Computer Methods and Programs in Biomedicine 221 (2022) 106825

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

Spatio-temporal convolution for classification of alzheimer disease and mild cognitive impairment

Gülce Turhan[a], Haluk Küçük[b,*], Esin Ozturk Isik[a]

[a] Institute of Biomedical Engineering, Bogaziçi University, Istanbul, Turkey
[b] IQS, Department of Industrial Engineering, Universitat Ramon Llull, Barcelona, Spain

ARTICLE INFO

Article history:
Received 22 April 2021
Revised 1 April 2022
Accepted 18 April 2022

Keywords:
CNN
Spatio-temporal convolution
Deep learning
MRI
Alzheimer's disease

ABSTRACT

Background and objective: Dementia refers to the loss of memory and other cognitive abilities. Alzheimer's disease (AD), which patients eventually die from, is the most common cause of dementia. In USA, %60 to %80 of dementia cases, are caused by AD. An estimate of 5.2 million people from all age groups have been diagnosed with AD in 2014. Mild cognitive impairment (MCI) is a preliminary stage of dementia with noticeable changes in patient's cognitive abilities. Individuals, who bear MCI symptoms, are prone to developing AD. Therefore, identification of MCI patients is very critical for a plausible treatment before it reaches to AD, the irreversible stage of this neurodegenerative disease.
Methods: Development of machine learning algorithms have recently gained a significant pace in early diagnosis of Alzheimer's disease (AD). In this study, a (2+1)D convolutional neural network (CNN) architecture has been proposed to distinguish mild cognitive impairment (MCI) from AD, based on structural magnetic resonance imaging (MRI). MRI scans of AD and MCI subjects were procured from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. 507 scans of 223 AD patients and 507 scans of 204 MCI patients were obtained for the computational experiments.
Results: The outcome and robustness of 2D convolutions, 3D convolutions and (2+1)D convolutions were compared. The CNN algorithms incorporated 2 to 6 convolutional layers, depending on the architecture, followed by 4 pooling layers and 3 fully connected layers. (2+1)D convolutional neural network model resulted in the best classification performance with 85% auc score, in addition to an almost two times faster convergence compared to classical 3D CNN methods.
Conclusions: Application of (2+1)D CNN algorithm to large datasets and deeper neural network models can provide a significant advantage in speed, due to its architecture handling images in spatial and temporal dimensions separately.

1. Introduction

Dementia entails reduction in mental abilities affecting the daily life of a patient. Damaged or destroyed neurons in certain regions of the brain associated with cognitive abilities ultimately leads to dementia. Alzheimer's disease (AD) is the most common cause of dementia and its most important symptoms are impaired communication, disorientation, confusion, poor judgment, behavioral changes, and eventually difficulties with visible motor functions such as speaking, swallowing, and walking [1]. AD eventually leads to death of the patient. Mild cognitive impairment is a syndrome due to several reasons among which an underlying reason is sometimes the Alzheimer's disease. Although the family, friends, and the MCI patients themselves are affected, MCI patients can sometimes maintain their autonomy. AD is not always easy to predict based on MCI symptoms of cognitive changes or mild cognitive dysfunction because, the patient histories are not always explicit or lower cognitive ability may be normal. All MCI patients have similar symptomatology but some of them tend to turn into mild dementia and ultimately AD. Therefore, identification of MCI patient is very critical for a plausible treatment before it reaches to AD, the irreversible stage of this neurodegenerative disease. Medical imaging plays a significant role in early diagnosis of AD or other types of dementia. Several machine learning methods have been proposed to assist the diagnostics of dementia from structural brain MRI, by helping to be more specific with respect to brain damage, differential diagnosis and the specific disease patterns.

* Corresponding author.
E-mail address: haluk.kucuk@iqs.url.edu (H. Küçük).

