Neuroscience Letters 730 (2020) 134971

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Research article

Deep learning based mild cognitive impairment diagnosis using structure MR images

Jingwan Jiang, Li Kang*, Jianjun Huang, Tijiang Zhang

College of Information Engineering, Shenzhen University, Shenzhen 518060, China

ARTICLE INFO

Keywords:
Early mild cognitive impairment
Convolutional neural network
Support vector machine
Transfer learning

ABSTRACT

Mild cognitive impairment (MCI) is an early sign of Alzheimer's disease (AD) which is the fourth leading disease mostly found in the aged population. Early intervention of MCI will possibly delay the progress towards AD, and this makes it very important to diagnose early MCI (EMCI). However, it is very difficult since the subtle difference between EMCI and cognitively normal control (NC). For improving classification performance, this paper presents a deep learning based diagnosis approach using structure MRI images for exploiting deeply embedded diagnosis features; then a feature selection strategy is performed to eliminate redundant features. A Support Vector Machine (SVM) is further employed to distinguish EMCI from NC. Experiments were performed on the publicly available ADNI dataset with a total of 120 subjects. The classification results demonstrate the superior performance of the proposed method with accuracy of 89.4% for EMCI versus NC.

1. Introduction

Alzheimer's disease, the most common form of dementia, is an irreversible progressive neurodegenerative disease happening to people over the age of 65. AD patients will suffer from memory loss, which disrupts daily life. In 2018, the estimated number of AD patients were 5.7 million in American, 0.5 million more than in 2014 [1,2]. However, if no precautions were taken, with the incidence rate increasing, there will be 13.8 million AD patients in 2050 [2]. At present, no effective prevention or treatment has been found. More importantly, it costs a lot of money and manpower to cure AD patients [2]. Mild cognitive impairment (MCI) is the transition state between age-related cognitive decline and AD or another dementia [3]. People with MCI have mild but measurable changes in thinking abilities and have high risk of conversion to AD [2]. Therefore early diagnosis of MCI is of vital importance for the early intervention in preclinical state of AD [4,5].

In Alzheimers Disease Neuroimaging Initiative (ADNI) project, the MCI stage is divided into early MCI (EMCI) and late MCI (LMCI). Compared with LMCI, EMCI patients have milder cognitive deficits; thus the diagnosis of EMCI is more challenging and has drawn much attention of researchers in the last decades [6]. In this study, we will review the advanced methods of previous literature about EMCI diagnosis and propose the novel classification approach to distinguish EMCI group from normal control (NC) group.

Nowadays, neuroimaging techniques have been widely applied to medical image analysis, such as structure magnetic resonance imaging (sMRI), functional MRI (fMRI), diffusion tensor imaging (DTI) and positron emission tomography (PET). Using these neuroimaging techniques, most researches utilized traditional machine learning methods to classify EMCI and NC.

Based on 68 cortical areas of diffusion weighted MR images, Prasad et al. [7] computed a 68×68 connectivity matric and a set of network measures as the input of SVM classifier and acquired a classification accuracy of 59.2% for EMCI versus NC. Rory Raeper et al.[8] constructed a cooperative correlational and discriminative ensemble learning framework using sMRI images where each individual brain was represented by a set of shallow convolutional brain multiplex (SCBM) used to train an ensemble of CCA-SVM and LDA-based classifiers, and an accuracy of 80.95% was reported. Parisa Forouzannezhad et al. [9] combined the features extracted from cortical region and subcortical region of MRI and PET images, neuropsychological test scores, age and education to train a deep neural network for EMCI classification and reported an accuracy of 84%. In addition, the authors trained a SVM classifier utilizing the same data in paper [9] and reported an accuracy of 81.1% [10].

The aforementioned approaches based on traditional machine learning try to find distinguishing features from neuroimaging data through complex feature engineering, then perform classification task. However, the above feature engineering not only needs accurate prior knowledge using for calculating brain region-based features, but also

* Corresponding author.
E-mail address: kangli@szu.edu.cn (L. Kang).

https://doi.org/10.1016/j.neulet.2020.134971
Received 18 November 2019; Received in revised form 24 March 2020; Accepted 6 April 2020
Available online 04 May 2020

needs considerable time and manpower. The low accuracy score reveals that it is very challenging to distinguish EMCI from NC just utilizing low-level and coarse-grained features based on prior knowledge [11,12]. Recently, researchers have shown an increasing interest in convolutional neural network (CNN) method for classification task [13,14]. CNN can extract low- and high-level features from complex high-dimensional image data in the form of end-to-end. Compared with the above brain region-based feature engineering, CNN is good at automatically seeking the most discriminating disease-related features from voxel values of image, which is beneficial to avoid errors introduced from feature engineering and retain the subtle but complex differences between EMCI and NC. There have been several relevant investigations into EMCI diagnosis using the popular CNN method.

