Automatic Recognition of Mild Cognitive Impairment from MRI Images Using Expedited Convolutional Neural Networks

Shuqiang Wang[1(✉)], Yanyan Shen[1], Wei Chen[2], Tengfei Xiao[3], and Jinxing Hu[1]

[1] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
sq.wang@siat.ac.cn

[2] Department of Diagnostic Radiology, Wake Forest University, Winston-salem, USA

[3] School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China

Abstract. Few studies have focused on the potential of applying deep learning algorithms into magnetic resonance imaging (MRI) for automatic recognition of subjects with mild cognitive impairment (MCI). In this work, we propose the expedited convolutional neural networks involving Tucker decomposition to recognize MCI using MRI images. We employ transfer learning and data augmentation to deal with limited training data. The effect of Tucker decomposition on saving computational time is discussed. The experimental results show that the proposed model outperforms the previous methods. The expedited convolutional neural networks can provide a good guidance for the applications of deep learning in real-world classification with large training dataset.

Keywords: Expedited convolutional neural network · Mild cognitive impairment · Tucker decomposition · Magnetic resonance imaging

1 Introduction

Mild cognitive impairment (MCI) is a syndrome defined as cognitive decline greater than expected for an individual's age and education level but that does not interfere notably with activities of daily life. It indicates an intermediate state between Alzheimers disease (AD) and healthy aging. For each year, about 10–15% of patients with MCI develop dementia while only 1–2% of healthy controls develop dementia [1,2]. Thus, MCI is often seen as a feasible aim for the early detection of Alzheimers disease. The detection of MCI is playing an important role in understanding of AD and AD drug development. However, the existing tools for the detection of MCI and AD, such as the CERAD [3], often take too much time and cost. Moreover, the objectivity of CERAD is contentious.

A. Lintas et al. (Eds.): ICANN 2017, Part I, LNCS 10613, pp. 373–380, 2017.
https://doi.org/10.1007/978-3-319-68600-4_43

In this work, we propose a method of combining magnetic resonance imaging (MRI) together with deep learning tools to automatically recognize MCI. Deep learning can be seen as an improvement of artificial neural networks, including more hidden layers that allow higher levels of abstraction and advanced predictions from a rich supply of data [4]. It is a growing trend in general data analysis and has been termed one of the 10 breakthrough technologies of 2013. Nowadays, deep learning is becoming a leading machine-learning tool in the general imaging and computer vision domains [5–8].

In particular, convolutional neural networks (CNNs) have presented outstanding effectiveness on medical image computing problems and made a lot of improvement for computer-aided detection [9]. Roth et al. [10] employed convolutional neural networks to improve three existing CAD systems for the recognition of colonic polyps on CT colonography, sclerotic spine metastases on body CT and enlarged lymph nodes on body CT. Dou et al. [11] used 3D CNN and weighted MRI scans to detect cerebral microbleeds. They address developed predictions with their 3D CNN compared to various classical and 2D CNN approaches. Sirinukunwattana et al. [12] employed CNNs to detect nuclei in histopathological images. In their work, they take small patches as input and model the output as a high peak in the vicinity of the center of each nucleus and flat elsewhere. Anthimopoulos et al. [13] employed CNNs to detect patterns of interstitial lung diseases from 2D patches of chest CT scans. Their results show that CNNs can outperform existing methods that use hand-crafted features. However, there is a big challenge for applying deep learning into medical images. The deep learning can only benefit from large amounts of training data while it is difficult in medical area. The reasons are mainly due to high cost and privacy issues [9,14].

In the current work, we propose an expedited convolutional neural network by involving tensor decomposition for classification of MCI and control subject. We employ transfer learning and data augmentation to deal with limited training data. The rest of this paper is organized as follows. In Sect. 2, the expedited convolutional neural networks are proposed, and the schemes for dealing with limited training data are given. The experimental results are presented in Sect. 3. Finally, Sect. 4 gives conclusions and future work.

