Improving Alzheimer's Disease Detection with Transfer Learning

Showkat A. Dar[1], Aafaq A. Rather[2], Mustafa Ibrahim Ahmed Araibi[3,*], I. Elbatal[4], Ehab M. Almetwally[4], Ahmed M. Gemeay[5], Sharvari R. Shukla[2], Faizan Danish[6] and Qaiser Farooq Dar[7]

[1]*Department of Computer Science and Engineering, GITAM University Bangalore Campus, India*

[2]*Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune, India*

[3]*Department of Business Administration, College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia*

[4]*Department of Mathematics and Statistics, College of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia*

[5]*Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt*

[6]*Department of Mathematics, School of Advanced Sciences, Vellore Institute of Technology-Andhra Pradesh (VIT- AP) University, Inavolu, Beside AP Secretariat, Amaravati AP-522237, India*

[7]*Department of Health Research, ICMR-National Institute of Virology Pune, North Zone Jammu-180001, J&K, India*

Abstract: Accurate and prompt diagnosis of Alzheimer's disease (AD) remains a challenge, with only a small percentage of patients receiving timely confirmation. Manual interpretation of MRI scans, the primary diagnostic tool, is time-consuming, subjective, and prone to error, particularly in differentiating between disease stages. This study aimed to develop a computer-aided diagnosis system (CAD) for AD classification using deep learning models. MobileNetV1 and Xception architectures were employed to classify AD into four stages: mild, normal, moderate, and severe. Transfer learning and layer freezing techniques were applied for feature extraction and classification. Model performance was evaluated using precision, recall score, and accuracy metrics. The Xception model achieved a higher accuracy (79%) compared to MobileNetV1 (73%) in classifying AD stages. Compared to MobileNetV1, this study shows that Xception-based CAD systems have the potential to diagnose AD more accurately, providing a promising path for future research and clinical application.

Keywords: Mobile NetV1, Alzheimer's Disease, Xception, Transfer Learning, MRI image.

1. INTRODUCTION

The growing population is making problems worse. The high death rate has gotten worse because there aren't enough medical professionals. It can be hard for doctors to figure out if a patient has a certain disease [1]. Many neurological diseases and delineating pathological regions have been analyzed, and the brain's anatomical structure has been researched with magnetic resonance imaging (MRI). The most prevalent form of dementia [2], is a progressive neurodegenerative condition that affects older populations [3]. Alzheimer's disease was first discovered in 1907 by a German scientist, Dr. Alois Alzheimer termed it a "neurodegenerative disease" [4]. continuous degeneration of the neural system, such as restlessness, irritability, disorientation, depression, and anxiety [5]. Neurodegenerative diseases are caused by the progressive degeneration of nerve cells [6]. Almost 50 million individuals worldwide are suffering from AD, and if no treatment or prevention measures are found soon, the number will rise significantly to 150 million by 2050. It is important to identify patients with Alzheimer's disease (AD) early so that preventative measures can be taken [7]. Alzheimer's Disease leads to mental disorders [8]. Mental disorders are neural issues that influence brain cognition and social connectivity.

Abundant medical image data is generated daily. The information available in such data is used for education, diagnosis and research. Automatic classification methods are commonly used for the analysis of neuroimaging studies. Several multiple-resolution methods have been proposed to use neighborhood information to detect significant changes in brain volume. Various computer-aided methods [9] have been projected to categorize brain images. Many computer-aided diagnoses [10] systems were developed using image processing & computational intelligence techniques. These systems usually have some steps, such as preprocessing, segmentation, feature extraction, and classification. Feature extraction and classification are the steps that best define the diagnosis performed by computer-aided diagnosis systems [11]. Several primary AI-aided applications, including healthcare, manufacturing, smart cities, and gaming, are being studied to be

*Address correspondence to this author at the Department of Business Administration, College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia;
Email: miarbi@imamu.edu.sa

promising deployment in virtual worlds [12]. AI has been presented as a technology that could revolutionize medical practices by improving decision-making in intricate and unpredictable systems [13]. The development of artificial intelligence enables the evaluation and detection process using computer vision now faster and more stable than the human eye. The disadvantages and limitations of human experts make it attractive to use technologies such as artificial intelligence [14], which learns very quickly and provides consistent responses to complex situations [15]. Innovative deep-learning techniques are used to compute the detection of brain disease [16]. To end the promise of early diagnosis of Alzheimer's Disease, we propose a novel model based on deep learning [17], which is a subset of machine learning that tries to mimic the human brain to provide a cheap, fast, yet accurate solution [18]. In order to validate further the advantages of deep learning, future extensions of this work could incorporate non-deep-learning baselines such as SVMs using handcrafted radiomic features. Figure **1** shows the sample images of various stages of Alzheimer's disease.

In more cases, the human eye becomes less sensitive when interpreting many images, especially when only a few slices are impacted. Through automated processes, a large number of cases can be processed with the same precision as, that is, results will not be affected by fatigue, data overload, or lack of manual steps [19]. Therefore, an automated system is needed to investigate and classify such medical images. Brain MRI is widely used to diagnose brain diseases [20]. Programmed classification of brain disorders using brain MRI images [21], how to evaluate clinical prognosis, and query image repositories using content-based queries could instruct radiologists

The literature review of some research articles is listed below with objectives, methods, limitations, and accuracy.

While previous studies utilized standard CNNs, our study contributes by evaluating the impact of layer freezing strategies under transfer learning, which is less explored in AD detection.

This research has been developed based on the following research questions.

- Can computers predict disease?
- What are the types of brain disorders?
- Can a deep learning model predict Alzheimer's disease using MRI images?
- How to implement transfer learning in deep learning models?
- How does this work support neurospecialists?

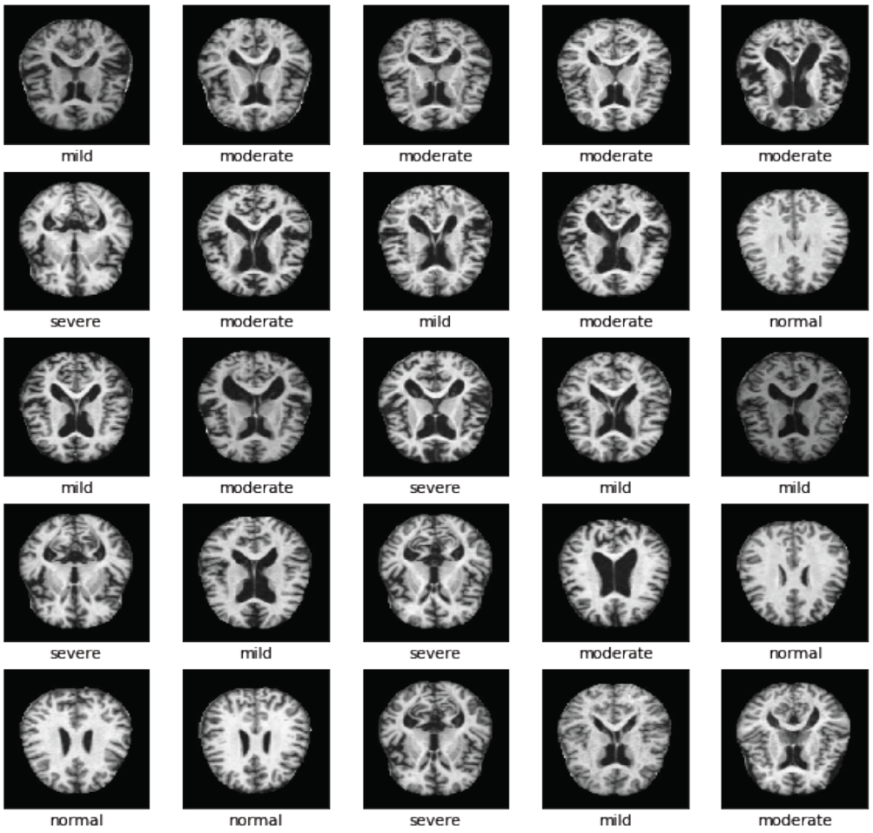

Figure 1: Sample of human brain MRI showing different stages of AD.