Tae-Eui Kam et al. [15] proposed a novel 3DCNN framework using fMRI data to extract deep embedded features from both static and dynamic brain functional networks for EMCI diagnosis and reported an accuracy of 76.07%. However, time-consuming, multi-channel and multi-model training did not exchange higher classification accuracy. Mukul Puranik et al. [16] employed Inception Resnet V2 model with transfer learning technique to classify NC, EMCI and AD, and obtained an accuracy of 98.41%. However, the input data are the 2D slices of fMRI images, which means the classification task isn't based on subject-level, deviating clinical needs.

In this study, in order to effectively solve the challenging binary classification problem for EMCI vs. NC, a hybrid diagnosis method based on deep CNN and support vector machine (SVM) was proposed, where SVM classifier has successfully been applied to EMCI classification in many researches [17–19]. Specifically, we selected sMRI as research data due to its universality in the clinical practice and convenience in the examination. Then, 120 sMRI 3D images (70 EMCI, 50 NC) acquiring from ADNI were decomposed into 3840 2D slice images for training VGG16 CNN with transfer learning technique. At last, all slice features of each subject were fused into a feature vector for LASSO feature selection and SVM classification. The experimental results showed that the proposed method greatly improved the classification performance thanking to the combination of CNN and SVM and the elimination of redundant features. What's more, transfer learning technique effectively alleviated the problems caused by small dataset and reduced a lot of training time. Most of EMCI diagnosis researches based on traditional machine learning utilized region-based features extracted from multi-step feature engineering for classification, the proposed method classified EMCI and NC using volumetric features of sMRI data with higher accuracy, which overcame the limitation of traditional machine learning and promoted the development of end-to-end computer-aided diagnosis of EMCI.

2. Material and methods

2.1. Participants

Data used in this study were obtained from Alzheimer's Disease Neuroimaging Initiative (ADNI) project launched in 2003 as a public-private partnership. The goal of ADNI study is to detect AD at the earliest possible stage and support advance in intervention, prevention and treatment through new diagnostic methods.

A total of 120 preprocessed MRI scans in NIfTI file format were downloaded from ADNIGO and ADNI2 database. The corresponding demographic information of the dataset is shown in Table 1. This study included 70 EMCI subjects and 50 age-matched Normal Control' (NC). All subjects from the baseline/screening visit have passed strict inclusion criteria. EMCI diagnostic criteria included 1) Mini Mental State Examination (MMSE) scores between 24-30, 2) a subjective memory concern reported by subject, informant or clinician, 3) objective memory loss of 0.5–1.5 SD (standard deviation) below normal measured by education adjusted scores on delayed recall of one paragraph from Wechsler Memory Scale Logical Memory II, 4) a Clinical Dementia Rating (CDR) of 0.5, 5) absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and absence of dementia. LMCI diagnostic criteria are the same as that of EMCI except that the objective memory loss score is more than 1.5 SD below normal.

Table 1
Corresponding demographic information of subjects.

	EMCI	NC
Number	70	50
Gender(F/M)	27/43	27/23
Age(year)	72.9 ± 8.3	72.5 ± 6.1
MMSE	27.86 ± 1.66	28.93 ± 1.18

2.2. sMRI acquisition

The origin T1-weighted structure MR images (sMRI) were acquired by 3-Tesla GE medical systems scanners at multiple sites with rigorous quality control to reduce site effect. The following imaging parameters were used: acquisition plane = sagittal, minimum full echo time, Tl = 400 ms, volume size = $256 \times 256 \times 196$, voxel size = $1.0 \times 1.0 \times 1.2$ mm^3, flip angle = 11°. More information about the parameters of the images can be searched on the website of ADNI (http://adni.loni.usc.edu/).

2.3. Preprocessing

All obtained data have been preprocessed through a series of standard preprocessing procedures. Using FSL and FreeSurfer software, the raw T1-weighted structure MR images are preprocessed with skull-stripping, intensity normalization and registration with a standard template Colin27 having the same coordinate system as MNI152. Finally, the sMRI image of each subject has a resolution of $110 \times 110 \times 110$ voxels.

In order to highlight most distinguishable features and improve the efficiency of classification task, 2D slices data of sMRI were studied. Each 3D sMRI image was decomposed into 2D slices along axial view and the slices with indices 37–68 were converted to JPEG image format by means of MATLAB (2018a). It is worth noting that some brain regions associating with memory, such as hippocampus and callosum, are contained in 32 slices. Then a subject-level dataset for SVM classification was built, which consists of 120 folders each of which includes 32 slices of one subject. In addition, another slice-level dataset was built for training 2DCNN model, which contains 2240 (32×70) EMCI slices and 1600 (32×50) NC slices split into train set and validation set with the ratio of 8:2.

