Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Deep Learning Techniques in Classification of Stages in Dementia: An Ensemble Approach

Mariappan Pandiyarajan[1,*], R. S. Valarmathi[1]

[1] Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India

ARTICLE INFO

Article history:
Received 21 July 2023
Received in revised form 4 October 2023
Accepted 17 November 2023
Available online 25 December 2023

Keywords:

Dementia; Contrast limited adaptive histogram equalization; Transfer learning; Ensemble classifier models

ABSTRACT

Dementia is the severe brain disease that affects the memory system of the humans which requires early diagnosis. It progresses through different stages from very mild dementia, which motivates to find the stage of dementia for treating the affected persons. A novel ensemble-based classification model to classify the stages in dementia such as mild-dementia, non-dementia, very-mild dementia has been proposed in this work. The MRI images were pre-processed with techniques such as resizing, data augmentation and Contrast Limited Adaptive Histogram Equalization (CLAHE) and different four forms of datasets such as Image Data-1, Image Data-2, Image Data-3, Image Data-4 are framed to justify preprocessing techniques. The pre-trained models namely ResNet, Inception, MobileNet, VGG, DenseNet, Xception, Inception Resnet, AlexNet, EfficientNet and NasNet Mobile are trained with all four forms of dataset and performance of each model is analysed. The models DenseNet201, MobileNetV2, VGG19, ResNet152 achieved accuracy of 93%,93%,93%.90% respectively with Image Data-4. The proposed model is the ensemble of DenseNet201, MobileNetV2, VGG19, ResNet152 which provides accuracy precision, recall and f1-score of 95%, 93%, 100% and 96% respectively. This ensemble model has undergone testing using a brain tumour dataset, and the experimental findings indicate that the proposed framework enhances the classifier's generalization capabilities for multiclass problems.

1. Introduction

The Dementia is a type of brain disease which causes memory loss that affects the brain and its initial stage is Alzheimer's, which causes memory, cognitive, and behavioral issues [1]. Dementia symptoms may not be severe in the early stages, it becomes severe as the disease damages the brain. Alzheimer's patients typically live about eight years after their first symptoms appear, while everyone's condition progresses at a different rate [2]. In the year 2020, approximately 5.8 million Americans experienced the effects of dementia. While Alzheimer's disease can affect individuals at a younger age, such cases are uncommon. However, as people grow older, their risk of developing

* *Corresponding author.*
E-mail address: pandiyarajan14@gmail.com

https://doi.org/10.37934/araset.36.1.131146

Alzheimer's disease increases significantly. Notably, for every 5-year interval beyond the age of 65, the number of individuals living with Alzheimer's doubles. Projections indicate that this number is expected to nearly quadruple, reaching 14 million by the year 2060. Symptoms of the disease typically manifest after the age of 60, and the likelihood of developing Alzheimer's disease continues to rise with advancing age [3]. Diagnosing dementia and determining its subtype is a complex process that requires the integration of information from multiple sources like clinical assessment, cognitive testing, laboratory workup, and neuroimaging [36]. Neuroimaging has transformed our ability to diagnose and understand neurodegenerative disorders that lead to dementia. The shift towards earlier detection and a deeper understanding of underlying neuropathology is a promising approach to tackle these devastating conditions and work towards more effective treatments and interventions [37].

2. Literature Review

Qiang *et al.,* examined the classification performance of AlexNet, GoogleNet, VggNet, and DenseNet using 27 image datasets, each created with three different image factors. Their goal was to assess how well these neural networks could classify various texture features through a series of texture-based experiments [4]. Sitara Afzal *et al.,* introduced a novel structure consisting of two distinct pre-trained convolutional models, each featuring four stages. They evaluated their proposed technique using the SWI dataset, which included data from 20 subjects. Notably, the performance accuracies achieved using ResNet50 and AlexNet, both with augmentation, reached an impressive 97% [5].

Chunfeng Lian *et al.,* proposed an attention-guided deep learning framework for dementia diagnosis using structural MRI. This framework combines a fully convolutional network (FCN) and a multi-branch hybrid network (HybNet) to jointly learn and fuse sMRI features for CAD model development [6] proposed the TrFEMNet model for medical image classification, showing comparable performance to other models, especially on complex datasets [7]. Mohammad Monirujjaman Khan *et al.,* explored the application of convolutional neural networks (CNNs) for diagnosing brain cancers using medical images. Their approach employed pre-trained deep CNN models, specifically VGG19 and MobileNetV2, to extract deep features from a dataset consisting of brain X-rays and CT scans from individuals diagnosed with malignant brain tumors [8].

Chetana Srinivas *et al.,* compared VGG-16, Inception-v3, and ResNet50 for brain tumor cell prediction. Their study, while centered on these models, offers potential for expansion with others like VGG-19, MobileNet, and EfficientNet to enhance brain tumor analysis [9]. Hiroaki Iwata *et al.,* used VGG16 and ResNet50 models with transfer learning to classify 10 different excipients based on particle morphology in SEM images of pharmaceutical raw materials [10]. Suriya Murugan *et al.,* introduced a dementia stage identification and prediction model. They employed a convolutional neural network (CNN) to extract distinctive features and mitigated class imbalance using the Synthetic Minority Oversampling Technique (SMOTE) [11].

Juan Miguel Valverde *et al.,* observed that most of the surveyed studies did not thoroughly investigate the interpretation of their strategies post-transfer learning application. Additionally, these studies often omitted comparisons with other transfer learning approaches [12]. Hongfei Wang *et al.,* emphasized the role of dense connections in improving 3D network performance. They presented an ensemble of 3D convolutional networks for Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI) diagnosis using brain MR images, enhancing performance through a probability-based ensemble method [13]. Li, Hejie *et al.,* proposed an attention-based micro-design (EB2) for Alzheimer's diagnosis. They used the Global Attention Module (GAM) to capture global lesion information and the Channel Attention (CA) mechanism for essential features. The Micro-Design (MD) reduced complexity for improved performance [14].

Al Shehri W *et al.,* proposed a deep learning-based solution utilizing DenseNet-169 and ResNet-50 CNN architectures for the diagnosis and classification of Alzheimer's disease [15]. According to the experimental results, the deep ResNet model demonstrated superior performance compared to other architectures [16-18]. Aulia *et al.,* introduced an image enhancement approach employing Clip Limit Adaptive Histogram Equalization (CLAHE) to enhance brain MRI segmentation. They then used the Visual Geometry Group-16 Layer (VGG-16) for brain tumor classification on MRI images [19]. Zahid Ullah *et al.,* employed histogram equalization techniques for image enhancement and deep neural network (DNN) for classification obtained improved results [20].

Sohaib Asif *et al.,* found that the Xception architecture outperformed other models for brain tumor detection, with a 99.67% accuracy on a 3-class dataset and 95.87% on a 4-class dataset. They compared it to DenseNet201, DenseNet121, ResNet152V2, and InceptionResNetV2 [21]. Clip Limit Adaptive Histogram Equalization (CLAHE) is used as image enhancement techniques in various bio medical applications like detection of Covid-19, prediction of diabetic retinopathy, diagnosis of breast cancer and also for removing noise in the images [22-25]. Ensemble of pretrained models increased the performance compared with other neural network model in the field of biomedical for disease classification and prediction and image recognition [26-31]. Multiples of research has been carried out with pretrained models like 3D-CNN, InceptionV3, Xception, MobilenetV2, VGG to improve the performance of predicting early stages of dementia [32-35].

