Research Article

An Improved Hybrid Transfer Learning-Based Deep Learning Model for Alzheimer's Disease Detection Using CT and MRI Scans

Sami Alshmrany[1], Gowhar Mohi ud din dar[2], Syed Immamul Ansarullah[3]

1. Faculty of Computer and Information Systems, Islamic University of Madinah, Medina, Saudi Arabia; 2. School of Computer Applications, Lovely Professional University, Phagwāra, India; 3. Government Degree College Bemina, Srinagar, India

Alzheimer's Disease (AD) is a neurological disorder that affects cognitive functions, including memory, thinking, and behavior. Early detection of Alzheimer's disease is critical for effective treatment and management of the condition. Deep Learning (DL) is a powerful tool that can be used for AD detection and diagnosis. DL algorithms can learn patterns and features in large datasets that can be used to classify and predict the presence of Alzheimer's Disease. The most common approach is to use brain imaging techniques, such as computed tomography and brain MRI scans, to extract features that are characteristic of Alzheimer's Disease. Transfer learning-based deep learning models can be effective in detecting Alzheimer's disease from medical images. Transfer learning involves using pre-trained neural network models as a starting point and fine-tuning them to suit a specific task, such as Alzheimer's disease detection. This paper focuses on classifying AD patients into various stages (early mental retardation, mild mental impairment, late mild mental impairment, and final Alzheimer's stage) by utilizing transfer learning with ResNet50, VGG16, and DenseNet121 along with CNN networks on a large dataset. The work classifies Alzheimer's patients into various stages using transfer learning with ResNet50, VGG16, and DenseNet121 along with CNN on a large dataset. The model is trained and tested on ADNI data using Keras API and divides the MRI images into: EMCI, MCI, LMCI, and AD. The performance of VGG16, DenseNet121, and ResNet50 outperformed other models significantly. The results demonstrate a significant improvement in accuracy compared to previous approaches, with a final accuracy of 96.6%.

 doi.org/10.32388/4RIOQC

Corresponding authors: Sami Alshmrany, s.alshmrany@iu.edu.sa; Syed Immamul Ansarullah, drsyedansar@kwintech-rlabs.org

1. Introduction

Alzheimer's disease is a global health crisis that affects millions of people around the world. This debilitating condition erodes the brain's ability to comprehend, remember, and perform basic functions, ultimately leading to death [1]. With projections indicating that the number of Alzheimer's patients will rise from 50 million to 152 million by 2050, it is imperative that we act now to address this growing health crisis [2]. The cost of treating Alzheimer's is staggering, with expected global expenses reaching nearly $186 billion in 2018 [3]. Unfortunately, this number is only expected to increase in the coming years, putting an enormous strain on the world's healthcare system. According to the established Clinical Dementia Rating (CDR) result, the disorder is split into four stages: early mild cognitive impairment, mild cognitive impairment, late mild cognitive impairment, and Alzheimer's (AD). Early diagnosis of dementia disorders is crucial for patient recovery and treatment expenses because the cost of treating patients with EMCI and LMCI is different. Diagnosis of Alzheimer's is best possible after the death of a patient since Alzheimer's pathology changes in patients could not be assessed early.

The initial diagnostic criteria for Alzheimer's disease were created in 1984 and relied solely on clinical symptoms. With the discovery of different biomarkers such as CSF, MRI, and PET data, the international working group devised a new approach in 2014, which served as the model for the National Institute on Aging and Alzheimer's Association's (NIA-AA) subsequent set of standards. Biomarker data are used to link the clinical condition of dementia or mild cognitive loss to intrinsic Alzheimer's pathological changes with high, moderate, or low risk in the NIA-AA criteria [4]. Imaging biomarkers are used to assess Alzheimer's disease, such as CT, fMRI, MRI, and PET scans. The hippocampus and entorhinal cortex have shown extremely early changes in Alzheimer's disease that are consistent with pathology, but it is still uncertain which structure would be best for an early diagnosis [5]. The physiology of dementia and its differential diagnosis have greatly benefited from structural and functional imaging, which also holds considerable potential for tracking the course of the disease [6]. Numerous articles have been written regarding how various imaging methods can be used to detect Alzheimer's disease. In volumetric MRI, patterns of sick and healthy subjects were identified using feature-based morphometry (FBM) [7]. In computerized medical image processing, convolutional neural networks (CNNs) have achieved major advancements. As a result, various CNN models, including VGG, MobileNet, AlexNet, and ResNet, are

available for object detection and segmentation. Despite the fact that CNNs are renowned deep learning techniques, their effectiveness is hampered by the absence of an extensive medical imaging dataset [8]. Transfer learning is one of the efficient methods for building deep convolutional neural networks without overfitting when the amount of data is minimal [8]. A pre-trained network serves as the foundation of transfer learning. The proposed method can learn the most useful features instead of training a specific CNN network from scratch. To categorize AD into five classes, the proposed research study has used four pre-trained networks, including VGG 16, ResNet, and DenseNet121. The main contribution of this research paper is to detect and classify the Alzheimer's stages, and it is done in the following stages:

- Identification of the image dataset, and the identified dataset is in ANN format.
- Conversion of this image dataset into JPEG format.
- Application of different normalization techniques on the dataset to remove ambiguities.
- Application of various data augmentation techniques on the normalized dataset.
- Ensemble of different deep learning approaches on the normalized dataset to detect and diagnose Alzheimer's stages.
- Finally, a comparison of the efficiency of deep learning models was performed, and it was found that VGG 16 and DenseNet121 outperform ResNet 50 and other models.

