

Detecting Alzheimer's disease Based on 4D fMRI: An exploration under deep learning framework



Wei Li, Xuefeng Lin, Xi Chen*

School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Luoyu Road 1037, 430074 Wuhan, China

ARTICLE INFO

Article history:

Received 3 May 2019

Revised 8 December 2019

Accepted 14 January 2020

Available online 20 January 2020

Communicated by Zhu Yu

Keywords:

Functional magnetic resonance imaging

Alzheimer's disease

Convolutional neural network

Long short-term memory network

ABSTRACT

Applying machine learning methods to various modality medical images and clinical data for early diagnosis of Alzheimer's disease (AD) and its prodromal stage has many significant results. So far, the image data input to classifier mainly focus on 2D or 3D images. Although some functional imaging technologies, such as functional magnetic resonance imaging (fMRI), generate 4D data which contain both spatial and time-varying information of the brain, for the lack of suitable 4D image processing algorithm, these 4D data were always used by transforming them into functional connectivity or slicing them into 2D/3D pictures which causing apparent information loss. In this paper, we present a 4D deep learning model (C3d-LSTM) for AD discrimination, which is able to utilize the spatial and time-varying information simultaneously by dealing with 4D fMRI data directly. The proposed C3d-LSTM combines a series of 3D convolutional neural networks (CNNs) to extract spatial features from each volume of 3D static image in fMRI image sequence. Then the feature maps obtained were put into the long short-term memory (LSTM) network to capture the time-varying information contained in the data. Because of the design of structure, C3d-LSTM became an end-to-end data-driven model, which was more convenient for processing 4D fMRI data. The model proposed conducted on the AD Neuroimaging Initiative dataset for algorithm evaluation compared with controlled experiments. Results showed that using 4D fMRI data directly with the proposed method did make a far better result for AD detection than the methods using functional connectivity, 2D, or 3D fMRI data. It demonstrated our assumption that making the most of the natural spatial and temporal information preserved in 4D fMRI data is significant for AD detection and can increase the performance of the classifier. Meanwhile the result implied that the C3d-LSTM model proposed is a suitable model for processing 4D fMRI data and extracting the spatio-temporal property of fMRI data fully for diagnosis of AD.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Alzheimer's disease (AD) is a complex progressive neurodegenerative disease that often occurs in people over 65 years old [1]. It affects human behavior, memory, and judgment. Because there is no cure for AD so far, many research works focus on how to effectively confirm the diagnosis of AD and start intervention as early as possible, which is also clinically meaningful.

Functional magnetic resonance imaging (fMRI) is an emerging neuroimaging method that can characterize the structure and function of the nervous system. fMRI detects the brain's metabolic activities by measuring the changes in blood flow and blood oxygen concentration. fMRI plays an important role in the study of AD [2].

In previous research, the fMRI data were usually used to construct functional networks [3]. And then graph theory and machine learning methods would be used to analyze the functional brain networks for figuring out the characteristics of AD [4]. For example, Wee et al. proposed a method combining structural and functional networks with several biomarkers to identify AD and mild cognitive impairment (MCI) individuals [5]. Chen et al. made brain networks which can extract both high-order and low-order features for classification to avoid the possible information loss in modeling [6]. As the brain network focuses on the description of the collaboration between regions of cerebral cortex, these researches encountered unavoidable information loss in modeling when processing 4D fMRI data into brain networks.

With the rise of artificial intelligence technology, the deep convolutional networks have been widely used in the field of computer vision and medical image processing [7–9]. Then many researchers tried to apply this method to fMRI processing area.

* Corresponding author.

E-mail addresses: liwei0828@hust.edu.cn (W. Li), [\(X. Lin\), \[chenxi@hust.edu.cn\]\(mailto:chenxi@hust.edu.cn\) \(X. Chen\).](mailto:M201772527@hust.edu.cn)

For example, Sarraf et al. used 2D fMRI image slices to train a CNN for AD and healthy people's classification [10]. Kam et al. used 3D fMRI image slices to train a PCA based 3D CNN [11]. These researches all tried to directly extract features using deep convolutional networks from 2D or 3D images sliced from original 4D fMRI data.

Until now, in the diagnosis of AD, most researchers used 4D fMRI data by transforming them into intrinsic functional connectivity [12] or slicing them into 2D or 3D images and then input the extracted data into the classifier or model. However, the former way neglected considerable spatial structure information and time information through coarse-grained modeling, and the latter way inevitably lose a variety of information on the time dimension of the 4D fMRI data. In fact, they didn't make the most of the information contained in the 4D fMRI data. We suppose that using 4D fMRI directly can preserve full time-varying and spatial information maybe significant for AD diagnosis. However, for the limitation of algorithm, this assumption has not been demonstrated before.

In this study, a deep learning model called as C3d-LSTM was proposed to solve this problem. It is a combination of 3D CNN and LSTM. A series of 3D CNNs were used to extract spatial features from each volume of 3D static image in fMRI image sequence. Then the feature maps obtained were put into the long short-term memory (LSTM) network to capture the time-varying information contained in the data. It takes the advantage of 3D CNN for extracting spatial structure information and the advantage of LSTM for processing time information from the data. It can process 4D fMRI data directly and utilize the time-varying information and structure information of fMRI fully and simultaneously for AD detection.

2. Methods

The C3d-LSTM model consists of two basic deep learning model structures, namely a 3D convolutional neural network (CNN) [13] and a long short-term memory (LSTM) [14] network. The 3D CNN is a generalization of traditional CNNs on 3D images. The biggest change is the conversion of convolution kernels from 2D to 3D, which makes the model more suitable for extracting features from 3D images. Therefore, the 3D CNN can be used to handle the spatial structure information contained in the fMRI data. The LSTM network is an improved form of the traditional recurrent neural network (RNN) that can solve the gradient vanish problem of long-dependence in time series more effectively than RNNs [15]. It is often used in natural language processing and speech signal processing; through its internal complex gate structure, time-varying information and correlation information in time series data can be well characterized.

