A Novel Framework for Classification of Different Alzheimer’s Disease Stages Using CNN Model

Gowhar Mohi ud din dar, Avinash Bhagat, Syed Immamul Ansarullah, Mohamed Tahar Ben Othman, Yasir Hamid, Hend Khalid Alkahtani, Inam Ullah, and Habib Hamam

School of Computer Applications, Lovely Professional University, Phagwara 144411, India
Kwintech-R Labs, Jammu & Kashmir, Srinagar 193501, India
Department of Computer Science, College of Computer, Qassim University, Buraydah 51452, Saudi Arabia
Abu Dhabi Polytechnic, Abu Dhabi 111499, United Arab Emirates
Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
BK21 Chungbuk Information Technology Education and Research Center, Chungbuk National University, Cheongju 28644, Republic of Korea
Faculty of Engineering, Uni de Moncton, Moncton, NB E1A3E9, Canada
Spectrum of Knowledge Production & Skills Development, Sfax 3027, Tunisia
International Institute of Technology and Management, Commune d’Akanda, Libreville BP 189, Gabon
School of Electrical Engineering, Department of Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg 2006, South Africa

Correspondence: maathaman@qu.edu.sa (M.T.B.O.); inam@chungbuk.ac.kr (I.U.)

Abstract: Background: Alzheimer’s, the predominant form of dementia, is a neurodegenerative brain disorder with no known cure. With the lack of innovative findings to diagnose and treat Alzheimer’s, the number of middle-aged people with dementia is estimated to hike nearly to 13 million by the end of 2050. The estimated cost of Alzheimer’s and other related ailments is USD321 billion in 2022 and can rise above USD1 trillion by the end of 2050. Therefore, the early prediction of such diseases using computer-aided systems is a topic of considerable interest and substantial study among scholars. The major objective is to develop a comprehensive framework for the earliest onset and categorization of different phases of Alzheimer’s. Methods: Experimental work of this novel approach is performed by implementing neural networks (CNN) on MRI image datasets. Five classes of Alzheimer’s disease subjects are multi-classified. We used the transfer learning determinant to reap the benefits of pre-trained health data classification models such as the MobileNet. Results: For the evaluation and comparison of the proposed model, various performance metrics are used. The test results reveal that the CNN architectures method has the following characteristics: appropriate simple structures that mitigate computational burden, memory usage, and overfitting, as well as offering maintainable time. The MobileNet pre-trained model has been fine-tuned and has achieved 96.6 percent accuracy for multi-class AD stage classifications. Other models, such as VGG16 and ResNet50 models, are applied to the same dataset while conducting this research, and it is revealed that this model yields better results than other models. Conclusion: The study develops a novel framework for the identification of different AD stages. The main advantage of this novel approach is the creation of lightweight neural networks. MobileNet model is mostly used for mobile applications and was rarely used for medical image analysis; hence, we implemented this model for disease detection and yielded better results than existing models.

Keywords: CNN; deep learning; MCI; EMCI; LMCI; AD; transfer learning; MRI; CSF

1. Introduction

Alzheimer’s is a condition of the brain’s central part that causes gradual memory loss, cognitive impairment, and emotional distress. Around 46.8 million individuals worldwide have dementia, with Alzheimer’s disease accounting for 60–70% of cases and costing more...
than USD 818 billion globally [1]. Given that growing old is the foremost basis of dementia, the figure of persons exaggerated is expected to rise to 151 million by 2050 [2]. Extensive neuronal impairments are caused by the formation of extracellular amyloid plaques and intracellular accretion of neurofibrillary tangles composed of hyperphosphorylated tau [3]. Because Alzheimer’s disease is presently incurable [4–7], disease-modifying medications are being explored that address not only amyloid Beta (Aβ) build-up and the tau pathway [8,9] and neuroinflammation [10–12] and nutritional pathways [13,14]. Recent medications’ inability to cure could be partly attributable to late-stage delivery and insensitive techniques of identifying cognitive changes, underscoring the clinical necessity for rigorous diagnostics [15]. The National Institute of Neurological and Communicative Disorders (NINCDS) and the Alzheimer’s Disease and Related Disorders Association currently set the standard to classify a person’s mental well-being (AD/RA). They discovered that the biochemical alterations that occur in Alzheimer’s disease last for decades. Differentiating between disease stages will demand the establishment of innovative biomarkers:

1. Early-symptomatic aliment;
2. Prevenient aliments, i.e., those with apparent early dementia AD;
3. AD or common variants of AD [16].