2.4. Proposed method

The pipeline of proposed approach for EMCI classification is shown in Fig. 1, where a CNN model of VGG16 [20] is chosen as feature extractor, LASSO algorithm is utilized for feature selection and SVM classifier performs classification task. Firstly, VGG16 network is fine-tuned by slice-level dataset using transfer learning technique through loading pre-trained weight, then the optimal VGG16 model is saved in term of the minimum loss value during training process. Secondly, the features of each slice in subject-level dataset are extracted by VGG16 optimal model whose output is a matrix of 1×256. Each subject has 32 slices; therefore 32 matrixes of 1×256 are concatenated into a total feature matrix as a feature representation of one subject. All feature representations of 70 EMCI and 50 NC subjects are integrated into a matrix dataset, as described in Fig. 1 (2). Thirdly, feature selection is performed by LASSO algorithm to reduce the dimension and irrelevant information of the above matrix dataset. Finally, the output of LASSO

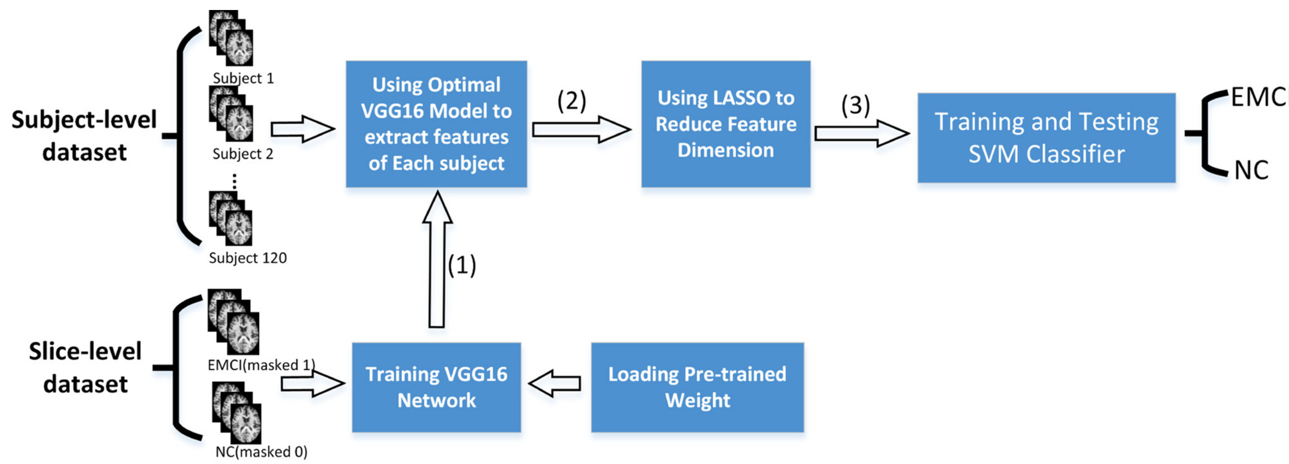

Fig. 1. An illustration of the pipeline of proposed method, where (1) is the optimal VGG16 model which is used as feature extractor in next step, (2) is a matrix with the size of 120 × 8192(32 × 256), (3) represents a matrix with the size of 120×*number of selected features of each subject*.

algorithm is used to train and test SVM classifier for binary classification of EMCI and NC.

2.4.1. Convolutional neural network and transfer learning

Convolutional neural network (CNN) is one of popular deep learning algorithm, which has gotten great success in computer vision and image processing applications in recent years. CNN commonly consists of convolution layer, pooling layer and fully connected layer [21]. In convolution layer, the convolution calculation is performed on an image of size $h \times w$ using a kernel size of k, padding of p and stride of s, then $2^n(n \in Z)$ feature maps with a size of $(\frac{h-k+2p}{s}+1) \times (\frac{h-k+2p}{s}+1)$ will be output. The convolution kernels play vital roles, just like feature detector which can learn general and fine-grained features, such as edge, shape and some hidden information [22]. With the similarity features among adjacent regions, the pooling layer can reduce redundant information through acquiring the maximum or mean of a region, and therefore high feature dimensions, vast network parameters and long training time will be reduced drastically. In the fully connected layer, all neurons have full connection to the output of the previous layer. The fully connected layer converges all learned features to classify, finally outputs the classification score and give the actual prediction of input data. Fig. 2 shows the network structure of VGG16 CNN model which consists of five convolution blocks and a fully connected layer containing one flatten layer and two dense layers.

However, the training of VGG16 network in Fig. 1 with small dataset will cause overfitting problem in a large probability. Therefore we used dropout layer and transfer learning technique to alleviate overfitting phenomenon. Dropout layer set the output of neurons in the hidden layers into 0 with a probability of r. Transfer learning technique is to utilize the pre-trained weights to initialize own network whose structure is the same as pre-trained model trained by a much larger dataset, thus realizes self-adaption from source to target domain [23,24]. In this study, we transferred VGG16 pre-trained weights trained by nature image dataset *Imagenet* with 1000 categories into own VGG16 network. Although brain image is very different from nature images, the first few layers in CNN can extract many generic features, such as side, angle, colour. According to the difference of category number and image attribute between source domain and target domain, we replaced fully connected layer and freezed the pre-trained weights of the first four convolution blocks of VGG16, then the pre-trained weights in the fifth convolution block and the initial weights of fully connected layer were continually updated during training process, the above courses are also called fine-tuning. The freezed layers will be used to extract generic feature while the fine-tuned layer extract high-level target-specific features.

