International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 24 August 2023 **Revised:** 12 October 2023 **Accepted:** 26 October 2023*

Classification of Alzheimer's and Parkinson's Disease Based on VGG19 Features with Batch Normalization

Suganya A[1] and Aarthy S L[2,*]

[1] Research scholar, School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India. email: suganya.a2019@vitstudent.ac.in

[2] Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India

[*]Corresponding author email: aarthy.sl@vit.ac.in

Abstract: Dementia is a condition when thinking, reasoning and memory skills are lost and patients have emotional instability and personality changes. Researchers are looking into how the underlying disease processes that lead to various kinds of dementia begin and interact. Additionally, they keep researching the various diseases and conditions that cause dementia. Alzheimer's and Parkinson's disease contribute to dementia development. Recently deep learning-based techniques have surpassed the performance of traditional algorithms in the field of machine vision, image detection, natural language handling, object detection, and medical image analysis. This study proposed a transfer learning-based model for Parkinson's and Alzheimer's disease classification from slices of MRI. Pretrained VGG19 with Batch normalization is used for feature extraction and the final dense (fully connected-FC) layers are fine-tuned to meet our requirements. The performance of the model is analyzed by varying hyperparameters. The proposed model outperformed other pre-trained CNN models by achieving an accuracy of 97.19%.

Keywords: Alzheimer's, Parkinson's, VGG, deep learning, transfer learning, batch normalization.

1. Introduction

Millions of individuals worldwide are affected by dementia, which is more prevalent in the elderly population and manifests as cognitive decline, personality changes, and behavioral alterations in daily life. The severity of dementia ranges from the least serious phase, when it is only beginning to affect a person's ability to function, to an extremely acute stage, when a person has to rely entirely on others for basic daily tasks like feeding themselves. It is caused by progressive and permanent loss of neurons and brain function caused by a variety of neurodegenerative disorders. The most prevalent senile form of dementia is Alzheimer's disease (AD), which is then followed by Parkinson's disease (PD). It is estimated that more than 44 million individuals worldwide are currently afflicted by this disease. By the year 2050, this number would have surpassed 100 million [1].

With the advancement of deep learning-based (DL) techniques, neuroimaging is now extremely effective at detecting functional and morphological alterations in the brain [2]. Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), and Single Photon Emission Tomography (SPECT) Single Photon Emission Tomography (SPECT) are the common neuroimaging techniques utilized in the diagnosis of various disorders. Using MRI technology, the inside physiology of the body of an individual is visualized. The speed of an MRI is substantially increased by the use of a quickly changing gradients magnetic field, which also improves the resolution of soft tissues without endangering human health from radiation [3]. When someone or a family member exhibits the initial symptoms of PD or AD, people could find it challenging to differentiate between the two conditions. The intensities of variation in medical imaging can be used to identify atrophy or lesions brought on by these disorders.

Transfer learning is the concept of transferring learned features (weights of a model) from a pre-trained model to apply for the desired new task instead of learning from scratch. It is more common to use a convolutional network (ConvNet) as initialization or a feature extractor for the desired task after pre-training it on a very massive database (like ImageNet data, containing around 1.2 million images with 1000 different classes). By adjusting the final fully connected layers, these pre-trained networks can be applied to smaller datasets using the transfer learning approach. Additionally, it has shown itself reliable for classifying cross-domain issues where CNN is trained on natural images and evaluated on medical images[4].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 24 August 2023 **Revised:** 12 October 2023 **Accepted:** 26 October 2023*

Many researchers have utilized the transfer learning approach for the prediction of these diseases. With the advancement of DL in the field of machine vision and image recognition, automated diagnosis of these diseases has been developed by many data scientists.[5] proposed finetuned VGG19 achieved an accuracy of 97.89% for the classification of several AD stages. [6] used a 3D CNN model from scratch and was trained to classify PD and AD from T1-w MRI images. ROC-AUC values of 0.878 and 0.789 on public datasets namely ADNI and OASIS test datasets respectively to classify AD from healthy subjects. Similarly for PD classification, this study attained average ROC-AUC scores of 0.743 and 0.667 on Upenn and PPMI public datasets respectively.