Cognitive assessments and imaging examinations, such as Neuro-imaging, rule out alternative factors of memory impairment, such as tumors, which are ruled out by cognitive assessments and imaging examinations, such as Neuro-imaging [17]. When integrated with these approaches and professional clinical expertise, medical criteria have an 80 percent good prognosis and a 60 percent diagnostic accuracy for clinical diagnosis. Recent visualization advances, such as Neuro-imaging scans (MRI) [18], PET scans [19,20], and single-photon emission computed tomography (SPECT) [21], have empowered the tracking of neurodegeneration, malformations in neuronal development, and edema.

The number of AD patients is expected to rise substantially, necessitating the use of a computer-aided diagnosis (CAD) system for early and precise AD diagnosis [22]. Furthermore, mild cognitive impairment (MCI) is an interim phase between sound perception and dementia. As per a previous study [23], MCI participants advance to clinical AD at a rate of 10–15 percent yearly. In recent years, research in detecting MCI patients who will develop clinical dementia has received much attention. Conversion of one stage into another is vital to identifying the different stages of Alzheimer’s disease.

The primary emphasis of this study is on identifying different stages of AD based on an image dataset. Deep learning techniques are frequently employed for time series classification, image identification, and multidimensional data processing [24]. These methods are widely applied to neuroimaging data to identify Alzheimer’s disease (AD) [25]. Potential genetic indicators of Alzheimer’s disease have also been explored using these methods [26].

In order to diagnose Alzheimer’s, Zhang et al. [27] combined neuroimaging data with clinical and neuropsychological evaluations using a multimodal deep learning model. The idea put forth by Spasov et al. [28] emphasizes the significance of deep learning designs in patients at risk of AD in stopping the progression of mild cognitive impairment. Deep learning-based models also use extensive genomic and DNA methylation data to anticipate AD, mitigating the symptoms of usual neurodegeneration (falls, memory loss, etc.).

Our work progresses these approaches (neuroimaging and clinical analysis) by applying a technique that uses an ensemble of CNN models to identify the different stages of AD. Different CNN models are applied to the same dataset and comparing results, and it is found that the MobileNet model is efficient for medical image analysis. The rest of the work is presented in this paper in the following subsections. Section 2 discusses previous findings for diagnosing Alzheimer’s disease. Our CNN-based method for determining the stage of Alzheimer’s disease based on MobileNet and image weights is described in Section 3. The experimental findings of our model are shown in Section 4. The research’s findings are discussed in Section 5, and in Section 6, we provided a research paper conclusion.
2. Related Work

AD detection is extensively investigated and encompasses many problems and complexities [29]. Payan et al. [30] used a sparse autoencoder and 3D convolutional neural networks. They developed an algorithm that analyses a brain MRI scan (MRI) to determine a person’s disease status. The primary innovation was 3D convolutions, which outperformed 2D convolutions in terms of results. The auto-encoder was used to train the convolutional layer, but it was not fine-tuned. Efficiency is supposed to enhance with fine-tuning [24].

Sarraf et al. [31] classified AD from the NC brain using a widely used CNN architecture, LeNet-5 (binary classification). The work presented in [30] was developed by Hosseini et al. [32]. A deeply supervised adaptive 3D-CNN (DSA-3D-CNN) classifier was used to predict AD. Three-layered autoencoder (3D-CAE) architectures are pre-trained without any skull stripping pre-processing on a CAD-Dementia dataset. Performance is analyzed using ten-fold cross-validation.

Gupta et al. [33] devised a sparse autoencoder model for the categorization of Alzheimer’s disease (AD), mild cognitive impairment (MCI), and healthy controls (HC). Payan et al. [34] used sparse autoencoders and a CNN architecture to diagnose Alzheimer’s. They also devised a two-dimensional CNN model that performed similarly. Brosch et al. [34] employed a deep belief network model with manifold learning to diagnose Alzheimer’s disease in MRI images, etc. [35–38].

