Penerbit Akademia Baru

Journal of Advanced Research in Computing and Applications

Journal homepage: www.akademiabaru.com/arca.html
ISSN: 2462-1927

Mri Brain Image Classification Using Convolutional Neural Networks and Transfer Learning

Khaw Li Wen [1,*], Shahrum Shah Abdullah[1]

[1] Department of iKohza BIO-IST, Department of Electronic System Engineering, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 5 March 2023
Received in revised form 14 April 2023
Accepted 16 May 2023
Available online 10 June 2023

ABSTRACT

Alzheimer's disease (AD) is a neurodegenerative disorder. There is no particular cure for Alzheimer's disease. Accurate and early diagnosis of AD could assist patients in receiving appropriate care. However, diagnosing AD in brain MRI images is a difficult task that depends on the presence of experienced radiologists or medical professionals. As one MRI exam might generate thousands of images, it typically takes several weeks for the results to be obtained. Many researchers use statistical and machine learning methods to diagnose Alzheimer's disease. Deep Learning algorithms have demonstrated human-level competence in a variety of fields. Deep learning, especially convolutional neural networks (CNN), is becoming popular because of its state-of-the-art performance in many computer vision tasks such as visual object classification, object detection, and segmentation. Transfer learning is a technique that can be used with CNNs to improve their performance. The purpose of this project is to develop a model of brain MRI image classification for Alzheimer's disease diagnosis by using CNN and transfer learning. In this project, the modified VGG16 model with fine-tuning was proposed, and MRI data from the OASIS database was used to classify Alzheimer's disease into three (3) different classes, which are AD, MCI and NC. The model is developed using Google Collaboratory and Adam's optimization algorithm. The proposed model has achieved a training accuracy of 98.56% and the validation accuracy of 90.24%.

Keywords:
Alzheimer's disease; Convolutional Neural Networks; Transfer Learning; MRI

1. Introduction

Alzheimer's disease (AD) is a brain disorder that worsens a person's daily functioning. There is no specific treatment for AD. Accurate and early diagnosis of Alzheimer's can assist patients in receiving appropriate care and taking precautions to control AD progression. The number of Alzheimer's disease patients worldwide is predicted to rise from 47 million to 152 million by 2050, having significant economic, medical, and societal effects. [1] Alzheimer's disease diagnosis progress can be sped up and helped by implementing deep learning techniques on magnetic resonance imaging [2].

[*] *Corresponding author.*
E-mail address: khawliwen@graduate.utm.my (Khaw Li Wen)

Magnetic Resonance Imaging (MRI) image was the most suitable method for diagnosing AD [3]. However, detecting AD in brain MRI images is a challenging task that relies on the availability of expert radiologists or doctors. It usually takes several weeks for the results of an MRI to come through as one exam of MRI scan can produce thousands of images [4]. Therefore, using artificial intelligence to automate illness diagnosis is becoming increasingly important. Deep learning, especially convolutional neural networks (CNN), is the most commonly used and practical approach for disease detection and classification with a huge quantity of input data. Therefore, a CNN-based system of Alzheimer's disease detection using magnetic resonance imaging (MRI) scans images is useful to identify the AD more quickly and accurately.

A typical CNN architecture consists of multiple layers like the input, convolutional, pooling, and fully connected layers. The convolutional layer applies filters to the input image to extract features, the pooling layer down samples the image to reduce computation, and the fully connected layer makes the final prediction.

Transfer learning is a technique that can be used with CNNs to improve their performance. Transfer learning refers to using a previously trained model on a different problem. Transfer learning can be used in various ways, such as feature extraction transfer learning and fine-tuning transfer learning. Feature extraction transfer learning is when you take the weights from a pre-trained model and adjust its output layers to be more suited to the new problem. Fine-tuning transfer learning is fine-tuning a pre-trained model by changing the model weights to fit the new problem better.

2. Methodology

For this project, the platform used to train the CNN model is Google Collaboratory. Besides, the deep learning framework used is Keras with TensorFlow as a backend. The sparse categorical cross-entropy loss function is used to train the model. The Adam optimizer is the optimization algorithm used for updating parameters, with a learning rate set to 0.0001. The model was fitted on training data in a batch size of 16 in 50 epochs and an additional 20 epochs in a fine-tuned model. For the training model of this project, the source of the dataset used to do training is taken from the Open Access Series of Imaging Studies (OASIS). The dataset consists of MRI images of Alzheimer's disease in 3 categories: AD, MCI and NC. The dataset consists of 2768 images and is split into 2215 images for training and 553 for validation with a split ratio of 80:20. In this project, the data augmentation transformations were applied, including randomly flipping images horizontally and randomly zooming into an image. After the input layers, the data augmentation was added as a layer in the transfer learning model.