[7] proposed transfer learning-based approach by freezing features of the pre-trained deep CNN model (VGG19-SVM) to classify AD, MCI(Mild Cognitive Impairment), and HC (Healthy controls) (MCI vs HC, HC vs AD, AD vs MCI). [8] developed a model assigned initially with pre-trained features of massive datasets containing real pictures and investigated finetuned layerwise learning by freezing groups of convolutional layers in different configurations by employing an intelligent way of selecting informative slices for the training dataset, achieving an accuracy of 95.19% for the classification of ADvsMCIvsNC.

[9] developed an early screening system using a capsule network architecture over the ADPP (Alzheimer's disease and Parkinson's progression) dataset to classify AD, PD, and normal subjects. This system achieved an encouraging accuracy of 97%, 98%, and 96% for the categorization of AD vs PD vs healthy. [10] proposed two different methods for multi-class AD stage detection. The first method used a simple CNN model with 2D and 3D mri samples of the ADNI dataset, achieving promising results of accuracy of 93.61% and 95.17% for muti-class AD detection. The second method employed transfer learning-based architecture using a pre-trained VGG19 model with an accuracy of 97%.

This study used a pre-trained VGG19 deep CNN model for feature extraction and the fully connected layers are fine-tuned to classify AD and PD from healthy controls. Additionally, Batch normalization of hidden activation functions is carried out to increase the speed of the training process and prevent overfitting issues. The objective of this study is listed below

- To propose a transfer-learning-based multi-class classification of AD, PD, and healthy controls.
- To analyze the effect of batch normalization on the efficiency of the proposed model.
- To analyze the comparative and performance analysis of the proposed model by varying parameters with the other pre-trained CNN models.

2. Materials and Methods

2.1 Data Set Description

The assessment of the proposed model's performance is done using a collection of PD and AD MRI images from the benchmark public datasets, Parkinson's Progression Markers Initiative shortly called PPMI and Alzheimer's Disease Neuroimaging Initiative (ADNI), respectively. The PPMI is an innovative project that is working in collaboration with worldwide partners to provide a comprehensive repository for unrestricted data collecting of biosamples. PPMI includes data from diagnostic, neuroimaging, omics, genetics, sensory tracker, reagents, and subject levels for the diagnosis of PD. Comparatively, ADNI is ongoing multi-modal research to create pharmacological, neuroimaging, genetic, and metabolic markers to monitor the disease progression through time and to identify methods for diagnosing AD before the onset of the ailment (pre-staged dementia). A detailed summary of the number of subjects (AD, PD, and Healthy) and MRI images involved in this study are tabulated below in the Table.

Table1. Dataset Description

Attributes		**ADNI**	**PPMI**
No. of subjects	Male	132	117
	Female	143	168
Classes	AD	206	-
	PD	-	207
	Healthy	69	78
Total no. of subjects		**275**	**285**

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 24 August 2023 **Revised:** 12 October 2023 **Accepted:** 26 October 2023*

2.2 Data Preprocessing

MRI images acquired from different modalities may possess diverse distortions caused by undesirable noise that results in an inaccurate diagnosis. This emphasizes the need for preprocessing techniques to strip noise and increase the quality of images for the accurate assessment of data. The MRI images can be downloaded in either DICOM or NIFTY format that contains images along with metadata. Here DICOM files are downloaded and images are converted to JPEG files. 2D MRI slices (of size 166x256) were obtained from the original MRI volume which was 166x256x256 in size. The input slice images were re-sized to 224x224 before being fed into the CNN model. A 2D Gaussian filter is used for image denoising in this study. Contrast enhancement technique, Contrast Limited Adaptive Histogram Equalization (CLAHE) is a variant type of AHE (Adaptive Histogram Equalization) for improving contrast in images. Here, the clip limit is fixed between 2 and 3. The computation to determine the newer pixel, P_n is expressed in equation (1).

$$Pn=Pm+Cpz\times PM-Pm \quad \textbf{(1)}$$

From (1), P_n indicates the pixel value, P_M refers to the highest pixel value, P_m refers to the lowest pixel value, and Cpz denotes the accumulative Probability Distribution. Figure 1 demonstrates the enhanced MRI image using CLAHE.

