8th International Conference on
Combinatorics, Cryptography, Computer Science and Computation
November. 15-16, 2023

Comparison of pre-training and classification models for early detection of Alzheimer's disease using magnetic resonance imaging

Amir Hossein Karami[1,*]
Sepehr Rezaee[1]
Elmira Mirzabeigi[2]
Kourosh Parand[1]

[1]Departemant of Computer Science and Data, Faculty of Mathematical Sciences,Shahid Beheshti University, Tehran, Iran

[2]Department of Applied Mathematics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran, Iran

Abstract

Alzheimer's disease (AD) is a devastating neurodegenerative disorder; early detection is paramount for effective intervention. This study conducts a comparative analysis of deep learning-based pre-training and classification models to enhance the early detection of AD using a comprehensive MRI dataset. Leveraging the Alzheimer's Disease Neuroimaging Initiative (ADNI) data, the research explores the potential of pre-trained convolutional neural networks (CNNs) and autoencoders for feature extraction from MRI scans, aiming to distinguish between individuals at different stages of AD progression. The study underscores the significance of accurate stage classification and showcases the utility of advanced deep-learning techniques for early AD diagnosis. Based on our findings, the effectiveness of pre-trained models in classification algorithms is evident, and this has significant implications for enhancing AD diagnosis and intervention strategies. This research contributes to the growing body of knowledge in the field, advancing the pursuit of early AD detection and ultimately improving patient care.

Keywords: Deep learning,Early detection, Alzheimer's disease, stage classification, MRI dataset

AMS Mathematical Subject Classification [2010]: 92C99, 68T20, 68T05, 92B20

1 Introduction

The history of Alzheimer's disease (AD) dates back to the early 20th century, when Alois Alzheimer, a German psychiatrist and neurologist, identified a disease characterized by severe cognitive and memory deficits. A brain post-mortem examination of the Alzheimer's patient revealed abnormal protein deposits and abnormal cell structures, which were hallmarks of the disease. During the following decades, researchers and clinicians delved deeper into the complexities of this disease and gradually revealed its pathophysiology and clinical manifestations. On the other hand, the development of neuroimaging techniques, especially magnetic resonance imaging (MRI), was a significant turning point in AD research and opened new avenues

*Speaker

for early diagnosis and monitoring of the disease. With the advancement of technology, MRI has become an essential tool in the diagnosis and study of AD. This enabled the researchers to examine the subtle changes in brain structure and connections associated with the disease and foster a better understanding of the disease's progression. In recent years, the convergence of artificial intelligence and machine learning has introduced new possibilities for the early detection of AD, as algorithms can analyze large amounts of MRI data to identify subtle patterns that may escape human observers.

Alzheimer's disease is a progressive neurodegenerative condition that manifests in various stages, reflecting the continuum of cognitive and functional decline. The continuum typically starts with the cognitively normal (CN) stage, where individuals exhibit no significant memory or cognitive impairment. Early mild cognitive impairment (EMCI) marks the next stage, characterized by subtle cognitive changes often detectable through assessments but still with relatively preserved daily functioning. This stage may progress to mild cognitive impairment (MCI), where cognitive deficits become more apparent, particularly in memory and problem-solving, often signaling a precursor to AD. Later mild cognitive impairment (LMCI) denotes a stage in which these deficits become more pronounced, and individuals may experience challenges with daily activities. Ultimately, AD signifies the severe and advanced stage, featuring profound cognitive and functional impairments, including memory loss, language difficulties, and an inability to carry out routine tasks. Understanding and diagnosing AD across these stages is critical for providing appropriate care and support tailored to the individual's specific needs at each phase of the disease.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) database stands as a pivotal resource in the quest to comprehend neurodegenerative disorders. Established to facilitate research on AD and its precursors, ADNI collates a wealth of clinical, imaging, genetic, and biochemical data from participants across various sites. In essence, the ADNI database, enriched with MRI data, serves as a dynamic canvas for researchers, painting a comprehensive picture of the structural intricacies altered by AD. It is within this wealth of information that the potential for deep learning methods to revolutionize our understanding and diagnostic capabilities lies.