https://doi.org/10.1016/j.cmpb.2022.106825
0169-2607/

Although there has been some progress in AD classification with classical machine learning algorithms, selecting hand crafted features has been challenging [2]. Therefore, recent approaches in AD classification have been towards implementing deep learning methods, such as autoencoders or convolutional neural networks (CNN). A sparse autoencoder followed by a 2D CNN was applied to specific classifications of AD, HC and MCI [3]. Moreover, a 3D CNN convolutional autoencoder was applied to diagnose AD vs MCI vs NC, and their binary combinations [4]. Although CNN was applied to classify stable MCI (sMCI), converted MCI (cMCI) and AD, low accuracies were obtained for AD vs cMCI and MCI vs sMCI classifications [5]. In another study, sMRI and functional positron emission tomography (PET) images were first segmented into descriptive small overlapping patches and then cross-combined before getting fed into a cascade 3D plus 2D CNN algorithm in classification of AD vs NC and progressive (pMCI) vs NC subjects [6]. Additionally, automatic identification of patch and region level spatial relationships in a hierarchical fully convolutional network (H-FCN) was proposed for AD vs NC and pMCI vs sMCI classifications [7]. Moreover, principal component analysis (PCA) and Lasso analysis of CNN features extracted from MRI patches, which were concatenated into a three channel RGB image to compose a colorful patch, was used for prediction of MCI to AD conversion [8]. More recently, CNN with dual learning and an ad hoc layer for 3D separable convolutions was proposed to classify sMCI, pMCI, and AD versus healthy control (HC) subjects [9]. Additionally, random forest feature selection followed by deep neural network classification resulted in a better performance for MCI vs cMCI classification, while a fuzzy model was more precise with AD versus HC classification [10]. More recently, a spectral graph CNN was designed incorporating cortical geometry for improved classification of CN vs. AD, early MCI (EMCI) vs. AD, CN vs. late MCI (LMCI), LMCI vs. AD, EMCI vs. LMCI, and CN vs. EMCI [11]. Later, long short term memory (LSTM) units were employed to capture the temporal dynamics for LMCI vs EMCI and AD vs NC classification [12]. DeepAd was proposed to diagnose AD [13], where LeNet [14] and GoogleNet [15] were implemented on sMRI and fMRI for AD versus NC classification. Finally, 3D subject level CNN, 3D ROI based CNN and 2D sliced based CNN was compared. Moreover, significance of data leakage was examined [16].

In this study, three different CNN architectures, including 2D, 3D and a spatio-temporal (2+1)D CNN model, were implemented for binary classification of AD and MCI using sMRI data obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, and their performances were compared in terms of auc score. Using spatio-temporal (2+1)D convolutional neural network will be valuable for several reasons. For one, this new model will result in comperable accuracy rate with previous approaches. In addition, spatio temporal convolutions will be more robust than 3D convolutional neural networks.

2. Methods

2.1. ADNI dataset and preprocessing

MRI image scans for both AD and MCI subjects were procured from ADNI database [adni.loni.usc.edu]. 3D T1-weighted MPRAGE images were used (TR=2s, TE=2.6ms, 256 × 256 matrix, 160 slices, 1.2 mm slice thickness). Although ADNI database incorporates much higher number of scans for MCI subjects, 507 scans of 223 AD patients and 507 scans of 204 MCI patients were selected, to avoid any possible imbalance in computational experiments. AD patients were diagnosed as AD and stayed stable during the follow-up. MCI patients were diagnosed as MCI, EMCI or LMCI and did not encounter multiple reversions and conversions and did not convert back to AD. For preprocessing, The N4ITK method was used for bias field correction [17]. Next, a linear (affine) registration was performed using the SyN algorithm from ANTs [18] to register each image to the MNI space [19] (ICBM 2009c nonlinear symmetric template). After pre-processing, MRI data was composed of evenly spaced 229 slices with a 193 × 193 in-plane image dimensions. The dataset then resized into 32 × 32 × 32 dimensions.

A 5-fold split was performed only once for all experiments; thus, the same subjects were utilized for such assessments. 15% of the data was split as test set at the very beginning. The test set was reserved until the end of the train/validation process and it was used for calculating the auc scores only. The remaining data was allocated for 5-fold cross validation.

2.2. Convolutional neural networks

CNN is one of the most prominent applications in deep learning, inspired from the brain's visual cortex. It involves four major image plane manipulations: convolutional, pooling, flattening and full connection layers. Typically, it starts with a few convolutional layers, followed by a pooling layer, then another convolutional layer and pooling before connection to the terminal output layer. Flattening layer converts all the images into a single long continuous linear vector. Finally, there is a final fully connected layer that outputs the prediction. CNNs are widely used in image recognition applications with several distinct models been developed, such as Lenet-5 [14], Alexnet [20], GoogleNet [15], and SqueezeNet [21] architectures. In this study three different CNN approaches were compared.

