitoring. *IEEE Trans. Instrum. Meas.* 70, 1–13. doi: 10.1109/tim.2021.3066542
- Liu, L., Huang, W., Liu, B., Shen, L., and Wang, C. (2018). Semisupervised hyperspectral image classification via Laplacian least squares support vector machine in sum space and random sampling. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 11, 4086–4100. doi: 10.1109/jstars.2018.2873051
- Liu, Y., Xu, Z., and Li, C. (2018). Distributed online semi-supervised support vector machine. *Inform. Sci.* 466, 236–257. doi: 10.1016/j.ins.2018.07.045
- Liu, Y., Zhang, L., Yang, Y., Zhou, L., Ren, L., Wang, F., et al. (2019). A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE Access* 7, 49088–49101. doi: 10.1109/access.2019.2909828
- Minerva, R., Awan, F. M., and Crespi, N. (2021). Exploiting digital twin as enablers for Synthetic Sensing. *IEEE Internet Comput.* 25, 1–1. doi: 10.1109/mic.2021.3051674
- Nguyen, H. X., Trestian, R., To, D., and Tatipamula, M. (2021). Digital twin for 5G and beyond. *IEEE Commun. Mag.* 59, 10–15. doi: 10.1109/mcom.001.2000343
- Ogino, M., Kanoga, S., Ito, S. I., and Mitsukura, Y. (2021). Semi-supervised learning for auditory event-related potential-based brain–computer interface. *IEEE Access* 9, 47008–47023. doi: 10.1109/access.2021.3067337
- Pan, Y. H., Qu, T., Wu, N. Q., Khalgui, M., and Huang, G. Q. (2021). Digital twin based real-time production logistics synchronization system in a multi-level computing architecture. *J. Manufact. Syst.* 58, 246–260. doi: 10.1016/j.jmsy.2020.10.015

- Qi, Q., and Tao, F. (2018). Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. *IEEE Access* 6, 3585–3593. doi: 10.1109/access.2018.2793265
- Shao, Y., and Chou, K. C. (2020). pLoc_Deep-mVirus: a CNN model for predicting subcellular localization of virus proteins by deep learning. *Nat. Sci.* 12, 388–399. doi: 10.4236/ns.2020.126033
- Shen, M., Tang, X., Zhu, L., Du, X., and Guizani, M. (2019). Privacy-preserving support vector machine training over blockchain-based encrypted IoT data in smart cities. *IEEE Internet Things J.* 6, 7702–7712. doi: 10.1109/jiot.2019.2901840
- Sun, J., Tian, Z., Fu, Y., Geng, J., and Liu, C. (2020). Digital twins in human understanding: a deep learning-based method to recognize personality traits. *Int. J. Comput. Integr. Manufact.* 33, 1–14. doi: 10.1080/0951192x.2020.1757155
- Tanveer, M., Sharma, A., and Suganthan, P. N. (2019). General twin support vector machine with pinball loss function. *Inform. Sci.* 494, 311–327. doi: 10.1016/j.ins.2019.04.032
- Tobon Vasquez, J. A., Scapaticci, R., Turvani, G., Bellizzi, G., Rodriguez-Duarte, D. O., Joachimowicz, N., et al. (2020). A prototype microwave system for 3D brain stroke imaging. *Sensors* 20, 2607–2613. doi: 10.3390/s20092607
- Wang, S., Guo, X., Tie, Y., Lee, I., Qi, L., and Guan, L. (2018). Graph-based safe support vector machine for multiple classes. *IEEE Access* 6, 28097–28107. doi: 10.1109/access.2018.2839187
- Wang, Z., Yao, L., Cai, Y., and Zhang, J. (2020). Mahalanobis semi-supervised mapping and beetle antennae search based support vector machine for wind turbine rolling bearings fault diagnosis. *Renew. Energy* 155, 1312–1327. doi: 10.1016/j.renene.2020.04.041
- Wu, Y., Luo, Y., Chaudhari, G., Rivenson, Y., Calis, A., De Haan, K., et al. (2019). Bright-field holography: cross-modality deep learning enables snapshot 3D imaging with bright-field contrast using a single hologram. *Light Sci. Appl.* 8, 1–7. doi: 10.1117/1.jatis.3.4.049002
- Yan, H., Wan, J., Zhang, C., Tang, S., Hua, Q., and Wang, Z. (2018). Industrial big data analytics for prediction of remaining useful life based on deep learning. *IEEE Access* 6, 17190–17197. doi: 10.1109/access.2018.2809681
- Zavrak, S., and İskefiyeli, M. (2020). Anomaly-based intrusion detection from network flow features using variational autoencoder. *IEEE Access* 8, 108346–108358. doi: 10.1109/access.2020.3001350
- Zhang, H., Li, Y., Lv, Z., Sangaiah, A. K., and Huang, T. (2020). A real-time and ubiquitous network attack detection based on deep belief network and support vector machine. *IEEE CAA J. Automatica Sinica* 7, 790–799. doi: 10.1109/jas.2020.1003099
- Zhao, H., Zheng, J., Deng, W., and Song, Y. (2020). Semi-supervised broad learning system based on manifold regularization and broad network. *IEEE Trans. Circuits Syst. I Regul. Pap.* 67, 983–994. doi: 10.1109/tcsi.2019.2959886
- Zheng, Y., Wang, S., Li, Q., and Li, B. (2020). Fringe projection profilometry by conducting deep learning from its digital twin. *Opt. Express* 28, 36568–36583. doi: 10.1364/oe.410428
- Zhou, M., Yan, J., and Feng, D. (2019). Digital twin framework and its application to power grid online analysis. *CSEE J. Power Energy Syst.* 5, 391–398.

Conflict of Interest: YD was employed by Qingdao Haily Measuring Technologies Co., Ltd.,

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

Copyright © 2021 Wan, Dong, Yu, Lv and Lv. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.