3. Methodology

Pre-trained deep learning models are neural networks that have undergone training on large datasets to learn feature representations directly from raw data. These models are designed with specific architectures tailored to different types of data, including convolutional neural networks (CNNs) for images, recurrent neural networks (RNNs) for sequences, and transformer models for text. The primary advantage of pre-trained models is that they capture a significant amount of information from the training data, making them useful for a wide range of tasks even with limited labelled data. Various pre-trained models like ResNet, Inception, MobileNet, VGG, DenseNet, Xception, Inception Resnet, AlexNet, EfficientNet and NasNet Mobile is used with different forms of input to choose the best model for this research work. Further, as shown in the flow diagram of proposed ensemble classifier model in Figure 1, the selected four models were ensembled in different combinations to find the model which increases the performance for the dataset used.

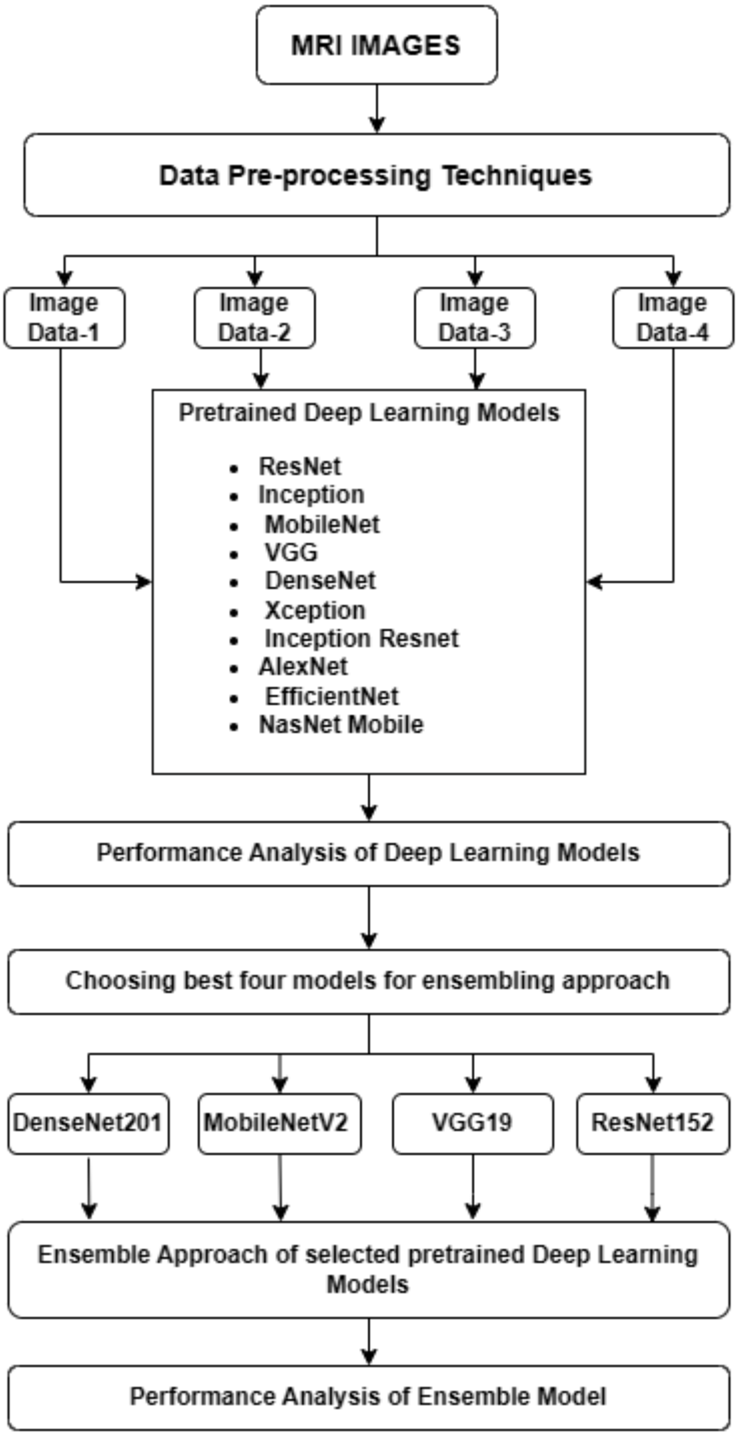


Fig. 1. Flow Diagram of Proposed Ensemble Classifier model

3.1 Data Preprocessing

The model's performance relies on both data quality and quantity. The brain MRI images are extracted from Open Access Series of Imaging Studies (OASIS). The MRI images are categorized in to three classes namely Mild Dementia (MD), Non-Dementia (ND)and Very Mild Dementia (VMD). The dataset contains MRI images of 235 patients under different stages of dementia. To improve the model's performance, various preprocessing techniques, including image augmentation, resizing, and histogram equalization, are applied to enhance the data. The dataset contains unbalanced number of subjects under each class. In order to achieve a balanced dataset, different data augmentation techniques are used. As a result, 235 MRI images were augmented in to 405 subjects, in which all three classes contains 135 MRI images as shown in Figure 2. The image enhancement technique is used for improving the visibility of details in images with varying illumination conditions. In this research work, by comparing various Histogram Equalization methods, CLAHE (Contrast Limited Adaptive Histogram Equalization) which enhance subtle features and textures in images that might be difficult to see due to poor lighting conditions is used. This is especially important in medical imaging, where details in X-rays, MRI scans, and microscopy images can be crucial for accurate diagnosis.

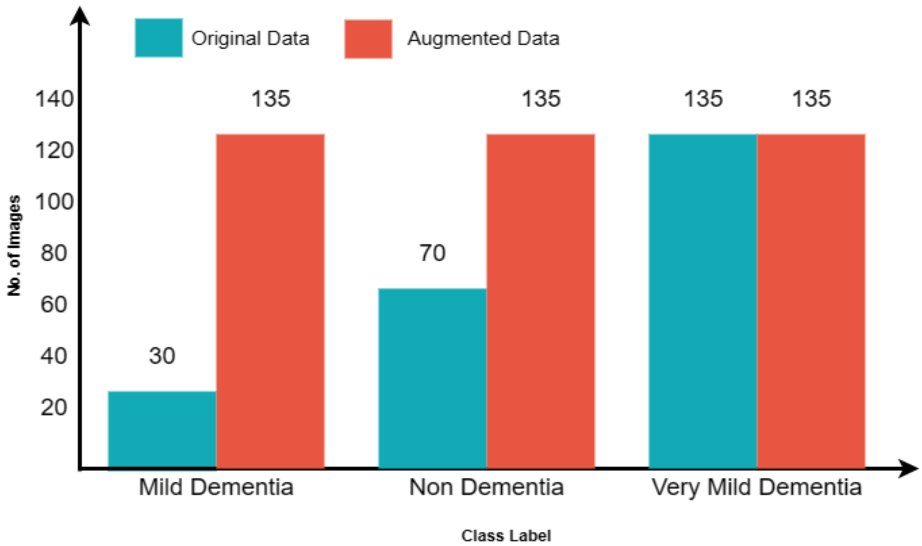


Fig. 2. Comparison of number of images after data augmentation

The size of the input image is 256 x 256 x 3. Further, based on the input size of the pre-trained deep learning models used, the image is resized. The pre-trained models were trained with four different forms of input images as shown in Figure 3. and the performance of the pre-trained network is analysed for each image data to justify the data preprocessing techniques and to choose the best models for assembling.