The rest of the paper is organized as follows: Section 2 illustrates the literature review, Section 3 discusses transfer learning (Section 3.1) VGG16, (Section 3.2) Resnet50, (Section 3.3) DenseNet121. Section 4 discusses the proposed work and its experimental evaluation.

2. Literature Review

A literature review on the use of machine learning techniques in Alzheimer's disease (AD) research shows a growing trend in the development of models that can assist in early diagnosis, predict disease progression, and improve the understanding of the underlying biological mechanisms of AD. One of the most common approaches in AD research is the use of magnetic resonance imaging (MRI) scans to study brain changes associated with the disease. Convolutional neural networks (CNNs) have been used to classify and differentiate between healthy brains and those with AD based on MRI scans. Some of the promising research in detecting early signs of AD, which can help in early intervention and improve patient outcomes, are described as follows.

An automated framework was developed by U. Rajendra et al. [9] to evaluate whether a baseline brain scan will detect any evidence of Alzheimer's disease. Lihua Wang et al. [10] integrated genomic data from six different brain areas using SVM machine learning techniques to find AD biomarkers. Martin Randles and Mohamed Mahyoub [11] proposed that, relying on characteristics including lifestyle, medical history, demography, and other considerations, Alzheimer's is predicted at various stages. Rueda et al. suggested a fusion-based image processing technique that identifies discriminative brain patterns connected to the presence of neurodegenerative disorders [12]. The effectiveness of classification using a support vector machine (SVM) was assessed on several datasets once the discriminative patterns had been identified. Li et al. [13] presented a classification approach based on multilayer brain divisions. Using SVM, histogram-based parameters from MRI data were used to categorize various brain levels.

Giraldo et al. [14] proposed an automated technique recently developed for identifying structural abnormalities in the thalamus, planum temporale, amygdala, and hippocampal areas. Hina Nawaz et al. [15] devised a framework based on the computer-aided system, which needs real-time AD diagnosis. They have suggested identifying the stages of AD. For certain deep feature modeling and extraction, researchers have used classification algorithms like KNN (K-nearest neighbor), RF (Random Forest), and SVM (Support Vector Machine). Large datasets were necessary for classification and extracting deep features to avoid overfitting problems. To attain the maximum accuracy in early Alzheimer's diagnosis, they have recommended on-time depth and propagation of learning techniques compared to previous approaches. There is currently no treatment for AD using any medical reasoning approaches, and early detection of Alzheimer's disease is complicated. To attain high accuracy, Ketki Tulpule et al. [16] focused on nonlinear SVM for the radial base purpose when developing a computerized machine learning approach for categorizing Alzheimer's phases.

Muazzam Maqsood et al. [17] devised a transfer learning approach to identify Alzheimer's disease. They suggested breaking down the AD category into different divisions. Since Alzheimer's is an incurable ailment, it is an emerging topic for research globally. The contribution of researchers across the globe for detection and diagnosis of this disorder is listed in Table 1.

References	Dataset	Classification	Results
[18]	MIAS dataset	Binary	95%
[19]	Retinal photographs	Binary	93%
[20]	MNIST	Binary	85%
[21]	ADNI	Binary	96%
[22]	ADNI	Binary	85%
[23]	ADNI	Binary	88%
[24]	ADNI	MULTI	96%

Table 1. Literature review for Alzheimer's detection

In addition to diagnosis and progression prediction, machine learning techniques have also been applied to understand the biological mechanisms underlying AD. This includes the analysis of genomic data, protein expression data, and other biological markers to identify potential drug targets and predict disease outcomes. This literature review highlighted the potential of machine learning techniques in advancing the understanding and treatment of Alzheimer's disease. While the field is still in its early stages, the results to date are promising, and continued research and development are necessary to fully realize the potential of these approaches.