Few years later, a new research group is formed to further investigate the flow structure on the blunt-edged delta wing, the team called as Vortex Flow Experiment (VFE-2). The main objective of the VFE-2 test was to validate the results of Navier-Stokes calculations and to obtain a more detailed experimental data. The VFE-2 experiments were carried out for both sharp and blunt leading edge shape delta wing [1-3].

In model architecture selection, experiments were done on different pre-trained CNN models (VGG16, InceptionV3 and MobileNetV2) to get the best performance base model. Besides, the modification was done on the selected VGG16 pre-trained base model by adding the fully connected layers and dropout layers to the base model to improve the base model's performance. Therefore, some experiments were done by varying the number of fully connected layers (1,2 and 3) and varying the ratio of dropout layers (0.2, 0.3 and 0.5) added after the VGG16 pre-trained base model to choose the optimal values for the number of fully connected layers and the ratio of dropout layers.

Among the conducted experiments, the VGG16 pre-trained model, originally trained on the ImageNet dataset, serves as the foundational model for processing brain MRI images to extract relevant features. To adapt this base model for Alzheimer's disease classification, we pruned its fully connected layers, initially designed for 1000 classes in ImageNet. Subsequently, we introduced new trainable layers appended to the base model. The sequence includes a global average pooling 2D layer, followed by a dense layer housing 256 neurons, a dropout layer with a dropout ratio of 0.5, and a softmax output layer that classifies Alzheimer's disease into three distinct categories. After training the new modified model on the new dataset, fine-tune the model by unfreezing the last convolution layers (Conv5) of the VGG16 pre-trained and further train the model on the dataset with a lower learning rate to improve the performance. Figure 1 shows the fine-tuned VGG16 model architecture.

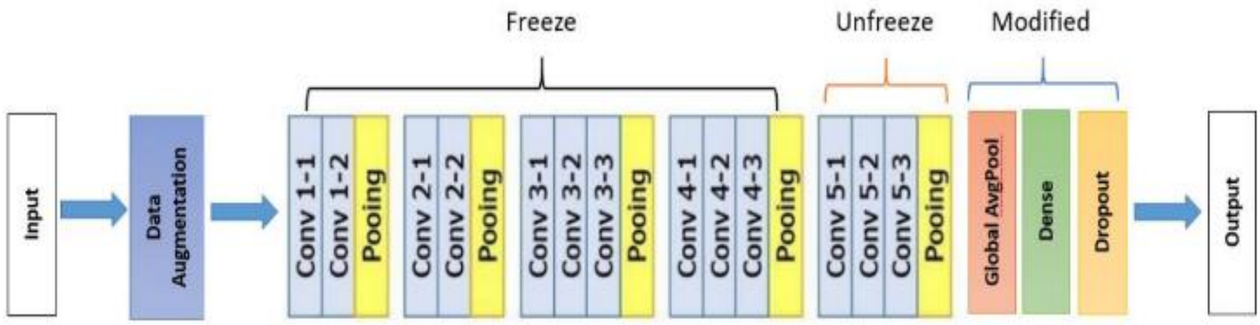


Fig. 1. Fine tune VGG16 model architecture

The training loss and accuracy measures are used to evaluate the model's performance. The training loss and accuracy on the training and validation dataset were plotted using the matplotlib library in Python to visualize the training history. In addition, confusion matrices were generated to evaluate the classification performance on the validation dataset. Accuracy has been used as the primary evaluation metric.

Accuracy calculates the percentage of predicted values that match with actual values shown in Eq. (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

True positives (TP) and true negatives (TN) are the results produced when the model correctly classifies the positive and negative classes. In contrast, false positives (FP) and false negatives (FN) are the outcomes obtained when the model wrongly classifies the positive and negative classes.

The loss function is a function that evaluates the difference between the neural network's predictions and the actual values of the observations used during learning. The neural network performs better as the result of this function is minimized.