2 Methods

2.1 Proposed Expedited Convolutional Neural Networks

Compared with the general CNNs, the architecture of our proposed tensor deep neural network mainly contains two aspects: (1) in each convolutional layer, decomposing its kernel and the training samples using Tucker decomposition [15], (2) Fine-tuning the entire networks using backpropagation. In the following, we first review the Tucker decomposition.

Tucker decomposition decomposes a tensor into a core tensor, multiplied by a matrix along each mode. The core tensor usually serves as the relationship/interaction between the modes. Let $\mathcal{X} \in R^{I_1 \times I_2 \times \cdots \times I_M}$ be a tensor and $\mathcal{X}$ can be written as:

$$\mathcal{X} = \mathcal{G} \times_1 A^{(1)} \times_2 \cdots \times_M A^M = [\mathcal{G}; A^{(1)}, A^{(2)}, \cdots, A^{(M)}] \tag{1}$$

We call the Eq. (1) the Tucker decomposition of tensor "$\mathcal{X}$" with r-rank, where $\mathbf{r} = (r_1, \cdots, r_d)$, $\mathcal{G} \in R^{r_1 \times r_2 \times \cdots \times r_d}$ is a core tensor, and $A^{(k)} \in R^{r_k \times I_k}$ is the factor matrix along to the kth mode, and the notation $\times_k$ is the k-mode multiply operation.

In the proposed model, the units within CNN are organized as a sequence of third-order tensors with two spatial dimensions and the third dimension corresponding to different "channels". The convolutional kernel and the corresponding training sample are decomposed respectively using Tucker decomposition before the convolution operation. The architecture of the proposed expedited convolutional neural networks is shown in Fig. 1.

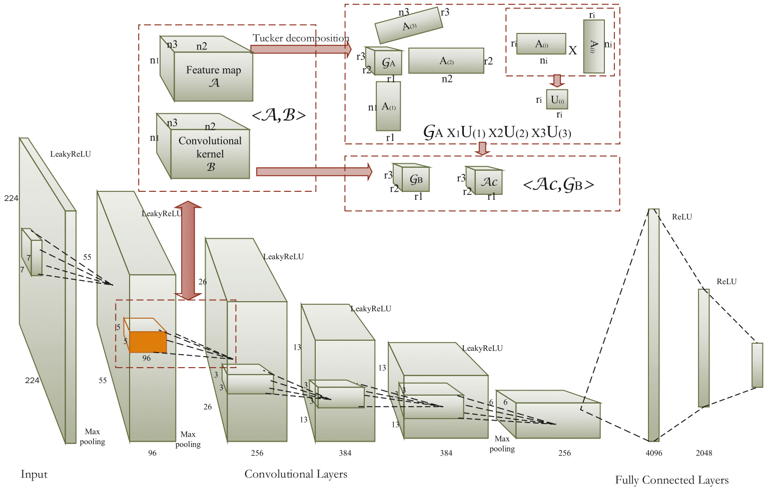

Fig. 1. Architecture of the proposed CNN for MCI recognition

The proposed networks consist nine layers including the input and output. There are five alternating convolutional and pooling layers followed by two fully connected layers and output layer. In convolutional layers, we employ five different filters to get feature maps. In this work, we choose Leaky-ReLU as activation functions. Leaky-ReLU can solve the dying ReLU problem that the tendency of ReLU for keeping a neuron constantly inactive as may happen after a large gradient update. In pooling layers, we employ max pooling. We choose max pooling due to the following two reasons: (1) By deleting non-maximal values, it can reduce computation for upper layers. (2) It provides a form of translation invariance and is a "smart" way of reducing the dimensionality of intermediate representations.