2.4.2. LASSO

With high-dimensional features of each subject, the feature selection algorithm of least absolute shrinkage and selection operator (LASSO) was performed to remove the irrelevant or redundant features so that the feature dimensions would be reduced drastically and overfitting phenomenon can be alleviated effectively. LASSO is performed through minimizing the penalized objective function with L1 regularization, which tends to give zero weight to irrelevant features; therefore the useful discriminative features can be saved [25]. The objective function

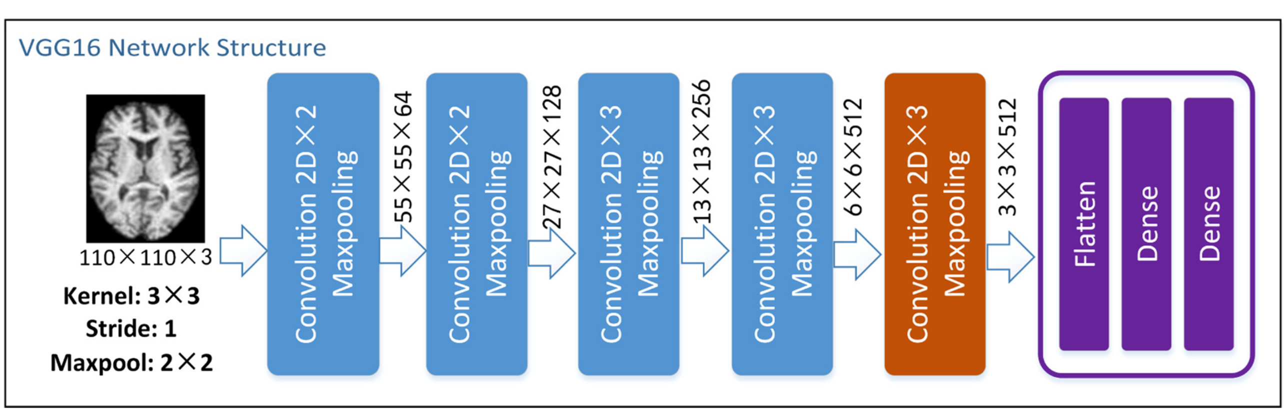

Fig. 2. VGG16 network structure, the gray-scale image in target domain with the size of 110 × 110 × 1 will be converted into the format of RGB in source domain with the size of 110 × 110 × 3.

of LASSO is defined as follows:

$$J(\theta) = \frac{1}{2}\|Y - X^T\theta\|_2^2 + \alpha\|\theta\|_1 \tag{1}$$

where $X = [x_1, x_2, ..., x_N] \in R^{N\times d}$ is a feature matrix. N is the number of subjects and d is the number of features of each subject, thus each x_i represents all features of one subject. $Y = \{y_i | y_i \in \{-1, +1\}\}_{i=1}^N$ is a set of corresponding class labels of subjects. θ represents the regression coefficient and α is the regularization parameter to balance the complexity of the model.

2.4.3. SVM

Support vector machine is suitable for the classification of high-dimensional small dataset [26], and Peng Yang et al. [27] and Xiaoke Hao et al. [28] found that SVM is superior to other classifiers in EMCI diagnosis. Given a training set $\{x_k, y_k\}_{k=1}^N$ with input data $x_k \in R^n$ and corresponding binary class labels $y_k \in \{-1, +1\}$, the output of primal SVM is presented as follows:

$$y(x) = \mathrm{sign}[w^T\varphi(x) + b] \tag{2}$$

where $\varphi(x)$ is a nonlinear function mapping the input space to higher dimensional feature space, which makes the input data linearly separable in hyperplane. The term b is a bias term. The optimization objective function is defined as follows [26]:

$$\min_{w,b,\xi} J(w, \xi) = \frac{1}{2}w^Tw + c\sum_{k=1}^{N}\xi_k \tag{3}$$

subject to:

$$y_k[w^T\varphi(x_k) + b] \geq 1 - \xi_k,\ k = 1, ..., N,\ \xi_k \geq 0 \tag{4}$$

ξ_k is slack variable which can allow model to appear misclassification. W is the weight applied for input data x. The positive constant c is a tuning parameter.

2.5. Implementation

The training of VGG16 network is implemented based on Keras with a single GPU (i.e. NVIDIA GTX TITAN 12GB). The network is optimized by root mean square propagation (RMSProp) with a learning rate of 10^{-4}. The weight update is performed in mini-batches of 32 samples per batch and stops after 100 epochs.

3. Results

3.1. Experimental setting

We first implemented all the steps in Fig. 1, then in order to highlight the advantage of transfer learning technique, we also trained VGG16 model from scratch without transferring pre-trained weights. In addition, to demonstrate the effectiveness of feature selection algorithm, we also made another contrast test where LASSO algorithm did not perform before training SVM classifier. Finally, we compared the proposed approach with other present methods based on sMRI and EMCI diagnosis. In the following, there are some corresponding experimental setup illustrated below.