However, the 4D fMRI data are different from the common 2D time series data in data's dimension, so it is not possible to directly use a LSTM network to process it. We decided to combine the 3D CNN with the LSTM network and obtained the C3d-LSTM deep learning model. The main advantage of the C3d-LSTM model was that it could directly use 4D fMRI data as input, but did not need to convert the fMRI data into functional brain network data or to slice them into 2D or 3D images. And we thought it may reduce the information loss of the fMRI data. The C3d-LSTM was an end-to-end model, which means it was a more general method and easier to use. It didn't need to rely on the prior information by the specialist.

An illustration of the C3d-LSTM model is shown in Fig. 1. The model can be seen as a series 3D CNN module connected to a LSTM network. The 3D CNN module was composed of repeated convolution layers and pooling layers, and it had a fully connected layer which also namely dense layer in the end. A more detailed description of 3D CNN module was introduced in the part 2.1. The

LSTM network was consist of a LSTM layer and a fully connected layer. And a more detailed description of LSTM networks was introduced in the part 2.2. The 4D fMRI image, preprocessed by Statistical Parametric Mapping (SPM) [16], can be seen as a time sequence of 3D images. Then, the 3D CNN of the C3d-LSTM model was used as a spatial feature extraction tool to obtain a feature map for each volume of 3D images. And the 3D CNN module was not shared across each volume of 3D images. The feature maps corresponding to all volumes of 4D image sequence for an individual were put into the LSTM of the C3d-LSTM model to capture the time-varying information contained in the data. The output of the LSTM were put into a fully connected layer, and the category label of the individual was obtained.

2.1. Using 3D convolutional neural networks to extract the image feature of a single point image of functional magnetic resonance imaging

The basic structure of the 3D CNN could be divided into a 3D convolutional layer, a 3D batch normalization layer, an activation function layer, and a 3D pooling layer. Among them, a convolutional kernel size was uniformly used (3, 3, 3) and the activation function used a rectified linear unit. The pooling layer had a step size of 1 and its kernel size was (2, 2, 2). The maximum pooling method was used to down-sample the results derived from the previous convolutional layer. Both the convolutional layer and the pooling layer used valid padding. The structure of the 3D convolutional network is illustrated in Fig. 1. The batch normalization method [17] was used between the convolutional layers to speed up the convergence of the model. At the end of the convolutional layers, flatten processing was applied to transform the 2D feature map into 1D data, and the dropout method [18] was used to randomly deactivate the output of the flatten layer, which could enhance the robustness and generalization ability of the model. The flatten layer was followed by a fully connected layer. The fully connected layer contained 256 neurons and used a linear activation function. The L2 regularization was added to the weight terms to prevent over-fitting of the model. The fully connected layer was followed by the LSTM layer. In the end of the model, a fully connected layer was used as the output layer. The number of neurons in the fully connected layer was the same as the number of labels of the fMRI data, which was the model that needed to be classified. For example, if the model were used to classify normal control (NC) and AD groups, then the number of neurons would be two. In addition, the activation function of the fully connected layer used softmax, which could obtain the probability of each label that the input data belonged to.

2.2. The structure of long short-term memory networks

A LSTM is an improved form of the RNN. Because of the addition of the memory unit, it can effectively solve the long-dependence that cannot be characterized because of the disappearance of the gradient in the traditional RNN forward propagation process. The LSTM has an internal structure called a gate, which can be used to regulate the flow of information inside the cell. These gate structures can determine which data in the sequence are important and need to be retained and which data need to be thrown away.

After deciding the retention of data in a single cell, the relevant sequence information is passed along the network for prediction. The gate structure contains a sigmoid function, the equation of which is shown as (1). Regardless of the input value, the output value of the sigmoid is between 0 and 1, which is fit to be used for updating or forgetting the information. Any number multiplied by 0 will be 0, and this information will be deleted. Similarly, for any