Liu and Shen [39] used unsupervised and supervised technology to develop a deep-learning model that categorized AD and MCI patients. Korolev et al. [40] showed that a corresponding finding might be accomplished. When the control neural network and basic 3D CNN designs were deployed to three-dimensional MRIs, the outcomes revealed that the deepness and complexity of the two networks were remarkably similar. They did not perform as well as they originally anticipated.

Sarraf and Tofghi [41] employed functional MRI data and the deep LeNet model for AD diagnosis. Suk et al. [42–45] used multiple complex SVM kernels for classification in an autoencoder network-based model for AD diagnosis. They used a multi-kernel classifier to classify small- to semi-characteristics extracted from magnetic current imaging, MCI-converter structural MRI, and PET data.

Wang et al. [46] developed a novel CNN approach that is based on a multimodal MRI analysis approach that involves diffusion tensor images or functional brain imaging data. Patients with Alzheimer’s ailment, dementia, and other related conditions were classified utilizing the framework. Despite the excellent classification accuracy, it is anticipated that employing 3D convolution rather than 2D convolution would enhance efficiency. A 3D multi-scale CNN (3DMSCNN) model was devised by Ge et al. [47]. The 3DMSCNN was a novel architecture for AD diagnosis. They also devised a multi-scale feature augmentation technique as well as a feature fusion.

Song et al. [48] postulated a Graph Convolutional Neural Network (GCNN) classifier based on graph techniques. By using structural connection graphs, which indicate a multi-class model, training and architecture evaluation divides the AD spectrum into four categories. Xu Y et al. [49] proposed a medical image segmentation method based on multi-dimensional statistical features. The main purpose of this paper is to integrate CNN and transformers to detect and diagnose brain tumors. This research has efficient results. Based on such techniques, this research can be further enhanced by integrating both data modalities and models to detect and diagnose brain disorders. We are also currently working on the detection and diagnosis of brain disorders using different data modalities.

3. Problem Description and Solution Strategy

As highlighted in the Section 2, numerous paradigms encompassing AD prognosis and clinical image assessment have recently been presented in the literature. However, most do not use transfer learning algorithms, multi-class clinical object detection, or an Alzheimer’s disease monitoring cloud service to assess AD distinct phases and provide
faraway guidance. These issues have received insufficient attention in literary works. Thus, the novelties of this research can be organized as follows following other cutting-edge techniques discussed in the Section 2:

A novel framework is devised to identify various Alzheimer’s ailment phases and the classification of medical images. The suggested method relies on CNN architectures for structural MRI images of the brain. Transfer learning is utilized to grab efficiencies of already trained architectures, such as VGG19, ResNet50, and DenseNet121.

Unbalanced datasets and notional size are the most problematic aspects of medical image analysis. Resampling techniques are employed to balance the datasets, while data expansion methods are utilized to enhance the dataset size and overcome the over-fitting issues. According to performance indicators, the experimental results predict a positive outcome.

4. Methods and Materials

The early diagnosis of Alzheimer’s ailment is critical for precluding and managing its progression. The data used in this research are taken from ADNI (Alzheimer’s Disease Neuroimaging Initiative). The Alzheimer’s Disease Neuroimaging Initiative (ADNI) is a longitudinal study designed to develop clinical, imaging, genetic, and biochemical biomarkers for the early detection and tracking of Alzheimer’s disease (AD). This initiative started in 2004 and is supported by multiple companies. Complete 1-year data were taken from ADNI to design this novel research. The main purpose of this research is to design a novel approach for the initial identification and tracking of Alzheimer’s stages. The proposed framework workflow, data preparation algorithms, and medical image classification techniques are thoroughly explained below.

5. The Proposed Framework

The proposed approach consists of below mentioned four steps:

Step1. Data Attainment

This approach uses the ADNI dataset in the T2w MRI format. It offers medical images in jpeg formats, such as Axial Coronal and Sagittal. The dataset comprises data from 300 subjects alienated into five classes that are cognitively normal, mild cognitive impairment, early mild cognitive impairment, late mild cognitive impairment (CN, MCI, EMCI, LMCI), and NC. The patient with LMCI symptoms can acquire more AD than the EMCI subject because the patient has undergone severe neuron damage in this stage. The total number of images available here is 1101. AD class comprises 145 images, EMCI comprises 204 images, LMCI comprises 61 images, MCI has 198, and NC has 493 images. We re-sized all the images into $224 \times 224$ pixels, and three channels (RGB) were used. A batch size of 32 images was transferred at each iteration during training to reduce the computational power. The batch size of 32 images transferred at each iteration of training to minimize the height = 224 width = 224 channels = 3 batch size = 32.