As described in section 2, the regularization parameter α of LASSO feature selection algorithm can balance the model complexity. we found from many trial experiments that the magnitude of α will influence the classification results to some extent. The variation curve of classification accuracy of EMCI and NC changing with α is shown in Fig. 3, where the range of α varies from 0 to 10 and the interval is 0.1. However, we did not select the α corresponding to the highest accuracy because the selected features will lead to overfitting in classification. In order to acquire a relatively robust classification model, an appropriate α would be selected through repeating experiment with different α.

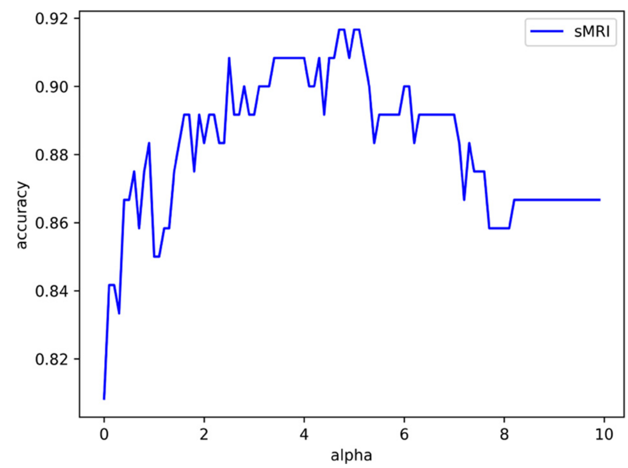

Fig. 3. The variation curve of classification accuracy changing with α.

To alleviate the effect of data abnormity as far as possible, the original data distribution of features was normalized as a normal distribution with unit standard deviation and zero mean before SVM classification. In order to obtain the optimal parameters of SVM classifier, we used *GridSesearchCV* [29] method to generate the training results of different parameter combinations through exhaustive search, and then acquired the optimal parameters according to validation accuracy. We repeated SVM classification task with optimal parameters 10 times using stratified 10 folds cross-validation strategy and acquired the mean of every metric as the final results. In addition, in order to alleviate the problem of data imbalance, class weights were imposed during training.

3.2. Features extraction and visualization

CNN is believed to have great ability of extracting class-discriminative features from images, and visualization of the activation map of features is a helpful approach to explore training progress of CNN model. We visualized the output of the first Maxpooling layer of VGG16 model, as shown in Fig. 4(a). Different filter (or convolution kernel) learns different features from various aspects, for example, some filters learn the brain shape, and others learn the interior structure of brain. In Maxpooling layer, after taking the maximum value within a 2×2 region in turn in whole feature map, the size of feature map in previous layer was decreased in half and the details became more obvious. The features will become more and more sparse and localized in deeper layers, at the same time, the embedded high-level abstract features can be learned.

Fig. 4(b) shows the feature matrix output of LASSO algorithm, which was also described in Fig. 1. There are apparent distinctions between 70 EMCI and 50 NC subjects in Fig. 4(b) where it is clearly seen that the color of the features of NC is more green while that of EMCI is more pink. The well discrimination between EMCI and NC benefited by the ability of feature extraction of VGG16 model to a large extent.

3.3. Performance evaluation

We quantitatively evaluated the classification performance of the proposed method based on accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under the receiver operating characteristic curve (AUC). The ROC curves of one 10-folds cross-validation are shown in Fig. 5. To our best knowledge, we are the first to combine CNN with traditional machine learning algorithm for EMCI diagnosis. In order to compare the classification performance with other recent methods reasonably, we chose these researches which used sMRI data from ADNI

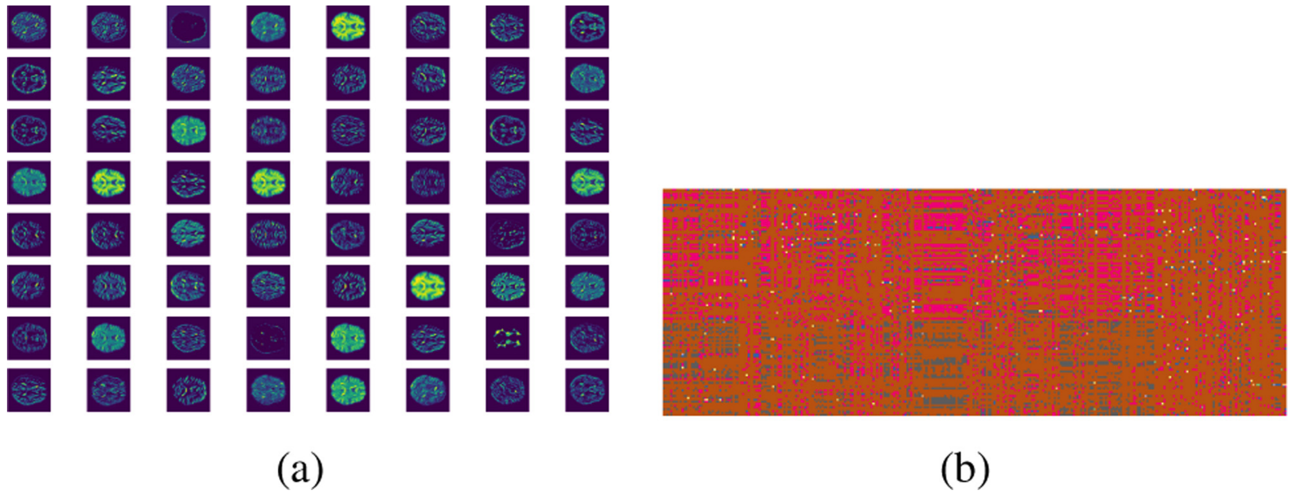

Fig. 4. (a) 64 feature activation maps of a sMRI slice with the size of 55 × 55 pixels, the output of the first Maxpooling layer of VGG16 model. (b) The feature matrix with the size of 120 × 330 from feature selection.