2.2.1. Convolution on sequence of NIFTI images

NIFTI file format was not appropriate for directly feeding into Tensorflow, therefore the data was converted into tf.record format for storing sequence of binary records. By using the tf.record, sequences of 2D axial image inputs were handled in a shape incorporating multiple channels, batch size, sequence length, image height, and image width. In our first experiment we applied 2D convolutions to the entire clip by ignoring the sequence length. In the second experiment, we applied 3D convolutions to whole image width, image height and sequence length dimensions. After converting to tf.record, the images possessed sequence length, image width and image height dimensions. Initially 2D CNN was applied using kernel size; (1, kernel_size, kernel_size). Afterwards, direct 3D CNN was applied incorporating the sequence with (kernel_size, kernel_size, kernel_size) dimensions.

In a recent study on spatio-temporal convolutions for action recognition, it was emphasized that 3D convolutional neural networks might not be appropriate for video understanding, proposing (2+1)D convolutional neural networks as an alternative solution [22]. (2+1)D convolutions splits the computation into 2 operations, in which, 2D convolution generates spatial features, while 1D convolution handles temporal features. Handling 3D CNN in temporal and spatial dimensions might significantly reduce the training time. As a final approach, (2+1)D convolutions were applied to the data increasing the speed of the training by a factor of two, possibly, due to decomposition of 3D convolution into a 2D convolution followed by a 1D convolution operation by the proposed method.

Neural network models were developed in Python programming language, using Keras library with Tensorflow 1.13.1 backend. The hyper-parameters were tuned as, Adam optimizer with 0.001 learning rate, fully connected layers with 0.4 dropout rate and a batch size of 32. For evaluating the performance of the model, binary accuracy and loss values were considered. Accuracy was defined as the percentage of the correctly classified images and the loss function was designated as a measure of binary cross-entropy.

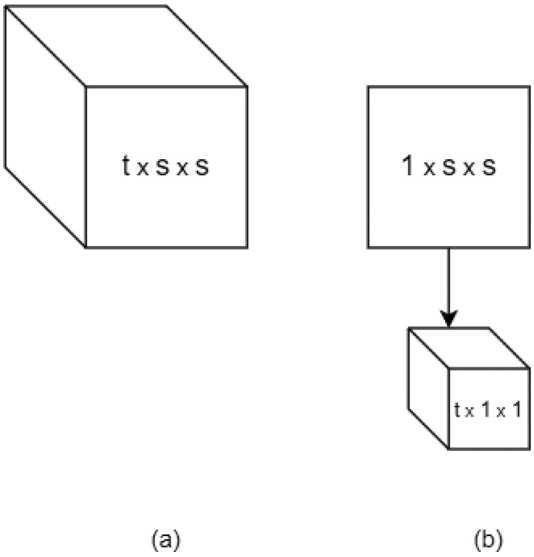

Fig. 1. Illustration of (2+1)D convolutions. a) 3D Convolution with filter size t x s x s where t is temporal extend, s is spatial width and height. b) (2+1)D Convolution where 3D convolution splits into a 2D (spatial) convolution, followed by 1D (temporal) convolution.

A 5-fold cross validation was performed. 20% of the data used for validation and the rest is used for training. A 5-fold split is performed only once for all experiments; thus the same subjects were used for all experiments. Reduce on plateau method was applied when the loss value did not decrease within a period of 10 epochs. In addition, early stopping was applied after repeated measurements of loss which indicates a stagnation for a number of 20 epochs.

The performance of the CNNs is usually expected to improve with the size of the dataset. Accordingly, the diversity of the training dataset was increased by augmentation with incorporating random rotations between 0-10 degrees, brightness adjustment between 0-5, zoom-in-out between 0-0.1, and adding Gaussian noise into the MR images. A similar model was used for all experiments with a different kernel size for 2D, 3D and (2+1)D methods as described in the following sections.

3. Results

3.1. Sequence based 2D CNN

In order to handle 3D images, they were first converted into tf.records and sequences of 2D images were obtained corresponding to each brain scan in NIFTI format initially. The input data shape was in sequence_length, img_height, img_width dimensions.

In this approach, 2D convolutions were applied to the entire clip by ignoring the sequence length (Fig. 2). This has been accomplished by applying 1 × 3 × 3 sized kernels to 3D CNN (Fig. 2), which implicitly ignores the sequence length. The network architecture involves 1 consecutive functional subunit with a set of 2 convolutional layers in each, applying 6 16, filters, respectively. Each functional subunit includes a max pooling layer with a (1 × 2 × 2) sized kernels (Fig. 2) followed by a 0.4 dropout rate. Final dimensions before flattening were 32 × 8 × 8 × 16. The first and the second fully connected layers incorporated 120 and 84 units, respectively. This model has 3,945,677 numbers of trainable parameters.