Confusion matrices were computed to evaluate the classification performance on the validation dataset. A confusion matrix is a table used to assess a classification model's performance, as shown in Figure 2. It summarizes the number of correct and incorrect predictions by comparing predicted and true labels.

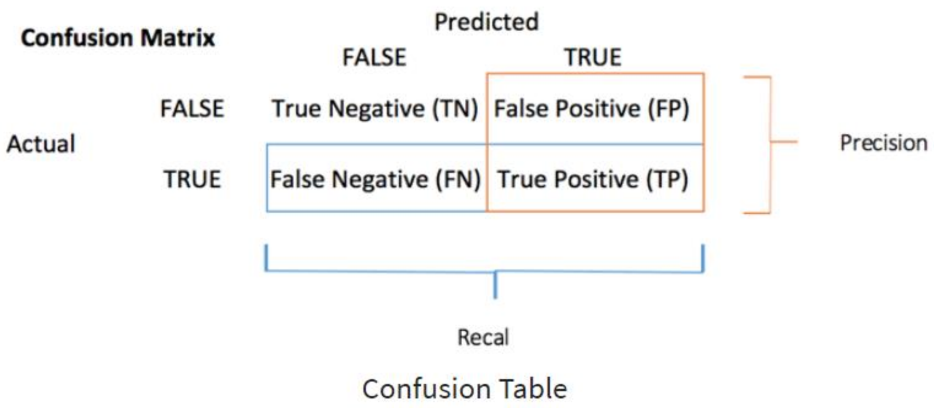
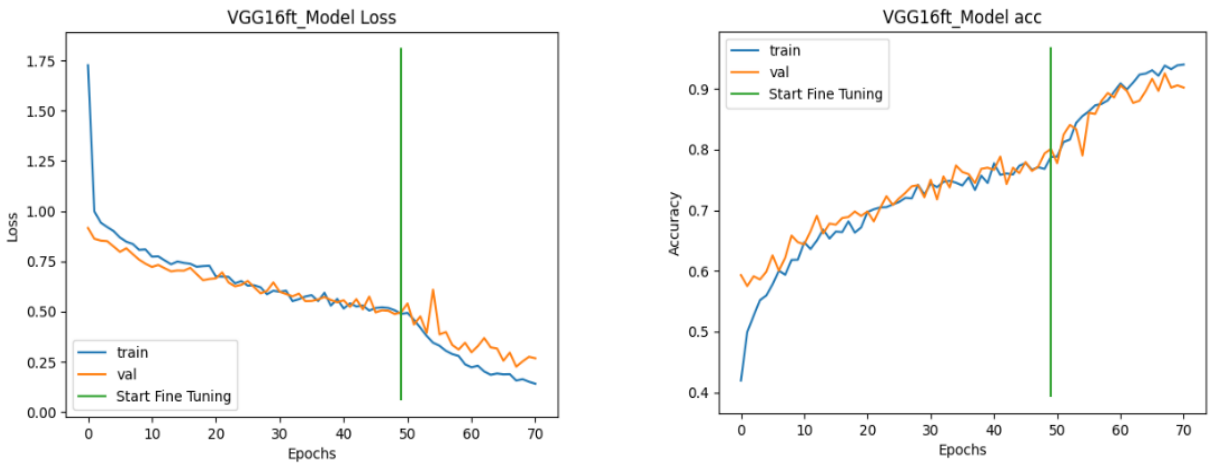


Fig. 2. Confusion matrix

3. Results

From all the experiment results, the best-modified model was the VGG16 model with an additional one (1) fully connected layer and one (1) dropout layer of 0.5 ratio. After modifying the base VGG16 model, the training accuracy of the model reached 89.80% on the training dataset and 80.11% on the validation dataset. After fine-tuning the model and continuing to train the model for 20 epochs, the model reached 98.56% accuracy on the training dataset and 90.24% on the validation set. Figure 3 shows the loss function and accuracy plots of the fine-tuned VGG16 model. Table 1 compares model loss and accuracy for modified VGG16 and Fine-tune VGG16 models.