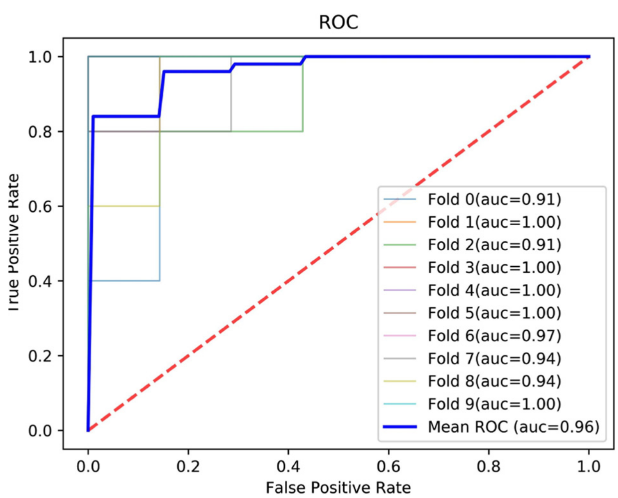

Fig. 5. The ROC curve of SVM classifier, where the blue-black ROC curve is the average of ten ROC curves of 10-folds cross-validation.

Table 2
Classification performance comparison with pervious researches using sMRI for distinguishing EMCI from NC. Where *from scratch* and *no LASSO* respectively represents transfer learning technique and LASSO algorithm did not performed in experiment. Fine-tune & LASSO is the complete process of the proposed method.

Method	ACC(%)	SEN(%)	SPE(%)
Forouzannezhad et al. [9]	61.1	66.5	58.7
Forouzannezhad et al. [10]	73.1	76.8	68.4
Raeper et al. [8]	81.0	83.3	78.57
ours (from scratch)	63.3	72.0	56.3
ours (no LASSO)	80.0	84.0	75.6
ours (fine-tune &LASSO)	**89.4**	**92.9**	**89.3**

website, as shown in Table 2. With the same metrics, it can be clearly seen that the proposed method yielded the best results whatever ACC, SEN or SPE. An mean AUC of 96% reported in Fig. 5 also illustrates that the classifier has fairly well classification capacity for positive samples and negative samples.

On the one hand, under the condition of using the same sMRI modality, our proposed method outperformed paper [8–10], achieving a accuracy of 89.4%, 8.4% higher than paper [8], 16.3 % higher than paper [10] and 28.3 % higher than paper [9]. It illuminates that the proposed CNN-based method is more effective than other methods just based on traditional machine learning algorithm. CNN plays a significant role in this study considering that it can efficiently extract useful features in different level and reduce the error resulting from incomplete prior hypothesis to some extent. For example, EMCI may have different pathosis from AD, such as degree of lesion, while the prior hypothesis of some researches about EMCI lesion still based on AD. On the contrary, CNN can ignore this difference between EMCI and AD better than prior hypothesis because CNN can automatically extract the most discriminative features no matter what the pathology is. On the other hand, the accuracy of using the features extracted from fine-tuned VGG16 model is much higher than that of using the features from the VGG16 model trained from scratch. The main reason is that VGG16 model trained from scratch cannot learn enough features with small dataset, conversely, the VGG16 model transferring pre-trained weights trained by much bigger source-domain dataset has good generalization for small dataset, which can extract many general features of target-domain data from the first few layers without updating weights. It also illustrates that model trained on mass of general images can be fine-tuned by few MRI images to extract specific target-domain features.

In addition, LASSO algorithm is a key step to improve classification accuracy in this study. As shown in Table 2, the LASSO algorithm improve the accuracy by 9.4%. It demonstrates that the LASSO algorithm can effectively eliminate redundant features that affect the reasonable distribution of data in SVM feature space. It also explains that there are only subtle but manifest changes for EMCI subjects; therefore these changes are sensitive to vast redundant features.