This model resulted in 0.5+/-0.012 average validation loss, (validation loss per fold: [0.51,0.5,0.49,05.,0.53]) and in 0.77+/-0.033 average validation accuracy, (validation accuracy per fold: [0.79,0.73,0.80,0.73,0.71]) and in 0.84+/-0.038 average auc score, (auc score per fold: [0.85,0.85,0.77,0.88.,0.87]).

3.2. Sequence based 3D CNN

In this approach, 3D CNNs application was carried out over the entire sequence of 2D images. 3D CNN preserves temporal information and propagates this information into the layers of the network [7].

The network architecture included 2 functional subunits each with 1 set of convolutional layers consisting of 6, 16 filters, respectively (Fig. 3). Each functional subunit included a max pooling layer with a kernel size of (2 × 2 × 2) (Fig. 3). The first and the second fully connected layers possessed 120 and 84 units, respectively. Dropout with a 0.4 rate was added after pooling layers. Final data format before flattening had (8 × 8 × 8 × 16) dimensions. The first and the second fully connected layers possessed 120 and 84 units, respectively. This model has 1,006,757 numbers of trainable parameters.

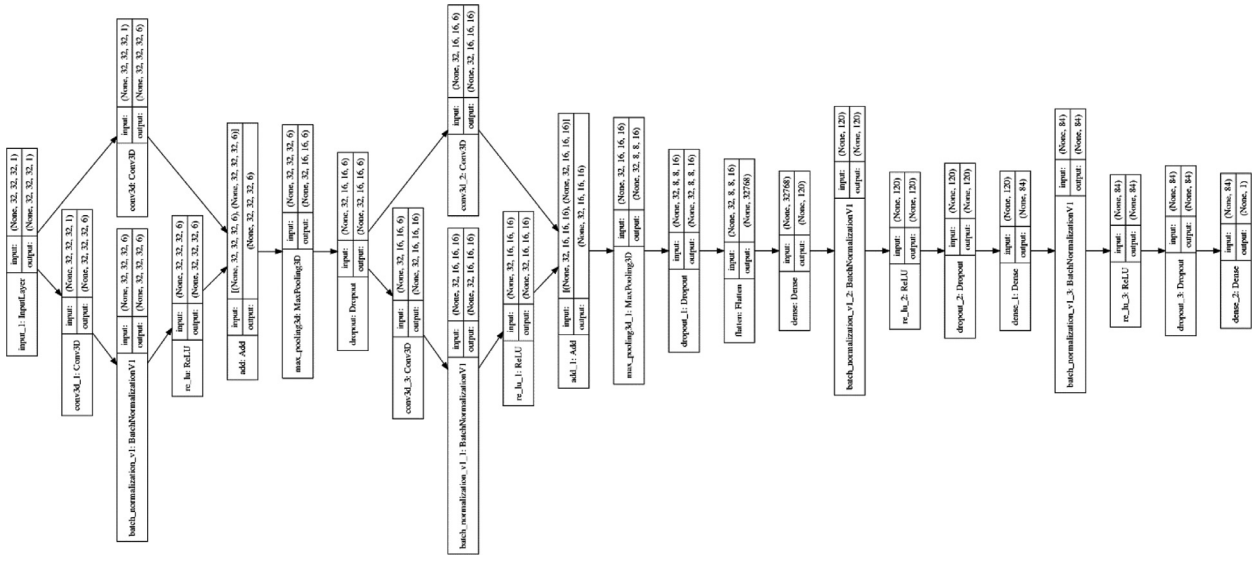

Fig. 2. Sequence based 2D CNN.