(a) (b)

Fig. 3. Loss function plot (a) and accuracy plots (b) for the fine-tuned VGG16 model

Table 1
Comparison of model loss and model accuracy for the modified VGG16 model and the fine-tuned VGG16 model

MODEL	MODEL LOSS		MODEL ACCURACY	
	Train	Validate	Train	Validate
Modified VGG16	0.2828	0.4972	0.8980	0.8011
Fine tune VGG16	0.0453	0.2679	0.9856	0.9024

For the results in the confusion matrix, the modified VGG16 model with fine-tuning shows the best results, evident from the stronger diagonal and lighter cells in the confusion matrix shown in Figure 4. However, it can also be observed from the confusion matrix that the model is most commonly misclassified in the MCI class.

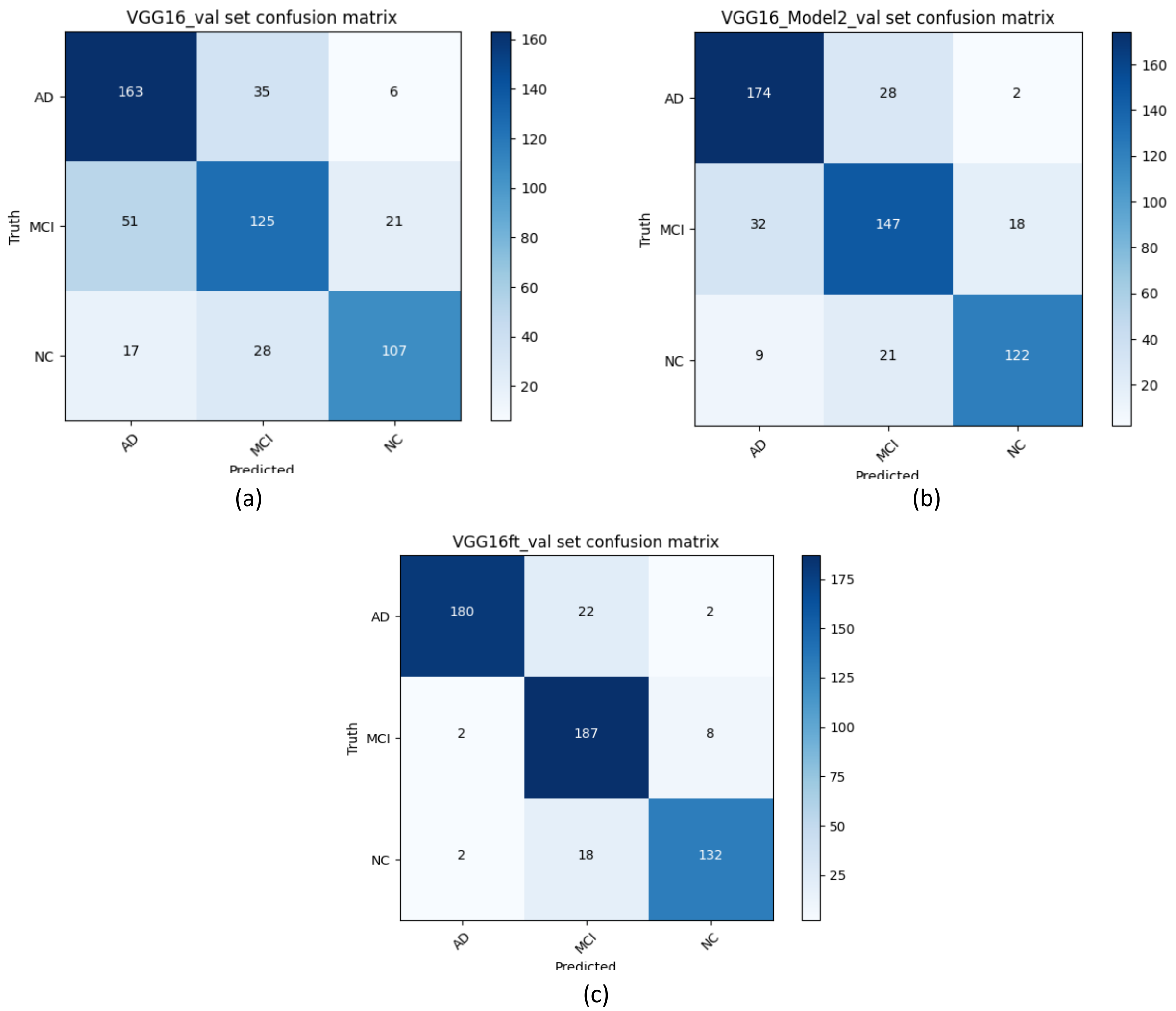


Fig. 4. Confusion matrix for VGG16 (a), Modified VGG16 (b) and Fine-tune VGG16 model (c)