Table 1: Literature Review

Ref	Objectives	Methods	Limitations	Accuracy
1	Classifying AD Stages	Multi-layer feedforward neural network	Limited dataset size	93%
2	To automate the classification of Alzheimer's disease and mild cognitive impairment (MCI) using a single MRI and deep neural networks	Utilized deep neural networks for classification	Limited to a single MRI modality	-
3	To investigate the use of Convolutional Neural Networks (CNNs) for medical image analysis and compare full training with fine-tuning approaches	Utilized CNNs for image analysis	Performance may depend on data characteristics	-
4	To explore Alzheimer's disease classification using Convolutional Neural Networks (CNNs)	Utilized CNNs for classification	Dataset size and diversity may impact results	-
5	Lung nodules localization from CT scans	Novel Machine Learning Approach	Limited sample size, dataset diversity	84.8%
6	To classify Alzheimer's disease using MRI and fMRI data with deep convolutional neural networks	Employed deep CNNs for classification	Dependency on data quality and preprocessing	97.79%
7	Detection of COVID-19 from chest X-ray images	Deep Learning Models Resnet 50, Inception V3, VGG16	Limited dataset size, generalization challenges	Resnet 50-97%, InceptionV3-98%, VGG16-98%
8	To classify Alzheimer's disease using a transfer learning approach	Employed transfer learning techniques using CNN	Dependence on the source domain for transfer	97.84%
9	To use deep learning for diagnostic classification and prognostic prediction in Alzheimer's disease using neuroimaging data	Employed deep learning techniques on neuroimaging data	Data availability and quality may affect results	-
10	To explore the use of machine learning and deep learning for skin cancer detection	Employed machine learning and deep learning techniques	Data quality and size may impact performance	-
11	To diagnose Alzheimer's disease using deep learning techniques	Utilized deep learning methods for diagnosis	Data availability and quality may impact results	-
12	Thorough analysis of deep learning techniques for AD early detection	CNN	Dependence on data quality and quantity	-
13	Develop an explainable deep-learning model for Alzheimer's disease diagnosis using PET and MRI images	CNN, Multimodal data fusion on PET and MRI images	Dependence on data quality and quantity	73.90%
14	To perform multi-class motor imagery EEG classification using Convolutional Neural Networks (CNNs)	Employed CNNs for EEG classification	Data variability and subject-specific patterns may affect performance	74%
15	To assess the fetus's health status using a hybrid deep learning algorithm (AlexNet-SVM) on cardiotocographic data	Utilized a hybrid deep learning approachRCNN	Data quality and availability	99%

The practical implication of this work is given below:

The finding of this study has practical implications in the field of medicine. This study is more helpful to neurologists and people with brain disorders. Nowadays, the number of neurologists is very low in rural areas. The early detection of AD is very important. Computer-aided diagnosis is more helpful for early disease detection and will save people from severe problems. AI technology is implemented in every field, especially in medical diagnosis. It analyzes huge MRI data and gives predictions consistently. The result will be accurate, and it is most helpful to medical professionals to find a rapid and accurate solution for the patients.

Another implication of this research is helpful for researchers who are doing research in the field of deep learning and also in Alzheimer's disease. This research provides knowledge of transfer learning in deep learning models.

The objectives of this study are summarized as follows:

- To implement MobileNetV1 CNN architecture to classify the brain MRI images into various stages of Alzheimer's disease, such as mild, moderate, normal and severe.
- To implement the novelty of the transfer learning technique that is freezing layers on the MobileNetV1 architecture. The layers freeze from the bottom up to four layers.
- Xception architecture layers are frozen from the bottom to achieve better accuracy.
- The Xception and MobileNetV1 performance are compared.

Section 1 contains the introduction and the literature survey, and also describes the framework for the proposed work. Section 2 describes the transfer learning techniques and also the architecture of transfer learning. Section 3 explained the MobilenetV1 architecture, and Section 4 describes the Xception architecture. In Section 5,the implementation of freezing of layers and performance of the models for Alzheimer's disease are discussed. Section 6 discusses about conclusion and future scope. The framework of the proposed diagram is shown in Figure **2**.

2. TRANSFER LEARNING

Creating a visual architecture diagram for transfer learning can be a bit challenging. In Transfer learning, a pre-trained neural network, often referred to as the "base model" or "source model," is adapted for a new task or domain. The architecture of a transfer learning model typically consists of two main parts: the pre-trained layers (base layers) and the task-specific layers (top layers) [22] where a pre-trained neural network, often referred to as the "base model" or "source model," is adapted for a new task or domain. The architecture of a transfer learning model typically consists of two main parts: the pre-trained layers (base layers) and the task-specific layers (top layers).

2.1. Pre-trained Layers (Base Layers):

- **Convolutional Base:** This part of the network consists of convolutional layers, pooling layers, and sometimes normalization layers. These layers have learned to extract hierarchical features from images and are responsible for capturing general patterns, textures, and shapes.
- **Feature Extractor:** The pre-trained layers act as a feature extractor by transforming the input data (e.g., images) into a more abstract and compact representation. These layers retain valuable knowledge learned from a large dataset during the original training process.
- **Transfer Learning Architectures:** Common pre-trained architectures include VGG, ResNet, Inception, MobileNet, and more. These architectures have different depths and complexities, depending on the specific use case and resource constraints.

2.2. Task-specific Layers (Top Layers):

- **Fully Connected Layers (Dense Layers):** These layers are added on top of the pre-trained base layers to adapt the network to the specific task. They can consist of one or more dense layers. The number of neurons in the final dense layer usually matches the number of classes in the new task (for classification tasks).
- **Activation Functions:** Activation functions like ReLU (Rectified Linear Unit) are applied to the outputs of dense layers to introduce non-linearity and enable the network to learn complex relationships.
- **Output Layer:** The output layer depends on the nature of the task. For example:
  - For image classification, the output layer typically consists of a softmax activation function to produce class probabilities.
  - For regression tasks, a single neuron without an activation function may be used to predict a continuous value.