Artificial Intelligence (AI) and deep learning methods have emerged as powerful tools in the early detection of AD. Their ability to analyze vast amounts of medical data, including MRI and PET scans, as well as genetic and cognitive information, has revolutionized the field of AD's research. Deep learning models, such as convolutional neural networks and recurrent neural networks, have shown remarkable capabilities in identifying subtle patterns and anomalies in medical imaging data. They enable the extraction of complex and hidden features, facilitating the detection of structural and functional brain changes associated with AD. Moreover, AI-based algorithms can predict the likelihood of disease progression, aiding clinicians in making more informed decisions about patient care and treatment strategies. With the potential to enhance the accuracy and efficiency of Alzheimer's diagnosis, AI and deep learning methods represent a promising frontier in the ongoing battle against this devastating disease.

In this study, we utilized pre-training and classification models to improve the diagnosis of all six stages of AD. We aim to improve early diagnosis of the disease by comparing models using MRI data from the ADNI database. By carefully examining the efficiency and performance of various pre-training models and classification algorithms, this research seeks to increase the accuracy and efficiency of AD diagnosis, a vital factor in facilitating timely intervention and improving patient care. The aim of this study is to utilize advanced machine learning techniques to create a powerful diagnostic tool from cognitive-medical data.

2 Materials and Methods

2.1 Dataset description

The data is taken from the ADNI database, which is made public on the website https://adni.loni.usc.edu/. The Food and Drug Regulatory Agencies (FDA), the National Institute on Aging (NIA), and the National Institute of Biomedical Imaging and Bioengineering (NIBIB) introduced ADNI in 2003. The main goal of ADNI is to check the sequence of MRI, positron emission tomography (PET), other biomarkers, and clinical and neuropsychological assessments that may be combined to measure the progression of MCI and early AD. Subjects were recruited from more than 50 locations across the United States and Canada, providing written information agreed upon at the time of registration for image and DNA sampling, and completed questionnaires were approved by the Institutional Review Board (IRB) of each participating site. Table 1Summary of participant demographics shows the number of datasets and stages used for the implementation. In this table A total of 2764 subjects, including 527 AD, 235 LMCI, 754 MCI, 477 EMCI, 165 SMC, and 1050 CN, were used in this work. In this research, we specifically focused on using T1-weighted images from the ADNI dataset. Furthermore, the respective numbers of cases in each category used for testing and training and examples of MRI data for the six groups are shown in Fig. 1MR image examples for different Dementia stages: (a) cognitively normal (CN), (b) Subjective Memory Complaints (SMC), (c) Early Mild Cognitive Impairment (EMCI), (d) Mild Cognitive Impairment (MCI) and (e) Lately Mild Cognitive Impairment (LMCI), and (f) Alzheimer disease (AD), from left to right, respectively..

	Subjects	Gender
CN	1050	545 F/ 505 M
SMC	165	86 F/ 79 M
EMCI	477	248 F/ 229 M
MCI	754	391 F/ 363 M
LMCI	235	122 F/ 113 M
AD	527	274 F/ 253 M
M: male, F: female		

Table 1: Summary of participant demographics

2.2 Data preprocessing

MRI images have deteriorated during data collection due to low variation owing to the optical equipment's inadequate brightness. Image improvement techniques are usually utilized to correct or enhance the distribution of pixels over a large variety of intensities to solve this problem for the enhancement of MRI scans. Thus, we initially employed image normalization to adjust the image pixel intensity values by reducing machinery and impulse noise. For image normalization, the pixel values are rescaled to [-1,1] using a pixel-wise multiplication factor of 0/255 as follows:

$$\widehat{IN} = (1 - \widehat{I}^T_{Min}) + \frac{\widehat{I}^T_{Max} - \widehat{I}^T_{Min}}{Max - Min}\widehat{I}^T_{Min} \tag{1}$$

where I and $\widehat{IN}$ represent the input and normalized brain image, respectively, $\widehat{I}^T_{Max}$, $\widehat{I}^T_{Min}$ are the normalized image's intensity range, and $Min = 0$; and $Max = 255$ represent the input brain image's pixel intensity range. For a reliable DL model, more training data should be used to avoid model overfitting and provide