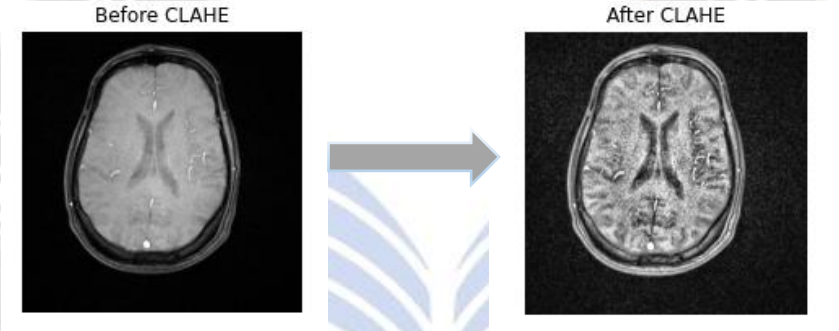

Figure1. Enhanced sample MRI image using CLAHE

2.3 Data Selection

Transfer learning offers the chance to use a smaller collection of training data, but the performance of the overall strategy still depends on selecting the optimum training data. In most of the studies, images are selected randomly for training. Instead in our study, we employed the image entropy strategy for extracting the more informative slices from the entire datasets. To determine the entropy for a single slice calculate its similar to the histogram. Entropy gives us a way to quantify variation in each slice. An image carries more information the higher its entropy. Entropy is not suitable for a general image collection since it is excessively noise-sensitive. But in this study, the images acquired have already undergone many pre-processing methods for denoising.

For an MRI slice, the Image entropy (I) can be calculated as below in equation (2).

$$I=-i=1Npilogpi \quad \textbf{(2)}$$

Here, the number of grey levels is denoted by N i.e., N=256 for an 8-bit image, and the probability of a pixel with grey level intensity by p_i. The normalization of input MR images is done to the extracted images to create an even contrast and intensity range across all the pictures. By choosing a general normalization criterion, each MR image in the input dataset is normalized to the range (0, 1).

2.4 Data Splitting

Data splitting is an important feature in Data science before building any model. It can be either a fixed or random selection of training and test data from the entire dataset collected. Most of the research studies have done slice-level data splitting. In that case, the brain slices belonging to a subject will be included in both the training and test data. This may lead to an issue called data leakage causing high accuracy from biased results. Here in our study, we employed subject-wise data selection. This approach is by which entire MRI slices belonging to a subject are included either in the training data or the test data.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 24 August 2023 **Revised:** 12 October 2023 **Accepted:** 26 October 2023*

2.5 Deep Neural Networks and Transfer Learning (TL)

Machine learning (ML) algorithms are a subset of Artificial Intelligence (AI) in which learning is made by machines without the intervention of human experts. Even though most ML models have surpassed the proficiency of human experts, especially in the field of machine vision and speech, their performance is not much satisfactory. These challenges give rise to the development of artificial neural networks caused the core of all DL techniques that have been rapidly developed in the late 1990s and gained enormous popularity in the domain of signal and image processing, machine vision, audio/video recognition, human language processing, and many more.

Convolutional Neural Networks (CNN) is one of the deep neural networks motivated by the biological structure of the human brain. Because of CNN's popularity, several known architectures for visual classification issues have been meticulously created by researchers over the previous few years. They gained enormous popularity since the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), to classify input images into 1000 different classes has to be done by the participants. The ILSVRC challenge's rigorous evaluation assures that models which score higher performance are fairly reliable and well-tested.