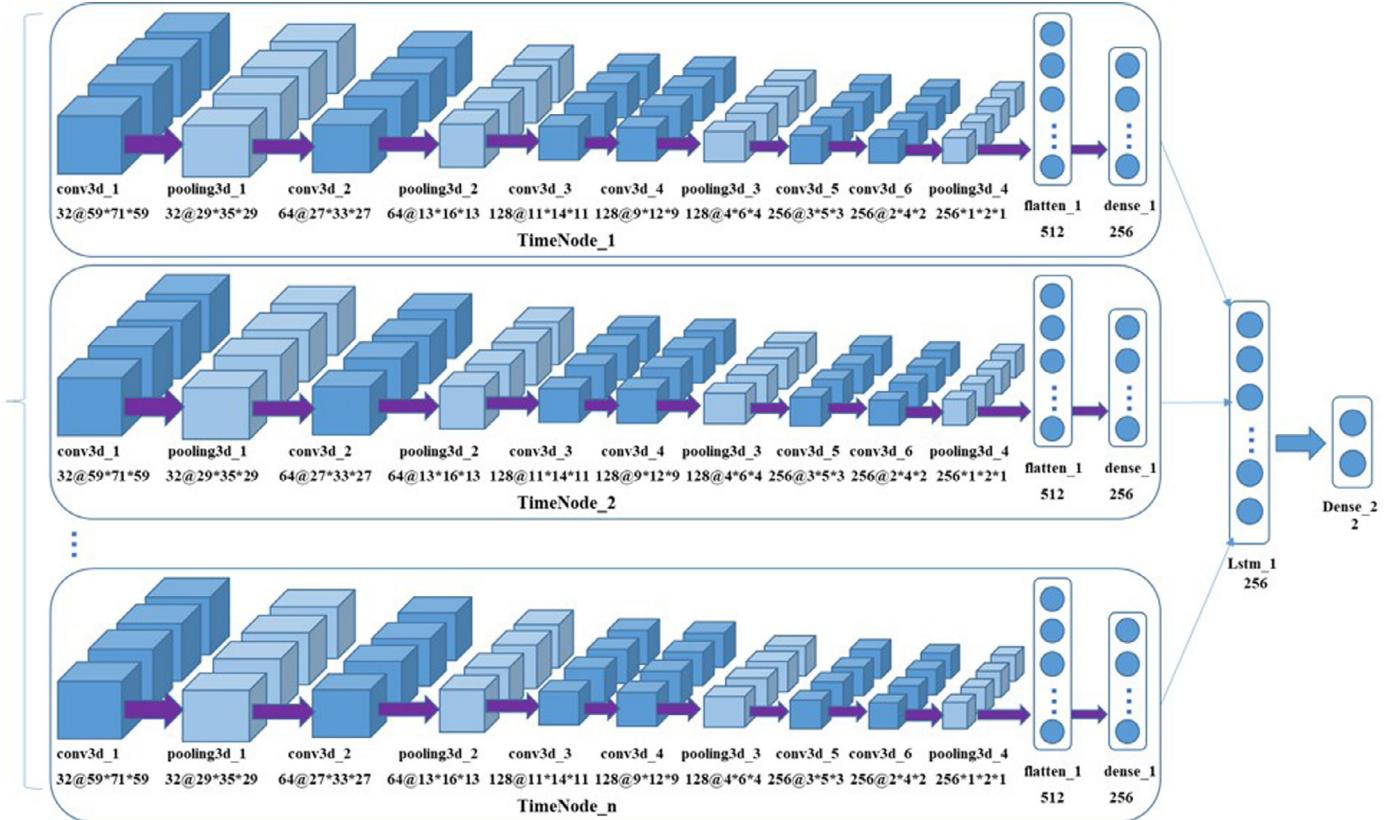


Fig. 1. The structure of the C3d-LSTM model.

number multiplied by 1, the result will be itself, so this information is completely preserved. With the sigmoid function, the network can control which information will be forgotten and which information will be retained.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

All types of RNNs, like the gated recurrent unit (GRU) and LSTM network [19], can be seen as a chained form of a repetitive neural network module. Thus, when studying the LSTM network structure, we only needed to discuss the cell module's structure of the LSTM network. The hidden layer cell module's structure of the LSTM network is shown in Fig. 2.

$$\text{Candidate value : } \tilde{c}^{<t>} = \tanh(W_c[a^{<t-1>}, x^{<t>}] + b_c) \quad (2)$$

$$\text{Update gate : } \Gamma_u = \sigma(W_u[a^{<t-1>}, x^{<t>}] + b_u) \quad (3)$$

$$\text{Forget gate : } \Gamma_f = \sigma(W_f[a^{<t-1>}, x^{<t>}] + b_f) \quad (4)$$

$$\text{Output gate : } \Gamma_o = \sigma(W_o[a^{<t-1>}, x^{<t>}] + b_o) \quad (5)$$

$$\text{Cell State : } c^{<t>} = \Gamma_u * \tilde{c}^{<t>} + \Gamma_f * c^{<t-1>} \quad (6)$$

$$\text{Output : } a^{<t>} = \Gamma_o * \tanh(c^{<t>}) \quad (7)$$

The LSTM network consists of a number of cell modules. The input of a current LSTM cell includes the value of the previous cell's output ($a^{<t-1>}$ and $c^{<t-1>}$) and the input value of the current time $x^{<t>}$. The output of a current cell is $a^{<t>}$ and $c^{<t>}$, which represents the output value of the current cell and the state value of the

current cell, respectively. The LSTM network contains three gates, namely Γ_u , Γ_f , and Γ_o , which represent the update gate, forget gate, and output gate, respectively. The cell state ($c^{<t>}$) is controlled by the update gate and the output gate. The update gate (Γ_u) controls the influence of the output value of the previous cell and the current input on the current cell state.

Specifically, the update gate first passes the output value of the previous cell ($a^{<t-1>}$), which contains the information of the hidden state of the previous cell and the input information of the current cell ($x^{<t>}$) to the sigmoid function. The update gate determines which information needs to be updated according to the output value of the sigmoid function. The sigmoid function's output value is between 0 and 1, with 0 meaning unimportant and 1 meaning important. Second, the output value of the previous cell ($a^{<t-1>}$) and the input value of the current cell ($x^{<t>}$) are also passed into the tanh function to obtain a new candidate value vector ($\tilde{c}^{<t>}$). This new candidate value is multiplied by the previously obtained sigmoid output value. The sigmoid output value determines which information of the candidate values is important and will be preserved from the tanh function's output.