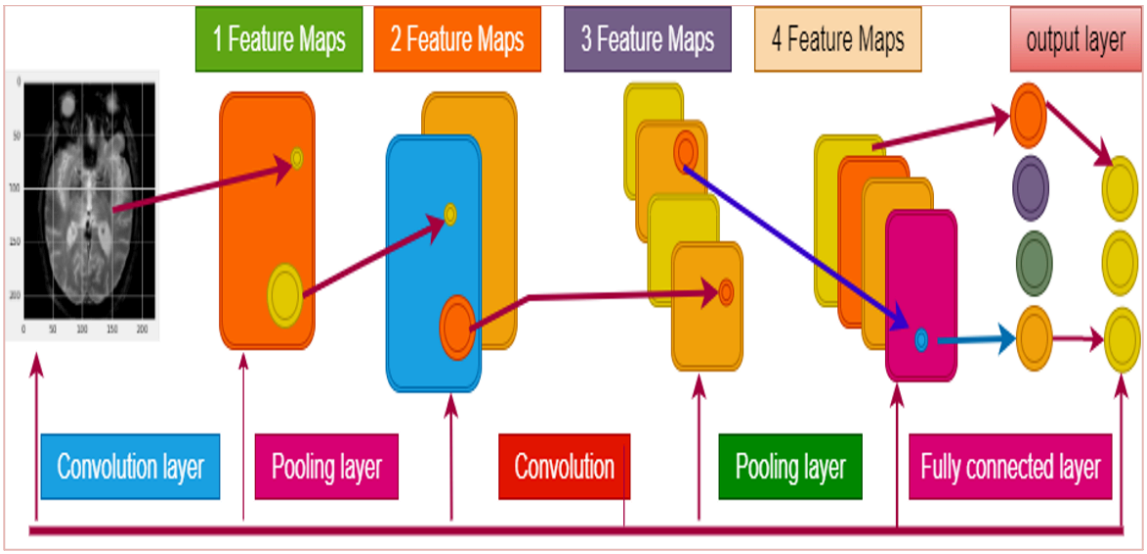

Figure 1. Basic CNN architecture for AD detection procedure

3. Transfer Learning

A model created for one task is used as the basis for another using the machine learning technique known as transfer learning. Deep learning tasks in computer vision and natural language processing are built on pre-trained models. Compared to building neural network models from scratch, they are both cheaper and faster, and they perform remarkably better on related tasks. Transfer learning is learning a new activity more effectively by applying what has already been learned about a related one [25]. For this approach to be practical, the features must be generic, i.e., applicable to both the base task and the target task [26]. Convolutional neural networks, often known as ConvNet, are a subset of Deep Neural Networks (DNN) and are most frequently applied to the processing of medical images. The fundamental structure of the CNN is shown in Figure 1. Various pre-trained deep learning models with transfer learning approaches have been explored in research. VGG 16, ResNet 50, and DenseNet 121 were used in this study.

3.1. VGG16: A Convolutional neural network with 16 layers is called VGG-16. The ImageNet database contains a pre-trained version of this network that has been trained on more than a million images [27]. The pre-trained model can categorize images into 1000 distinct object groups. The network has, therefore, acquired rich feature representations for a variety of images.

3.2. ResNet50: A ResNet model version called ResNet50 contains forty-eight Convolution layers, one MaxPool layer, and one Average Pool layer. There are 3.8 x 10^9 floating-point operations available. It is a

commonly used architecture, and we thoroughly examined the ResNet50 design [28].

3.3. DenseNet121: In densely connected Convolutional networks, each layer is linked to every other layer. There are L (L+1)/2 direct connections between 'L' layers. DenseNet resolves the vanishing gradient problem by altering the typical CNN architecture and streamlining the connectivity between layers [29].

4. Proposed Work and Its Experimental Evaluation

MRI images from the ADNI dataset are used in this study (adni.loni.usc.edu). There are 3400 images in this dataset (680 images from each class), each measuring 224 × 224. The research flow of the proposed work is shown in the flowchart below (Figure 2). Data balancing is essential for the Model to predict with optimal accuracy. Unbalanced data leads to overfitting and underfitting; thus, data needs to be balanced. Here in this study, we use downsampling techniques to balance the data.

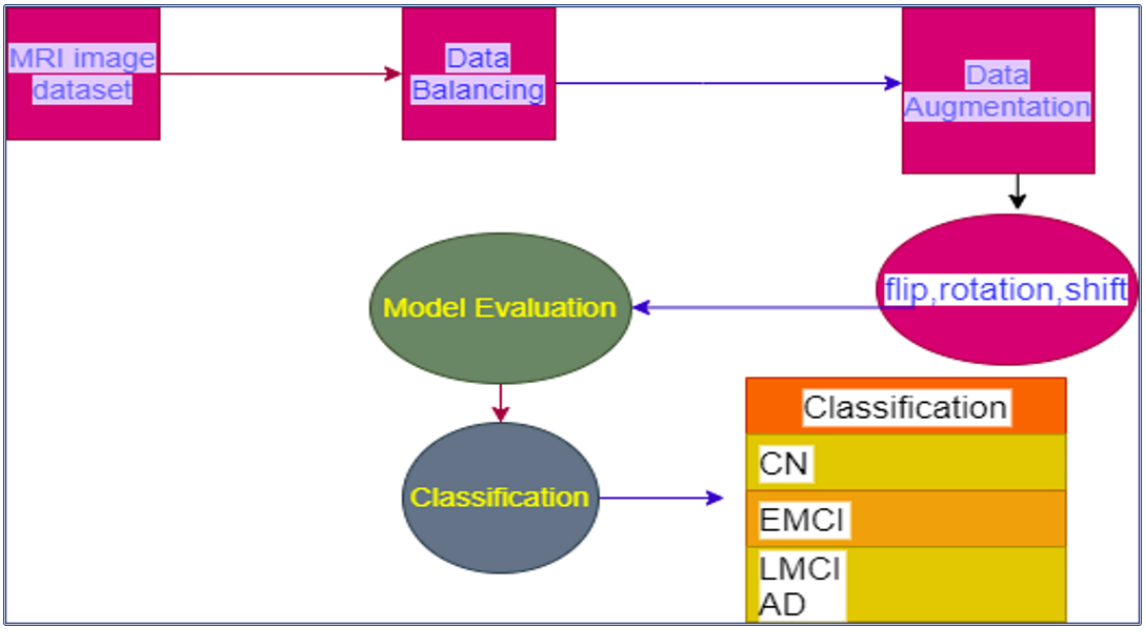

Figure 2. The basic flowchart of this proposed work

The images from each AD stage are selected and given input to the specified models. The data is divided into training, validation, and testing. The complete information regarding each stage is listed in Table 2.