Recently, CNN algorithms have shown exemplary performance in object detection, restoration, registration, segmentation, and categorization of medical images in the area of healthcare computer-assisted diagnosis. The most attractive feature of CNN is the utilization of temporal or spatial correlation in the input data. CNN is a multi-layered architecture that includes ConvNet (convolutional filters +activation function) layers, pooling layers, and dense (fully connected-FC) layers. Here Feature extraction is done by the convolutional layers where output is computed by kernels initially by splitting the input into small segments and performed convolution using weights obtaining distinct feature maps. A nonlinear activation function is used to evaluate and obtain new feature maps facilitating the learning of different semantics in images. Not only does this aid in the learning of abstractions, but it also embeds non-linearity in the feature space. Relu and its variants, TanH, and Sigmoid are the different activation functions used for learning complicated patterns. Following each block of convolution layers is a pooling layer that is used as feature reduction to downsample kernel output by summarising similar information from the adjacent receptive region and outputting the dominant response from the data entered. Different varieties of pooling like average, L2, overlapping, and max pooling can be used. Finally, the last layer is a set of a dense layer and a softmax or regression layer to evaluate class scores. Different CNN models have been developed recently, including AlexNet, LeNet, GoogLeNet, InceptionV3, VGG, Inception-ResNetV2, DeneNet, mobileNet, and many others.

CNNs are trained using backpropagation [19], which entails incrementally updating each layer's unknown weights to maximize a predetermined cost function. The initialization of the weights typically starts with an arbitrary set of values. Due to the vast number of weight parameters that are often connected with a CNN model, an enormous number of training samples are needed for the repeated backpropagation technique to attain convergence appropriately. A method may become trapped at local minima when given such huge data samples for training, leading to subpar classification results. The CNN's weights are taken from a neural network that was previously trained on a larger set of data in a process known as fine-tuning or transfer learning as alternatives. Transfer learning has produced an extensive examination and encouraging outcomes of training deep models from scratch as opposed to fine-tuning medical images, such as the classification of precancerous diseases, heart and brain imaging, and lung diseases. Fine-tuned CNNs have outperformed training from scratch in most cases. All of these methods show that transfer learning has a lot of utility in the medical imaging arena and the ability to achieve better accuracy in PD and AD diagnosis by having small training samples than learning from scratch[11].

The benchmark challenge, ILSVRC is commonly used to evaluate the best architectures. TL concept aims to use elegant architectures for desired new tasks. As a result, we looked into the recent ILSVRC competition finalists to find an architecture that would be acceptable for AD and PD diagnosis.

We closely adhere to the VGG design [12] offered by the Oxford Visual Geometry Group, which asserted an accuracy rating of 92.7% in the ILSVRC 2014 competition. In addition to having a higher degree of accuracy, the VGG design was chosen for its efficiency and, more crucially, its flexibility to use outside of ImageNet. The architecture has recently been proven to help with computer-aided diagnosis issues. The main idea behind the architecture is to maintain other network characteristics while adding more convolutional layers to the network to enhance its depth. The size of the convolution filter is fixed to a minimum (3x3) across all layers to deal with the number of parameters for training.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 24 August 2023 **Revised:** 12 October 2023 **Accepted:** 26 October 2023*

2.5.1 Fine-tuned VGG19 with Batch normalization

We proposed a single deep model that is finetuned from a pre-trained VGG19 model to classify AD, PD, and Healthy controls from MRI images. The architectural design of the proposed study is depicted in Figure 2. This design is very similar to the concept of VGG-19 proposed in the actual publication, with minor alterations in the fully connected layers matching our needs. Six blocks of 16 CONV2d (convolutional) layers, 5 Maxpool2d layers with stride 2, 4 FC (completely connected) layers, and a Logsoftmax layer are included in the finetuned VGG-19 network. The input MRI pictures are routed through a stack of 3x3 convolutional filters.