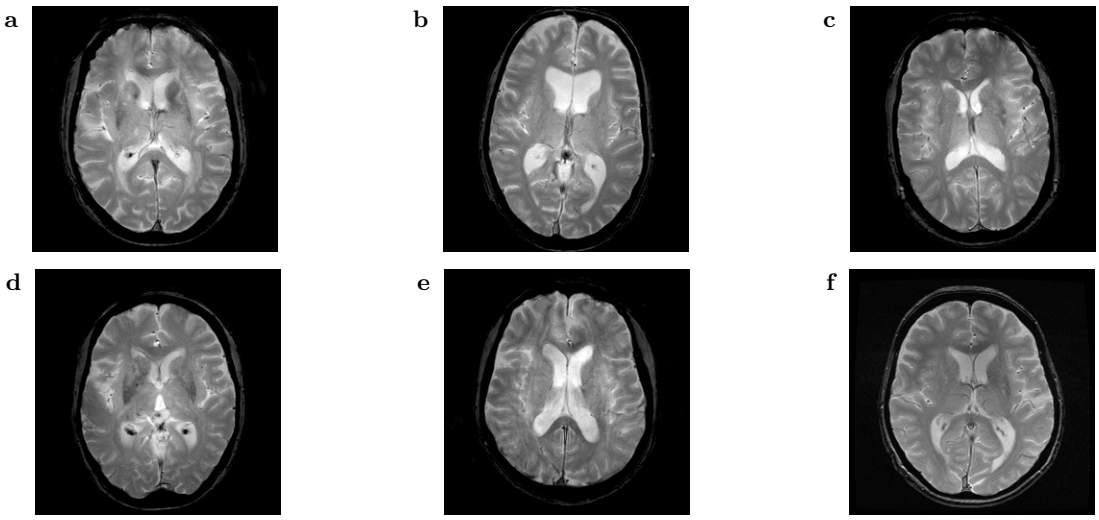

Figure 1: MR image examples for different Dementia stages: (a) cognitively normal (CN), (b) Subjective Memory Complaints (SMC), (c) Early Mild Cognitive Impairment (EMCI), (d) Mild Cognitive Impairment (MCI) and (e) Lately Mild Cognitive Impairment (LMCI), and (f) Alzheimer disease (AD), from left to right, respectively.

a more generalized model. We used unbalanced data to design software for AD diagnosis and evaluated different models to strengthen it. The image augmentation approach was not used because this would cause bias despite balancing the data.

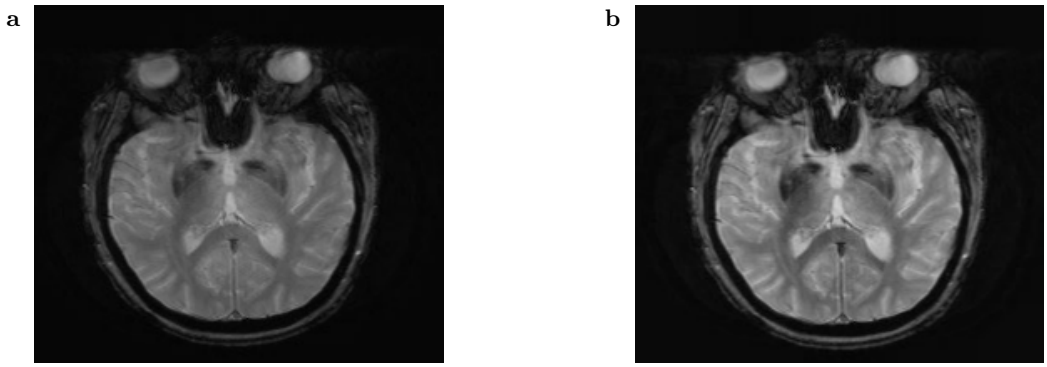

Figure 2: Example of an MR image preprocessing: (a) before and (b) after preprocessing.

2.3 Model selection

The suggested framework's analytical flowchart is shown in Fig. 3Flowchart representation for the proposed analysis framework. and consists of multiple processing stages. The initial phase involved preprocessing, followed by the creation of a convolutional neural network (CNN) architecture, encompassing transfer learning, model training, parameter configuration, and ultimately culminating in classification. Next, the details of those stages are fully described.