4. Discussion

The incidence of AD is increasing rapidly every year; therefore it is urgent to find effective methods to delay and prevent AD. Diagnosis of EMCI is helpful for early intervention of AD. We reviewed previous researches about EMCI diagnosis and proposed a more effective method combining CNN and SVM to distinguish EMCI from NC. The results suggest that the classification performance is significantly improved compared with the previous researches, and the average accuracy achieves an unprecedented 89.4%. Two factors are identified as being potentially important: 1) LASSO can deal with a small number of subjects with high-dimensional features and get rid of a large number of redundant features. 2) Transfer learning based on CNN can greatly enhance the learning ability of small dataset and helps excavate more target-domain high-level features. Additionally, transfer learning can reduce training time due to the usage of pre-trained model realizing faster convergence with less data. Similarly, Wee et al. [30] used the model pre-trained on the ADNI-2 cohort (including NC, EMCI, LMCI, AD) to test the ADNI-1 cohort (including NC, MCI, AD), and the experimental results illustrated that the classification performance of the fine-tuned model were better than the learn-from-scratch model, which is consistent with our experimental results. Actually, they constructed cortical thickness graphs using sMRI data and input them into a cortical graph neural network including three graph convolutional layers and a fully connected layer, achieving 51.8% accuracy for EMCI vs. NC. The possible reason of low classification accuracy is that the selected cortical thickness feature is too single for identifying EMCI and NC.

In recent years, CNN gradually becomes an important tool of analyzing medical image with small dataset. Lulu Yue et al. [21] and Gorji et al. [31] made 3D gray matter images of sMRI data decompose into many 2D slices to train the CNN including three convolutional layers and sub-sampling layers. Although they achieved the high accuracy of more than 90% for identifying EMCI, the classification results are based on slice-level. Similarly, Yosra Kazemi et al. [32] and Saman Sarraf et al. [33] acquired the high slice-level accuracy in AD classification using 2D slices images. Apparently, the subject-level accuracy can be further acquired through majority voting mechanism in the above studies, however, we acquired the subject-level accuracy through integrating all slice features of each subject for SVM classification, which can obtain disease-related information from all slices for prediction rather than from one slice.

In this study, the classification results using CNN method still have a certain gap with clinical diagnosis. The possible reason is that there are no significative difference between some EMCI subjects and NC in brain structure. Although sMRI is one of the commonest neuroimaging tool for disease diagnosis, more and more researchers turn their attention from structure change of brain to functional change [17–19] in recent years. They used fMRI data to construct brain functional networks of EMCI group and NC group for classification and achieved over 80% accuracy. It is worth noting that they found that temporal lobe is the discriminating disease-related region, in this study, we have intentionally selected all 2D slices including temporal lobe. In addition, there are many researches illustrating that multi-modality data are more effective than single modality data for EMCI classification [28,34,35]. Therefore, we will combine other neuroimaging techniques to further improve the classification performance utilizing the proposed method and use bigger dataset to enhance the robustness of model in the following study.

Acknowledgment

We are grateful for the continued support from the National Natural Science Foundation of China (No. 81960312) and the Natural Science Foundation of Guangdong Province (No. 2018A030310511).

Data collection and sharing for this study 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).