Each block can be summarised as follows

B1: 2 layers of 3x3 convolution filters and 64 channels. (64x224x224)

B2: 2 layers of 3x3 convolution filters and 128 channels. (128x112x112)

B3: 4 layers of 3x3 convolution filters and 256 channels. (256x56x56)

B4: 4 layers of 3x3 convolution filters and 512 channels.(512x28x28)

B5: 4 layers of 3x3 convolution filters and 512 channels.(512x14x14)

B6: Four Fully connected layers. (FC1 and FC2-1x1x4096, FC3-1x1x256, FC4-1x1x3)

The network depth makes the VGG model more robust and hence the above deepest architecture is followed. The kernel size is fixed as 3x3 with a stride of 1 and the number of channels specifying the width gradually increases to utilize the complex features. Keeping a small stride makes receptive fields overlap by which most of the important features can be utilized. Batch normalization is a process used to improve the speed of the training process by normalizing activations of the hidden layers. Every CONV block with a stride value of 2 is combined with a maxpool2d layer. As the Relu activation function is computationally more efficient and lowers the impact of the vanishing gradient problem, it is used here by all the hidden layers. The B5 CONV block is followed by a 2D global avg pooling operation to get a 1-dimensional tensor of size (1x1x4096) as the output to the four dense layers and the final output log-softmax layer performs the disease classification process. The proposed design used only FC layers for training and the convolutional blocks have been kept fixed.

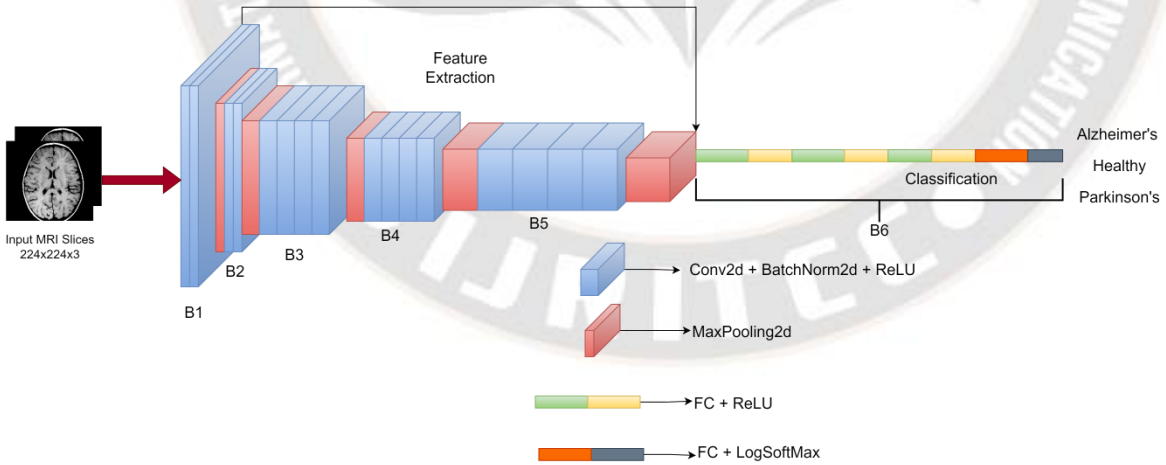

Figure2. The architecture of the Proposed model-Fine tuned VGG19 With Batch Normalization

2.5.2 Batch Normalisation

Overfitting is one of the most common problems encountered while training deep models by data scientists. Regularisation is one of the solutions used to avoid overfitting and also improves the performance of the model. Several regularisation techniques like dropout, early stopping, and batch normalization are available. We employed batch normalization [13] in every CNN block of the training model and analyzed the effectiveness of the learning process. Initially, the normalization

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 24 August 2023 **Revised:** 12 October 2023 **Accepted:** 26 October 2023*

of inputs in each hidden block is processed and followed by the rescaling and offsetting process. Normalization is the process to transform input data (in batches) such that the mean is zero and the standard deviation is one. By preventing division by a zero value, the smoothing term () ensures the numerical stability of the neural networks. Below is the formula to calculate normalization. Finally, the parameters gamma and beta are used for rescaling and offsetting respectively. These two learnable parameters are optimally assigned to allow for accurate batch normalization.