Step2. Pre-processing

Data are unbalanced, and training on an unbalanced dataset leads us toward data underfitting or overfitting issues, and in the end, our model would not be able to classify the images correctly. The solution is to balance the data, for which we use the upsampling technique, in which the labels with a smaller number of images are increased or unsampled. After resampling, all the classes become 580 MRI images, and thus the entire dataset size is 2900. The data are refined, standardized, scaled, denoised, and formatted appropriately. In the future, other techniques, such as downsampling, will be used. Figure 1 illustartes the resampling technique of MRI images.
unsampled. After resampling, all the classes become 580 MRI images, and thus the entire dataset size is 2900. The data are refined, standardized, scaled, denoised, and formatted appropriately. In the future, other techniques, such as downsampling, will be used. Figure 1 illustrates the resampling technique of MRI images.

Figure 1. Example of resampling technique applied on MRI image [50,51].

Step3. Data Augmentation

The primary objective of employing data augmentation methods is to (1) expand the data size and (2) solve the issue of overfitting. Data augmentation approaches are used in the following way.

The input images are pre-processed by using the pre-processing function of the pre-trained model, horizontal flipping of the images, rotation of images by 5 degrees, and width and shift in the images.

We used the Keras API of the image data generator to apply the data augmentation. We can observe that some images are 5 degrees rotated, and some are flipped, as shown in Figure 2.

Figure 2. The different augmentation techniques are rotation, flipping, shift, etc. In our approach, some images are 5 degrees rotated, some are flipped, etc. [52].
As a result, the dataset expands to 2900 images, partitioned into 580 images in each class. After that, the balanced dataset of 2900 MRI scans is reconfigured and randomly fragmented into training, validation, and test groups, with an 80:10:10 split ratio for each class. The division of data testing, training, and validation groups for 5-way classification is summarized in Table 1 (CN vs. MCI vs. EMCI vs. LMCI vs. AD).

Table 1. Shows the images inputting the model from different AD classes. Two thousand images (400 from each class) are used for training purposes, 450 for validation (90 from each class), and 450 for testing (90 from each class) [53].

<table>
<thead>
<tr>
<th>Class</th>
<th>Training Data</th>
<th>Validation Data</th>
<th>Testing Data</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN</td>
<td>400</td>
<td>90</td>
<td>90</td>
<td>580</td>
</tr>
<tr>
<td>MCI</td>
<td>400</td>
<td>90</td>
<td>90</td>
<td>580</td>
</tr>
<tr>
<td>EMCI</td>
<td>400</td>
<td>90</td>
<td>90</td>
<td>580</td>
</tr>
<tr>
<td>LMCI</td>
<td>400</td>
<td>90</td>
<td>90</td>
<td>580</td>
</tr>
<tr>
<td>AD</td>
<td>400</td>
<td>90</td>
<td>90</td>
<td>580</td>
</tr>
<tr>
<td>TOTAL</td>
<td>2000</td>
<td>450</td>
<td>450</td>
<td>2900</td>
</tr>
</tbody>
</table>

6. Proposed Classification Methods and Techniques

The three critical components of machine learning algorithms are feature extraction, reduction, and classification. All three steps are performed manually or separately while implementing machine learning algorithms. The beauty of deep learning algorithms such as CNNs is that there is no need for manual feature extraction. These three stages are performed in combination with CNN architectures. CNN architectures have high classification performance than traditional models. The three layers of CNN architectures are the convolution layer, the pooling layer, and the entire connected layer [54]. Extraction of features is the responsibility of the convolution layer, dimension reduction by the pooling layer, and classification by fully connected layers. Conversion of two-dimensional metrics into one-dimensional vectors is also performed by a fully connected layer [55].