AD stage	Total images in the dataset			
	Training data	Test data	Validation data	Total
NC	500	90	90	680
EMCI	500	90	90	680
MCI	500	90	90	680
LMCI	500	90	90	680
AD	500	90	90	680

Table 2. The images given as inputs to the model

4.1. Data Augmentation: The size of the dataset is significant for deep learning models. These models predict more accurately and yield better accuracy results with large datasets. The major drawback of image datasets is their limited size. Therefore, it needs to be augmented to enlarge the dataset for the models. We applied different data augmentation techniques to the datasets, such as horizontal flipping of the images, rotating the images by 5 degrees, and adjusting the width and shift in the images. In this study, we applied data augmentation with the help of an image data generator from the Keras API. Figure 3 below shows the effect of data augmentation techniques on brain MRI images.

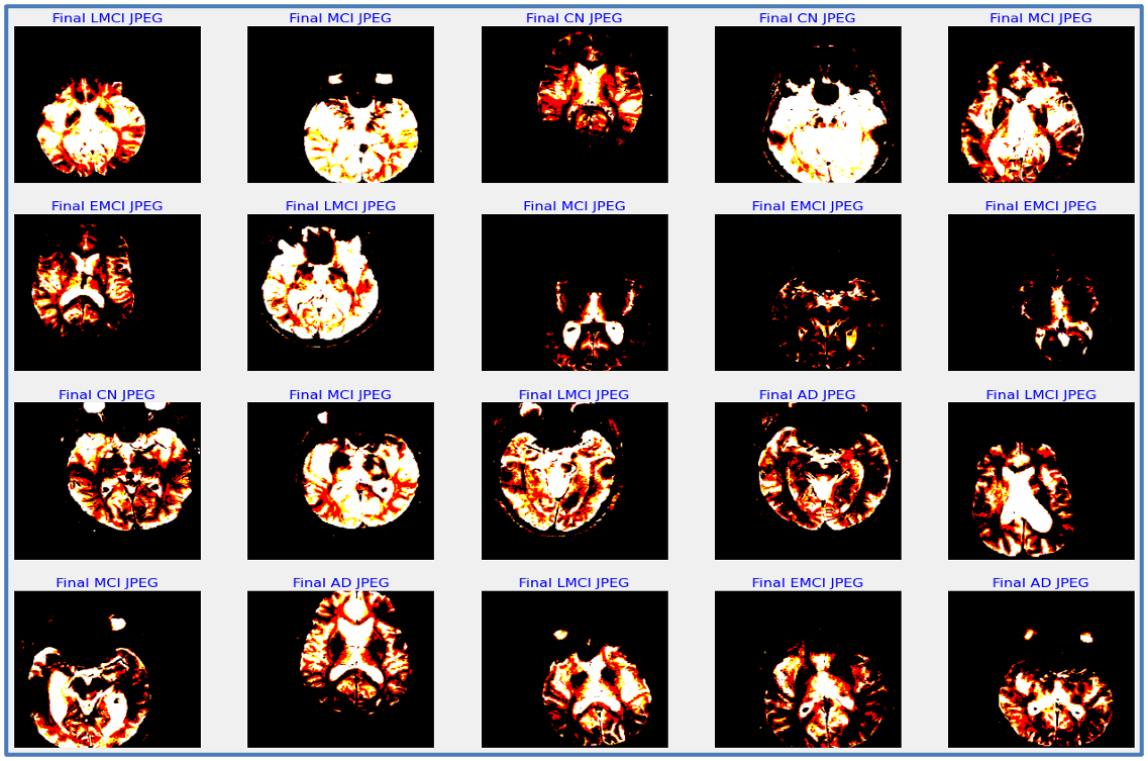

Figure 3. Application of data augmentation techniques on the public dataset

5. Result Evaluation

The dataset used in this paper is divided into testing, training, and validation data. A total of 2900 images are used in this research: 2000 images for training (400 from each class), 450 for testing (90 from each category), and 450 for validation (90 from each type). We applied transfer learning by using pre-trained CNN models such as DenseNet121 and VGG16 with ImageNet weights. For multiclass classification, we are utilizing RMSProp as our optimizer with a learning rate of 0.00001 and categorical cross-entropy as the loss metric while monitoring accuracy metrics to provide training and validation results as well as loss and accuracy values.

5.1. DenseNet121: DenseNet121 comprises one 7x7 Convolution layer, fifty-eight 3x3 Convolution layers, sixty-one 1x1 Convolution layers, four AvgPool layers, and one Fully Connected Layer (Figure 4).