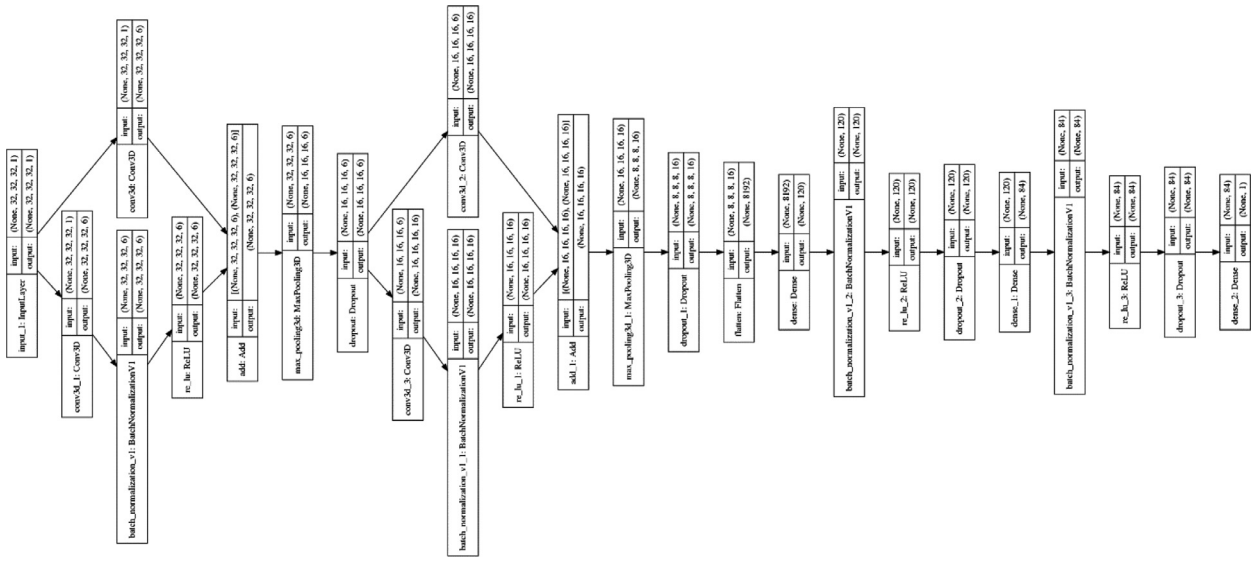

Fig. 3. Sequence based 3D CNN.

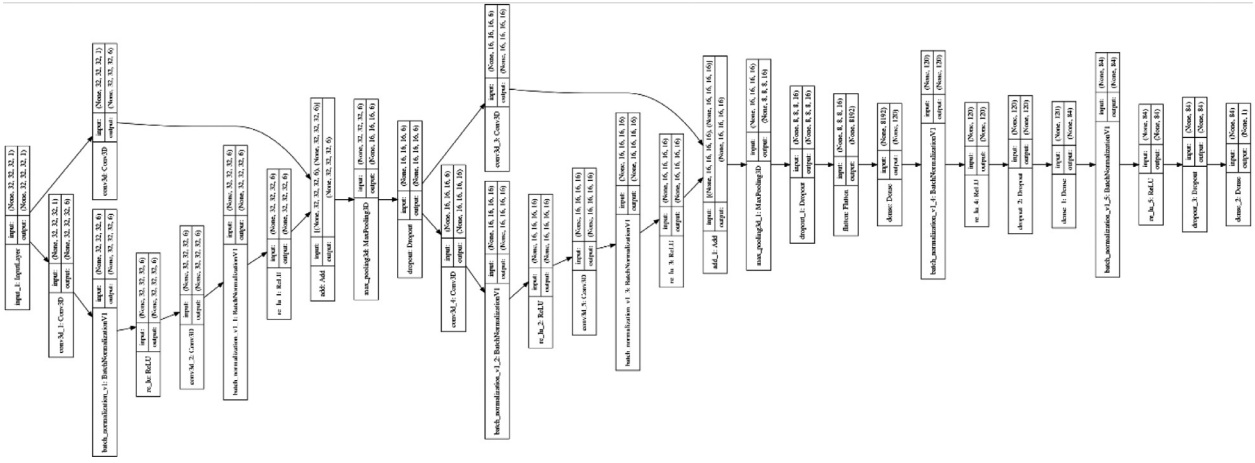

Fig. 4. Sequence based (2+1)D CNN.

Performing convolutions over the whole sequence resulted in 0.56+/-0.047 average validation loss, (validation loss per fold: [0.53,0.66,0.56,056.,0.53]) and in 0.7+/-0.058 average validation accuracy, (validation accuracy per fold: [0.79,0.66,0.62,0.73,0.71]) and in 0.81+/-0.087 average auc score, (auc score per fold: [0.87,0.65,0.80,088.,0.86]).

Interpretation of the whole sequence length did not improve the progression of learning. Although training of 3D CNN took a longer time compared to 2D CNN, accuracy and auc score values 2D CNN was higher, which indicates the lack of 3D CNN in the improvement of temporal features [22].

3.3. Sequence based (2+1)D CNN

Full 3D convolutions were approximated by a 2D convolution followed by a 1D convolution [22]. To perform this type of decomposition, 2D (spatial) convolutional filters were resized to (1 x kernel_size x kernel_size) and a 1D (temporal) convolution in a modified dimension (kernel_size x 1 $\times$ 1). This spatio-temporal decomposition can be applied to any 3D convolutional layer.

