ISSN: 1813-4890

Multi-scale dynamic functional connectivity networks for Alzheimer's disease classification

Dengsong Yang [1], Haojie Xu[2]

[1]School of Information Technology and Engineering, Tianjin University of Technology and Education, Tianjin 300222, China;

[2]School of Information Technology and Engineering, Tianjin University of Technology and Education, Tianjin 300222, China.

Abstract

Recently, research applications have applied machine learning methods, such as the CNN and Transfomer architectures, to extract features from the functional connectivity network (FC) for brain disease analysis and transfomer classification. However, the Pearson Correlation Coefficient (PCC) method, which uses a fixed window size to build a dynamic Functional connection (dFC) network, has limitations because it is challenging to extract potentially high-level features from the dFC and determine local feature correlations. Transformer-based methods usually extract the global features of dfc, and Transformer is improved in this paper to be more flexible and efficient in handling multi-scale data and dynamic input. A multi-scale Convolutional neural network (MsCNN-Tran) for learning and analyzing various scales of dFC in Alzheimer's disease fMRI data is proposed. By fusing the features of multiple branches, feature learning can be carried out from multiple scales and angles, and finally higher quality features can be output for classification. Finally, a deep fusion of features learned at different scales is used for the diagnostic classification of brain diseases. In the rs-fMRI dataset applied to ADNI, the classification accuracy of AN/CN was 97%.

Keywords

Transformer, convolutional neural network, Alzheimer's disease.

1. Introduction

Deep learning methods, using CNN and BiLSTM in our previous research, are applied to functionally connected networks (FCS). Among them, CNN-based methods learn highly complex feature representations from FC, showing advantages in the analysis and classification of brain diseases [4], [5]. However, these methods struggle to capture global representations (i.e., long-range sequence features) that are critical to dynamic dFC analysis tasks. Compared with CNN, BiLSTM introduces a unique gating mechanism, which can selectively remember some important information requiring long-term memory while forgetting some minor information. However, BiLSTM can only alleviate the long-term sequence feature problem to a certain extent, and cannot fundamentally solve it [6]. By using the self-attention mechanism in transformer, compared with BiLSTM, transformer can effectively avoid information loss and inherent sequence structure, and provide a wide receptive field [7]. At the same time, sliding window data of different scales are used for experiments, which may help to explore the changes and functional connectivity of injured brain regions observed in patients throughout the progression of AD, thereby revealing its pathogenesis.

To this end, we propose a feature learning framework named MsCNN-Tran, which is designed as three branch structures to process sliding window data of different scales and finally carry out feature fusion to obtain richer representation. The structure of each branch is to combine

ISSN: 1813-4890

CNN-based local features with transformer's global representation. In addition, the feature alignment unit (FAU) is designed to solve the semantic divergence problem (i.e. local and global) caused by the inconsistent feature dimensions in the coupling process. Dynamic position coding can be dynamically generated according to the sequence length of input features according to the position coding, instead of fixed use of predefined encoding, which enhances the adaptability of the model to different input lengths. Residual connections are used in each branch, between the convolutional layer and Transformer layer, and inside Transformer to help gradient flow and retain the original feature information. The experimental results of the ADNI dataset show that our proposed learning framework not only significantly improves the classification performance of brain diseases, but also facilitates the search for neuroimaging biomarkers associated with diseases.

2. fMRI data preprocessing

Data used in this study came from fMRI data from 118 AD patients and 127 normal control older adults (CN) in the Alzheimer's Disease Neuroimaging Program (ADNI) database. fMRI data were collected by 3T Philips fMRI scanner. Specific scanning parameters were as follows: TR 3000ms, TE 30ms, rollover Angle 80°, imaging matrix 64×64, voxel size 3.31mm× 3.31mm× 3.31mm, 48 sections. Data preprocessing was performed using the Matlab toolbox SPM12 and DPARSFA (2.2). The pre-processing steps are as follows: data format conversion (DICOM to NIFTI), deletion of the first 10 unstable time points, time layer slice correction, head motion correction, space normalization, time filtering to reduce drift and noise, and removal of excessive head motion time points. The pre-processing process in this study was not smoothed, because network analysis requires high spatial accuracy, so smoothing is not needed, so as not to affect the activation of adjacent regions of interest (ROIs).