2.4 Training the model

We utilized a total of four distinct models namely VGG-16, Inception-V3, EfficientNet-B4, and ResNet-101. This section outlines the rigorous process we undertook in training these models. The VGG-16 model was trained using a two-step process. Initially, the model was pre-trained on a large dataset to learn generic

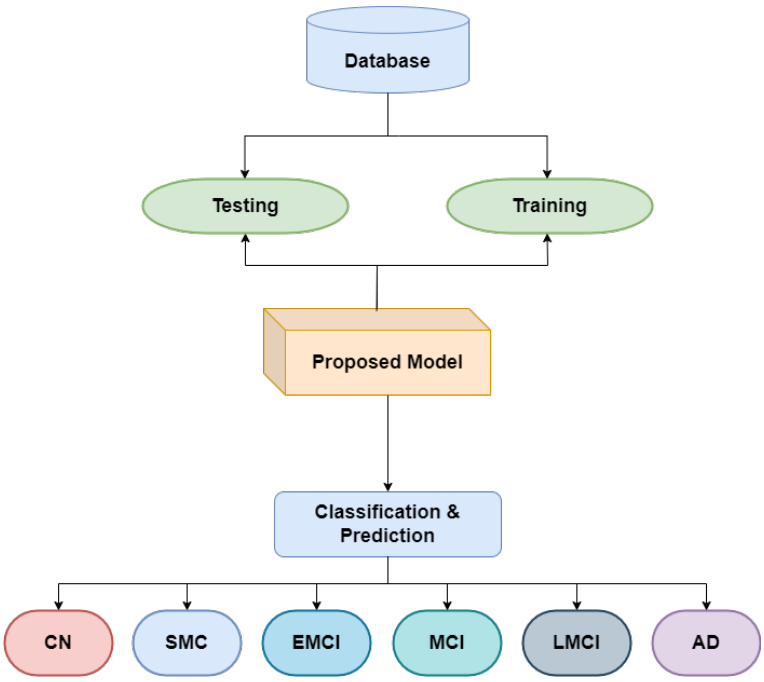

Figure 3: Flowchart representation for the proposed analysis framework.

features. Subsequently, the model was fine-tuned on our specific task of early detection of Alzheimer's disease using MRI data. The learning rate was set to 0.001 and reduced by a factor of 10 every 30 epochs. The Inception-V3 model was trained in a similar manner to the VGG-16 model. However, due to its complex architecture, we used a smaller learning rate of 0.0001. The learning rate was reduced by half every 20 epochs. EfficientNet-B4 was trained using a compound scaling method that uniformly scales all dimensions of depth, width, and resolution. The initial learning rate was set to 0.01 and reduced by a factor of 10 every 15 epochs. ResNet-101 was trained using a unique approach that leverages skip connections or shortcuts to jump over some layers, which helps solve the vanishing gradient problem in deep neural networks. The initial learning rate was set to 0.1 and reduced by a factor of 10 every 30 epochs.

In all models, we used the Adam optimizer for its adaptive learning rate capabilities. We also employed early stopping with a patience of 10 epochs to prevent overfitting. The models were trained until convergence, and the model weights that achieved the highest validation accuracy were saved for evaluation.

2.5 Evaluation metrics

In this section, we describe the metrics that we used to evaluate the performance of the pre-training and classification models for early detection of Alzheimer's disease using magnetic resonance imaging. We used four models: VGG, InceptionV3, EfficientNet, and ResNet. We compared them on two tasks: binary classification (AD vs. NC) and multi-class classification (AD vs. MCI vs. NC).

2.5.1 Multi-Class Classification

For the multi-class classification task, we used the following metrics:

- Accuracy: This is the proportion of correctly classified samples out of the total number of samples. It is calculated as:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

- Precision: This is the proportion of samples that are correctly classified as a certain class out of the total number of samples that are predicted as that class. It is also known as the positive predictive value. It is calculated as:

$$\text{Precision}_i = \frac{\text{True Positives}_i}{\text{True Positives}_i + \text{False Positives}_i}$$

where i denotes the class label.

- Recall: This is the proportion of samples that are correctly classified as a certain class out of the total number of samples that belong to that class. It is also known as the sensitivity or true positive rate. It is calculated as:

$$\text{Recall}_i = \frac{\text{True Positives}_i}{\text{True Positives}_i + \text{False Negatives}_i}$$

where i denotes the class label.