3. MOBILE NETV1

MobileNets are based on depth-wise separable convolution layers [23]. Each layer of convolution with depthwise separable consists of a depth wise convolution and a pointwise convolution [24]. The 28 convolutional layers in a MobileNetV1 are counted independently in both depth and point wise directions. By properly adjusting the width multiplier hyperparameter, the 4.2 million parameters in a traditional MobileNetV1 can be reduced. The size of the provided image is 224 × 224 × 3. A standard

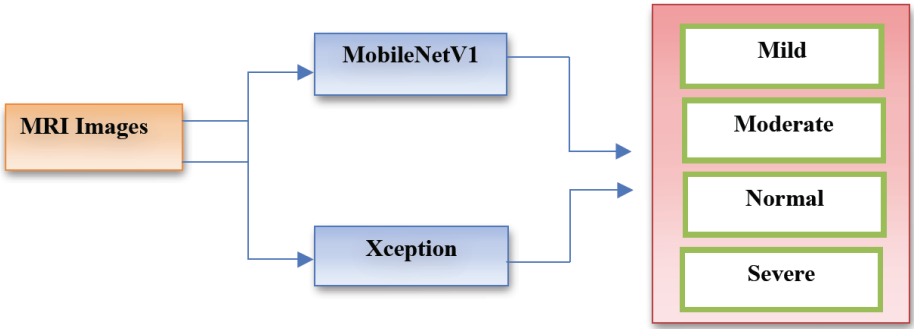

Figure 2: Framework of the Proposed Classification Model for AD.

convolution unit, a depth-wise separate convolution unit, a width multiplier, and a resolution multiplier make up this system. Figure **3** depicts the MobileNetV1 architecture [25].

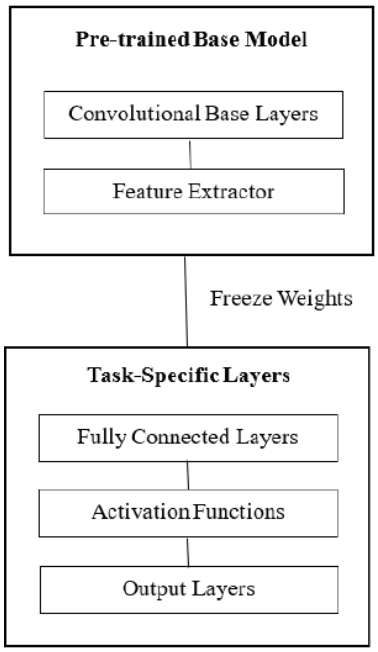

Figure 3: Architectural Diagram for Transfer Learning.

Depth-wise Separable Convolution is used in this MobileNetV1 architecture to reduce the complexity and size of the model. It should be mentioned that Batch Normalization (BN) and ReLU come after each depth-wise convolution [26]. A depth-wise separable convolution is a depth-wise convolution that has been tailed by a pointwise convolution, as shown in Figure **4**.

4. XCEPTION

It is possible to think of Xception as an extreme version of the Inception architecture that uses depth-wise separated convolution. Xception widened the first inception block to control computation complexity and swapped out for a single dimension (3x3) and a 1x1 convolution. Figure **5** depicts the architecture of the Xception block. It initially utilizes 1x1 convolutions to translate the convolved output to low-dimensional embeddings. It then undergoes n spatial transformations, where n is a cardinality specifying the width and represents the number of transformations.

$$f_{l+1}^{K}(p,q) = \sum_{xy} f_{l}^{k}(x,y) \cdot e_{l}^{k}(u,v) \tag{1}$$

$$F_{l+2}^{k} = g_{c}(F_{l+1}^{k}, K_{l+1}) \tag{2}$$

Kl is a k[th] kernel of the l[th] layer with depth one that is spatially convolved across the k[th] feature map in equation (2). F_{l}^{k}, where the spatial indices of the feature map and kernel are represented by (x, y) and (u, v), respectively. It should be emphasized that in depth-separable convolution, the number of input feature maps is denoted as K, as opposed to traditional convolutional layers, where K is independent of the feature maps from the preceding layer.While the l+1[th] layer, which conducts depth wise convolution across the output feature, uses the (k l+1)[th] kernel of the (1x1) spatial feature-maps $[F_{l+1}^{1},\ldots,F_{l+1}^{k}]$ of l[th] layer, used as input of l+1[th] layer.

Xception simplifies the computation by doing independent cross-axes convolutions on each featuremap before performing pointwise convolutions (1x1 convolutions) to conduct cross-channel correlation. CNN architectures only use one transformation segment for the convolutional operation and three for the inception block. The transformation technique Xception uses improves performance while increasing learning efficiency, even if it does not decrease the total number of parameters [27].

Initially, the data passes over the entry flow, then completes the middle flow, where it replications itself eight times, and to the end, concludes the exit flow [28].

5. FREEZING OF LAYERS

To study the influence of layer freezing, ablation studies were conducted. These involved progressively

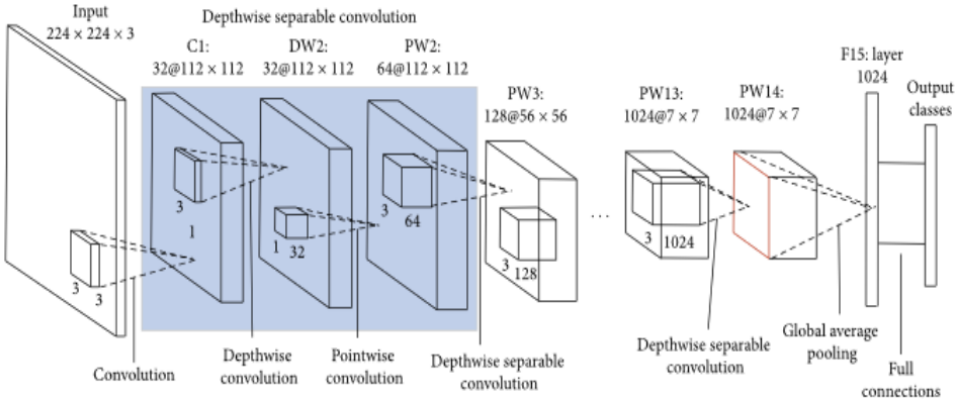

Figure 4: Architecture of MobileNetV1.

unfreezing the last 1 to 4 layers and comparing model performance. Xception responded well to full training, while MobileNetV1 performed optimally when the last three layers were trainable.

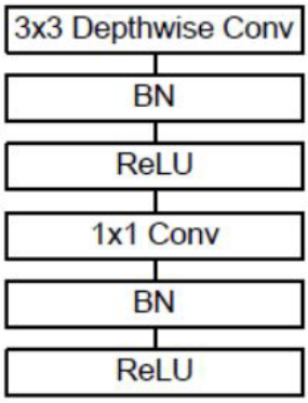

Figure 5: Working of Depth-wise Separable Convolution.