2.6 Training protocols

Fine-tuning a pre-trained model (VGG19) in the transfer learning approach is carried out by freezing the weights of some layers and varying hyper-parameters accordingly. To minimize the loss function over training data and improve the performance of the model, we used categorical cross-entropy loss (CEL). The fast convergence of the model is obtained using the Adam optimizer[14]. The open-source deep learning framework, PyTorch is used to implement the deep learning methods. Fine-tuning of various hyperparameters is tabulated in Table 2.

Table 2. Hyper-parameters proposed in this study

Batch Size	8,16,32,64,72
Learning rate	0.00001
Epochs	100
Loss	Cross Entropy loss
Optimiser	Adam
dropout	0.5

3. Performance Evaluation Metrics

The following efficiency metrics are used for evaluation of the proposed model.

$$Accuracy=TP+TN(TP+TN+FP+FN)$$

$$Precision=TP(TP+FP)$$

$$Recall=TP(TP+FN)$$

$$F1\text{-}Score=2*Precision*RecallPrecision+Recall$$

Here, TP-TruePositive, TN-TrueNegative, FP-FalsePositive, FN-FalseNegative

4. Results and Discussion

Data samples collected from the two public datasets had undergone several preprocessing techniques as discussed earlier. Once the preprocessed and most informative slices are extracted, the entire datasets are separated into three subsets: training data, validation data, and test data as shown in Table 3.

Table 3. Prevalence of datasets acquired	
Total no. of MRI images=7748 AD controls:2561 PD controls:2560 Healthy controls:2627	Training data set(70%): 5422 images Validation data set (20%): 1549 images Test data set (10%): 777 images

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *24 August 2023* ***Revised:*** *12 October 2023* ***Accepted:*** *26 October 2023*

Experimental results based on varying hyperparameters of VGG19 (with and without batch normalization) are tabulated in Table 4. It is clearly shown that the average accuracy of VGG19 with batch normalization for all three classes is 97.19% from the table while the average accuracy obtained by using VGG19 without Batch normalization is 96.56%.

Table 4. Comparative Analysis of varying hyperparameters for proposed fine-tuned VGG19 with and without Batch normalization(BN)

CNN Model	Batch size	Best epoch	Loss method	Training Acc (%)	Training loss (%)	Val Acc (%)	Val loss (%)	Training Time (m)
Fine-tuned VGG19 Without BN	16	93	CEL	96.15	0.0836	96.32	0.1145	65.12
	32	93	CEL	96.15	0.0865	95.80	0.1695	62.69
	64	93	CEL	95.39	0.0963	95.61	0.1710	58.09
	72	99	CEL	94.96	0.1086	94.58	0.2042	59.91
Fine-tuned VGG19 With BN	8	74	CEL	96.48	0.0787	94.96	0.1971	91.84
	16	78	CEL	96.40	0.0779	96.32	0.1363	72.59
	32	**74**	**CEL**	**97.33**	**0.0809**	**94.58**	**0.2949**	**69.40**
	64	78	CEL	95.66	0.1015	94.71	0.2026	59.82

Acc- Accuracy, Val- Validation

The efficiency of the proposed model is further compared and analyzed with other pre-trained CNN models like AlexNet, ResNet50, and VGG16 and tabulated in Table 5. ResNet50 has the least accuracy and does not perform well in the classification process. The accuracy of VGG19 with or without batch normalization is almost nearer but the training time is less when batch normalization is performed. It can be concluded that the batch normalization procedure just speeds the model training process; it does not affect the model's accuracy.