References

[1] A. Association, Alzheimer’s disease facts and figures, Alzheimers Dementia 10 (2) (2014) e47–e92.
[2] A. Association, Alzheimer’s disease facts and figures, Alzheimers Dementia 14 (3) (2018) 367–429.
[3] L.S. Schneider, Mild cognitive impairment, Lancet 367 (9518) (2006) 1262.
[4] E.Y. Lim, Comparison of diagnostic accuracy of volumetry or diffusion tensor imaging in mild cognitive impairment, Alzheimers Dementia 12 (7) (2016) P728–P729.
[5] W. Wen, W. Li, The mathematical model of mild cognitive impairment based on brain network analysis, in 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference, March, 2019, pp. 2234–2238.
[6] Y. Zhang, Z. Dong, et al., Magnetic resonance brain image classification via stationary wavelet transform and generalized eigenvalue proximal support vector machine, Journal of Medical Imaging and Health Informatics 5 (7) (2015).
[7] G. Prasad, S.H. Joshi, et al., Brain connectivity and novel network measures for Alzheimer’s disease classification, Neurobiol. Aging 36 (3) (2015) S121–S131.
[8] R. Raeper, A. Lisowska, I. Rekik, Cooperative correlational and discriminative ensemble classifier learning for early dementia diagnosis using morphological brain multiplexes, IEEE Access PP (99) (2018) 1–1.
[9] P. Forouzannezhad, A. Abbaspour, et al., A deep neural network approach for early diagnosis of mild cognitive impairment using multiple features, in 2018 17th IEEE International Conference on Machine Learning and Applications, Dec (2018) 1341–1346.
[10] P. Forouzannezhad, A. Abbaspour, et al., Early diagnosis of mild cognitive impairment using random forest feature selection, in 2018 IEEE Biomedical Circuits and Systems Conference, Oct (2018) 1–4.
[11] Z. Qi, M. Goryawala, et al., Multivariate analysis of structural mri and pet (FDG and 18F-AV-45) for Alzheimer's disease and its prodromal stages, Conf Proc IEEE Eng Med Biol Soc (2014) 1051–1054.
[12] Y. Cabrera-Len, P.G. Bez, et al., Classification of mild cognitive impairment stages using machine learning methods, In 2018 IEEE 22nd International Conference on Intelligent Engineering Systems, June 000067-000072, 2018.
[13] J. Islam, Y. Zhang, An Ensemble of Deep Convolutional Neural Networks for Alzheimer's Disease Detection and Classification, (2017) arXiv e-prints, p. arXiv:1712.01675, Dec.
[14] L. Yue, X. Gong, et al., Hierarchical feature extraction for early Alzheimers disease diagnosis, IEEE Access 7 (2019) 93752-93760.
[15] T. Kam, H. Zhang, et al., Deep learning of static and dynamic brain functional networks for early MCI detection, IEEE Transactions on Medical Imaging (2019) 1–1.
[16] M. Puranik, H. Shah, et al., Intelligent Alzheimer's detector using deep learning, in 2018 Second International Conference on Intelligent Computing and Control Systems 06 (2018) 318–323.
[17] T. Zhang, Z. Zhao, et al., Classification of early and late mild cognitive impairment using functional brain network of resting-state fmri, Front. Psychiatry 10 (2019) 572.
[18] a. Bi, Q. Xu, et al., Analysis of progression toward Alzheimer’s disease based on evolutionary weighted random support vector machine cluster, Front. Neurosci. 12 (2018) 716.
[19] X. Cui, J. Xiang, et al., Integrating the local property and topological structure in the minimum spanning tree brain functional network for classification of early mild cognitive impairment, Front. Neurosci. 12 (2018) 701.
[20] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, Comput. Sci. (2014).
[21] L. Yue, X. Gong, et al., Auto-detection of Alzheimer's disease using deep convolutional neural networks, in 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, July, 2018, pp. 228–234.
[22] S.P.K. Karri, D. Chakraborty, J. Chatterjee, Demnet: A convolutional neural network for the detection of alzheimers disease and mild cognitive impairment, Biomed. Opt. Express Feb 8 (2) (2017) 579–592.
[23] S.J. Pan, Y. Qiang, A survey on transfer learning, IEEE Trans. Knowledge Data Eng. 22 (10) (2010) 1345–1359.
[24] M. Hon, N. Khan, Towards Alzheimer's disease classification through transfer learning, in 2017 IEEE International Conference on Bioinformatics and Biomedicine, 11 (2017).
[25] Z. Jiao, Z. Xia, et al., Multi-scale feature combination of brain functional network for eMCI classification, IEEE Access PP 06 (2019) 1-1.
[26] J.A.K. Suykens, Support Vector Machines: A nonlinear modelling and control perspective, Eur. J. Control 7 (2–3) (2001) 311–327.
[27] P. Yang, F. Zhou, et al., Fused sparse network learning for longitudinal analysis of mild cognitive impairment, IEEE Trans. Cybernet. (2019) 1–14.
[28] X. Hao, Y. Bao, et al., Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer’s disease, Med. Image Anal. 60 (2020) 101625.
[29] F. Pedregosa, G. Varoquaux, et al., Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[30] C.Y. Wee, C. Liu, et al., Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations, NeuroImage: Clinical 23 (2019) 101929.
[31] G.H. Taheri, K. Naima, A deep learning approach for diagnosis of mild cognitive impairment based on MRI images, Brain Sci. 9 (9) (2019) 217.
[32] S. Sarraf, G. Tofighi, Deep learning-based pipeline to recognize Alzheimer's disease using fmri data, in 2016 Future Technologies Conference (FTC), Dec (2016) 816–820.
[33] Y. Kazemi, S. Houghten, A deep learning pipeline to classify different stages of Alzheimer's disease from fmri data, in 2018 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), May (2018) 1–8.
[34] P. Forouzannezhad, A. Abbaspour, et al., A gaussian-based model for early detection of mild cognitive impairment using multimodal neuroimaging, J. Neurosci. Methods 333 (2020) 108544.
[35] N. Amoroso, D. Diacono, et al., Deep learning reveals Alzheimer’s disease onset in MCI subjects: Results from an international challenge, J. Neurosci. Methods 302 (2018) 3–9.

Jingwan Jiang received the B.S. degree in Electronic Information Science and Technology from Wuhan University of Technology in 2018. She is now a post-graduate in College of Information Engineering, Shenzhen University, China. Her research interests include medical image processing and intelligent information processing.

Li Kang received the PhD degree in signal and Information Processing from Xidian University in 2008. She was a Postdoc with the Optical Engineering at Shenzhen University. She is now a lecturer in College of Information Engineering, Shenzhen University, China. Her research interests include medical image processing and intelligent information processing.

Jianjun Huang (M'02) received the Ph.D degree in Signal and Information Processing from Xidian University in 1997. He was a Postdoc in Computer Sciences with Northwestern Polytechnical University from 1997 to 1999. He is now a professor with College of Information Engineering, Shenzhen University, China. His current research interests include intelligent information processing, computer vision and medical image processing.

Tijiang Zhang received the Ph.D.degree of Medical Imaging in West China Medical School of Sichuan University in 2010. The director of Medical Imaging Center of Guizhou Province and Department of Radiology, the Affiliated Hospital of Zunyi Medical University. The 11th batch Youth technology talented person of Guizhou province. He is now professor with Radiology, the Affiliated Hospital of Zunyi Medical University. His current research interests include imaging diagnosis of Central nervous system and clinical application of high field MRI.