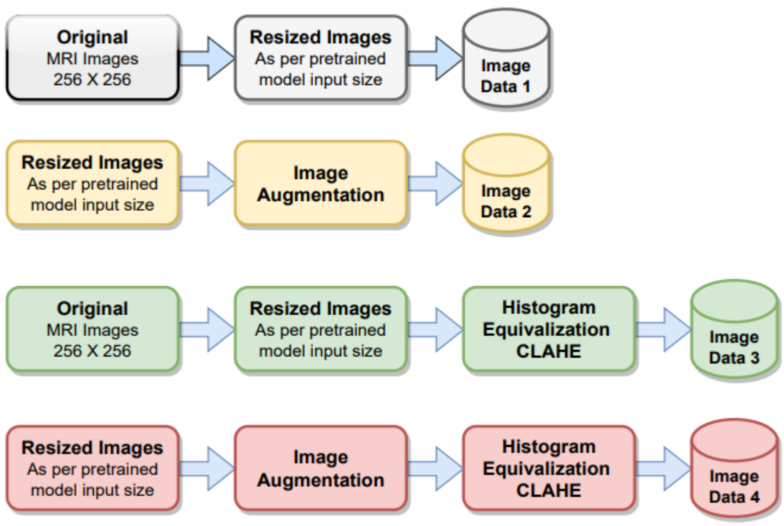


Fig. 3. Preprocessing Techniques for Image Data

3.2 Pre-Trained Deep Learning Model

In this research work, pre-trained deep learning models like ResNet, Inception, MobileNet, VGG, DenseNet, Xception, Inception Resnet, AlexNet, EfficientNet and NasNet Mobile is used with different forms of input to choose the best model.

3.2.1 DenseNet201

DenseNet-201 is a deep neural network with 201 layers, emphasizing maximum feature reuse. It establishes connections between all layers in a feedforward fashion. Each dense block includes bottleneck layers that use 1x1 convolutions to reduce feature map dimensions before applying 3x3 convolutions. It uses a growth rate parameter that determines the number of feature maps added to the output of each layer. This parameter controls the network's width. Between dense blocks, there are transition layers that perform down sampling using 1x1 convolutions followed by average pooling. It uses batch normalization and ReLU activation functions to stabilize and accelerate training. At the end of the network, global average pooling is typically applied to reduce the spatial dimensions to a single vector, which is then connected to a fully connected layer for classification.

3.2.2 MobileNetV2

MobileNetV2, designed for efficient and lightweight neural network inference on mobile and embedded devices, utilizes a unique "inverted residual with linear bottleneck" building block. This building block comprises three key components: depth-wise separable convolution, a linear bottleneck layer, and shortcut connections. Linear bottlenecks, characterized by 1x1 convolutional layers with a limited number of channels, are an essential part of this architecture. It incorporates the concept of channel expansion and the squeeze-and-excitation (SE) mechanism to improve feature representation. It provides various architectural variants, allowing users to choose the trade-off between model size and performance.

3.2.3 VGG19

VGG19 is a convolutional neural network (CNN) architecture from the VGG family. It consists of 19 layers, including 16 convolutional layers and 3 fully connected layers. The "19" in its name signifies the total layers. VGG19 employs 3x3 convolutional filters with a stride of 1 and padding of 1, maintaining feature map dimensions. Max-pooling layers with 2x2 windows and a stride of 2 are used for down sampling. After the convolutional layers, there are three fully connected layers, commonly used for classification. The last fully connected layer uses SoftMax for multi-class classification. ReLU is the activation function for most layers, while SoftMax is used for the output layer. Batch normalization is applied after each convolutional layer for training stability and speed.

3.2.4 ResNet152

ResNet-152 is a member of the ResNet (Residual Network) family of convolutional neural network (CNN) architectures, distinguished by its depth with 152 layers. This depth surpasses earlier CNN models like VGG and traditional deep networks. ResNet-152 achieves this depth by employing a multitude of residual blocks, each containing multiple convolutional layers and a skip connection (identity mapping) that adds the input to the block with the output. This innovative architecture effectively addresses the vanishing gradient problem, enabling the training of very deep networks. Within each residual block, a bottleneck architecture is used, incorporating 1x1, 3x3, and 1x1 convolutional layers. This design reduces computational complexity while preserving model capacity. Instead of fully connected layers, ResNet-152 typically employs global average pooling at the end of the network. This choice aids in reducing overfitting and enhances the network's suitability for transfer learning. To ensure stability and faster training, batch normalization is applied after each convolutional layer in ResNet-152. Additionally, Rectified Linear Units (ReLU) serve as activation functions after each convolutional layer in the network.

3.3 Ensemble Deep Learning Model

The Proposed Ensemble Classifier model for classification of Dementia is shown in Figure 4. The pretrained models were assessed for their performance, leading to the motivation for proposing an ensemble stacked classifier. This ensemble combines four deep learning models: DenseNet201, MobileNetV2, VGG19, and ResNet152. The base neural networks are kept frozen to retain the ImageNet weights during training. After the concatenation of these models, two dense layers with ReLU activation, each consisting of 64 neurons, are introduced, along with dropout layers. Finally, a dense layer with SoftMax activation is added for multiclass classification.