(2+1)D network architecture consists of 4 major subunits each with 6 convolutional layers and a single max pooling layer (Fig. 4). Dropout with a 0.4 rate was added after pooling layers. Final data format before flattening had (8 $\times$ 8 $\times$ 8 $\times$ 16) dimensions. The first and the second fully connected layers possessed 120 and 84 units, respectively. This model has 998,083numbers of trainable parameters.

Deep learning of the data using (2+1)D model resulted in a 0.47+/-0.033 average validation loss, (validation loss per fold: [0.51,0.51,0.46,0.48,0.42]) and in 0.78+/-0.032 average validation accuracy, (validation accuracy per fold: [0.74,0.79,0.78,0.84.,0.78]) and in 0.85+/-0.046 average auc score, (auc score per fold: [0.86,0.94,0.86,0.80.,0.83]). When different approaches employed in this study were compared, (2+1)D method demonstrated the best auc score with comparable accuracy values and a less amount of parameters (Table 1).

Considering the minimal preprocessing and subject level data split, the proposed model performed comparably well to the recent studies for similar classification problems (Table 2). Although the previous studies did not mention speed values, considering the models including 3D CNN, the proposed (2+1)D model offers a much faster iteration.

4. Discussions

Alzheimer's disease is a debilitating neurodegenerative disorder of the central nervous system with a growing economic and social impact in today's aging population. MCI is one of the risk factors

Table 1
A comparision of different approaches employed in this study.

Model	AUC Scores	Validation Accuracy	Number of Parameters
Sequential 2D CNN	0.84+/-0.038	0.77+/-0.033	3,945,677
Sequential 3D CNN	0.81+/-0.087	0.7+/-0.058	1,006,757
Sequential (2+1)D CNN	0.85+/-0.046	0.78+/-0.032	998,083

Table 2
A comparision with current literature results.

Author	Modalities	Method	Classification	Accuracy
Gupta [3]	MRI	SAE + 2D CNN	AD-MCI	88.10%
Hosseini-Asl [4]	MRI	AE + 3D CNN	AD-MCI	95%
Payan [23]	MRI	SAE + 3D CNN	AD-MCI	86.8%
Sarraf [13]	MRI	2D CNN	AD-CN	98.84%
Sarraf [13]	FMRI	2D CNN	AD-CN	96.86%
Islam [24]	MRI	2D CNN	AD-MCI	93%
Bäckström [2]	MRI	3D CNN	AD-NC	98.74%
Present work	**MRI**	**(2+1)D CNN**	**AD-MCI**	**78%**

for developing dementia later in life, and there has been several studies for identifying MCI for possibly slowing down its progression into dementia. The purpose of this study was to employ different CNN architectures, including a spatio-temporal (2+1)D CNN model, to classify AD vs MCI. Our results indicated that (2+1)D CNN model resulted in a high classification accuracy and auc score with a shorter training time than a 3D CNN approach.

There have been several studies for MCI and AD classification using classical machine learning approaches. Feature extraction methods combined with classical machine learning methods and recent implementations of deep learning methods often revealed lower success rates for AD vs MCI classification problems. Higher performance metrics were usually possible for distinguishing AD patients from HC subjects. On the other hand, the accuracy values were lower for the identification of progressive MCI (pMCI) cases.

4.1. Deep learning based methods

Application of CNNs have resulted in slightly better accuracy values than the classical machine learning methods, not to mention the advantage of minimal preprocessing and hence the collateral benefit of utilizing the inherent features of the raw image information. The use of CNN to classify AD vs HC and AD vs MCI resulted in 99% and 75% success rates, respectively [5]. Additionally, structural MRI combined with positron emission tomography (PET) images evaluated in a cascade 3D plus 2D CNN algorithm indicated a 93.26% success for AD vs. NC and 82.95% for progressive (MCI) vs. NC classification [7]. Moreover, identification of patch and regional relationship in a hierarchical fully convolutional network (H-FCN) classified AD vs NC and pMCI vs sMCI with 90% and 80.9% rates, respectively [7]. On the other hand, principal component analysis (PCA) and Lasso analysis of CNN features predicted MCI to AD conversion with a 79.9% success rate [8]. Additionally, a dual learning based CNN and an ad hoc layer in 3D separable convolutions resulted in 72% accuracy for sMCI vs pMCI classification [9].