2.2 Schemes for Dealing with Limited Training Data

Deep learning enables highly representative features that have been demonstrated to be a very strong and robust representation in many application domains. However, it only can benefit from large amounts of training data. A key challenge in applying deep convolutional neural networks is that sufficient training data are not always available in medical images. In this work, we employ two schemes to deal with this issue: (1) Data augmentation enables generating new training data from a smaller data set such that the new data set represents the real-world data one may see in practice. Augmentation can extract as much information from data as possible. In this work, we employ the following data augmentation techniques: brightness augmentation, horizontal and vertical shifts, shadow augmentation and flipping. One example of data augmentation is shown in Fig. 2. (2) Transfer learning and fine tuning: we first employ different medical dataset or natural image dataset to pre-train the proposed CNN model. We take the pre-trained model as initialization of the network. Then, we further conduct supervised training on several or all the network layers using the new medical data for the task.

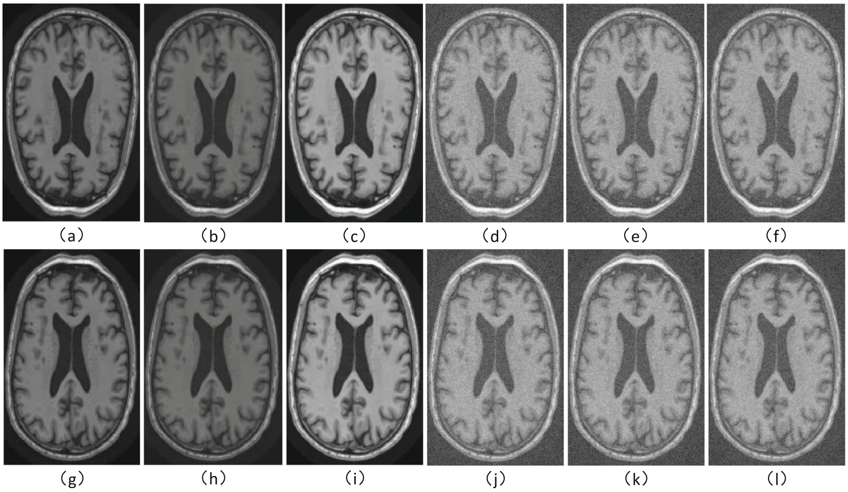

Fig. 2. Data augmentation for MRI image of MCI. (a) is an original image and (b)–(l) are the images generated by data augmentation.

3 Experiment and Results

3.1 Datasets

In this work, we employ the MRI dataset from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu/methods/mri-analysis/adni-standardized-data/). The selected dataset includes serial brain MRI scans from 400 individuals with MCI (age: 74.8 ± 7.4 years, 257 Male/143 Female), and 229 healthy elderly controls (age: 76.0 ± 5.0 years, 119 Male/110 Female). Subjects were scanned at screening and followed up at 6, 12, 18 (MCI), 24, and 36 months (MCI and normal). For more details, please refer to Ref. [16]. After data augmentation, we obtain 8000 images including 4000 images of MCI and 4000 images of healthy control. We extract 5000 images for training, 1500 images for validation, 1500 images for testing.

3.2 Evaluation

The proposed CNN model is evaluated by a train-validation-test scheme. The actual training of the method is carried-out on the training set, while the validation set is used for fine tuning the hyper-parameters; the overall performance of each system is assessed on the test set. The performance of the proposed model is evaluated by F-score and accuracy. F-score is given by

$$\text{F-score} = \frac{2\text{precision} \times \ \text{recall}}{\text{precision} + \text{recall}},$$

where

$$\text{precision} = \frac{\text{samples correctly classified as c}}{\text{samples classified as c}}$$

and

$$\text{recall} = \frac{\text{samples correctly classified as c}}{\text{samples of class c}}.$$

The accuracy is given by

$$\text{Accuracy} = \frac{\text{correctly classified samples}}{\text{total number of samples}}.$$

3.3 Implementation

The proposed method was implemented using the Theano [17] framework. All the experiments are performed on a computer with Intel(R) Core(TM) i5-4200M CPU @2.50 GHz, GPU NVIDIA GeForce Titan X, and 64 GB of RAM.