Table 5. Comparative analysis of the proposed model with futuristic models

Pre-trained Model	Acc (%)	Precision (%)	Recall (%)	F1-Score (%)	Training time (m)
AlexNet	93.82	81.23	86.25	83.66	47.30
ResNet50	66..92	56.30	62.36	59.17	90.70
VGG16	94.85	96.54	97.12	96.82	73.65
VGG19 (without BN)	96.56	93.56	88.79	91.11	61.45
VGG19 (with BN)	**97.19**	**94.36**	**96.21**	**95.27**	**53.24**

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *24 August 2023* ***Revised:*** *12 October 2023* ***Accepted:*** *26 October 2023*

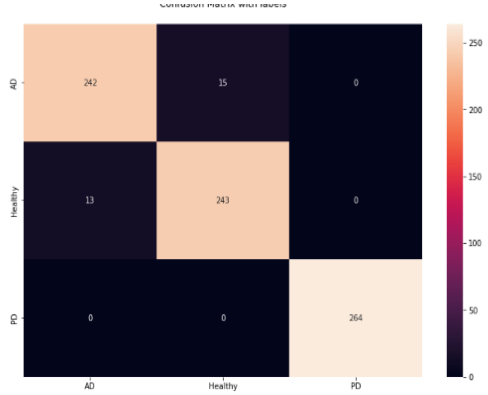

Figure3.Confusion Matrix for VGG19 Without BN

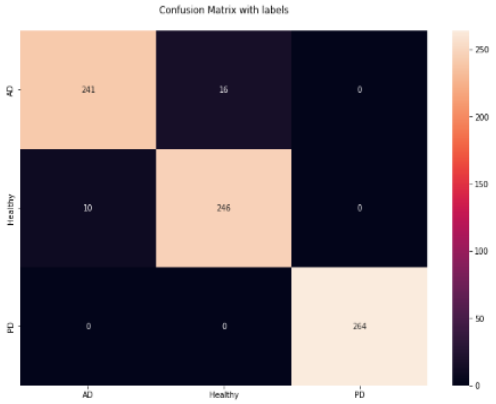

Figure 4. Confusion Matrix for VGG19 with BN

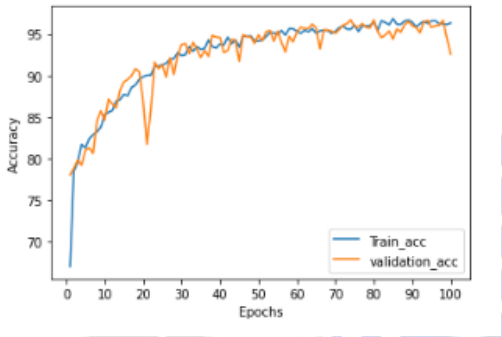

Figure 5. Accuracy graph of fine-tuned VGG19 with BN

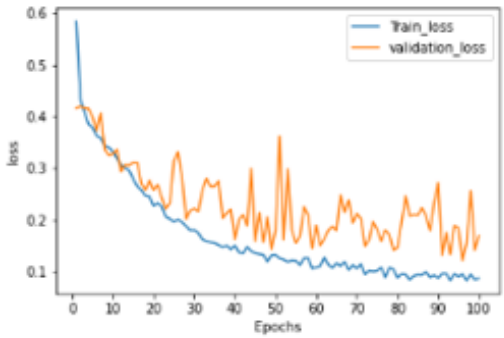

Figure 6. Loss graph of fine-tuned VGG19 with BN

Figures 3 and 4 demonstrate the confusion matrices for the VGG19 with and without batch normalization, respectively. The suggested fine-tuned VGG19 model's accuracy, recall, F1-score, and precision are shown in Table 5, along with those of other pre-trained models including VGG-16, AlexNet, and ResNet50. Figures 5 and 6 display the accuracy and loss graphs for the suggested model. The findings obtained demonstrate that for the categorization of AD, PD, and healthy controls, the fine-tuned VGG19 model outperforms the other pre-trained CNN models.