In our previous study, the size of the fixed sliding window was set to 30. In this paper, the scale used consists of three transition scales of 30, 50, and 70 to better describe the dynamic changes in brain regions. A three-scale sliding window is performed on the subject's ROIs data to build the dFC, and the generated dFC is used as input to the model.

3. Method

Fig.1 shows the detailed process of our model, including multi-scale dFC construction, feature filtering, each branch and the module feature coupling in the branch to realize the diagnosis of Alzheimer's disease.

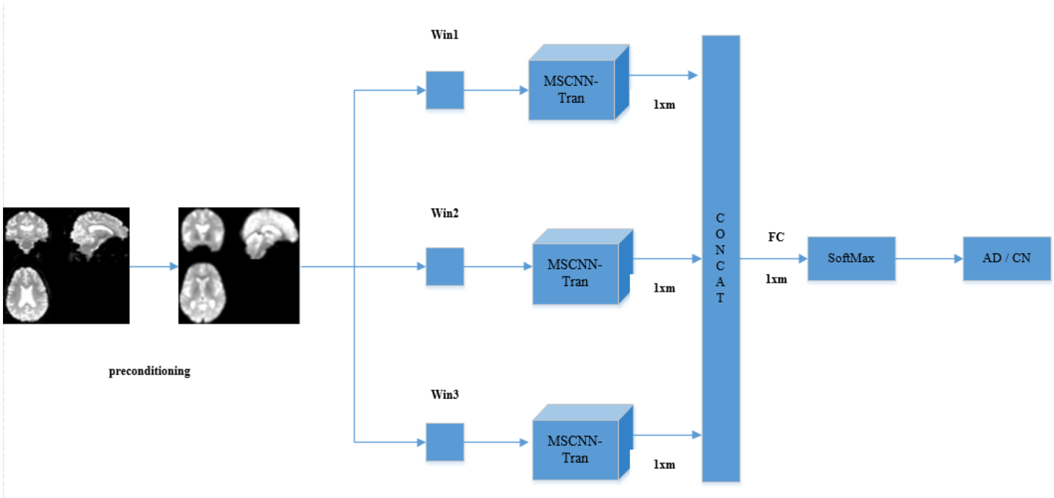

Fig. 1 Model frame drawing

ISSN: 1813-4890

3.1. Multi-scale dFC structure

(1) Division of multi-scale dFC

Our first task was to standardize ROI per subject using the AAL116 template [8] for the BOLD time series of 116 ROIs. To describe the dynamic changes in brain regions and construct the dFC, we divide the time series of ROI obtained by rs-fMRI into overlapping sliding Windows. Based on sliding Windows of different scales, dFC of corresponding scales is constructed. Construct M overlapping subsequences$S = [S_1$、 $S_2, ..., S_m]$ 与 $S_i = [t_{2i-1}$、 $t_{2i-1+1}, ..., t_{2i-1+29}$ $i = 1,2, ..., m$. The window scales are 30, 50, and 70 time points, and the step size is 2, corresponding to m of the number of 54, 44, and 34 subsequences.

(2) Construction of multi-scale dFC

Specifically, we first segment the entire time series of ROI for each scale individually into T continuous and overlapping Windows of length L. Then, $F^t \in R^{N\times N}(t = 1, ..., T)$ is constructed by calculating the Pearson correlation coefficient (PCC) between the average time series of any pair of regions of interest (ROIs) at the t th time window, as follows:

$$\mathbf{F}^t(i,j) = \mathrm{corr}\left(x_i^t, x_j^t\right) \tag{1}$$

corr represents the correlation between the two mean time series of ROI, and xti and xtj represent the blood oxygen level dependence of the i and j brain regions in the t time window, respectively. To better capture the association between adjacent FC, we flatten each FC to dFC $G=\{F^t \in R^{N\times N}\}^T_{t=1}$ to get the new representation $G \in R^{T\times N2}$, which will be treated as the input to the frame.

3.2. Multi-branch model based on multi-scale

In the model we built, the convolutional Neural network (CNN) and Transformer are two key modules, each taking on specific tasks and collaborating in the model to take advantage of each other.