While independent treatment of MRI and PET data in a two stack deep polynomial network (SPDN) predicted MCI vs NC with a 87.24% success rate, the same study reported 97.13% accuracy for AD vs NC classification [25]. On the other hand, random forest feature selection combined with deep neural network classified MCI vs. cMCI with %51.2 accuracy, while the fuzzy model performed 78.6% accuracy for AD vs HC classification [10]. On the other hand, a study using cortical geometry incorporated into a spectral graph neural network reported lower classification accuracies of 85.8% for CN vs. AD, 79.2% for early MCI (EMCI) vs. AD, 69.3% for CN vs. late MCI (LMCI), 65.2% for LMCI vs. AD, 60.9% for EMCI vs. LMCI, and 51.8% for CN vs. EMCI [11]. Additionally, temporal dynamics captured by long short term memory (LSTM) units revealed 79.36% and 90.28% classification accuracies for LMCI vs EMCI and AD vs NC groups [12].

The combination of CNN and autoencoders have also been reported in the literature. A 3D CNN convolutional autoencoder was able to diagnose AD vs MCI vs NC, AD vs NC, AD vs MCI, and MCI vs NC with 89.1%, 97.6%, 95% and 90% success rates [4]. Sparse autoencoder combined with a 2D CNN resulted in performance measures of 94.74% for AD vs HC, 86.35% for MCI vs HC, 88.10% for AD vs MCI and 85% for AD vs MCI vs HC [3]. LeNet and GoogleNet applied to sMRI and fMRI [13] resulted in a 98.84% performance for AD versus NC classification.

4.2. (2+1)D CNN method

Conversion of NIFT images into tf.records allowed the handling of 3D MR images as sequences of 2D images all compiled into one single labelled data. Interpretation of 3D images as 2D image frames of a video stream resulted in a significant improvement with classification indices, despite, the fact that the training time needed for the new model increased by a factor of 2. Decomposing images into spatial and temporal dimensions in (2+1)D CNN provided a more plausible method for video understanding yielding 85.8% auc score for classification of AC versus MCI subjects. The training was as rapid as 2D CNN and the result has improved significantly.

4.3. Comparison with literature

Considered minimal preprocessing and a small convolutional neural network architecture, the proposed spatio-temporal CNN model achieved a high accuracy rate. Gupta et al. used sparse autoencoder (SAE) followed by a 2D CNN to classify AD and MCI and had 88.10% test accuracy [3]. Hosseini et al. proposed an adaption of 3D CNN (3D ACNN), where 3D convolutional autoencoder (CAE) was used for pretraining followed by 3D CNN for classification that resulted in 95% accuracy [4]. Payan et al. used 3D sparse autoencoders for pretraining, followed by 3D CNN and had 86.84% test accuracy for AD vs MCI classification [23]. Islam et al. used 2D CNN for AD-MCI classification and had 93% accuracy [24]. In addition, although the previous studies did not mention speed values, compared to the models including 3D CNN, the (2+1)D model proposed here offered a faster solution [24]. With more data and deeper neural networks, the training speed will be more important in future studies. Application of the method to complex deep neural network problems is expected to result in higher validation accuracy and higher auc score, and a faster training period, which could be considered as a future study.

The experimental studies were conducted on a workstation with 32 GB RAM. It has not been possible to fine tune the hyperparameters such as filter sizes and batch size, due to out of memory errors. It could have been possible to obtain much better performance measures with a deeper CNN architecture.

5. Conclusions

In this study, 3D MR images were assessed as a whole sequence of slices by converting them to tf.record followed by the application of 2D, 3D and (2+1)D CNN algorithms. (2+1)D CNN performance was superior in terms of accuracy and auc score, while attaining the training results in a shorter period of time. Specifically, compared to 3D CNN, the training was accomplished two times faster with a slightly better accuracy, due to handling of images in separate spatial and temporal dimensions. The advantage of (2+1)D CNN algorithm is expected to be even more evident for handling larger datasets.

In this experimental work, the proposed amount of layers with a minimum batch size of 32 was attained due to computational limitations. Hence, the speed of (2+1)D algorithm is expected to be even higher in deeper neural network models (Fig. 1).

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgment

Data used in this work were acquired from the Alzheimer's Disease Neuro Imaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI).