3.4 Results

We first demonstrate the effects of pre-training and data augmentation on classification MCI and control subjects. The experimental results are shown in Table 1. LIDC [18] dataset is employed for Pre-training-1 and OASIS [19] dataset is used for Pre-training-2. The best performance with accuracy of 90.6%, recall of 92.8% and F-score of 89.4% is given by the proposed CNNs combining with Pre-training and data augmentation, while the proposed CNNs without any tuning options produce the lowest F-score of 72.8%, accuracy of 74.3%, recall of 74.7% and precision of 71.1%. From Table 1, the following findings can be given: (1) The proposed CNNs combining with pre-training outperform the proposed CNNs

Table 1. Performance of the proposed model with different tuning options.

Model tuning	F-Score (%)	Accuracy (%)	Recall (%)	Precision (%)
No pre-training & no augmentation	72.8	74.3	74.7	71.1
No Pre-training & augmentation	80.2	80.9	81.4	79.0
Pre-training-1 & no augmentation	81.7	82.1	80.8	82.6
Pre-training-2 & no augmentation	83.5	84.2	85.4	81.7
Pre-training-1 & augmentation	87.1	88.8	86.5	87.7
Pre-training-2 & augmentation	89.4	90.6	92.8	87.2

combining with data augmentation. This is mainly because the proposed model can learn more features using different kinds of datasets in the pre-training stage. (2) The effect of transfer learning is influenced by the dataset employed for pre-training. The proposed CNNs combining with pre-training using OASIS dataset outperform the proposed CNNs combining per-training using LIDC dataset.

Table 2. Comparison of the proposed CNNs with previous methods.

Methods	F-Score (%)	Accuracy (%)	Recall (%)	Precision (%)
Wee et al. [20]	81.1	81.5	82.3	79.9
Liu et al. [21]	78.6	79.1	81.5	75.9
Wolz et al. [22]	82.1	82.5	83.7	80.6
Our methods	89.4	90.6	91.2	87.7

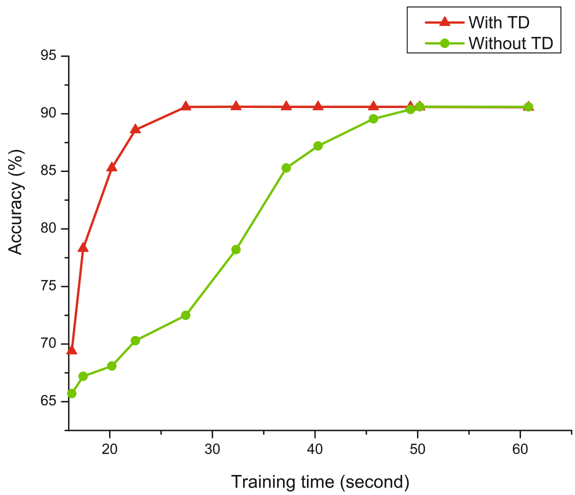

Fig. 3. Testing accuracy versus training time for the proposed CNNs with/without Tucker decomposition. TD is short for Tucker decomposition.

Table 2 shows the comparison of the proposed model with previous methods. It can be seen that our proposed CNNs perform best in classification of MCI with accuracy of 90.6%, recall of 91.2%, precision of 87.7% and F-score of 89.4%, while the best previous method produces an accuracy of 82.5%, a recall of 83.7%, a precision of 83.6% and a F-Score of 82.1%. Figure 3 shows the effect of Tucker decomposition on training time. From Fig. 3, it can be seen that Tucker decomposition can shorten the training time without loss of testing accuracy. In this work, the employed medical dataset is relatively small. Therefore, it seems that the effect of involving Tucker decomposition takes subtler forms. However, it is not a trivial if the dataset of training is large.

4 Conclusion

In this work, we propose the expedited convolutional neural networks for the automatic diagnosis of MCI using MRI images. To address the problem of limited labeled data, we introduce transfer learning and data augmentation into the proposed model. The experimental results demonstrate that the proposed model outperforms the previous methods. In future work, we will focus on the application of the proposed model in classification with large training dataset.