6.1. Convolution Layer

It acts as a base for the CNN architecture. It comprises a set of filters, also called kernels, which are learned through the training process. The filter dimensions are smaller than the real image. Filters convolve with images and create activation maps. The convolution layer extracts all the features. A learnable filter that retrieves features out of a given image is represented by the convolutional layer. For a three-dimensional image with the dimensions \( H \times W \times C \), \( H \) denotes height, \( W \) width, and \( C \) is the total count of channels. Applying a 3D filter-sized \( F \times F \times F \times F \times F \times C \), where \( FC \) is the number of filter channels, \( FW \) denotes the filter width, and \( FH \) denotes the filter height. Hence, the output activation map size...
must be \(AH \times AW\), where \(AH\) stands for activation height and \(AW\) activation width. The following equations are used to calculate activation height and width values.

\[
AH = 1 + \frac{H - FH + 2P}{S} \tag{1}
\]

\[
AW = 1 + \frac{W - Fw + 2P}{S} \tag{2}
\]

\(P\) signifies padding, \(S\) represents stride, and there are \(n\) filters, so the activation map dimensions must turn out to be \(AH \times AW \times n\). Figure 3 illustrates the complete convolution.

![Figure 3](image-url)

**Figure 3.** The complete convolution process is shown in the figure above, where different filters are applied to images before obtaining the final output [56].

### 6.2. Polling Layer

The pooling layer’s primary purpose is to lower the size of the feature maps. Therefore, there are fewer parameters to learn and fewer computations to be made by the network. The different polling layers are max pooling, average pooling, and global pooling. By applying a non-linear conversion to the given inputs, the activation function addresses non-linearity in the network. Our proposed multi-classifier uses the SoftMax activation function in the output layer. The main function of the SoftMax function is to calculate relative probabilities. The general equation of the SoftMax equation function is given below in Equation (3).

\[
\text{Softmax} \ (z_i) = \frac{\exp(z_i)}{\sum_j \exp(Z_j)} \tag{3}
\]

In this case, \(Z\) stands for the values of the output layer neurons, with the exponent being a non-linear function. These values are then normalized and transformed into probabilities by dividing them by the sum of exponent values. For all hidden layers, we applied the ReLU activation function, the most familiar implicated function in CNNs. There are different variants of ReLU activation functions, such as parametricReLU, leaky ReLU, exponential linear (ELU, SELU), and concatenated ReLU (CReLU). We applied leaky ReLU since it has some benefits over other variants, such as it fixes the problem of “dying ReLU” because it has no zero-slope parts. It also speeds the training process because it is more balanced and therefore learns faster. However, it should be kept in mind that leaky ReLU is not superior to simple ReLU and should be considered as an alternative. The general equation for the ReLU and leaky ReLU activation functions are given below in Equation (4) and Equation (5), respectively.

\[
F(x) = x^+ = \max(0, x) \tag{4}
\]
When \( z \) is less than 0, leaky ReLU allows a small nonzero, constant gradient \( \alpha \). Generally, the value of \( \alpha = 0.01 \). In our study for medical image analysis, we ensemble MobileNet with ImageNet weights along with transfer learning for image classification. These architectures can easily handle two-dimension and three-dimension brain neuroimages built on 2D, 3D, and convolutions. The general flow of our novel framework is shown in Table 2.

Table 2. Shows the general architecture of the MobileNet model with total trainable and non-trainable parameters [57].