Figure 5 shows the results of predicted output images of the proposed model, which is the modified VGG16 fine-tuning model. To ensure the model is validated reasonably, all images picked for testing are not repeated from the training datasets. The proposed model shows satisfying classification results on the testing image as it can correctly classify it into three (3) different classes**.**

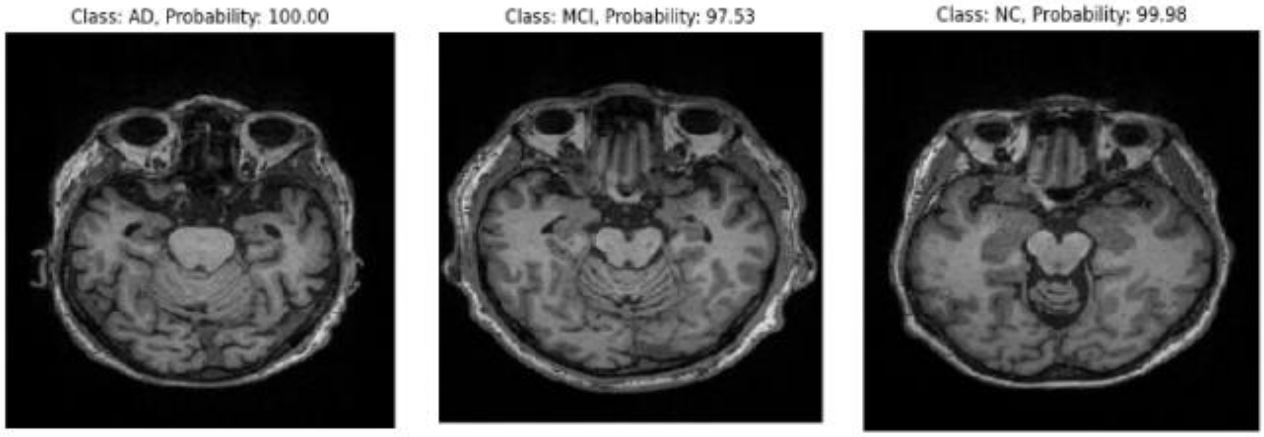


Fig. 5. The results of predicted output images

4. Conclusions

Deep learning algorithms applied to MRI data have significantly accelerated Alzheimer's disease diagnosis, aiding clinicians. Researchers have developed systems utilizing various machine learning techniques, particularly Convolutional Neural Networks (CNNs), to improve the classification accuracy of medical images. Effective image classification is crucial for clinical care. Our VGG16 Fine Tuning model achieved an outstanding 98.56% accuracy in classifying MRI images into AD, MCI, and NC, showcasing the potential of deep learning for swift and precise AD detection.

This project showed that it is challenging to categorize MCI classes in comparison to AD or NC. Therefore, further research is required by using a multi-modal approach of including PET scans along with MRI scans and implementing more recent state-of-the-art networks such as ResNet and EfficientNet to achieve better results. Moreover, some visualization tools such as Layerwise-Relevance Propagation (LRP) and Gradient-weighted Class Activation Mapping (Grad-CAM) can be added to visualize and analyze the regions on the MRI images that the trained CNN model utilized to identify that class. In the future, this model could be developed to include a web-based interface or application to create real-time classification results for the users.

Acknowledgement
This study was supported by the Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM).

References

[1] Pan, Dan, An Zeng, Longfei Jia, Yin Huang, Tory Frizzell, and Xiaowei Song. "Early detection of Alzheimer's disease using magnetic resonance imaging: a novel approach combining convolutional neural networks and ensemble learning." *Frontiers in neuroscience* 14 (2020): 259. https://doi.org/10.3389/fnins.2020.00259

[2] Islam, Jyoti, and Yanqing Zhang. "Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks." *Brain informatics* 5 (2018): 1-14. https://doi.org/10.1186/s40708-018-0080-3

[3] Mayo Clinic. Alzheimer's disease - Diagnosis and treatment - Mayo Clinic. [online] (2018).

[4] Medline. medlineplus.gov. (n.d.). MRI: MedlinePlus Medical Encyclopedia. [online] Available at: https://medlineplus.gov/ency/article/003335.htm