(1). Convolutional Neural Network (CNN) part

The main role of convolutional neural networks in this model is to extract local features from input data. Each input branch is first processed through multiple convolution layers. These convolution layers slide over the input data by applying convolution kernels to extract local spatial or temporal features. After the convolution operation, we apply the ReLU activation function to introduce nonlinearity so that the network can better fit complex patterns. In addition, to reduce the risk of overfitting, we add Dropout layers after each convolutional layer to randomly drop some neurons.

After the convolution layer, we also use the Feature Adjustment Unit (FAU) to adjust the features by 1x1 convolution, so that the features generated by different convolution layers are aligned in dimension, providing consistent input for subsequent processing.

(2). Transformer part

Later in the model, we introduced the Transformer encoder to capture global features and long distance dependencies. The Transformer encoder includes Multi-Head Self-Attention and Feed-Forward neural networks and is an important tool for processing time series data and capturing long-term dependencies.

In each Transformer encoder, the input data is first normalized through Layer Normalization and then passed into the multi-head self-attention mechanism for computation. The self-attention mechanism can effectively capture the interrelationship between different positions by weighted summation of each position in the input sequence, so as to learn the long-term step dependence feature. The output of the self-attention layer is then residually connected with the input to alleviate the gradient disappearance problem.

The Transformer then continues processing through a feedforward neural network consisting of two fully connected layers, non-linear mapping via ReLU activation functions. Finally, the features processed by the Transformer encoder are residually connected to the input again, ensuring that the information is retained.

(3). Combination of CNN and Transformer

In this model, CNN and Transformer work in parallel to take advantage of each other. Convolutional neural networks are responsible for extracting local low-level features, while Transformer focuses on capturing global timing dependencies. Through this combination, the model is able to process both local and global information at the same time, thus improving the modeling ability of complex data patterns.

The convolutional layer is fused with the output of the Transformer encoder via a residual connection, ensuring that the information from both can be effectively combined without loss. In addition, the Feature Adjustment Unit (FAU) helps align the feature dimensions of the different layers, providing a consistent input format for subsequent Transformer processing. This multi-level and multi-scale feature learning method enables the model to consider both local details and global structure when processing time series data, thus improving the expressiveness and accuracy of the model.

4. Experiment

In the experiment, we built and trained this model and other comparative network models using python and Keras on NVIDA A100-PICE-40GB. A binary classification task (AD and NC) is performed using five-fold cross-validation, which divides subjects into five roughly equal subsets. In each cross-validation, one subset is selected as the test set and the remaining four subsets as the training set. In addition, during each cross-validation process, we further split 20% of the data from the training set into validation sets. And we use accuracy, precision, recall, and F1 scores corresponding to formulas 2, 3, 4, and 5, respectively, to evaluate the effect of classification. Among them, true positive, false negative, true negative and false positive are respectively represented by TP, FN, TN and FP [11].

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FN+FP} \tag{2}$$

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

$$\text{Precision}=\frac{TP}{TP+FP} \tag{4}$$

$$\text{F - Score}=2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \tag{5}$$

4.1. Comparison method

We will use the following four methods to compare with our experiment.

(1) CNN: This method is a method that directly uses 2DCNN to extract features from the functional connection matrix for disease diagnosis [9].

(2) CNN-LSTM: Long memory network (LSTM) solves the defects of the original recurrent neural network. LSTM is responsible for processing the temporal characteristics of sequence

data. By combining the feature extraction capability of CNN and the long-term memory capability of LSTM, CNN-LSTM performs well in sequence modeling tasks [10].

(3) CNN-BI-LSTM: Compared with the proposed method, this variant does not use multi-head attention mechanism, but directly uses CNN and Bi-LSTM for feature extraction, and Bi-LSTM combines forward and backward information flow to process temporal features.

(4) MsCNN-Tran: Different from others, our method directly carries out CNN learning and transformer learning for dFC of different scales. Specifically, for dFC whose data structure is N×N×C, MsCNN-Tran directly learns from both local features and global features to obtain advanced fusion features, and finally performs multi-scale feature fusion and classification.