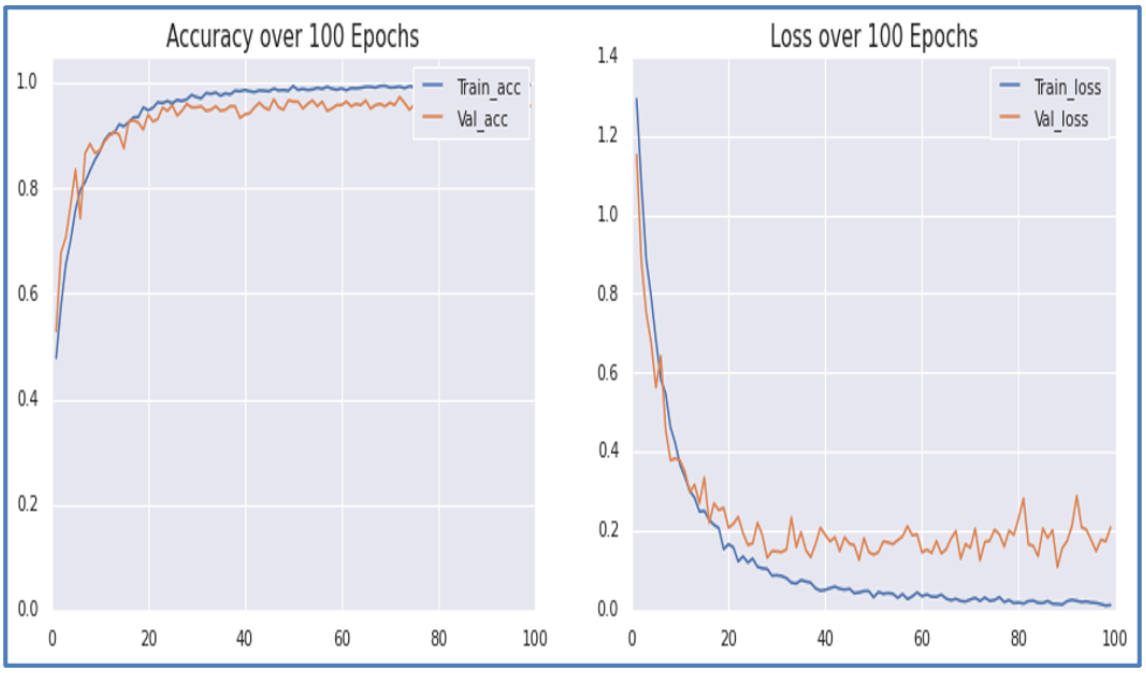

Figure 4. DenseNet121 model architecture for the prediction of Alzheimer's stages

The performance of the classification models for a particular set of test data is assessed using a confusion matrix. Figure 5 below shows the accuracy and loss plot generated by the DenseNet121 model.

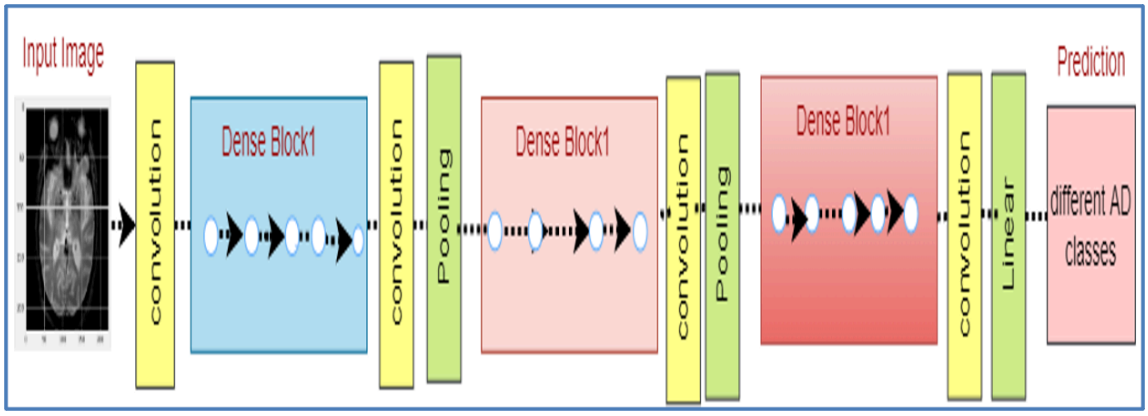

Figure 5. Accuracy and loss plot generated by the DenseNet121 model over 100 epochs.

5.2. VGG16: The VGG16 model comprises 16 layers and is implemented on an input image with dimensions (224x224) and converts it into (7x7) and five dense layer feature matrices as output. The overall accuracy of the model is 96.0. The loss and accuracy over 100 epochs are shown in Figure 6.