<table>
<thead>
<tr>
<th>Layer</th>
<th>Output Shape</th>
<th>Param</th>
</tr>
</thead>
<tbody>
<tr>
<td>input1 (Input Layer)</td>
<td>(None, 224, 224, 3)</td>
<td>0</td>
</tr>
<tr>
<td>conv1 (Conv2D)</td>
<td>(None, 112, 112, 32)</td>
<td>864</td>
</tr>
<tr>
<td>convolution1 (Batch Normalization)</td>
<td>(None, 112, 112, 32)</td>
<td>128</td>
</tr>
<tr>
<td>convolution1 (ReLU)</td>
<td>(None, 112, 112, 32)</td>
<td>0</td>
</tr>
<tr>
<td>convolution1_dw (DepthwiseConv2D)</td>
<td>(None, 112, 112, 32)</td>
<td>288</td>
</tr>
<tr>
<td>convolution1_bn (Batch Normalization)</td>
<td>(None, 112, 112, 32)</td>
<td>128</td>
</tr>
<tr>
<td>convolution1_dw_relu (ReLU)</td>
<td>(None, 112, 112, 32)</td>
<td>0</td>
</tr>
<tr>
<td>convolution1_pw (Conv2D)</td>
<td>(None, 112, 112, 64)</td>
<td>2048</td>
</tr>
<tr>
<td>convolution1_pw_1_bn (Batch Normalization)</td>
<td>(None, 112, 112, 64)</td>
<td>256</td>
</tr>
<tr>
<td>convolution1_pw_relu (ReLU)</td>
<td>(None, 112, 112, 64)</td>
<td>0</td>
</tr>
<tr>
<td>convolution2_pad (ZeroPadding2D)</td>
<td>(None, 113, 113, 64)</td>
<td>0</td>
</tr>
<tr>
<td>convolution2_dw (DepthwiseConv2D)</td>
<td>(None, 56, 56, 64)</td>
<td>576</td>
</tr>
<tr>
<td>convolution2_dw_bn (Batch Normalization)</td>
<td>(None, 56, 56, 64)</td>
<td>256</td>
</tr>
<tr>
<td>convolution2_dw_relu (ReLU)</td>
<td>(None, 56, 56, 64)</td>
<td>0</td>
</tr>
<tr>
<td>convolution2_pw (Conv2D)</td>
<td>(None, 56, 56, 128)</td>
<td>8192</td>
</tr>
<tr>
<td>convolution2_pw_bn (Batch Normalization)</td>
<td>(None, 56, 56, 128)</td>
<td>512</td>
</tr>
<tr>
<td>convolution2_pw_relu (ReLU)</td>
<td>(None, 56, 56, 128)</td>
<td>0</td>
</tr>
<tr>
<td>convolution3_dw (DepthwiseConv2D)</td>
<td>(None, 56, 56, 128)</td>
<td>1152</td>
</tr>
<tr>
<td>flatten (Flatten)</td>
<td>(None, 50176)</td>
<td>0</td>
</tr>
<tr>
<td>dense (Dense)</td>
<td>(None, 512)</td>
<td>25,690,624</td>
</tr>
<tr>
<td>batch_normalization_1 (normalization)</td>
<td>(Batch (None, 512))</td>
<td>2048</td>
</tr>
<tr>
<td>dense1 (Dense)</td>
<td>(None, 512)</td>
<td>262,656</td>
</tr>
<tr>
<td>dropout (Dropout)</td>
<td>(None, 512)</td>
<td>0</td>
</tr>
<tr>
<td>dense2 (Dense)</td>
<td>(None, 5)</td>
<td>2565</td>
</tr>
</tbody>
</table>

Total params: 29,190,853
Trainable params: 25,958,917
Non-trainable params: 3,231,936

We applied the MobileNet model and ImageNet weights to categorize distant phases of Alzheimer’s. It uses depth-wise separable convolutions. The main advantage of this model is that it reduces the parameter number compared to other networks and generates lightweight deep neural networks [58,59]. It is a class of CNN and gives us the optimum initial point for training our classifier to be insanely small and fast. Mobile Nets are built on depthwise separable convolution layers. Each dw Conv consists of depthwise convolution and pointwise convolution [60–62]. There are almost 4.2 million parameters in a MobileNet architecture. The size of the input image is 224 × 224 × 3. The convolution kernel shape is 3 × 3 × 3 × 32, with Avg pool size of 7 × 7 × 1024. Dropout layers are succeeded by a flattened layer and entirely connected layers. The final fully connected layer with SoftMax as the activation function is implemented to manage five classes of Alzheimer’s (AD), while ReLU is the activation function for hidden layers.

The general architecture of the MobileNet model we applied is shown below in Figure 4. The total trainable parameters are 25,958,917, and the non-trainable are 3,231,936. We used the RMSProp as our optimizer with a learning rate of 0.00001 and used the loss.
as categorical cross entropy for multi-class classification and keeping the metrics such as accuracy, which give the results of training and authentication, loss, and accuracy values during the training. The evaluation of the novel approach with other developed approaches is revealed in the tables below.

Figure 4. Shows the basic design of the proposed model [59].