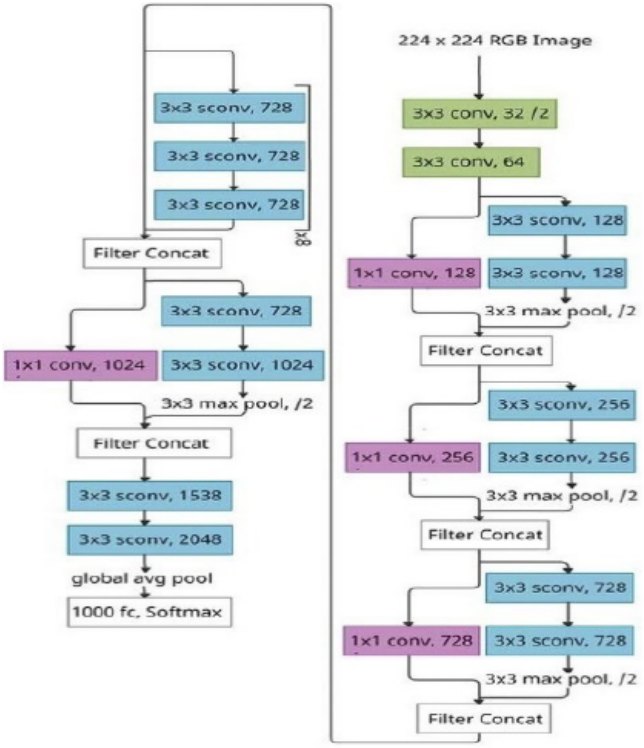

Figure 6: Architecture of Xception Model.

Freezing of MobileNetV1 Layers

In MobileNetV1, the last four layers play a crucial role in the final feature extraction and classification stages. MobileNetV1 performs image classification tasks; the layers are responsible for producing the final class probabilities. Here's a breakdown of the role of the last four layers in MobileNetV1:

1. **Global Average Pooling (Global_avg_pool):** This layer comes after the final depth-wise separable convolution layer. Global average pooling preserves crucial channel-wise information while shrinking the feature maps' spatial dimensions to 1x1. A feature vector summarizing the most notable aspects of the entire image is produced by this operation.

2. **Fully Connected Layer (FC)**: The Global Average Pooling layer is followed by a fully connected layer, sometimes referred to as the "FC" layer. This layer maps the feature vector obtained from global average pooling to the final class scores or logits. The number of units (neurons) in this layer typically corresponds to the number of classes in the classification task.

3. **Normalized function**: Finally, the output of FC is passed to a normalized activation function. This function ensures that the class probabilities sum to one across all classes.

4. **Output**: The final layer in MobileNetV1 is the output layer, which consists of class probabilities. The network produces a vector of probabilities, with each element corresponding to a different class. The class with the highest probability is considered the predicted class for the input image during inference.

These last four layers collectively form the final classification and prediction stage of MobileNetV1, where the network takes the feature representations learned from the earlier layers and transforms them into class probabilities for image classification. During training, the network learns to adjust its parameters, including those in these final layers, to minimize the classification error and make accurate predictions on the given dataset.

In the proposed work, tuningis performed in five ways. Initially, keep all layers static except the dense layer. Then, make all but the last two layers non-trainable. Similarly, keep all layers frozen except the final three. By tuning in different ways, trainable parameters and performance vary. The total parameters of MobileNetV1are 3,228,864, without adding the dense layer, 3,206,976 are trainable, and 21,888 are non-trainable. Once the dense layer is added, the Alzheimer's MRI images are classified into four classes Viz: mild, moderate, normal, and severe, where the parameters are increased by 200708. The total parameters are 3,429,572, the trainable parameters are 3,407,684, and 21,888 are non-trainable. The parameters after the dense layer are shown in Table **1**.

5.1. Freezing of Xception Layers

Xception (short for "Extreme Inception") is a deep convolutional neural network architecture that was

introduced as a variation of the Inception architecture. It has a distinctive structure, and its last four layers play important roles in the final feature extraction and classification stages. Here's a breakdown of the role of the last four layers in the Xception architecture:

1. **Entry Flow**: Before discussing the last five layers, it's important to mention that Xception is organized into two main parts: the Entry Flow and the Middle Flow. The Entry Flow consists of initial layers that perform feature extraction and dimensionality reduction.

2. **Middle Flow**: The Middle Flow is a repeated block of convolutional layers that preserves spatial dimensions and allows for feature extraction across different scales. **Exit Flow**:

a. **Global Average Pooling (GAP)**: After the last separable convolution block in the Exit Flow, Xception employs a global average pooling layer. Global average pooling reduces the spatial dimensions of the feature maps to a 1x1 spatial size while retaining important channel-wise information. This operation generates a feature vector that summarizes the most salient features across the entire image.

b. **Fully Connected Layer (FC)**: The output of the global average pooling layer is connected to a fully connected layer (FC). The number of units (neurons) in this layer typically corresponds to the number of classes in the classification task.

c. **Softmax Activation**: The FC layer is followed by a softmax activation function. This activation function converts the class scores into class probabilities. Each probability represents the likelihood of the input image belonging to a particular class. The softmax function ensures that the class probabilities sum to one across all classes.

d. **Output**: The final layer in Xception is the output layer, which consists of class probabilities. The network produces a vector of probabilities, with each element corresponding to a different class. The class with the highest probability is considered the predicted class for the input image during inference.

In summary, the last four layers in the Xception architecture, located in the Exit Flow, are responsible for final feature extraction, spatial reduction, and classification. They play a crucial role in transforming

Table 2: Parameters of MobileNetV1

Freezing Layer Except	Total Parameters	Trainable Parameters	Non-Trainable Parameters
Last one layer	3,429,572	202,756	3,226,816
Last two layer	3,429,572	202,756	3,226,816
Last three-layer	3,429,572	1,251,332	2,178,240
Last four layer	3,429,572	1,251,332	2,178,240

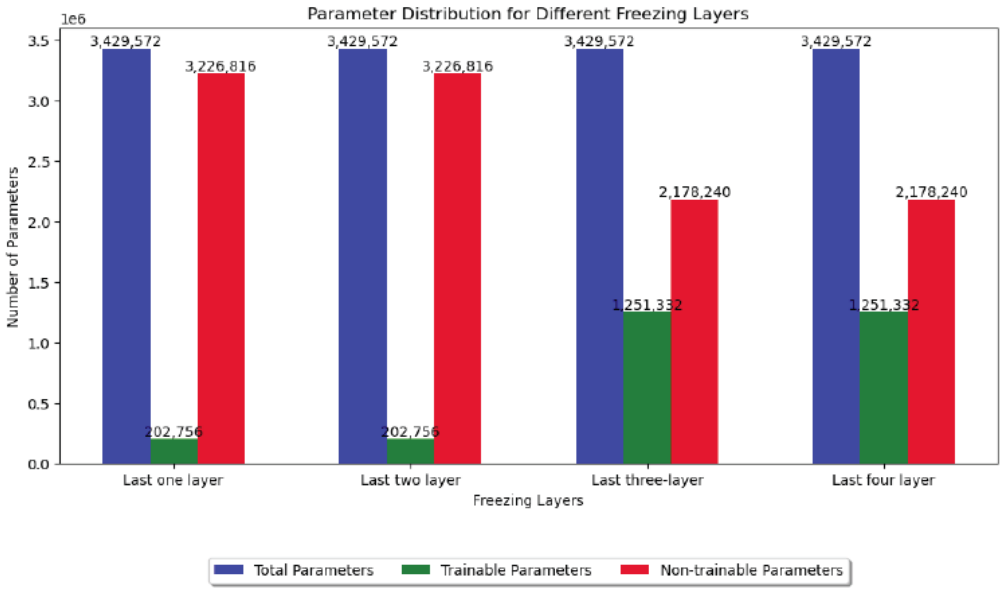

Figure 7: Parameter Distribution for Xception.