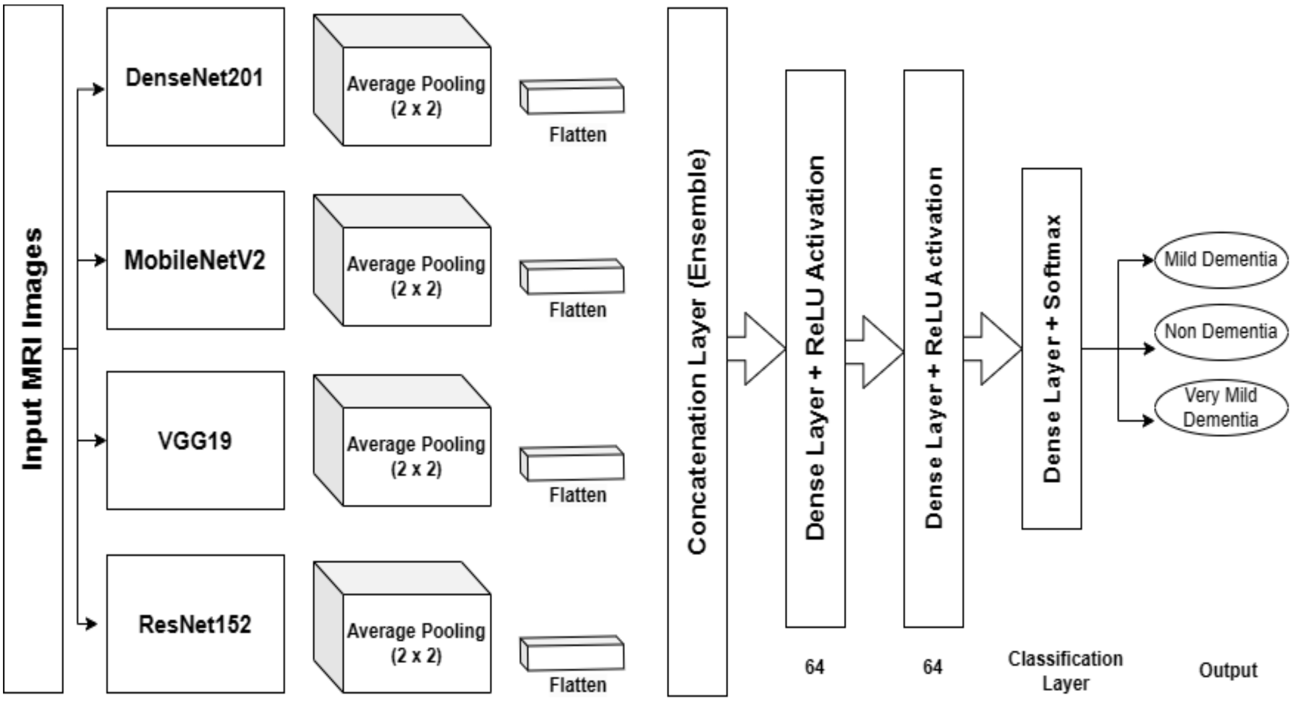


Fig. 4. Proposed Ensemble Classifier model for classification of Dementia

3.4 Performance Measures of Multiclass Classification for Deep Learning Models

To analyse the performance of neural networks the performance metrices like accuracy, precision, recall, f1-score, confusion matrix can be very useful and they are defined in this section with mathematical expression.

3.3.1 Accuracy

It measures the proportion of correctly classified instances out of the total number of instances in the dataset.

$$\text{Accuracy} = \frac{\text{Total Number of Instances}}{\text{Number of Correctly Predicted Instances}} \tag{1}$$

In the context of multiclass classification, accuracy considers all classes and calculates the percentage of correctly classified instances across all classes.

3.3.2 Precision

It measures the ability of a model to correctly classify instances for a specific class out of all the instances that the model predicted as that class. In the context of multiclass classification, precision can be calculated for each class separately, and you can then compute various averages (e.g., micro-average, macro-average) to get an overall understanding of model performance. precision is calculated for a single class (class i) in multiclass classification as defined in Eq. (2).

$$Precision_i = \frac{\text{Total Predicted Positives for class}}{\text{Total Predicted Positives for class i}} \tag{2}$$

Macro-average precision is calculated by computing the precision for each class separately and then taking the average of these class-wise precision scores.

$$Macro - average\ Precision = \frac{1}{n}\sum_{i=1}^{n} Precision_i \tag{3}$$

Where, n is the total number of classes.

3.3.3 Recall

Recall, also known as sensitivity or true positive rate, is another important evaluation metric for multiclass classification in deep learning. It measures the ability of a model to correctly identify all relevant instances of a specific class out of all the instances that actually belong to that class. In the context of multiclass classification, recall can be calculated for each class separately, and you can then compute various averages (e.g., micro-average, macro-average) to assess the overall performance of the model.

$$Recall_i = \frac{\text{Total Predicted Positives for class}}{\text{Total Actual Positives for class i}} \tag{4}$$

Macro-average recall is calculated by computing the recall for each class separately and then taking the average of these class-wise recall scores.

$$Macro - average\ Recall = \frac{1}{n}\sum_{i=1}^{n} Recall_i \tag{5}$$

Where n is the total number of classes.

3.3.4 F1-score

The F1 score is a harmonic mean of precision and recall. For multiclass classification the score of each class is calculated and then aggregated as defined in Eq. (7). to assess the overall performance of the model.

$$F1\ score_i = \frac{2*Precision_i*Recall_i}{Precision_i+Recall_i} \tag{6}$$

$$Macro - average\ F1\ Score = \frac{1}{n}\sum_{i=1}^{n} F1\ score_i \tag{7}$$

Where, n is the total number of classes. For the input image data 1 and image data 3, the metrices will follow weighted average score, as it contains unbalanced number of images and input image data 2, image data 3 follows macro average score ,as it contains balanced number of images.

4. Result and Discussion

The pre-trained models were trained with four types of image data set as described in section 3.1. The performance is assessed for choosing models which fits for ensemble. From Table 1, it is evident that the pre-trained deep learning models DenseNet201, EfficientNetB4 has accuracy of 62% while the models EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB5, NASNetMobile, DenseNet169, MobileNetV2, Xception has accuracy of 58% for the dataset of Image Data-1.

The accuracy of 93% is obtained with Image Data-4 for the models DenseNet201, MobileNetV2, VGG19 and 90% for the models InceptionV3, ResNet152.

Table 1
Performance Measures of Pre-Trained Deep Learning Models with input as Image Data 1

S.NO.	Model	Model Image Size	Predicted			Weighted Average			Accuracy (Valid)	Accuracy
			MD (3)	ND (14)	VMD (7)	precision	recall	f1-score		
1	AlexNet	224 X 244	0	7	2	0.27	0.26	0.26	0.375	0.38
2	DenseNet121	224 X 244	3	0	0	0.04	0.33	0.08	0.2174	0.12
3	DenseNet169	224 X 244	0	14	0	0.19	0.33	0.25	0.5652	0.58
4	DenseNet201	224 X 244	1	14	0	0.38	0.44	0.39	0.3478	0.62
5	EfficientNetB0	224 X 244	0	14	0	0.19	0.33	0.25	0.5652	0.58
6	EfficientNetB1	240 X 240	0	14	0	0.19	0.33	0.25	0.5652	0.58
7	EfficientNetB2	260 X 260	0	14	0	0.19	0.33	0.25	0.5652	0.58
8	EfficientNetB3	300 X 300	0	6	6	0.38	0.43	0.36	0.4783	0.5
9	EfficientNetB4	380 X 380	0	14	1	0.54	0.38	0.34	0.5652	0.62
10	EfficientNetB5	456 X 456	0	14	0	0.19	0.33	0.25	0.5833	0.58
11	InceptionResNetV2	229 X 229	0	12	1	0.31	0.33	0.3	0.6087	0.54
12	InceptionV3	229 X 229	0	0	7	1	0.3	0.15	0.3043	0.29
13	MobileNet	224 X 244	0	14	0	0.19	0.33	0.25	0.5652	0.58
14	MobileNetV2	224 X 244	0	13	0	0.2	0.31	0.24	0.5652	0.54
15	NASNetMobile	224 X 244	0	14	0	0.19	0.33	0.25	0.5652	0.58
16	ResNet101	224 X 244	0	0	7	0.1	0.33	0.15	0.2174	0.29
17	ResNet152	224 X 244	0	0	7	0.1	0.33	0.15	0.3043	0.29
18	ResNet50	224 X 244	0	0	7	0.1	0.33	0.15	0.3043	0.29
19	VGG16	224 X 244	0	0	7	0.1	0.33	0.15	0.3043	0.29
20	VGG19	224 X 244	0	8	3	0.29	0.33	0.31	0.5217	0.46

21	Xception	229 X 229	0	14	0	0.19	0.33	0.25	0.5652	0.58

* Final precision, recall, F1-score and accuracy of the best models with image data 2 are highlighted in bold.

For the dataset of Image Data-2, the models DenseNet201, InceptionResNetV2, InceptionV3, MobileNetV2, ResNet152, VGG19 gives accuracy of 74% respectively.