The forget gate (Γ_f) controls the influence of the state value of the previous cell ($c^{<t-1>}$) on the current cell state. Specifically, the forget gate determines which information of the previous cell's state ($c^{<t-1>}$) is to be thrown away or preserved. The output value from the previous cell ($a^{<t-1>}$) and the input information of the current unit ($x^{<t>}$) are simultaneously passed into the sigmoid function. The output value of the sigmoid function is between 0 and 1. The closer the output value is to 0, the more information should be forgotten. The closer the output value is to 1, the more information should be retained. The output value of the forgetting gate (Γ_f) is multiplied by the state value of the previous cell ($c^{<t-1>}$) and is then pointwise added to the

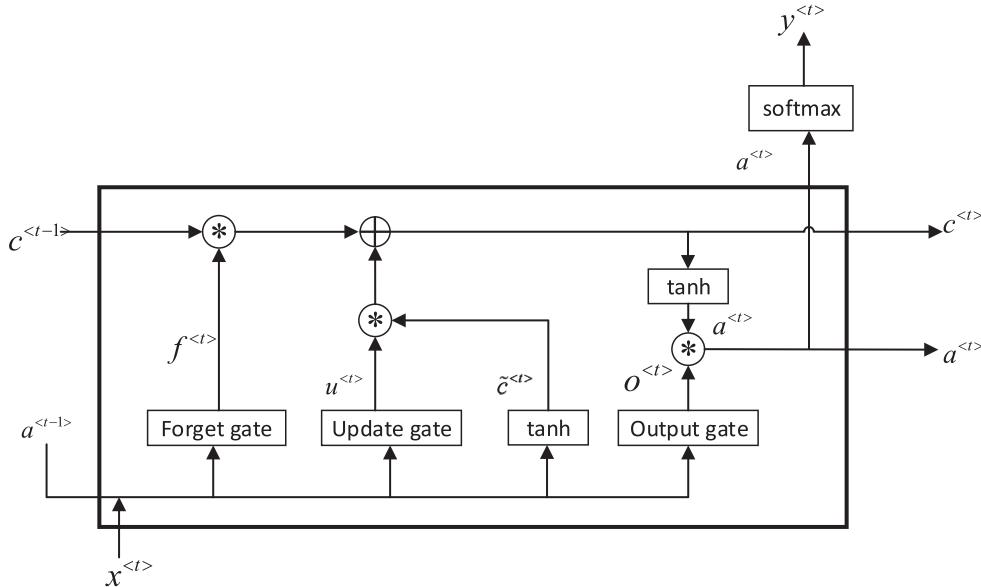


Fig. 2. The structure of long short-term memory network

result from multiplying the update gate output value (Γ_u) by the candidate value ($\tilde{c}^{<t>}$). Finally, the cell state value of the current cell ($c^{<t>}$) is updated. This is a new cell state.

The output gate (Γ_o) controls the influence of the current cell state value on the final output value of the current cell ($a^{<t>}$). The value of the output gate is obtained by simultaneously inputting the output value of the previous cell ($a^{<t-1>}$) and the input value of the current cell ($x^{<t>}$) into the sigmoid function. The cell state value of the current cell ($c^{<t>}$) is passed into the tanh function, and the output value of the tanh function is multiplied by the output value of the sigmoid function (Γ_o). Finally, the output value of the current unit $a^{<t>}$ is obtained.

In the formula, W_o , W_u , and W_f represent the weight value learned by the corresponding control gate after training, respectively, and b_o , b_u , and b_f are the corresponding offset values, which are also obtained by training the network. σ represents the sigmoid function. tanh is the tanh function, and $*$ represents the element multiplication. The LSTM network used in this study was a multiple-input multiple-output structure (many-to-many) where each cell corresponded to one output ($a^{<t>}$) and did not contain a hidden layer.

The output of the LSTM layer was input into a fully connected layer consisting of the same number of neurons as the number of categories. In the fully connected layer with softmax as the activation function, the label of the input data was determined according to the output value of the fully connected layer.

3. Experiment results

3.1. Data processing

In this study, a public dataset, namely the AD Neuroimaging Initiative dataset (ADNI), was used to study and verify the method. The acquisition was performed using a 3T Philips MRI scanner manufactured by Philips Medical Systems. The specific scanning parameters of the experimental data included a TR/TE of 3000 ms/30 ms, flip angle of 80°, imaging matrix of 64 × 64, voxel size of 3.31 mm × 3.31 mm × 3.31 mm, and 48 slices. The volume of each series was 140. The data were processed using the Data Processing Assistant for Resting-State fMRI (DPARSF) toolbox [20] and the RESTing-state fMRI data analysis Toolkit (REST) [21]. For sig-

Table 1

Demographic information and other sample characteristics.

	AD	MCI	NC
Amount	116	99	174
Age	74.6 ± 7.5	73.4 ± 16.0	75.5 ± 6.1

nal equilibration, each series' first 10 volumes were removed. The post-processing included slice timing, head motion correction, normalizing to an EPI template, a Gaussian kernel with 6 mm × 6 mm × 6 mm full width at half maxima (FWHM) spatial smoothing, and band-pass filtering within the interval of 0.01–0.08 Hz. Finally, nuisance signals, including six head motion parameters, a global mean signal, white matter signal, and cerebrospinal fluid sign, were regressed out. When processing the fMRI data, the data with a large head motion were excluded, and the linear trends of time courses were removed using REST. After all data processing work was completed, the dimensions of the data obtained were (61, 73, 61, and 130), where the fourth dimension was the time dimension. The data statistics information is shown in Table 1. The sample was divided into three parts, namely the AD group, the MCI group, which is the prodromal stage of AD, and the normal control (NC) group. The number of people in the AD group was 116, that in MCI was 99, and that in NC was 174. For different controlled experiments, the same fMRI data which have been preprocessed will be sliced into 2D pictures and 3D brain images.