Table 3: Parameter of Xception Model

Freezing Layer Except	Total Parameters	Trainable Parameters	Non-Trainable Parameters
Last one layer	20,869,676	8,196	20,861,480
Last two layer	20,869,676	8,196	20,861,480
Last three-layer	20,869,676	8,196	20,861,480
Last four layer	20,869,676	12,292	20,857,384

the learned feature representations into class probabilities for image classification tasks.

In this proposed work, five kinds of tuning are done. Initially, freeze entire layers except for the dense layer. Secondly, freeze all layers except the final two. Like the last four layers, freeze all layers except the last three. Trainable parameters and performance continue to fluctuate as a result of various tuning techniques. The total parameters of Xception are 20,861,480 without adding the dense layer; the trainable parameters are 20,806,952, and the non-trainable parameters are 54,528. After adding the dense layer, the Alzheimer's MRI images are classified into four types of classes: mild, moderate, normal and Normal, where the parameters are increased by 8196. The total parameters are 20,869,676 and 20,815,148 are trainable, and 54,528 are non-trainable. Four classes of Alzheimer's images are created after the dense layer is added: mild, moderate, normal, and severe. Analysis is done on the trainable and non-trainable parameters. The parameter is shown in Table **3**.

5.2. Dataset Description

The dataset was composed of Kaggle online database. Anoverall of 7679 MRI images were collected in which 6400 (896 Moderate, 64 Severe, 3200 Normal, 2240 Mild) were used for training and 1279 (896 Moderate, 12 Severe, 60 Normal, 448 Mild) were used for Validation. The models were trained with a batch size of 32, learning rate of 0.0001, and Adam optimizer. Each model was run for 3 epochs on a Tesla V100 GPU with 16 GB memory. Hyperparameters were selected via grid search. Given the significant class imbalance (e.g., only 64 severe cases), class weighting was applied during training. We also experimented with SMOTE for balancing minority classes, which showed modest improvements in recall for underrepresented categories. Per-class sensitivity and specificity metrics were calculated to assess class-wise performance.

5.3. Performance Measures

The Accuracy, Precision, Recall, and F-score are evaluation parameters utilized in this work to assess performance. These measurements are established using the confusion matrix obtained from the classification process's results [29].

Accuracy

The quality of measurement that produces genuine (no systematic mistakes) and consistent (no random errors) results is known as a measuring system's accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

Precision

The number of TP (i.e., the total number of objects properly categorized as fitting to the positive class) divided by the total number of objects categorized as fitting to the positive class in classification work determines the precision for the class (i.e. the summation of true positives and false positives, which are objects in accurately categorized as fitting to the class).

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Recall

The division of the total number of TP to the total number of components that really fall into the positive class is known as recall (i.e., the summation of true positives and false negatives, which are objects which are not categorized as fitting to the positive class but ought to have been).

$$Recall = \frac{TP}{TP + FP} \tag{5}$$

F-Measure

In order to calculate the score, the F-Measure accuracy metric considers both the precision and recall of the test (Harmonic mean).

$$F - Measure = 2\frac{Precision * Recall}{Precision + Recall} \tag{6}$$

5.3. Performance

The higher Performance based on the comparison of the classification of Alzheimer's disease using

Xception and MobileNetV1, shown in Table **3** and Table **4** are taken as the accuracy. The Xception Pre-trained model provides an accuracy of 79%, and MobileNetV1 provides an accuracy of 62%. The deep learning models [30] are trained with 6400 samples and validated with 1279 samples for three epochs. The transfer learning is applied here to freeze the layers in Xception and MobileNetV1 architecture [31]. First, all layers in the architecture are trained, and then the last layer is trained; all other layers are frozen.

The Xception provides higher accuracy when it's trained by all the layers than when freezing layers. The Mobilentv1 [32] provides higher accuracy when trained by all the layers. The Xception model performed with an accuracy of 79%, whereas the MobileNetV1 [33]

Table 4: Performance for Classification of Alzheimer's Disease using Xception

Freezing of layers except in Xception	Alzheimer level	Precision (in %)	Recall (in %)	F- Score (in %)	Accuracy (in %)
All layers	Moderate	76.01	71.11	73.08	79.34
	Severe	88.21	81.42	84.11	
	Normal	74.03	67.00	70.44	
	Mild	68.32	89.23	77.45	
Last one layer	Moderate	83.00	93.11	88.12	73.14
	Severe	56.12	78.21	65.37	
	Normal	95.56	55.41	70.56	
	Mild	77.21	62.11	69.09	
Last two layers	Moderate	51.00	43.00	47.00	72.28
	Severe	62.12	71.29	70.28	
	Normal	34.26	30.00	32.11	
	Mild	42.28	48.35	44.19	
Last three layers	Moderate	52.32	46.17	49.51	74.34
	Severe	68.00	77.24	72.53	
	Normal	32.65	28.21	30.19	
	Mild	42.41	47.15	44.19	
Last four layers	Moderate	80.19	81.11	83.21	73.54
	Severe	75.26	84.22	80.27	
	Normal	80.11	65.40	74.21	
	Mild	70.17	70.15	69.21	

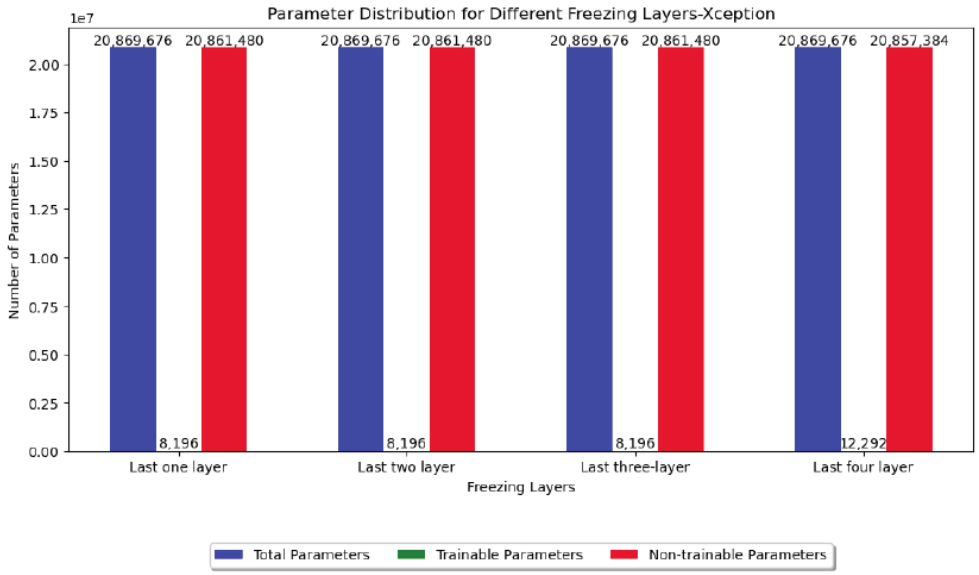

Figure 8: Parameter Distribution for MobileNetV1.