- F1-score: This is a harmonic mean of precision and recall, which balances both metrics. It ranges from 0 to 1, with 0 being the worst and 1 being the best. It is calculated as:

$$\text{F1-score}_i = \frac{2 \times \text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}$$

where i denotes the class label.

- Macro-average: This is an average of a metric across all classes, giving equal weight to each class. It is calculated as:

$$\text{Macro-average} = \frac{\sum_{i=1}^{n} \text{Metric}_i}{n}$$

where n denotes the number of classes and Metric denotes any of the above metrics.

- Weighted-average: This is an average of a metric across all classes, giving weight to each class according to its frequency in the data. It is calculated as:

$$\text{Weighted-average} = \frac{\sum_{i=1}^{n} w_i \times \text{Metric}_i}{\sum_{i=1}^{n} w_i}$$

where w_i denotes the weight or frequency of class i and Metric denotes any of the above metrics.

3 Result and Discussion

In this section, we present and discuss the results of our experiments on applying four different methods to classify six stages of Alzheimer's disease using MRI data. The methods are: VGG-16, ResNet-101, InceptionNet-V3, and EfficientNet models fine-tuned on the MRI dataset.

We used the ADNI database as our source of MRI data, which contains 20230 images from 2754 patients at various stages of the disease. We split the data into 80% training, 10% validation, and 10% testing sets. We trained each method for 100 epochs using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. We evaluated each method on the test set using accuracy as the performance metric.

According to Table 2Proposed Model evaluation against other CNN that employs transfer learning., the EfficientNet model achieved the highest accuracy among all the methods, with a remarkable score of 98.90. The comparative analysis of our methods with those implemented in other articles is presented in Table 3Comparative accuracy of the proposed pipeline with other state-of-the-art literature techniques.. This indicates that the EfficientNet model can effectively extract and learn the features that are relevant for distinguishing the different stages of Alzheimer's disease from MRI data. The EfficientNet model also has the advantage of being more efficient and scalable than other models, as it uses a compound scaling method that optimally balances network depth, width, and resolution. Figure 4The best CNN model architecture used for AD diagnosis. shows the architecture of this model. The VGG-16 model had the lowest accuracy, which may be due to its large number of parameters and its tendency to overfit complex data. The ResNet-101 model improved upon the VGG-16 model by using residual connections that enable deeper networks and mitigate the vanishing gradient problem. The InceptionNet-V3 model further improved upon the ResNet-101 model by using inception modules that allow for multiple filter sizes and reduce computational costs. We

Method	Train Accuracy	Test Accuracy	Train f1-score*	Test f1-score
VGG16	85.73	83.98	84.5	82.78
ResNet101	88.66	88.08	87.43	86.85
EfficientNet	99.54	98.90	98.31	97.67
InceptionV3	97.47	96.54	96.24	95.31
*Macro-average f1-score				

Table 2: Proposed Model evaluation against other CNN that employs transfer learning.

can conclude that pre-training and fine-tuning deep learning models on MRI data is a promising approach for the early detection of Alzheimer's disease. Among the four methods we tested, the EfficientNet model showed superior performance and efficiency in classifying the six stages of Alzheimer's disease.

Methods	Labels	Data Source	ACC	SEN	SPE	AUC
Zhang et al.[1]	MCI vs. AD vs. NC	ADNI	88.58%	-	-	-
Chiyu et al. [2]	AD vs. NC vs. s-MCI	ADNI	94.82%	-	-	-
Liu et al. [3]	MCI vs. AD vs.NC	ADNI	88.9%	-	90.8%	92.5%
Zhenyu et al. [4]	MCI vs. AD vs.NC	ADNI	85.7%	100%	93%	-
Sarraf et al. [5]	AD vs. NC	ADNI	96%	97.39%	84.27%	-
E. Jabason et al. [6]	AD vs. NC vs. MCI	OASIS	95%	-	-	-
"ACC", "SEN", "SPE", "AUC", stand for Accuracy, Sensitivity, Specificity, and Area Under the Curve, respectively.						

Table 3: Comparative accuracy of the proposed pipeline with other state-of-the-art literature techniques.