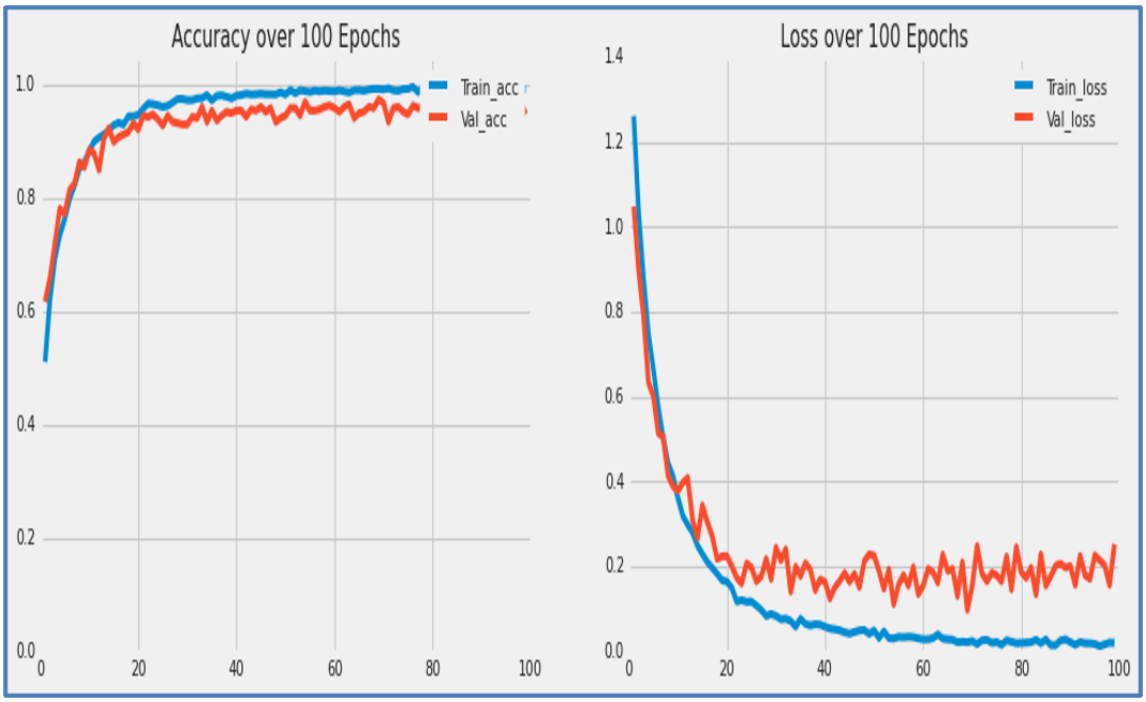

Figure 6. Accuracy and loss plot generated by the VGG 16 model over 100 epochs.

Classification Report	Precision	Recall	F1-score	support
Final AD JPEG	0.90	1.00	0.95	90
Final CN JPEG	0.94	0.89	0.91	90
Final EMCI JPEG	0.98	0.92	0.95	90
Final LMCI JPEG	0.97	0.99	0.97	90
Final MCI JPEG	0.98	0.96	0.97	90
Accuracy			0.95	450
Macro Avg.	0.95	0.95	0.95	450
Weighted Avg.	0.95	0.95	0.95	450

Table 2. Classification report generated by the VGG16 model

doi.org/10.32388/4RIOQC

5.3. ResNet50: The input image size (224x224) is converted to (7x7) by applying the ResNet50 model, which has fifty layers of convolution, and the output feature matrix has five dense layers. The Model's accuracy is measured based on different parameters such as Recall, score, Precision, etc. The basic architecture and loss plot are shown in Figure 7 and Figure 8, respectively.