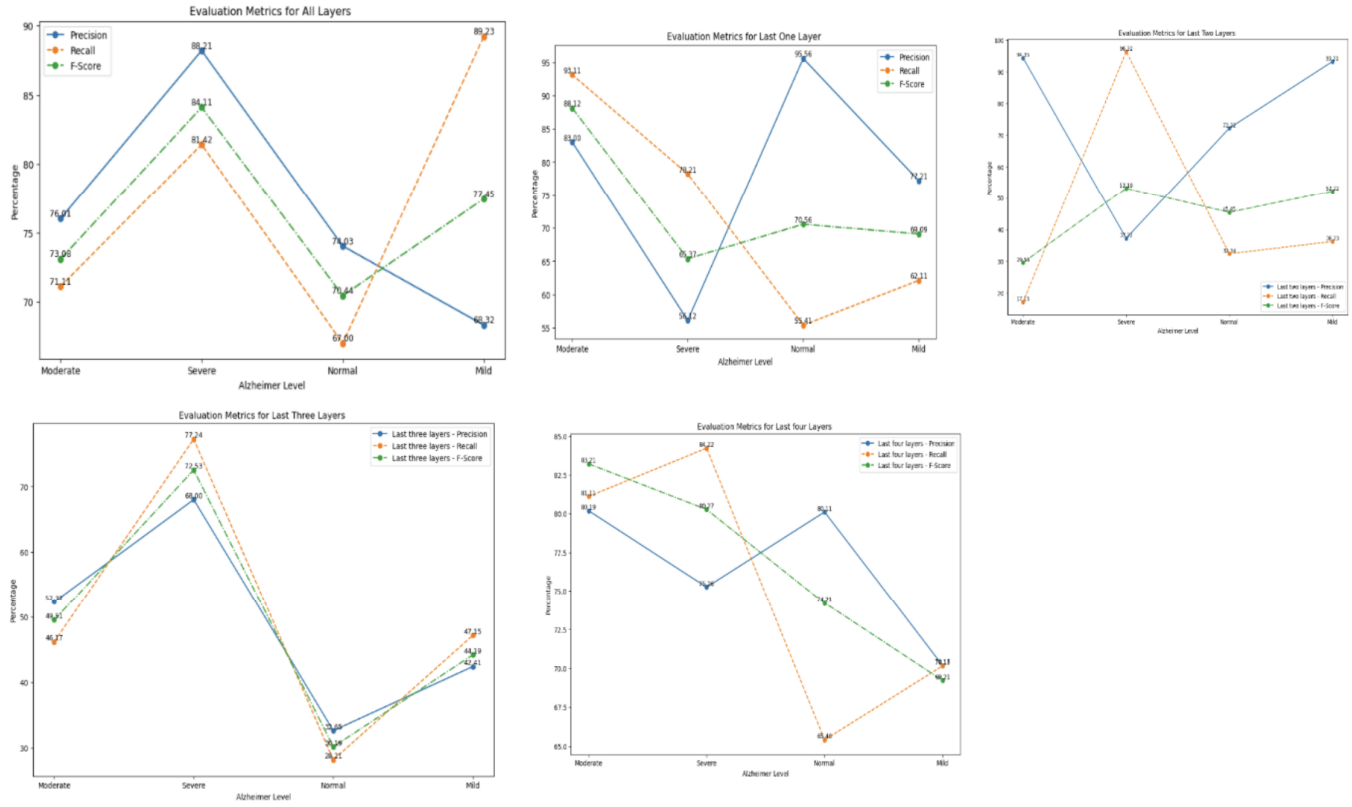

Figure 9: Evaluation Metrics for Xception.

Table 5: Accuracy for Xception

Freezing of Layers Except Freezing of layers except	Percentage of Accuracy Accuracy (in %)
All layers	79.34
Last one layer	73.14
Last two layers	72.28
Last three layers	74.34
Last four layers	73.54

performed with 62%. Compared to these two models, the Xception outperforms well.

For broader benchmarking, future comparisons should include state-of-the-art models like Vision Transformers, Swin Transformers, and Efficient Net. These newer architectures have shown strong performance on similar classification tasks.

5.4. Sample output

The sample output of Alzheimer's Disease classification is shown in the following Figure **8**.

CONCLUSION

In the quest for faster (AD) diagnosis, this study explored the potential of deep learning. We employed pre-trained models, Xception and MobileNetV1, to classify AD into four stages: mild, normal, moderate, and severe. By leveraging transfer learning and strategically freezing model layers, we achieved impressive results. Xception, with its deeper architecture, excelled, reaching an accuracy of 79%, surpassing MobileNetV1's 62%. This suggests Xception's ability to capture the subtle nuances of AD progression. Notably, optimal training configurations differed between models. Xception thrived with all layers trained, while MobileNetV1 preferred focusing

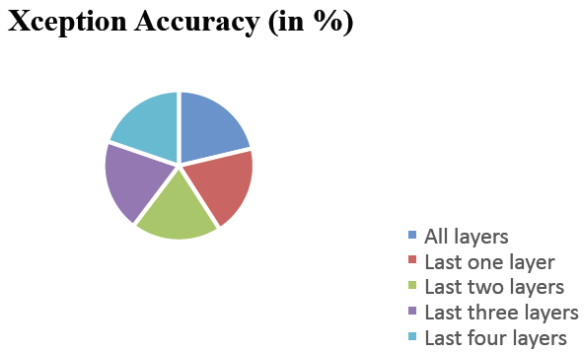

Figure 10: Accuracy for Xception.

Table 6: Performance for Classification of Alzheimer's Disease using MobileNetV1

Freezing of layers except MobileNetV1	Alzheimer level	Precision (in %)	Recall (in %)	F- Score (in %)
All layers	Moderate	49.34	40.20	44.00
	Severe	34.12	42.20	32.12
	Normal	52.11	75.26	45.44
	Mild	78.42	26.21	39.00
Last one layer	Moderate	70.23	14.41	23.46
	Severe	54.36	35.41	42.41
	Normal	41.23	89.46	56.32
	Mild	62.40	59.26	61.35
Last two layers	Moderate	94.35	17.15	29.56
	Severe	37.22	96.32	53.18
	Normal	72.22	32.34	45.45
	Mild	93.21	36.23	52.22
Last three layers	Moderate	56.34	26.32	35.21
	Severe	34.23	54.22	42.45
	Normal	58.23	76.50	73.00
	Mild	73.28	24.34	37.17
Last four layers	Moderate	88.51	52.11	20.21
	Severe	53.20	38.23	50.11
	Normal	69.43	52.11	68.21
	Mild	45.22	68.21	21.19