References

[1] Alzheimer's Association, Alzheimer's disease facts and figures (2020) https://www.alz.org/search?. searchtext=Alzheimer's+disease+facts+and+figures.
[2] K. Bäckström, M. Nazari, I.Y.H. Gu, A.S. Jakola, An efficient 3D deep convolutional network for Alzheimer's disease diagnosis using MR images, IEEE Int. Symp. Biomed. Imaging (2018) 149–153 ISBI.
[3] A. Gupta, M. Ayhan, A. Maida, Natural image bases to represent neuroimaging data, Int. Conf. Mach. Learn. (2013) 987–994 PMLR.
[4] E. Hosseini-Asl, M. Ghazal, A. Mahmoud, A. Aslantas, A.M. Shalaby, M.F. Casanova, G.N. Barnes, G. Gimel'farb, R. Keynton, A. El-Baz, Alzheimer's disease diagnostics by a D deeply supervised adaptable convolutional network, Front. Biosci. 23 (2018) 584–596.
[5] S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi, Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks, NeuroImage:Clin 21 (2019) 101645.
[6] M. Liu, D. Cheng, K. Wang, Y. Wang, Alzheimer's Disease Neuroimaging, Multimodality cascaded convolutional neural networks for Alzheimer's disease diagnosis, Neuroinformatics 16 (3) (2018) 295–308.
[7] C. Lian, M. Liu, J. Zhang, D. Shen, Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI, IEEE Trans. Pattern Anal. Mach. Intell. 42 (4) (2018) 880–893.
[8] W. Lin, T. Tang, Q. Gao, D. Guo, X. Du, Y. Yang, G. Guo, M. Xiao, M. Du, X. Qu, Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment, Front. Neurosci. 12 (2018) 777.
[9] S. Spasov, L. Passamonti, A. Duggento, P. Liò, N. Toschi, Alzheimer's Disease Neuroimaging, A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease, Neuroimage 189 (2019) 276–287.
[10] N. Amoroso, D. Diacono, A. Fanizzi, M. La Rocca, A. Monaco, A. Lombardi, C. Guaragnella, R. Belloti, S. Tangaro, Deep learning reveals Alzheimer's disease onset in24 MCI subjects: Results from an international challenge, J. Neurosci. Methods 302 (2018) 3–9.
[11] C.Y. Wee, C. Liu, A. Lee, J.S. Poh, H. Ji, A. Qiu, Alzheimers Disease Neuraimage Initative, Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations, NeuroImage:Clin 23 (2019) 101929.
[12] M. Wang, C. Lian, D. Yao, D. Zhang, M. Liu, D. Shen, Spatial-temporal dependency modeling and network hub detection for functional MRI analysis via convolutional recurrent network, IEEE Trans. Biomed. Eng. 67 (8) (2019) 2241–2252.
[13] S. Sarraf, G. Tofighi, Alzheimer's Disease Neuroimaging, DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI, BioRxiv (2016) 070441.
[14] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, in: Proc. IEEE, 86, 1998, pp. 2278–2324.
[15] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit, 2015, pp. 1–9.
[16] J. Wen, E. Thibeau-Sutre, M. Diaz-Melo, J. Samper-González, A. Routier, S. Bottani, D. Dormont, S. Durrleman, N. Burgos, O. Colliot, Alzheimer's Disease Neuroimaging Initiative, Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation, Med. Image Analysis 63 (2020) 101694.
[17] N.J. Tustison, B.B. Avants, P.A. Cook, Y. Zheng, A. Egan, P.A. Yushkevich, J.C. Gee, N4ITK: improved N3 bias correction, IEEE Trans. Med. Imaging 29 (6) (2010) 1310–1320.
[18] B.B. Avants, C.L. Epstein, M. Grossman, J.C. Gee, Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain, Med. Image Anal. 12 (1) (2008) 26–41.
[19] V.S. Fonov, A.C. Evans, R.C. McKinstry, C.R. Almli, D.L. Collins, Unbiased nonlinear average age-appropriate brain templates from birth to adulthood, NeuroImage (47) (2009) S102.
[20] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inform. Process. Syst. 25 (2012) 1097–1105.
[21] F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally, K. Keutzer, SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size, arXiv preprint (2016) arXiv:1602.07360.
[22] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun, M. Paluri, A closer look at spatiotemporal convolutions for action recognition, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit, 2018, pp. 6450–6459.
[23] A. Payan, G. Montana, Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks, arXiv preprint (2015) arXiv:1502.02506.
[24] J. Islam, Y. Zhang, Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks, Brain Inform. 5 (2) (2018) 1–14.
[25] J. Shi, X. Zheng, Y. Li, Q. Zhang, S. Ying, Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease, IEEE J. Biomed. Health Inform. 22 (1) (2017) 173–183.