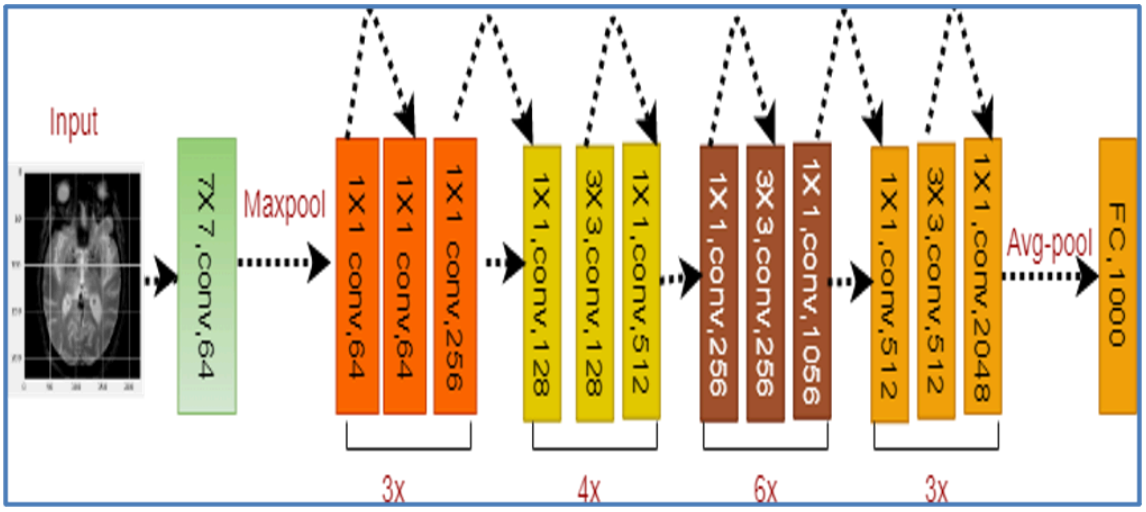

Figure 7. Basic architecture of the ResNet50 Mode for the detection of AD stages

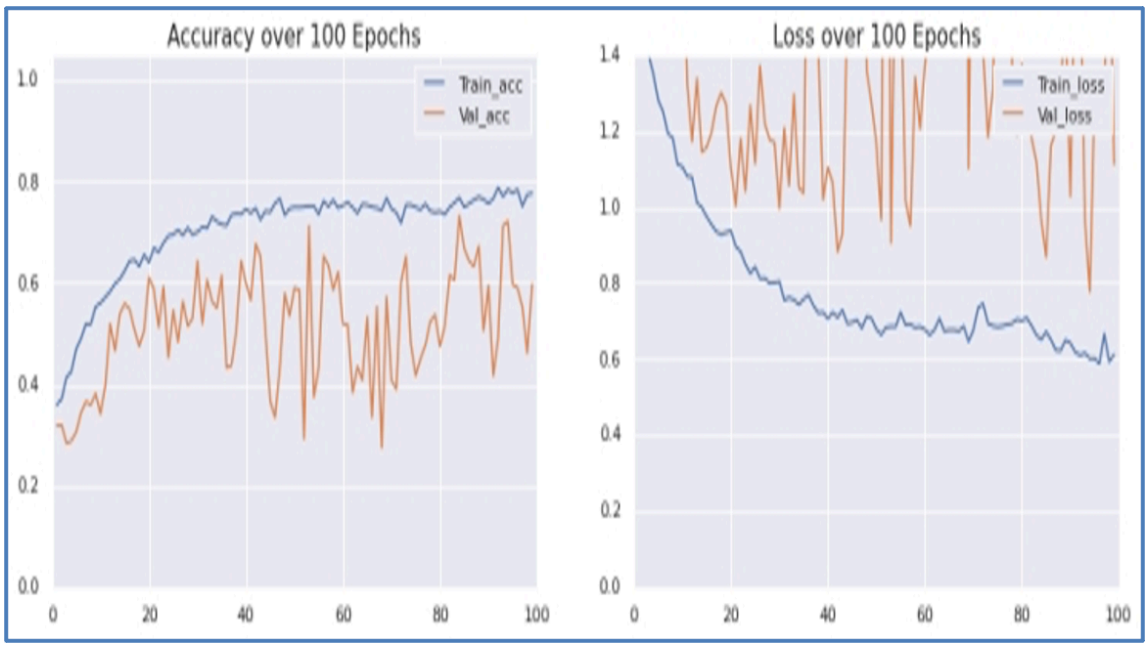

Figure 8. Accuracy and loss plot generated by the ResNet50 model over 100 epochs

doi.org/10.32388/4RIOQC

Finally, the classification report generated by the model on the specified dataset is shown in Table 3.

Classification Report	Precision	Recall	F1-score	support
Final AD JPEG	0.77	0.74	0.76	90
Final CN JPEG	0.52	0.64	0.57	90
Final EMCI JPEG	0.86	0.47	0.60	90
Final LMCI JPEG	0.49	1.00	0.66	90
Final MCI JPEG	1.00	0.22	0.36	90
Accuracy			0.62	450
Macro Avg.	0.73	0.62	0.59	450
Weighted Avg.	0.73	0.62	0.59	450

Table 3. Classification report generated by the ResNet50 model

6. Discussion and Significance of the Proposed Work

The proposed model evaluates the efficiency of models in different performance metrics, such as the confusion matrix, accuracy, loss, F1 Score, precision, recall, ROC, and sensitivity. The general formulas to calculate different parameters are calculated by the following equations.

Accuracy = (Number of Correct Predictions) / (Total Number of Predictions) (1)

Precision = (No of True Positives) / (No of True Positives + No of False Positives) (2)

Recall = (No of True Positives) / (No of True Positives + No of False Negatives) (3)

*F1-Score = 2 * (Precision * Recall) / (Precision + Recall) (4)*

For efficient classification results, precision and recall should always be high. In the present study, 3400 images from the ADNI dataset are split into groups based on the stages of Alzheimer's. For evaluation, the whole data is divided into training, testing, and validation (500, 90, 90 images from each class). The performance analysis comparison of the applied models is shown in Figure 9.

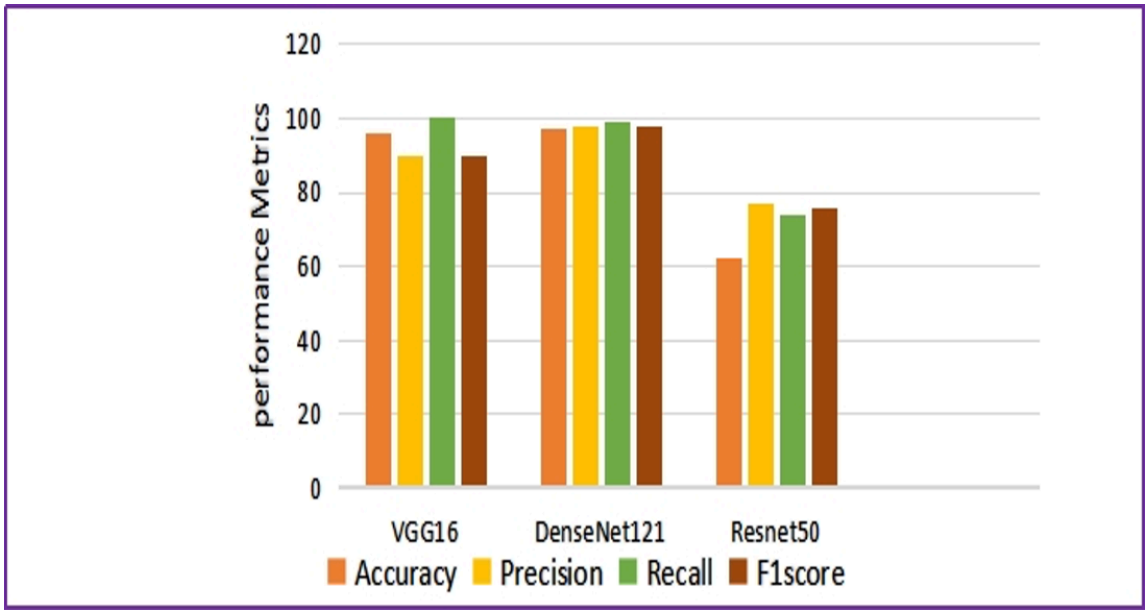

Figure 9. Comparative performance analysis generated by pre-trained deep learning models on the dataset

6. Conclusion and Future Scope

This study examined pre-trained strategies for predicting the phase of Alzheimer's disease. The highest accuracy achieved by the Model is 97.23 percent. The proposed Model operates on ADNI data using Keras API, where the MRI image is divided into five categories: EMCI, MCI, LMCI, and AD. The analysis has addressed underfitting and overfitting issues, their solutions, and the impact of Model adjustments on our application's performance. In this research, three advanced networks, VGG16, DenseNet121, and ResNet50, were used, and the results were compared. The suggested model significantly outperformed the others. In future studies, we will investigate applying the same Model to other disorders using the same data modality. The primary priority will be the enhancement of classification results during training and testing of the data.