3.2. Model parameters and training strategy

All the deep learning models used in this study were built using the Keras [22] framework with TensorFlow [23] as the backend. The models were trained by cross-validation and the data were divided into a training set, validation set, and test set according to the proportions of 70 %, 10 %, and 20 %. In the training process, the techniques of early stopping [24] and weight decay were used by monitoring the accuracy of the validation set, and the best performing models on the validation set were saved as the final results. Using the cross-entropy loss [25] between the predicted value and the true value (ground-truth) as the loss function, the Adam optimizer was used for parameter optimization. The

best learning rate used for training C3d-LSTM model was 0.00001. The Adam optimizer parameters for training C3d-LSTM model were $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 1e-08$.

Limited by graphics memory, the maximum batch size for training C3d-LSTM model could only be four. The drop rate for training C3d-LSTM model was 0.2, that is, 20 % of the input value was randomly deactivated each time. The dropout of the LSTM layer was 0.5, and the l2 regularization coefficient value of the fully connected layer was 0.0001.

3.3. Method comparison and analysis

Most of recent studies have revealed that under the same conditions, the deep learning methods always have got results far better than those traditional hand-craft designed features methods [26] in computer-aided diagnosis of AD. In order to evaluate the experiment results from different data usage methods, we constructed several controlled experiments for comparison. First, we used the most popular 2D deep learning models, respectively based on VGG19 [27], Resnet [28] and DenseNet [29], to conduct controlled 2D fMRI experiments. Second, a 3D CNN model which has a same structure as the CNN in our proposed C3d-LSTM model was also applied to this study. In addition, a functional brain network modeling method using RBF kernel SVM as classifier was conducted to show a baseline accuracy.

For the controlled functional brain network experiment, there are two steps. First, warping the automated anatomical labeling (AAL) atlas for the fMRI data to get the time series of 90 ROIs (region of interests) of each sample. Then, calculating pairwise Pearson correlation coefficients between any two time-series of 90 ROIs of one scan. The functional brain networks data are used as the input of the SVM classifier.

For controlled 2D fMRI experiments, the VGG19, Resnet and DenseNet models adopt a transfer learning approach that uses model weights trained by ImageNet [30] images for initialization. Because it is a 2D neural network, the data used are a 2D slice of the fMRI data after preprocessing. These slices are taken from the origin plane of the axial (or transverse) plane of the 3D brain image corresponding to each time node according to the time dimension.

For controlled 3D fMRI experiments, the 3D data are sliced from the original fMRI data after preprocessing; that is, the 4D fMRI data are segmented from the time dimension, and the 3D data corresponding to each time node are sequentially saved as the 3D matrix data of the corresponding category.

Our proposed C3d-LSTM network model uses the preprocessed fMRI 4D data directly. In order to avoid the adverse effects of data imbalance on the experiment results, for each experiment, the number of the class with the smallest amount of data participating in the experiment was taken as the standard, and the corresponding amount of each remaining class data was randomly selected from the class dataset for the experiment.

The classification results of the control experiments were all verified by a five-fold cross-validation, with the average of the five experiment results as the final value. The evaluation metric was accuracy (ACC), which is the number of correct classifications divided by the total number of samples. The formulation for its calculation is shown as (8):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (8)$$

In this study, AD, NC, and MCI were tested in one-versus-one, two-class classification, and three-class classification experiments. The classification accuracies are shown in Table 2.

As shown in Table 2, all the deep learning models have an obvious improvement in accuracy with comparison to the functional

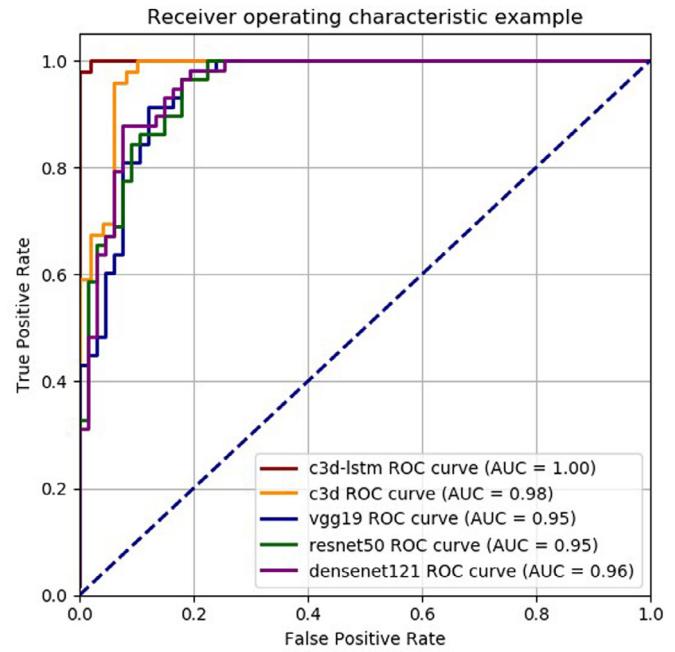


Fig. 3. The receiver operating characteristic curves of each model's Alzheimer's disease and normal control classification results.

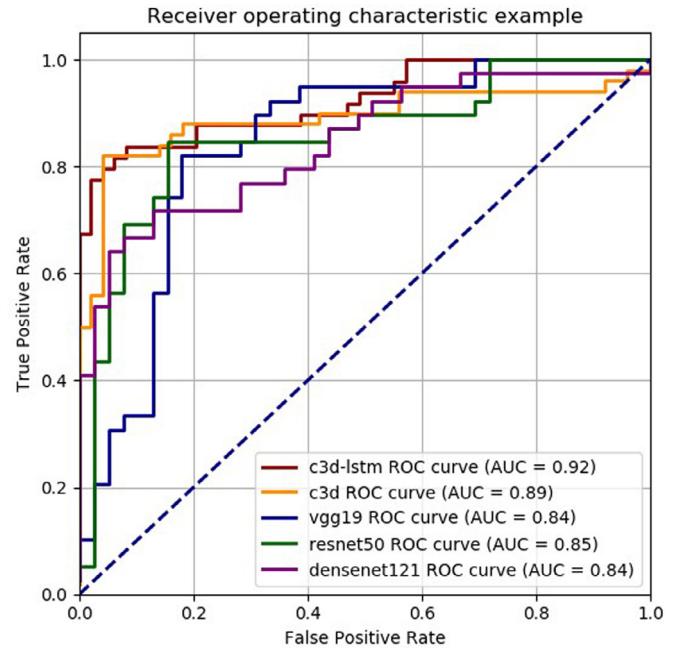


Fig. 4. The receiver operating characteristic curves of each model's Alzheimer's disease and mild cognitive impairment classification results.