Acknowledgment. This work was supported by Shenzhen Basic Research Projects (Grant No. JCYJ 20160531184426303 and JCYJ 20150401150223648), National Natural Science Foundations of China (Grants No. 61502473 and No. 61503368), and Natural Science Foundation of Guangdong Province (Grant No. 2014A030310154 and No. 2016A030313176).

References

1. Bischkopf, J., Busse, A., Angermeyer, M.C.: Mild cognitive impairment–a review of prevalence, incidence and outcome according to current approaches. Acta Psychiatr. Scand. **106**, 403–414 (2002)
2. Reese, L.C., Laezza, F., Woltjer, R., et al.: Dysregulated phosphorylation of Ca2+/calmodulin-dependent protein kinase II-α in the hippocampus of subjects with mild cognitive impairment and Alzheimers disease. J. Neurochem. **119**, 791–804 (2011)
3. Mirra, S.S., Heyman, A., McKeel, D., et al.: The consortium to establish a registry for Alzheimer's disease (CERAD) part II, standardization of the neuropathologic assessment of Alzheimer's disease. Neurology **41**, 479–486 (1991)
4. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**, 436–444 (2015)
5. Ciodaro, T., Deva, D., De Seixas, J.M., et al.: Online particle detection with neural networks based on topological calorimetry information. J. Phys. Conf. Ser. **368**, 012030 (2012)
6. Turaga, S.C., Murray, J.F., Jain, V., et al.: Convolutional networks can learn to generate affinity graphs for image segmentation. Neural Comput. **22**, 511–538 (2010)

7. Tompson, J., Goroshin, R., Jain, A., et al.: Efficient object localization using convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 648–656 (2015)
8. Taigman, Y., Yang, M., Ranzato, M.A., et al.: Deepface: closing the gap to human-level performance in face verification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1701–1708 (2014)
9. Greenspan, H., van Ginneken, B., Summers, R.M.: Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. IEEE Trans. Med. Imaging **35**, 1153–1159 (2016)
10. Roth, H., Lu, L., Liu, J., et al.: Improving computer-aided detection using convolutional neural networks and random view aggregation. IEEE Trans. Med. Imaging **35**, 1170–1181 (2016)
11. Dou, Q., Chen, H., Yu, L., et al.: Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans. Med. Imaging **35**, 1182–1195 (2016)
12. Sirinukunwattana, K., Raza, S.E.A., Tsang, Y.W., et al.: Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans. Med. Imaging **35**, 1196–1206 (2016)
13. Anthimopoulos, M., Christodoulidis, S., Ebner, L., et al.: Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans. Med. Imaging **35**, 1207–1216 (2016)
14. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
15. Kolda, T.G., Bader, B.W.: Tensor decompositions and applications. Soc. Ind. Appl. Math. Rev. **51**, 455–500 (2009)
16. Wyman, B.T., Harvey, D.J., Crawford, K., et al.: Standardization of analysis sets for reporting results from ADNI MRI data. Alzheimer's Dement. **9**, 332–337 (2013)
17. Bastien, F., Lamblin, P., Pascanu, R., et al.: Theano: new features and speed improvements. arXiv preprint arXiv:1211.5590 (2012)
18. Armato, S.G., McLennan, G., Bidaut, L., et al.: The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Med. Phys. **38**, 915–931 (2011)
19. Marcus, D.S., Wang, T.H., Parker, J., et al.: Open access series of imaging studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J. Cogn. Neurosci. **19**, 1498–1507 (2007)
20. Wee, C.Y., Yap, P.T., Shen, D.: Prediction of Alzheimers disease and mild cognitive impairment using cortical morphological patterns. Hum. Brain Mapp. **34**, 3411–3425 (2013)
21. Liu, X., Tosun, D., Weiner, M.W., Schuff, N.: Locally linear embedding (LLE) for MRI based Alzheimers disease classification. Neuroimage **83**, 148–157 (2013)
22. Wolz, R., Julkunen, V., Koikkalainen, J., et al.: Multi-method analysis of MRI images in early diagnostics of Alzheimers disease. PLoS ONE **6**, e25446 (2011)