Table 2
Performance Measures of Pre-Trained Deep Learning Models with input as Image Data 2

S.NO.	Model	Model Image Size	Predicted			Macro Average			Accuracy (Valid)	Accuracy
			MD (14)	ND (14)	VMD (14)	precision	recall	f1-score		
1	AlexNet	224 X 244	5	6	0	0.26	0.25	0.25	0.2857	0.26
2	DenseNet121	224 X 244	11	12	2	0.60	0.58	0.55	0.5952	0.60
3	DenseNet169	224 X 244	1	13	1	0.36	0.29	0.26	0.2857	0.36
4	DenseNet201	224 X 244	11	10	10	0.74	0.74	0.74	0.7380	0.74
5	EfficientNetB0	224 X 244	6	14	0	0.48	0.48	0.41	0.4762	0.48
6	EfficientNetB1	240 X 240	6	14	0	0.48	0.48	0.41	0.4762	0.48
7	EfficientNetB2	260 X 260	7	14	0	0.50	0.48	0.43	0.4762	0.50
8	EfficientNetB3	300 X 300	6	14	0	0.48	0.48	0.41	0.4524	0.48
9	EfficientNetB4	380 X 380	7	14	0	0.50	0.48	0.43	0.4524	0.50
10	EfficientNetB5	456 X 456	6	14	0	0.48	0.48	0.41	0.4524	0.48
11	InceptionResNetV2	229 X 229	11	10	10	0.74	0.74	0.74	0.7380	0.74
12	InceptionV3	229 X 229	11	10	10	0.74	0.74	0.74	0.7380	0.74
13	MobileNet	224 X 244	11	12	2	0.60	0.58	0.55	0.5952	0.60
14	MobileNetV2	224 X 244	11	10	10	0.74	0.74	0.74	0.7380	0.74
15	NASNetMobile	224 X 244	5	6	0	0.26	0.25	0.25	0.3571	0.26
16	ResNet101	224 X 244	1	13	1	0.36	0.29	0.26	0.3571	0.36
17	ResNet152	224 X 244	11	10	10	0.74	0.74	0.74	0.7380	0.74
18	ResNet50	224 X 244	11	12	2	0.60	0.58	0.55	0.5952	0.60
19	VGG16	224 X 244	11	12	2	0.60	0.58	0.55	0.5952	0.60
20	VGG19	224 X 244	11	10	10	0.74	0.74	0.74	0.7380	0.74
21	Xception	229 X 229	11	12	2	0.60	0.58	0.55	0.5952	0.60

* Final precision, recall, F1-score and accuracy of the best models with image data 2 are highlighted in bold.

The models DenseNet201, MobileNetV2, ResNet152, VGG19, NASNetMobile has accuracy of 83% for the Image Data-3.

Table 3
Performance Measures of Pre-Trained Deep Learning Models with input as Image Data 3

S.NO.	Model	Model Image Size	Predicted			Weighted Average			Accuracy (Valid)	Accuracy
			MD (3)	ND (14)	VMD (7)	precision	recall	f1-score		
1	AlexNet	224 X 244	1	7	2	0.42	0.48	0.43	0.3913	0.42
2	DenseNet121	224 X 244	3	7	5	0.63	0.84	0.67	0.5217	0.63
3	DenseNet169	224 X 244	1	14	2	0.71	0.73	0.66	0.6957	0.71
4	DenseNet201	224 X 244	1	14	5	0.83	0.87	0.82	0.8260	0.83
5	EfficientNetB0	224 X 244	1	13	2	0.67	0.70	0.63	0.6521	0.67
6	EfficientNetB1	240 X 240	1	13	2	0.67	0.70	0.63	0.6521	0.67
7	EfficientNetB2	260 X 260	1	13	2	0.67	0.70	0.63	0.6521	0.67
8	EfficientNetB3	300 X 300	1	13	2	0.67	0.70	0.63	0.6521	0.67
9	EfficientNetB4	380 X 380	3	14	2	0.79	0.85	0.75	0.6957	0.79
10	EfficientNetB5	456 X 456	1	13	1	0.63	0.59	0.56	0.5217	0.63
11	InceptionResNetV2	229 X 229	3	12	4	0.79	0.81	0.79	0.6957	0.79
12	InceptionV3	229 X 229	2	13	4	0.79	0.80	0.79	0.6957	0.79

13	MobileNet	224 X 244	2	13	4	0.79	0.79	0.78	0.6957	0.79
14	MobileNetV2	224 X 244	2	13	5	0.83	0.84	0.83	0.8260	0.83
15	NASNetMobile	224 X 244	2	13	5	0.83	0.84	0.83	0.8260	0.83
16	ResNet101	224 X 244	1	7	5	0.54	0.67	0.59	0.3913	0.54
17	ResNet152	224 X 244	1	14	5	0.83	0.87	0.82	0.8260	0.83
18	ResNet50	224 X 244	3	11	5	0.79	0.88	0.81	0.6957	0.79
19	VGG16	224 X 244	1	11	5	0.71	0.76	0.73	0.6957	0.71
20	VGG19	224 X 244	2	11	7	0.83	0.86	0.83	0.8260	0.83
21	Xception	229 X 229	1	13	5	0.79	0.81	0.78	0.6957	0.79

* Final precision, recall, F1-score and accuracy of the best models with image data 3 are highlighted in bold.

The pretrained models' performance is increased with image Data-4 which justifies the applied preprocessing techniques like data augmentation and histogram equalization (CLAHE) for the dataset used in this research.

Table 4
Performance Measures of Pre-Trained Deep Learning Models with input as Image Data 4

S.NO.	Model	Model Image Size	Predicted			Macro Average			Accuracy (Valid)	Accuracy
			MD (14)	ND (14)	VMD (14)	precision	recall	f1-score		
1	AlexNet	224 X 244	11	11	9	0.74	0.74	0.74	0.7143	0.74
2	DenseNet121	224 X 244	12	12	10	0.81	0.81	0.81	0.7857	0.81
3	DenseNet169	224 X 244	12	12	10	0.81	0.81	0.81	0.8095	0.81
4	DenseNet201	224 X 244	13	14	12	0.93	0.94	0.93	0.9047	0.93
5	EfficientNetB0	224 X 244	12	11	9	0.76	0.76	0.76	0.7143	0.76
6	EfficientNetB1	240 X 240	12	11	9	0.76	0.76	0.76	0.7143	0.76
7	EfficientNetB2	260 X 260	12	11	9	0.76	0.76	0.76	0.7143	0.76
8	EfficientNetB3	300 X 300	12	11	9	0.76	0.76	0.76	0.7143	0.76
9	EfficientNetB4	380 X 380	13	11	9	0.79	0.78	0.78	0.7143	0.79
10	EfficientNetB5	456 X 456	12	11	9	0.76	0.76	0.76	0.7143	0.76
11	InceptionResNetV2	229 X 229	12	12	10	0.81	0.81	0.81	0.8095	0.81
12	InceptionV3	229 X 229	13	13	12	0.90	0.91	0.91	0.9047	0.90
13	MobileNet	224 X 244	12	12	10	0.81	0.81	0.81	0.8095	0.81
14	MobileNetV2	224 X 244	13	14	12	0.93	0.94	0.93	0.9047	0.93
15	NASNetMobile	224 X 244	12	12	10	0.81	0.81	0.81	0.8095	0.81
16	ResNet101	224 X 244	12	12	10	0.81	0.81	0.81	0.8095	0.81
17	ResNet152	224 X 244	13	13	12	0.90	0.91	0.91	0.9047	0.90
18	ResNet50	224 X 244	12	12	10	0.81	0.81	0.81	0.8095	0.81
19	VGG16	224 X 244	12	12	10	0.81	0.81	0.81	0.8095	0.81
20	VGG19	224 X 244	13	14	12	0.93	0.94	0.93	0.9047	0.93
21	Xception	229 X 229	12	12	10	0.81	0.81	0.81	0.8095	0.81

* Final precision, recall, F1-score and accuracy of the best models with image data 4 are highlighted in bold.