brain network modeling method using SVM as classifier. And the proposed C3d-LSTM model, which used 4D fMRI data, showed a significant increase in accuracy over other deep learning models in all four classification tasks, including the discrimination of AD and MCI, MCI and NC, AD and NC, and AD, NC and MCI.

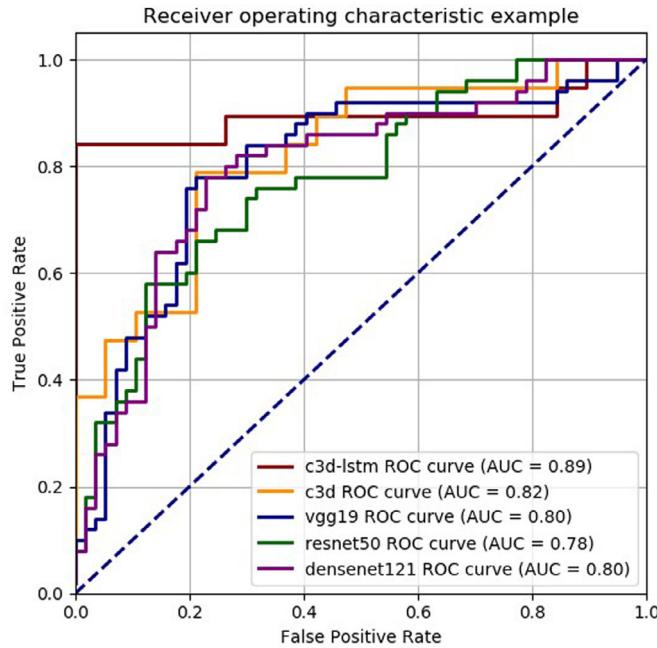
Tables 3 and Figs. 3–5 compare the classification results of different deep learning models using metrics of area under curve (AUC) [31,32], and receiver operating characteristic (ROC) curves, respectively.

From the experiment results of Tables 2 and 3 and Figs. 3–5, we can see that the proposed method which used 4D fMRI data directly did make a much better result for AD detection than the

Table 2

The classification experiment results of each model (accuracy (%)) and standard deviation.

Model	AD/MCI	MCI/NC	AD/NC	AD/NC/MCI
Brain Networks Model [2]	73.94 ± 1.21	70.42 ± 1.42	79.97 ± 0.83	62.31 ± 1.24
VGG19 [27]	83.88 ± 0.74	79.52 ± 0.86	94.99 ± 0.58	79.21 ± 0.72
Resnet50 [28]	84.23 ± 0.84	76.75 ± 0.68	94.78 ± 0.60	78.70 ± 0.71
Densenet121 [29]	82.32 ± 0.93	78.88 ± 0.82	95.12 ± 0.55	81.58 ± 0.76
C3d	88.47 ± 0.78	81.17 ± 0.72	96.47 ± 0.68	82.46 ± 0.74
C3d-LSTM	92.11 ± 0.54	88.12 ± 0.74	97.37 ± 0.56	89.47 ± 0.62

**Fig. 5.** The receiver operating characteristic curves of each model's normal control and mild cognitive impairment classification results.**Table 3**

The classification experiment results of each model (area under curve).

Model	AD/MCI	MCI/NC	AD/NC
VGG19	0.84	0.80	0.95
Resnet50	0.85	0.78	0.95
Densenet121	0.84	0.80	0.96
C3d	0.89	0.82	0.98
C3d-LSTM	0.92	0.89	1.00

methods using 2D, 3D fMRI data or functional brain networks. It demonstrated our assumption that the intact time-varying and spatial information preserved in 4D fMRI data are significant for AD detection. In addition, the C3d-LSTM model proposed in this paper has also been proved to be an effective method to dispose 4D fMRI data and extract the spatio-temporal property of fMRI data fully for diagnosis of AD.

Furthermore, to figure out the impacts of structure and parameters' selections on the performance of the C3d-LSTM model, some experiments were conducted as below.

Considering the specificity of LSTM in the model, the structure of the LSTM network was likely to have an important influence on the performance of the model. Therefore, the influence of the number of LSTM hidden cells and the number of LSTM layers on the results was also studied. As shown in Table 4, a single-layer LSTM network of 64, 128, 256, and 512 hidden cells was used for the controlled experiments. From 64 to 128 to 256, the accuracy of the two-category and three-category classification experiment results

Table 4

The experiment results of using different numbers of long short-term memory hidden cells (accuracy (%)).

Number	AD/MCI	AD/NC	NC/MCI	AD/NC/MCI
64	86.84	94.74	84.21	82.46
128	89.47	92.11	84.21	87.72
256	92.11	97.37	88.12	89.47
512	81.58	86.84	86.84	84.21

Table 5

The experiment results of using different numbers of long short-term memory layers (accuracy (%)).