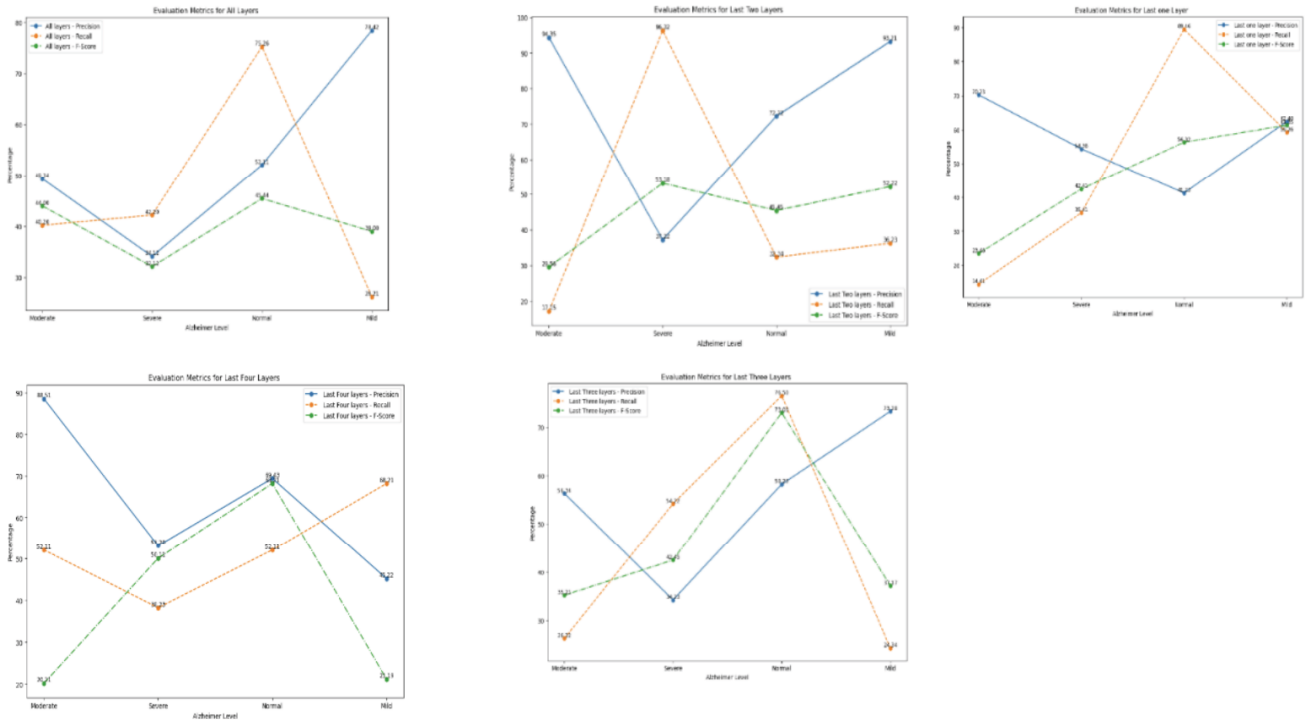

Figure 11: Evaluation Metrics for MobileNetV1.

on the last three. These findings highlight the importance of model-specific tuning for peak performance.

The primary drawback of the proposed research computer-based diagnostic tool being developed for Alzheimer's [34] will solely focus on classifying the disease into various stages without providing any medical suggestions. It should be emphasized that the neuro specialist's role remains essential in offering guidance regarding the disease and its treatment. A critical limitation is the lack of external validation. Future work will focus on testing the trained models on datasets like ADNI and OASIS to assess

Table 7: Accuracy for MobileNetV1

Freezing of all Layers Except Freezing of layers except	Accuracy in (%) Accuracy (in %)
All layers	62.01
Last one layer	48.25
Last two layers	49.10
Last three layers	60.22
Last four layers	58.15

generalizability across different populations. Additional limitations include dataset bias in terms of age and ethnicity, and high computational costs that may restrict deployment in low-resource clinical settings. Furthermore, the model's ability to distinguish AD from other types of dementia remains limited and is a subject for future study.

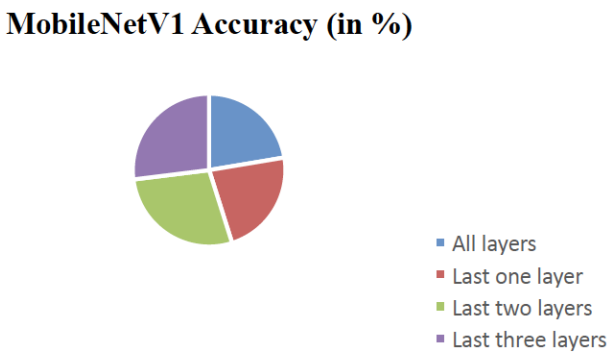

Figure 12: Accuracy for MobileNetV1.

The future scope of this research involves improving the accuracy of the classification model. While the current research has achieved an accuracy of 79%, further modifications and refinements to the methodology are expected to enhance the accuracy in future iterations of work.

This research can be further extended by considering Multi-Modal Transfer Learning and Domain Adaptation. Alzheimer's disease research often involves multiple modalities, such as neuroimaging data (MRI, PET), genetic data, and clinical information. Transfer learning can be extended to incorporate these diverse modalities by jointly learning from multiple pre-trained models or developing fusion strategies to combine information from different domains.

COMPLIANCE WITH ETHICAL STANDARDS ETHICAL APPROVAL

The authors certify that they have no affiliation with or involvement with human participants or animals performed by any authors in any organization or entity with any financial or non-financial interest in the subject matter or materials discussed in this paper.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest among them.

FUNDING DETAILS

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2503).

AVAILABILITY OF DATA AND MATERIAL

The data supporting this study's findings are available on request from the first author.

REFERENCES

[1] Tang C, *et al.* A Novel Machine Learning Technique for Computer-Aided Diagnosis. Engineering Applications of Artificial Intelligence 2020; 92: 103627. https://doi.org/10.1016/j.engappai.2020.103627

[2] McLaughlin Trent *et al.* Dependence as a Unifying Construct in Defining Alzheimer's Disease Severity. Alzheimer's and Dementia 2010; 6(6): 482-93. https://doi.org/10.1016/j.jalz.2009.09.004

[3] Acharya RU, *et al.* Automated Detection of Alzheimer's Disease Using Brain MRI Images- A Study with Various Feature Extraction Techniques. Journal of Medical Systems 2019; 43(9). https://doi.org/10.1007/s10916-019-1428-9

[4] Khan NM, Abraham N, *et al.* Transfer learning with intelligent training data selection for prediction of Alzheimer's Disease. IEEE Access 2019; 7: 72726-72735. https://doi.org/10.1109/ACCESS.2019.2920448

[5] Cilia ND, *et al.* Diagnosing Alzheimer's Disease from online Handwriting: A Novel Dataset and Performance Benchmarking. Engineering Applications of Artificial Intelligence 2022; 11: 104822. https://doi.org/10.1016/j.engappai.2022.104822

[6] Nagaraj Y, Choi JY, Lee B. MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer's Disease: A Survey. Sensors (Switzerland) 2020; 20(11): 1-31. https://doi.org/10.3390/s20113243