From Figure 5 it is observed that the accuracy of DenseNet201, MobileNetV2, VGG19, ResNet152 is 93%,93%,93%.90% respectively. The proposed ensemble architectures consist of DenseNet201, MobileNetV2, VGG19, ResNet152 which achieves 0.959 as precision indicating that the correct predictions of stages in dementia.

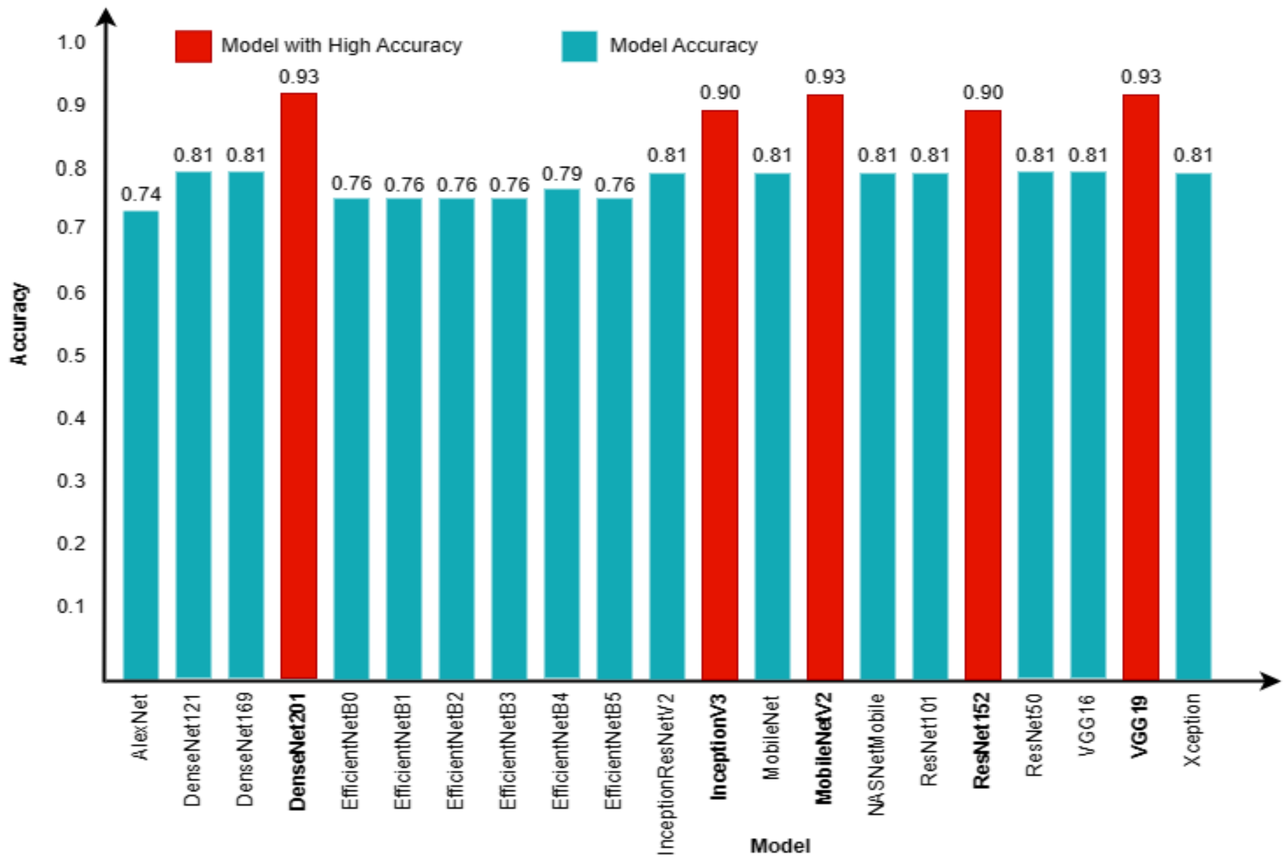


Fig. 5. Comparison of accuracy of pre-trained models with Image Data-4

Various combinations of selected pretrained models is tested for the image data -4 dataset and found that the proposed ensemble architecture which uses four models has high accuracy in classification as given in Table 5.

Table 5
Performance metrices of Proposed Ensemble Architecture (DenseNet201 + MobileNetV2 + VGG19 + ResNet152)

Label	Precision	Recall	F1-Score	Accuracy
VMD	0.93	0.93	0.93	
[Macro Average]	0.93	1	0.96	95%
MD	0.93	1	1	
ND	0.93	0.93	0.93	

From Table 4, the models DenseNet201, MobileNetV2, VGG19, InceptionV3, ResNet152 has highest accuracy compared to other models. InceptionV3 is not included in ensemble approach since the image size 229 x 229 compared to input size of all other models considered for ensemble approach which is 224 x 224.

The confusion Matrix of the proposed ensemble model is given in Figure 6 which clearly indicates that the model has high true positive rates which maximizes the use of the proposed ensemble model to classify the stages in dementia. Figure 6 shows the sample images from each stage from image data-4 used in the proposed model.