Layers	AD/MCI	AD/NC	NC/MCI	AD/NC/MCI
1	92.11	97.37	88.12	89.47
2	84.21	92.17	89.34	87.72
3	86.84	86.84	86.64	85.96

of the model generally increased with the increase in the number of hidden cells. However, when the number of cells increased to 512, the accuracy of the two-category and three-category classification experiment results decreased significantly. It is likely that when the number of hidden cells was less than 256, the model was restricted by the number of cells; this was not enough to fully characterize the information contained in the fMRI data, and an under-fitting occurred. In addition, when the number of hidden cells was increased to 512, over-fitting occurred because the parameters of the cells were too redundant.

In addition to discussing the number of hidden cells in the LSTM network, a controlled experiment on the effect of the number of layers of the LSTM network on the results was also conducted. Under the premise of keeping the other structures of the model unchanged, the single-layer, two-layer, and three-layer LSTM networks were used in the model, respectively. The models with different numbers of LSTM layers were tested in two-category and three-category classification experiments, and the results were also measured using ACC. The experiment results are shown in Table 5. It can be seen that with the increase in the number of layers of LSTM, the accuracy of both the two-category and three-category experiment results showed different degrees of decline. When the number of layers of LSTM was increased to three, the performance of the model decreased significantly, and over-fitting was likely to occur.

Furthermore, under the premise of keeping the number of hidden cells and other parameters unchanged, the LSTM was replaced with a GRU, which has a simpler structure, and a controlled experiment was performed. The GRU is another improvement to the traditional RNN. Its structure is very similar to that of LSTM, but compared to LSTM, the GRU is simpler because it removes the cell state unit. The GRU only contains update and reset gates and uses hidden state values for information transfer. The GRU's update gate is similar to the LSTM's output gate and forget gate. It can select which information in the cell is forgotten and which new information needs to be retained. The GRU's reset gate determines which previous information will be forgotten. Compared to LSTM,

Table 6

The experiment results of using a gated recurrent unit (GRU) and long short-term memory (LSTM) (accuracy (%)).

	AD/MCI	AD/NC	NC/MCI	AD/NC/MCI
GRU	85.53	94.74	86.84	81.58
LSTM	92.11	97.37	88.12	89.47

the GRU's structure is simpler and the tensor operation is less, so it takes less time to train, which is also a large advantage. The experiment results are shown in Table 6.

It can be seen from the tables and figures above that the C3d-LSTM model was significantly better than the other 2D or 3D models in the classification tasks involving MCI types, regardless of whether it was two-category or three-category classification. The 3D model was better than the 2D model in the classification task. This could be attributed to the use of 3D data. The information content was larger than that of the 2D data. More information helped the model to distinguish the AD data of different stages. Because the C3d-LSTM model introduced the neural network structure of LSTM, the input data of the model became 4D. In addition to extracting the 3D image features, the model could also extract the feature information of the input data time dimension, thereby making full use of the information of fMRI data so that it could obtain better classification results than those of other models.

4. Conclusion

In this study, we mainly have two contributions. First, so far, the image data input to classifier for AD detection mainly focus on 2D or 3D images. Although some functional imaging technologies, such as fMRI, generate 4D data which contain both spatial and temporal information of the brain, these 4D data were always used by transforming them into functional connectivity or slicing them into 2D/3D pictures. We suppose that this operation apparently causes information loss for classification. This work demonstrated our assumption that making the most of the natural spatial and temporal information preserved in 4D fMRI data is significant for AD detection and also can increase the performance of classifiers under the same condition. Second, this paper developed a 4D deep learning model (C3d-LSTM) for AD discrimination, which is able to utilize the spatial and time-varying information simultaneously by dealing with 4D fMRI data directly. The experiment result showed that this algorithm is effective and make a far better result for AD detection than the methods using functional connectivity, 2D, or 3D fMRI data under the same condition. It makes it possible to make full use of all the information of various 4D data in AD detection, not only limited to 4D fMRI. Some experiments about the parameters' selections of the C3d-LSTM also have been done to show how we made a suitable model for the 4D fMRI data processing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Wei Li: Conceptualization, Funding acquisition, Investigation, Writing - review & editing. **Xuefeng Lin:** Data curation, Methodology, Software, Writing - original draft. **Xi Chen:** Project administration, Supervision.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (61473131, 60905024). 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 is 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