7. Experimental Findings and Model Evaluation

The novel model considers various scenarios. We examined the empirical results in terms of many performance benchmarks, including confusion matrix, accuracy, loss, F1 score, precession, recall, ROC, sensitivity, and AUC. Table 3 below provides a summary of the novel model.

Table 3. Shows us the evaluation of the presentation metrics of the devised model [60].

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final AD JPEG</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>90</td>
</tr>
<tr>
<td>Final CN JPEG</td>
<td>0.95</td>
<td>0.90</td>
<td>0.93</td>
<td>90</td>
</tr>
<tr>
<td>Final EMCI JPEG</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>90</td>
</tr>
<tr>
<td>Final LMCI JPEG</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>90</td>
</tr>
<tr>
<td>Final MCI JPEG</td>
<td>0.93</td>
<td>0.98</td>
<td>0.95</td>
<td>90</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td>0.96</td>
<td>450</td>
</tr>
<tr>
<td>Macro Avg.</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>450</td>
</tr>
<tr>
<td>Weighted Avg.</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>450</td>
</tr>
</tbody>
</table>

Model Evaluation

For the multi-classification, we used MobileNet architecture, a version of CNN networks. The efficacy of the planned model is equated with prevailing models; as depicted in Table 3, the proposed model shows better accuracy results than existing models. Our model achieves an accuracy of 96.22%, as shown in Figure 5, while Juan Ruiz et al. [49] achieved 66.67%, Spasov et al. [55] achieved 88%, and Sahumbaiev et al. [56] achieved 89.47%. The training and validating accuracy and loss of the suggested approach are shown in Figure 6.
Figure 5. Shows the accuracy and loss values during training over 100 Epochs.

![Accuracy and Loss Values](image)

Figure 6. Depicts the normalized confusion matrix generated by the proposed model.

The number of patients diagnosed with each type of AD stage (NC/MCI/AD/LMCI/EMCI) is shown in the confusion matrix. The normalized confusion matrix for the suggested framework is shown in Figure 6.

The comparative analysis of the proposed method with existing approaches is shown graphically in the below figure. Our approach shows better results than existing results. Figure 7 shows the assessment of the devised architecture with other.
Figure 7. Assessment of the devised architecture with the existing frameworks. Our model shows better performance results than existing architectures [61,62].

8. Conclusions

This study presents a system for medical image categorization and Alzheimer’s ailment recognition. Deep-learning CNN architectures support the proposed approach. Alzheimer’s disease has five stages. We employ the MobileNet model with ImageNet weights. The beauty of MobileNet architecture is that it uses depthwise separable convolutions, which reduces parameter numbers compared to other models with regular convolutions and results in lightweight neural networks. The other important characteristic of MobileNet architecture is that instead of having a single $3 \times 3$ convolution layer in traditional networks, it has batch norm and ReLU. Additionally, it divides convolution into $1 \times 1$ pointwise convolutions and $3 \times 3$ depth-wise convolutions. For detection, embedding, segmentation, and classification, MobileNet models are essential. ImageNet is a standard for image classification. It provides a standard measure of how efficient a model is for classification. Different performance metrics are implemented for the assessment of the model. Our model achieves an accuracy of 96.22%. In the impending, it is premeditated to implement other pre-trained models for classification purposes and check whether a patient can convert from one AD stage into another. In the imminent, the dataset’s size will also be increased to improve accuracy, and different augmentation approaches, such as downsampling, will be used. The author is currently working on multi-modal fusion data fusion techniques to detect and diagnose AD.


Funding: The researchers would like to thank the Deanship of Scientific Research, Qassim University for funding the publication of this project.

Data Availability Statement: Not applicable.
Acknowledgments: The researchers would like to thank the Deanship of Scientific Research, Qassim University for funding the publication of this project.

Conflicts of Interest: The authors declare no conflict of interest.

References


51. Khan, R.; Yang, Q.; Ullah, I.; Rehman, A.U.; Bin Tufail, A.; Noor, A.; Rehman, A.; Cengiz, K. 3D convolutional neural networks based automatic modulation classification in the presence of channel noise. IET Commun. 2021, 16, 497–509. [CrossRef]


62. Zhu, Z.; He, X.; Qi, G.; Li, Y.; Cong, B.; Liu, Y. Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal MRI. Inf. Fusion 2023, 91, 376–387. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.