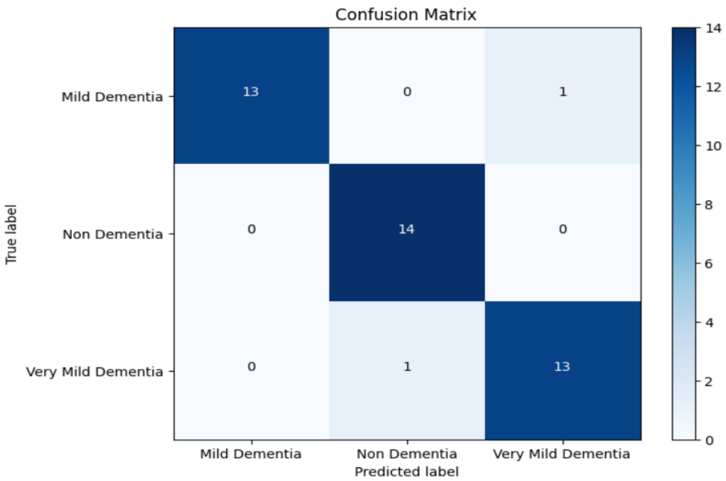
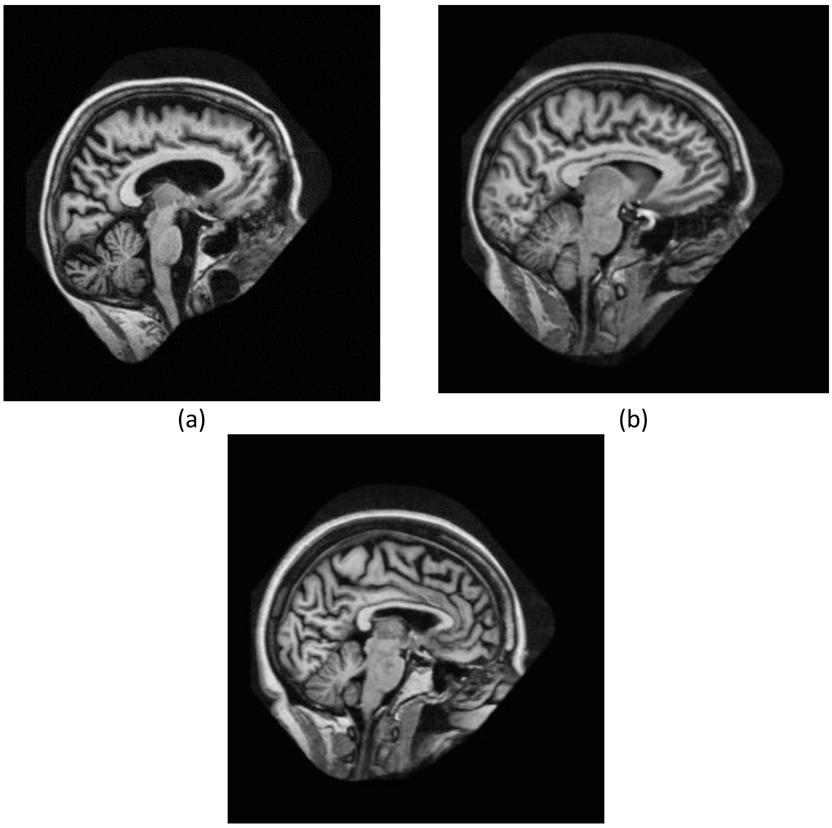


Fig. 6. Confusion Matrix of Proposed Ensemble Architecture (DenseNet201 + MobileNetV2 + VGG19 + ResNet152)



(a) (b)

(c)

Fig. 7. (a) Mild Dementia (b) Non-Dementia (c) Very Mild Dementia

5. Conclusion

In this paper, an ensemble classifier model leveraging pretrained deep learning models is introduced for dementia stage classification. The model's performance is evaluated on a standard brain tumour dataset consisting of MRI images, achieving an impressive accuracy of 96.21%. This proposed model has the potential to aid health professionals and researchers in accelerating the early prediction stages of dementia, which can be valuable for diagnosis and intervention. There are some limitations of the proposed framework, which can be improved in future. The number of images taken for training the dense model is limited which may have impact on the accuracy of the model. In addition, the MRI images are not segmented to concentrate on the portion of brain where there are changes in structure for different stages. In future segmentation may be carried out for the input image and the proposed model may be trained with the output of segmented images for improving performance.

Acknowledgement
This research was not funded by any grant.

References

[1] Prabha, S., K. Sakthidasan@ Sankaran, and D. Chitradevi. "Efficient optimization based thresholding technique for analysis of alzheimer MRIs." *International Journal of Neuroscience* 133, no. 2 (2023): 201-214. https://doi.org/10.1080/00207454.2021.1901696
[2] Alzheimer's Association India. “Alzheimer's & Dementia Help.” https://www.alz.org/in/dementia-alzheimers-en.asp
[3] CDC. “What is Alzheimer's Disease?.” https://www.cdc.gov/aging/alzheimers-disease-dementia/about-alzheimers.html#:~:text=Alzheimer's%20disease%20is%20the%20most,thought%2C%20memory%2C%20and%20language.
[4] Li, Qiang, Yingjian Yang, Yingwei Guo, Wei Li, Yang Liu, Han Liu, and Yan Kang. "Performance evaluation of deep learning classification network for image features." *IEEE Access* 9 (2021): 9318-9333. https://doi.org/10.1109/ACCESS.2020.3048956
[5] Afzal, Sitara, Imran Ullah Khan, and Jong Weon Lee. "A Transfer Learning-Based Approach to Detect Cerebral Microbleeds." *Computers, Materials & Continua* 71, no. 1 (2022). https://doi.org/10.32604/cmc.2022.021930
[6] Lian, Chunfeng, Mingxia Liu, Yongsheng Pan, and Dinggang Shen. "Attention-guided hybrid network for dementia diagnosis with structural MR images." *IEEE transactions on cybernetics* 52, no. 4 (2020): 1992-2003. https://doi.org/10.1109/TCYB.2020.3005859
[7] Jha, Ritesh, Vandana Bhattacharjee, and Abhijit Mustafi. "Transfer Learning with Feature Extraction Modules for Improved Classifier Performance on Medical Image Data." *Scientific Programming* 2022 (2022). https://doi.org/10.1155/2022/4983174
[8] Khan, Mohammad Monirujjaman, Atiyea Sharmeen Omee, Tahia Tazin, Faris A. Almalki, Maha Aljohani, and Haneen Algethami. "A novel approach to predict brain cancerous tumor using transfer learning." *Computational and Mathematical Methods in Medicine* 2022 (2022). https://doi.org/10.1155/2022/2702328
[9] Srinivas, Chetana, Nandini Prasad KS, Mohammed Zakariah, Yousef Ajmi Alothaibi, Kamran Shaukat, B. Partibane, and Halifa Awal. "Deep transfer learning approaches in performance analysis of brain tumor classification using MRI images." *Journal of Healthcare Engineering* 2022 (2022). https://doi.org/10.1155/2022/3264367
[10] Iwata, Hiroaki, Yoshihiro Hayashi, Aki Hasegawa, Kei Terayama, and Yasushi Okuno. "Classification of scanning electron microscope images of pharmaceutical excipients using deep convolutional neural networks with transfer learning." *International Journal of Pharmaceutics: X* 4 (2022): 100135. https://doi.org/10.1016/j.ijpx.2022.100135
[11] Murugan, Suriya, Chandran Venkatesan, M. G. Sumithra, Xiao-Zhi Gao, B. Elakkiya, Muthuramalingam Akila, and S. Manoharan. "DEMNET: a deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images." *Ieee Access* 9 (2021): 90319-90329. https://doi.org/10.1109/ACCESS.2021.3090474
[12] Valverde, Juan Miguel, Vandad Imani, Ali Abdollahzadeh, Riccardo De Feo, Mithilesh Prakash, Robert Ciszek, and Jussi Tohka. "Transfer learning in magnetic resonance brain imaging: A systematic review." *Journal of imaging* 7, no. 4 (2021): 66. https://doi.org/10.3390/jimaging7040066
[13] Wang, Hongfei, Yanyan Shen, Shuqiang Wang, Tengfei Xiao, Liming Deng, Xiangyu Wang, and Xinyan Zhao. "Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer’s disease." *Neurocomputing* 333 (2019): 145-156. https://doi.org/10.1016/j.neucom.2018.12.018
[14] Li, Hejie, Ying Tan, Jiaqing Miao, Ping Liang, Jinnan Gong, Hui He, Yuhong Jiao, Fan Zhang, Yaolin Xing, and Donghan Wu. "Attention-based and micro designed EfficientNetB2 for diagnosis of Alzheimer’s disease." *Biomedical Signal Processing and Control* 82 (2023): 104571. https://doi.org/10.1016/j.bspc.2023.104571
[15] Al Shehri, Waleed. "Alzheimer’s disease diagnosis and classification using deep learning techniques." *PeerJ Computer Science* 8 (2022): e1177. https://doi.org/10.7717/peerj-cs.1177
[16] Fulton, Lawrence V., Diane Dolezel, Jordan Harrop, Yan Yan, and Christopher P. Fulton. "Classification of Alzheimer’s disease with and without imagery using gradient boosted machines and ResNet-50." *Brain sciences* 9, no. 9 (2019): 212. https://doi.org/10.3390/brainsci9090212