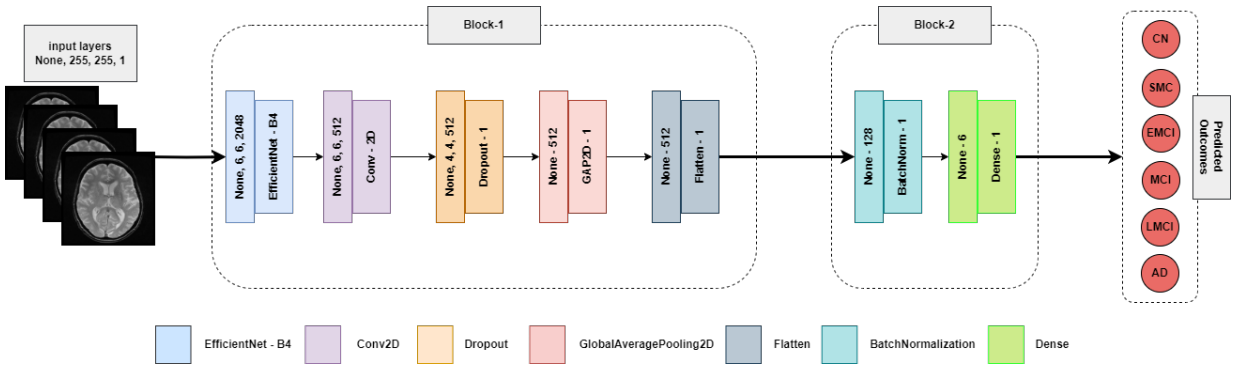

Figure 4: The best CNN model architecture used for AD diagnosis.

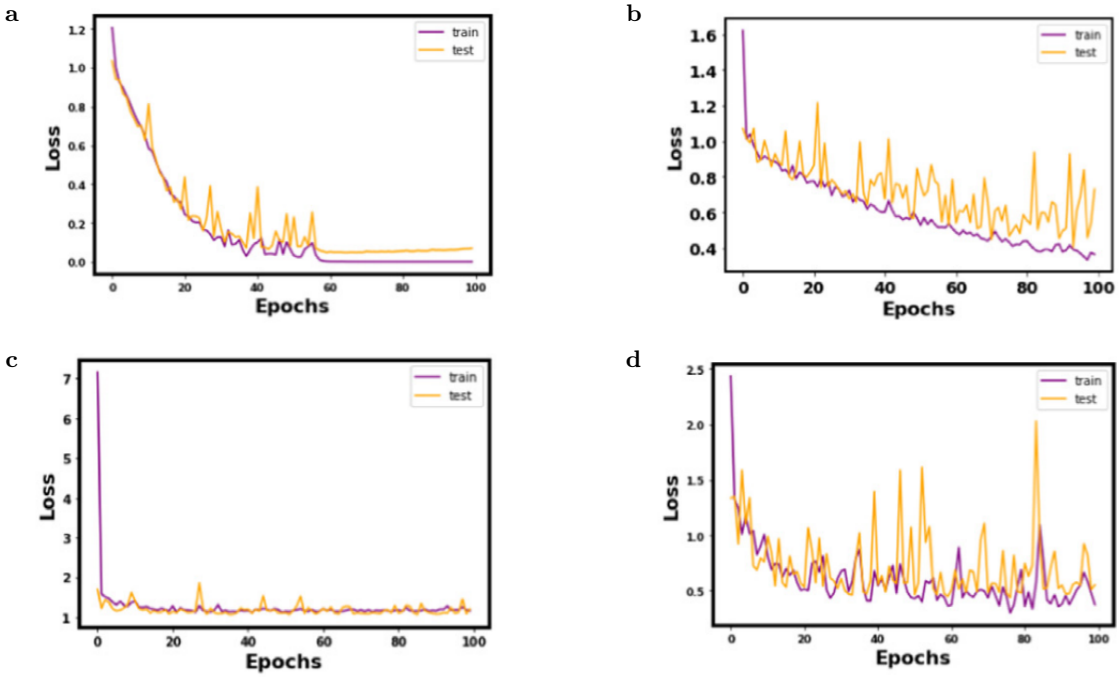

Figure 5: Training and validation loss against the number of epochs for (a) VGG18, (b) ResNet101 (c) EfficientNet, and (d) InceptionV3 models.