Conclusion

This study performed the classification of Parkinson's and Alzheimer's disease from MRI images by a transfer learning-based approach. Nowadays, deep learning techniques have given outperformed results than traditional machine learning methods in the field of medical image recognition. The transfer learning approach is very helpful when the dataset is small.MRI samples acquired from two public datasets ADNI and PPMI are undergone several preprocessing techniques. The most informative slices are selected based on the image entropy method. To avoid data leakage issues, subject-level splitting of data is done by which MRI slices of a subject are completely included in either the training or test dataset. This study leverages the VGG19 model pre-trained to extract features and fine-tune the last dense layers for the categorization of MRI images. The proposed finetuned VGG19 with batch normalization has a better performance compared to other CNN models. The effect of batch normalization on the accuracy of the training model has been analyzed in this study. The proposed finetuned VGG19 with batch normalization model achieved an overall accuracy of 97.19% for the multi-classification of AD, PD, and Healthy subjects.

References

Data Dashboard

Parkinson's Progression Markers Initiative. (n.d.). Www.ppmi-Info.org. Retrieved December 28, 2022, from https://www.ppmi-info.org/access-data- specimens/data

ADNI | ACCESS DATA. (n.d.). Retrieved December 28, 2022, from https://adni.loni.usc.edu/data-samples/access-data/

[1].Vijayan, M., & Reddy, P. H. (2016). Stroke, vascular dementia, and Alzheimer's disease: molecular links. Journal of Alzheimer's disease, 54(2), 427-443.

[2].Ebrahimighahnavieh, M. A., Luo, S., & Chiong, R. (2020). Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review. Computer methods and programs in biomedicine, 187, 105242.

[3].Duyn, J. H. (2012). The future of ultra-high field MRI and fMRI for study of the human brain. Neuroimage, 62(2), 1241-1248.

[4].Shi, J., Zheng, X., Li, Y., Zhang, Q., & Ying, S. (2017). Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. IEEE journal of biomedical and health informatics, 22(1), 173-183.

[5].Mehmood, A., Yang, S., Feng, Z., Wang, M., Ahmad, A. S., Khan, R., ... & Yaqub, M. (2021). A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images. Neuroscience, 460, 43-52.

[6].Dhinagar, N. J., Thomopoulos, S. I., Owens-Walton, C., Stripelis, D., Ambite, J. L., Ver Steeg, G., ... & Thompson, P. M. (2021, December). 3D convolutional neural networks for classification of Alzheimer's and Parkinson's disease with T1-weighted brain MRI. In 17th International Symposium on Medical Information Processing and Analysis (Vol. 12088, pp. 277-286). SPIE.

[7].Naz, S., Ashraf, A., & Zaib, A. (2022). Transfer learning using freeze features for Alzheimer neurological disorder detection using ADNI dataset. Multimedia Systems, 28(1), 85-94.

[8].Khan, N. M., Abraham, N., & Hon, M. (2019). Transfer learning with intelligent training data selection for prediction of Alzheimer's disease. IEEE Access, 7, 72726-72735.

[9].Bhatele, K. R., Jha, A., Kapoor, K., & Tiwari, D. (2022). Neurodegenerative diseases-Caps: a capsule network based early screening system for the classification of neurodegenerative diseases. Cognitive Neurodynamics, 16(6), 1361-1377.

[10].Helaly, H. A., Badawy, M., & Haikal, A. Y. (2021). Deep learning approach for early detection of Alzheimer's disease. Cognitive computation, 1-17.

[11].Suganya, A., & Aarthy, S. L. (2022). Alzheimer's And Parkinson's Disease Classification Using Deep Learning Based On MRI: A Review. International Journal of Communication Networks and Information Security, 14(1s), 9-21.

[12].Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[13].Tuvshinjargal, B., & Hwang, H. (2022). VGG-C transform model with batch normalization to predict Alzheimer's disease through MRI dataset. Electronics, 11(16), 2601.

[14].Kingma, D.P.; Ba, J.L. (2015). Adam: A method for stochastic optimization. International Conference on Learning Representations (ICLR), 1–15.