[17] Carcagnì, Pierluigi, Marco Leo, Marco Del Coco, Cosimo Distante, and Andrea De Salve. "Convolution Neural Networks and Self-Attention Learners for Alzheimer Dementia Diagnosis from Brain MRI." *Sensors* 23, no. 3 (2023): 1694. https://doi.org/10.3390/s23031694

[18] Subramoniam, Manu, T. R. Aparna, P. R. Anurenjan, and K. G. Sreeni. "Deep learning-based prediction of Alzheimer's disease from magnetic resonance images." In *Intelligent vision in healthcare*, pp. 145-151. Singapore: Springer Nature Singapore, 2022. https://doi.org/10.1007/978-981-16-7771-7_12

[19] Aulia, Suci, and Dadi Rahmat. "Brain Tumor Identification Based on VGG-16 Architecture and CLAHE Method." *JOIV: International Journal on Informatics Visualization* 6, no. 1 (2022): 96-102. https://doi.org/10.30630/joiv.6.1.864

[20] Ullah, Zahid, Muhammad Umar Farooq, Su-Hyun Lee, and Donghyeok An. "A hybrid image enhancement based brain MRI images classification technique." *Medical hypotheses* 143 (2020): 109922. https://doi.org/10.1016/j.mehy.2020.109922

[21] Asif, Sohaib, Ming Zhao, Fengxiao Tang, and Yusen Zhu. "An enhanced deep learning method for multi-class brain tumor classification using deep transfer learning." *Multimedia Tools and Applications* (2023): 1-28. https://doi.org/10.1007/s11042-023-14828-w

[22] Siddhartha, Manu, and Avik Santra. "COVIDLite: A depth-wise separable deep neural network with white balance and CLAHE for detection of COVID-19." *arXiv preprint arXiv:2006.13873* (2020).

[23] Alwakid, Ghadah, Walaa Gouda, and Mamoona Humayun. "Deep Learning-based prediction of Diabetic Retinopathy using CLAHE and ESRGAN for Enhancement." In *Healthcare*, vol. 11, no. 6, p. 863. MDPI, 2023. https://doi.org/10.3390/healthcare11060863

[24] Pamula Pullaiah, Nagaraja Rao, Dorai Venkatasekhar, Padarthi Venkatramana, and Balaraj Sudhakar. "Detection of Breast Cancer on Magnetic Resonance Imaging Using Hybrid Feature Extraction and Deep Neural Network Techniques." *International Journal of Intelligent Engineering & Systems* 13, no. 6 (2020). https://doi.org/10.22266/ijies2020.1231.21

[25] Chakraverti, Sugandha, Pankaj Agarwal, Himansu Sekhar Pattanayak, Sanjay Pratap Singh Chauhan, Ashish Kumar Chakraverti, and Manoj Kumar. "De-noising the image using DBST-LCM-CLAHE: A deep learning approach." *Multimedia Tools and Applications* (2023): 1-26. https://doi.org/10.1007/s11042-023-16016-2

[26] Rajaraman, Sivaramakrishnan, Jenifer Siegelman, Philip O. Alderson, Lucas S. Folio, Les R. Folio, and Sameer K. Antani. "Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays." *Ieee Access* 8 (2020): 115041-115050. https://doi.org/10.1109/ACCESS.2020.3003810

[27] Abayomi-Alli, Adebayo, Olabode Atinuke, Saidat Adebukola Onashoga, Sanjay Misra, and Olusola Abayomi-Alli. "Facial image quality assessment using an ensemble of pre-trained deep learning models (EFQnet)." In *2020 20th international conference on computational science and its applications (ICCSA)*, pp. 1-8. IEEE, 2020. https://doi.org/10.1109/ICCSA50381.2020.00013

[28] Kang, Jaeyong, Zahid Ullah, and Jeonghwan Gwak. "Mri-based brain tumor classification using ensemble of deep features and machine learning classifiers." *Sensors* 21, no. 6 (2021): 2222. https://doi.org/10.3390/s21062222

[29] Tanveer, M. "Random vector functional link neural network based ensemble deep learning." (2021).

[30] Vallabhajosyula, Sasikala, Venkatramaphanikumar Sistla, and Venkata Krishna Kishore Kolli. "Transfer learning-based deep ensemble neural network for plant leaf disease detection." *Journal of Plant Diseases and Protection* 129, no. 3 (2022): 545-558. https://doi.org/10.1007/s41348-021-00465-8

[31] Gianchandani, Neha, Aayush Jaiswal, Dilbag Singh, Vijay Kumar, and Manjit Kaur. "Rapid COVID-19 diagnosis using ensemble deep transfer learning models from chest radiographic images." *Journal of ambient intelligence and humanized computing* (2020): 1-13. https://doi.org/10.1007/s12652-020-02669-6

[32] Bae, Jinhyeong, Jane Stocks, Ashley Heywood, Youngmoon Jung, Lisanne Jenkins, Virginia Hill, Aggelos Katsaggelos *et al.,* "Transfer learning for predicting conversion from mild cognitive impairment to dementia of Alzheimer's type based on a three-dimensional convolutional neural network." *Neurobiology of aging* 99 (2021): 53-64. https://doi.org/10.1016/j.neurobiolaging.2020.12.005

[33] Hussain, Emtiaz, Mahmudul Hasan, Syed Zafrul Hassan, Tanzina Hassan Azmi, Md Anisur Rahman, and Mohammad Zavid Parvez. "Deep learning based binary classification for alzheimer's disease detection using brain mri images." In *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 1115-1120. IEEE, 2020. https://doi.org/10.1109/ICIEA48937.2020.9248213

[34] Savaş, Serkan. "Detecting the stages of Alzheimer's disease with pre-trained deep learning architectures." *Arabian Journal for Science and Engineering* 47, no. 2 (2022): 2201-2218. https://doi.org/10.1007/s13369-021-06131-3

[35] Herzog, Nitsa J., and George D. Magoulas. "Deep learning of brain asymmetry images and transfer learning for early diagnosis of dementia." In *International Conference on Engineering Applications of Neural Networks*, pp. 57-70. Cham: Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-80568-5_5

[36] Duong, Silvia, Tejal Patel, and Feng Chang. "Dementia: What pharmacists need to know." *Canadian Pharmacists Journal/Revue des Pharmaciens du Canada* 150, no. 2 (2017): 118-129. https://doi.org/10.1177/1715163517690745

[37] Lövdén, Martin, Weili Xu, and Hui-Xin Wang. "Lifestyle change and the prevention of cognitive decline and dementia: what is the evidence?." *Current opinion in psychiatry* 26, no. 3 (2013): 239-243. https://doi.org/10.1097/YCO.0b013e32835f4135