References

[1] F. Zhang, Z. Li, B. Zhang, H. Du, B. Wang, X. Zhang, Multi-modal deep learning model for auxiliary diagnosis of alzheimer's disease, Neurocomputing 361 (2019) 185–195.

[2] F. CHIYU, A. ELAZAB, Y. PENG, T. WANG, F. ZHOU, Deep learning framework for alzheimer's disease diagnosis via 3d-cnn and fsbi-lstm.

[3] M. Liu, F. Li, H. Yan, K. Wang, Y. Ma, L. Shen, M. Xu, A.D.N. Initiative, et al, A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in alzheimer's disease, NeuroImage 208 (2020) 116459.

[4] Z. Cui, Z. Gao, J. Leng, T. Zhang, P. Quan, W. Zhao, Alzheimer's disease diagnosis using enhanced inception network based on brain magnetic resonance image, in: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2019, pp. 2324–2330.

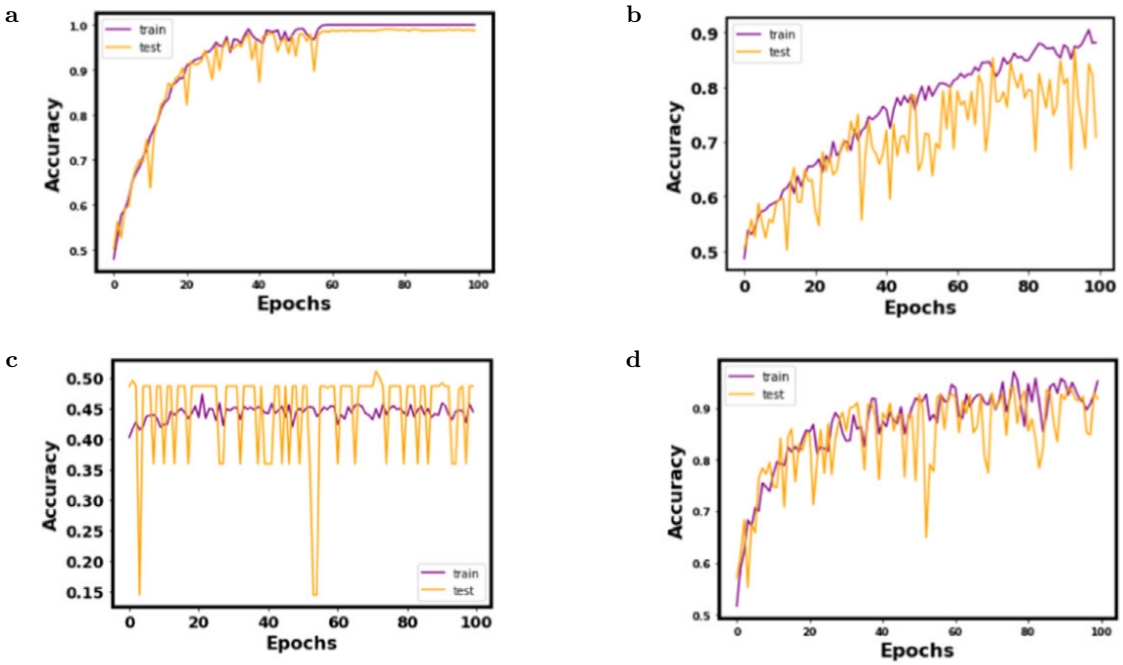

Figure 6: Training and validation accuracy against the number of epochs for (a) VGG18, (b) ResNet101 (c) EfficientNet, and (d) InceptionV3 models.

[5] S. Sarraf, G. Tofighi, Classification of alzheimer's disease using fmri data and deep learning convolutional neural networks, arXiv preprint arXiv:1603.08631.

[6] E. Jabason, M.O. Ahmad, M. Swamy, Classification of alzheimer's disease from mri data using an ensemble of hybrid deep convolutional neural networks, 2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS), IEEE, 2019, pp. 481–484.

e-mail: k.parand1394@gmail.com
e-mail: elmiramirzabeigi7@gmail.com