- [1] C. Behl, The search for novel avenues for the therapy and prevention of Alzheimer's disease, *Drug News Perspect* 19 (2006) 5–12.
- [2] E.L. Dennis, P.M. Thompson, Functional brain connectivity using fMRI in aging and Alzheimer's disease, *Neuropsychol. Rev.* 24 (2014) 49–62.
- [3] Rong Fang, Disrupted Structural Brain Network in AD and aMCI: A Finding of Long Fiber Degeneration[J], *Current Alzheimer Res.* 12 (6) (2015).
- [4] E.J. Sanz-Arigita, M.M. Schoonheim, J.S. Damoiseaux, S.A. Rombouts, E. Maris, F. Barkhof, P. Scheltens, C.J. Stam, Loss of 'small-world' networks in Alzheimer's disease: graph analysis of fMRI resting-state functional connectivity, *PLoS One* 5 (2010) e13788.
- [5] C.Y. Wee, P.T. Yap, D.Q. Zhang, K. Denny, J.N. Browndyke, G.G. Potter, K.A. Welsh-Bohmer, L.H. Wang, D.G. Shen, Identification of MCI individuals using structural and functional connectivity networks, *Neuroimage* 59 (2) (2012) 2045–2056 2012.
- [6] X. Chen, H. Zhang, Y. Gao, et al., High-order resting-state functional connectivity network for MCI classification[J], *Human Brain Mapping* 37 (9) (2016) 3282–3296.
- [7] H. Chen, Q. Dou, L. Yu, J. Qin, P.-A. Heng, 'VoxResNet, Deep voxelwise residual networks for brain segmentation from 3D MR images, *NeuroImage* 170 (Apr. 2018) 446–455.
- [8] J. Li, Z.L. Yu, Z. Gu, H. Liu, Y. Li, 'MMAN, Multi-modality aggregation network for brain segmentation from MR images, *Neurocomputing* 358 (Sep. 2019) 10–19.
- [9] H. Wang, et al., 'Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease, *Neurocomputing* 333 (Mar. 2019) 145–156.
- [10] S. Sarraf, G. Tofighi, Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data[C], *Future Technologies Conference*, IEEE, 2017.
- [11] T. Kam, H. Zhang, D. Shen, A Novel Deep Learning Framework on Brain Functional Networks for Early MCI Diagnosis, *Medical Image Computing and Computer Assisted Intervention*, 2018.
- [12] W. Yan, H. Zhang, J. Sui, D. Shen, Deep Chronnectome Learning via Full Bidirectional Long Short-Term Memory Networks for MCI Diagnosis, *Medical Image Computing and Computer Assisted Intervention*, 2018.
- [13] S. Ji, W. Xu, M. Yang, K. Yu, 3D Convolutional Neural Networks for Human Action Recognition, *IEEE T. Pattern Anal.* 35 (2013) 221–231.
- [14] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, *Neural. Comput.* 9 (8) (1997) 1735–1780.
- [15] Z.C. Lipton, J. Berkowitz, C. Elkan, "A critical review of recurrent neural networks for sequence learning," *arXiv:1506.00019*, 2015.
- [16] K.J. Friston, A.P. Holmes, K.J. Worsley, J.P. Poline, C.D. Frith, R.S.J. Frackowiak, Statistical parametric maps in functional imaging: A general linear approach, *Hum. Brain Mapp.* 2 (1994) 189–210.

- [17] S. Ioffe, C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," arXiv:[1502.03167](https://arxiv.org/abs/1502.03167), 2015.
- [18] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *J. Mach. Learn. Res.* 15 (2014) 1929–1958.
- [19] J. Chung, C. Gulcehre, K.H. Cho, Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," arXiv:[1412.3555](https://arxiv.org/abs/1412.3555), 2014.
- [20] Y. Chao-Gan, Z. Yu-Feng, DPARSF: A MATLAB Toolbox for "Pipeline" Data Analysis of Resting-State fMRI, *Front Syst. Neurosci.* 4 (2010) 13.
- [21] X.W. Song, Z.Y. Dong, X.Y. Long, S.F. Li, X.N. Zuo, C.Z. Zhu, Y. He, C.G. Yan, Y.F. Zang, REST: a toolkit for resting-state functional magnetic resonance imaging data processing, *PLoS One* 6 (2011) e25031.
- [22] F. Chollet, Keras, 2015, Available: <https://github.com/fchollet/keras>.
- [23] M. Abadi, A. Agarwal et al., "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," arXiv:[1603.04467](https://arxiv.org/abs/1603.04467), 2016.
- [24] L. Prechelt, Automatic early stopping using cross validation: quantifying the criteria, *Neural Networks* 11 (1998) 761–767.
- [25] P.T. De Boer, D.P. Kroese, S. Mannor, R.Y. Rubinstein, A tutorial on the cross-entropy method, *Ann. Oper. Res.* 134 (2005) 19–67.
- [26] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A.W.M. van der Laak, B. van Ginneken, C.I. Sanchez, A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88.
- [27] S. Han, H. Mao, W.J. Dally et al., "Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding" arXiv:[1510.00149](https://arxiv.org/abs/1510.00149), 2015 – arxiv.org.
- [28] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [29] G. Huang, Z. Liu, L. van der Maaten, K.Q. Weinberger, Densely Connected Convolutional Networks, *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
- [30] J. Deng, W. Dong, R. Socher, L. Li, K. Li, F. Li, ImageNet: A Large-Scale Hierarchical Image Database, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.
- [31] J. Li, Z.L. Yu, Z. Gu, W. Wu, Y. Li, L. Jin, 'A Hybrid Network for ERP Detection and Analysis Based on Restricted Boltzmann Machine', *IEEE Trans. Neural Syst. Rehabil. Eng.* 26 (3) (2018) 563–572.
- [32] Haibo He, E.A. Garcia, 'Learning from Imbalanced Data', *IEEE Trans. Knowl. Data Eng.* 21 (9) (Sep. 2009) 1263–1284.



Wei Li (M'08) received the B.S. degree in automatic control in 1997, and the M.S. and Ph.D. degrees in control science and engineering respectively in 2000 and 2003, from Huazhong University of Science and Technology, Wuhan, China. She joined School of Automation, Huazhong University of Science and Technology, Wuhan, China, as an Associate Professor in 2004. Her research interests include neural signal analysis, neural plasticity, neuromotor rehabilitation, and brain network analysis.



Xuefeng Lin received the B.S. degree in control science and engineering in 2017 from Huazhong University of Science and Technology, Wuhan, China. Now he is studying for the M.S. degree in pattern recognition and intelligent system in Huazhong University of Science and Technology. His research interests include brain network analysis, medical image processing, and machine learning.



Xi Chen received the B.S. degree in automatic control in 1997, and the M.S. and Ph.D. degrees in system engineering respectively in 2002 and 2007, from Huazhong University of Science and Technology, Wuhan, China. He joined Huazhong University of Science and Technology, Wuhan, China, as an Associate Professor in 2008, where is now a Professor of School of Automation. His research interests include signal processing and analysis, system modeling and simulation, and brain network analysis.