[7] Freitas, Francisco A. da S, *et al.* IoT System for School Dropout Prediction Using Machine Learning Techniques Based on Socioeconomic Data. Electronics (Switzerland) 2020; 9(10): 1-14. https://doi.org/10.3390/electronics9101613

[8] McFarland Dennis J, Jonathan RW. Brain-Computer Interfaces for Communication and Control. Communications of the ACM 2011; 54(5): 60-66. https://doi.org/10.1145/1941487.1941506

[9] Ogado, Luis Henrique Silva, *et al.* Diagnosing Leukemia in Blood Smear Images Using an Ensemble of Classifiers and Pre-Trained Convolutional Neural Networks. Proceedings - 30th Conference on Graphics, Patterns and Images, SIBGRAPI 2017; 2017: 367-73. https://doi.org/10.1109/SIBGRAPI.2017.55

[10] Huynh-The T, *et al.*. Artificial Intelligence for the Metaverse: A Survey. Engineering Applications of Artificial Intelligence 2023; 117(October 2022): 105581. https://doi.org/10.1016/j.engappai.2022.105581

[11] Pranjal K, Chauhan S, Kumar Awasthi L. Artificial Intelligence in Healthcare: Review, Ethics, Trust Challenges & Future Research Directions. Engineering Applications of Artificial Intelligence 2023; 120(January): 105894. https://doi.org/10.1016/j.engappai.2023.105894

[12] Parlak İE, Erdal E. Deep Learning-Based Detection of Aluminum Casting Defects and Their Types. Engineering Applications of Artificial Intelligence 2023; 118(October 2022). https://doi.org/10.1016/j.engappai.2022.105636

[13] Kang Z, *et al.* Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography. Cell 2020; 181(6): 1423-1433.e11. https://doi.org/10.1016/j.cell.2020.04.045

[14] Muhammad Hussain, Nadia *et al.* Accessing Artificial Intelligence for Fetus Health Status Using Hybrid Deep Learning Algorithm (AlexNet-SVM) on Cardiotocographic Data. Sensors 2022; 22(14). https://doi.org/10.3390/s22145103

[15] Pakize E, Kabakus AT. The Promise of Convolutional Neural Networks for the Early Diagnosis of the Alzheimer's Disease. Engineering Applications of ArtificialIntelligence 2023; 123(April): 106254. https://doi.org/10.1016/j.engappai.2023.106254

[16] Okur E, Mehmet T. A Survey on Automated Melanoma Detection. Engineering Applications of Artificial Intelligence 2018; 73(March): 50-67. https://doi.org/10.1016/j.engappai.2018.04.028

[17] Guefrechi S, *et al.* Deep Learning Based Detection of COVID-19 from Chest X-Ray Images. Multimedia Tools and Applications 2021; 80(21-23): 31803-20. https://doi.org/10.1007/s11042-021-11192-5

[18] Goyal J, Padmavati K, Trilok CA. Classification, Prediction, and Monitoring of Parkinson's Disease Using Computer Assisted Technologies: A Comparative Analysis. Engineering Applications of Artificial Intelligence 2020; 96(September): 103955. https://doi.org/10.1016/j.engappai.2020.103955

[19] Peifang G. Brain Tissue Classification Method for Clinical Decision-Support Systems. Engineering Applications of Artificial Intelligence 2017; 64: 232-41. https://doi.org/10.1016/j.engappai.2017.05.015

[20] Fuzhen Z, *et al.* A Comprehensive Survey on Transfer Learning. Proceedings of the IEEE 2021; 109(1): 43-76. https://doi.org/10.1109/JPROC.2020.3004555

[21] Amira E, *et al.* Multi-Class Motor Imagery EEG Classification Using Convolution Neural Network. ICAART 2021 - Proceedings of the 13th International Conference on Agents and Artificial Intelligence 2021; 1: 591-600. https://doi.org/10.5220/0010425905910595

[22] Inayatul H, *et al.* Lung Nodules Localization and Report Analysis from Computerized Tomography (CT) Scan Using a Novel Machine Learning Approach. Applied Sciences (Switzerland) 2022; 12(24). https://doi.org/10.3390/app122412614

[23] Phan H, *et al.* MoBiNet: A Mobile Binary Network for Image Classification. Proceedings - 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020; 2020: 3442-51. https://doi.org/10.1109/WACV45572.2020.9093444

[24] Dai X, *et al.* Dynamic Head: Unifying Object Detection Heads with Attentions. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2021; 7369-78. https://doi.org/10.1109/CVPR46437.2021.00729

[25] Raju B, Kumari A, Dasari CM, Amilpur S. An Attention-Based Hybrid Deep Neural Networks for Accurate Identification of Transcription Factor Binding Sites. Neural Computing and Applications 2022; 34(21): 19051-60. https://doi.org/10.1007/s00521-022-07502-z

[26] Erik W, Seitz H. Machine Learning for the Intelligent Analysis of 3D Printing Conditions Using Environmental Sensor Data to Support Quality Assurance 2022.

[27] Gliner JA, *et al.* Measurement Reliability and Validity. Research Methods in Applied Settings 2021; 319-38.

[28] Johnson JM, Khoshgoftaar MT. Survey on Deep Learning with Class Imbalance. Journal of Big Data 2019; 6(1). https://doi.org/10.1186/s40537-019-0192-5

[29] Chen Y, *et al.* Mobile-Former : Bridging MobileNet and Transformer Microsoft University of Science and Technology of China 2 . Related Work. Cvpr 2022; 5270-79.

[30] Tehseen M, *et al.* The Role of Machine Learning and Deep Learning Approaches for the Detection of Skin Cancer. Healthcare (Switzerland) 2023; 11(3). https://doi.org/10.3390/healthcare11030415

[31] Takeshi I, *et al.* Japanese and North American Alzheimer's Disease Neuroimaging Initiative Studies: Harmonization for International Trials. Alzheimer's and Dementia 2018; 14(8): 1077-87. https://doi.org/10.1016/j.jalz.2018.03.009

[32] Oskar H, *et al.* CSF Biomarkers of Alzheimer's Disease Concord with Amyloid-β PET and Predict Clinical Progression: A Study of Fully Automated Immunoassays in BioFINDER and ADNI Cohorts. Alzheimer's and Dementia 2018; 14(11): 1470-81. https://doi.org/10.1016/j.jalz.2018.01.010

[33] der Lee K, Sven J, *et al.* Circulating Metabolites and General Cognitive Ability and Dementia: Evidence from 11 Cohort Studies. Alzheimer's and Dementia 2018; 14(6): 707-22. https://doi.org/10.1016/j.jalz.2017.11.012

[34] Niccolò T, *et al.* Centenarian Controls Increase Variant Effect Sizes by an Average Twofold in an Extreme Case-Extreme Control Analysis of Alzheimer's Disease. European Journal of Human Genetics 2019; 27(2): 244-53. https://doi.org/10.1038/s41431-018-0273-5

Received on 07-06-2025 Accepted on 06-07-2025 Published on 01-08-2025

https://doi.org/10.6000/1929-6029.2025.14.39