Tables 1-1 shows the classification performance of the five methods. It can be seen from the table that the MsCNN-Tran model proposed by us has the best performance in the classification task, with an accuracy rate of 97.06%, accuracy of 97.37%, recall rate of 96.88%, and F1 score of 97.04%. The model using only CNN has the worst performance, and the CNN-BILSTM model is second only to the model proposed by us.

Table 1-1 Comparative Values of IGD Indicators

Methods	Accuracy	Precision	Recall	F1-Score
CNN	0.8333	0.8346	0.8082	0.8218
CNN-LSTM	0.8455	0.8467	0.8408	0.8434
CNN-Bi-LSTM	0.9273	0.9231	0.9231	0.9231
MsCNN-Tran	0.9706	0.9737	0.9688	0.9704

5. Conclusion

In this paper, a hybrid model combining convolutional neural networks (CNN) and Transformer is proposed to capture both local features and global dependencies to improve performance in multiple tasks. By extracting local features with CNN and modeling long-distance dependencies with Transformer, our model can understand input data more comprehensively and demonstrate strong feature learning capability. In terms of information fusion, the model effectively combines the advantages of both through residual connection and feature alignment unit (FAU), thus avoiding information loss and improving training efficiency. In addition, our model shows good generalization ability and adaptability to multiple tasks such as images, time series data and text. The experimental results show that the architecture combined with CNN and Transformer has significantly improved performance compared to the traditional single model, especially when dealing with complex data patterns. In a word, this model provides new ideas and methods for multi-task learning and complex data analysis, and has high application value.

References

[1] Rawan Tarawneh and David M Holtzman, "The clinical problem of symptomatic alzheimer disease and mild cognitive impairment", Cold Spring Harbor perspectives in medicine, vol. 2, no. 5, pp. 1-16, 2012.

[2] Martin Prince et al., "The global prevalence of dementia: a systematic review and metaanalysis", Alzheimer's & dementia, vol. 9, no. 1, pp. 63-75, 2013.

[3] Sun H, Wang A, He S. “Temporal and spatial analysis of alzheimer’s disease based on an improved convolutional neural network and a resting-state FMRI brain functional network”, International Journal of Environmental Research and Public Health, vol. 19, no. 8, pp. 4508-4524, 2022.

[4] Mingliang Wang, Chunfeng Lian, Dongren Yao, Daoqiang Zhang, Mingxia Liu and Dinggang Shen, "Spatial-temporal dependency modeling and network hub detection for functional mri analysis via

ISSN: 1813-4890

convolutional-recurrent network", IEEE Transactions on Biomedical Engineering, vol. 67, no. 8, pp. 2241-2252, 2019.

[5] Biao Jie, Mingxia Liu, Chunfeng Lian, Feng Shi and Dinggang Shen, "Designing weighted correlation kernels in convolutional neural networks for functional connectivity based brain disease diagnosis", Medical image analysis, vol. 63, pp. 1-14, 2020.

[6] Kai Lin, Biao Jie, Peng Dong, Xintao Ding, Weixin Bian and Mingxia Liu, "Convolutional recurrent neural network for dynamic functional mri analysis and brain disease identification", Frontiers in Neuroscience, vol. 16, pp. 1-14, 2022.

[7] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, et al., "Attention is all you need", Advances in neural information processing systems, vol. 30, pp. 5998-6008, 2017.

[8] Nathalie Tzourio-Mazoyer, Brigitte Landeau, Dimitri Papathanassiou, Fabrice Crivello, Octave Etard, Nicolas Delcroix, et al., "Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain", Neuroimage, vol. 15, no. 1, pp. 273-289, 2002.

[9] Wang R, He Q, Han C, et al. "A deep learning framework for identifying Alzheimer's disease using fMRI-based brain network", Frontiers in Neuroscience, vol. 17, pp. 1177424, 2023.

[10] Yang D, Wang R, Song C. Classification of Alzheimer's Disease Using fMRI-Based Brain Functional Network Data[C]//2024 6th Asia Symposium on Image Processing (ASIP). IEEE, 2024: 97-101.

[11] Huang Z, Ma Y, Wang R, et al. "DSCNN-LSTMs: A lightweight and efficient model for epilepsy recognition", Brain Sciences, vol. 12, no. 12, pp. 1672-1689, 2022.