Statements and Declarations

Author Contributions: All authors contributed equally to the conceptualization, formal analysis, investigation, methodology, and writing and editing of the original draft. All authors have read and agreed to the published version of the manuscript.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used during the current study are available from the corresponding author upon reasonable request.

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

References

1. ^*Nawaz, H., et al. (2021). A deep feature-based real-time system for Alzheimer disease stage detection. Mult imedia Tools and Applications, 80(28), 35789-35807.*
2. ^*Maurer, K., Volk, S., & Gerbaldo, H. (1997). Auguste D and Alzheimer's disease. The Lancet, 349(9064), 1546 -1549.*
3. ^*Richards, S. S., & Hendrie, H. C. (1999). Diagnosis, management, and treatment of Alzheimer's disease: A gu ide for the internist. Archives of Internal Medicine, 159(8), 789-798.*
4. ^*Jellinger, K., et al. (1990). Clinicopathological analysis of dementia disorders in the elderly. Journal of the N eurological Sciences, 95(3), 239-258.*
5. ^*Scheltens, P., et al. (2016). Alzheimer's disease. The Lancet, 388, 505-517.*
6. ^*O'Brien, J. T. (2007). Role of imaging techniques in the diagnosis of dementia. The British Journal of Radiol ogy, 80(special_issue_2), S71-S77.*
7. ^*Toews, M., et al. (2010). Feature-based morphometry: Discovering group-related anatomical patterns. Neu roimage, 49(3), 2318-2327.*
8. a, b*Xiao, T., Liu, L., Li, K., Qin, W., & Yu, S. (2018). [Online] Role of imaging techniques in the diagnosis of dem entia. BioMed Research International, 2018.*
9. ^*Acharya, U. R., et al. (2019). Automated detection of Alzheimer's disease using brain MRI images—a study with various feature extraction techniques. Journal of Medical Systems, 43(9), 1-14.*
10. ^*Wang, L., & Liu, Z.-P. (2019). Detecting diagnostic biomarkers of Alzheimer's disease by integrating gene ex pression data in six brain regions. Frontiers in Genetics, 10, 157.*
11. ^*Mahyoub, M., et al. (2018). Effective use of data science toward early prediction of Alzheimer's disease. In 2 018 IEEE 20th International Conference on High-Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (H PCC/SmartCity/DSS) (pp. 1-7). IEEE.*
12. ^*Rueda, A., Gonzalez, F. A., & Romero, E. (2014). Extracting salient brain patterns for imaging-based classifi cation of neurodegenerative diseases. IEEE Transactions on Medical Imaging, 33(6), 1262-1274.*

13. ^Li, T., & Zhang, W. (2016). Classification of brain disease from magnetic resonance images based on multi-level brain partitions. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1759-1762). IEEE.
14. ^Giraldo, D. L., et al. (2018). Characterization of anatomical brain patterns by comparing region intensity distributions: Applications to the description of Alzheimer's disease. Brain and Behaviour, 8(4), e00942.
15. ^Gupta, Y., et al. (2019). Prediction and classification of Alzheimer's disease based on combined features from apolipoprotein-E genotype, cerebrospinal fluid, MR, and FDG-PET imaging biomarkers. Frontiers in Computational Neuroscience, 13, 72.
16. ^Vickers, N. J. (2017). Animal communication: When I'm calling you, will you answer too? Current Biology, 27(14), R713-R715.
17. ^Maqsood, M., et al. (2019). Transfer learning assisted classification and detection of Alzheimer's disease stages using 3D MRI scans. Sensors, 19(11), 2645.
18. ^Chowdhary, C. L., et al. (2020). An efficient segmentation and classification system in medical images using intuitionist possibilistic fuzzy C-mean clustering and fuzzy SVM algorithm. Sensors, 20(14), 3903.
19. ^Cheung, C. Y., et al. (2022). A deep learning model for detection of Alzheimer's disease based on retinal photographs: A retrospective, multicentre case-control study. The Lancet Digital Health, 4(11), e806-e815.
20. ^Nagabushanam, D. S., Mathew, S., & Chowdhary, C. L. (2022). A study on the deviations in performance of FNNs and CNNs in the realm of grayscale adversarial images. arXiv e-prints. arXiv:2209.
21. ^Amoroso, N., et al. (2018). Deep learning reveals Alzheimer's disease onset in MCI subjects: Results from an international challenge. Journal of Neuroscience Methods, 302, 3-9.
22. ^Mirabnahrazam, G., et al. (2022). Machine learning based multimodal neuroimaging genomics dementia score for predicting future conversion to Alzheimer's disease. Journal of Alzheimer's Disease Preprint, 1-21.
23. ^Hashemifar, S., et al. (2022). DeepAD: A robust deep learning model of Alzheimer's disease progression for real-world clinical applications. arXiv preprint. arXiv:2203.09096.
24. ^Ning, Z., et al. (2021). Relation-induced multi-modal shared representation learning for Alzheimer's disease diagnosis. IEEE Transactions on Medical Imaging, 40(6), 1632-1645.
25. ^Yosinski, J., et al. (2014). How transferable are features in deep neural networks? Advances in Neural Information Processing Systems, 27.
26. ^Rayar, F. (2017). ImageNet MPEG-7 Visual Descriptors-Technical Report. arXiv preprint. arXiv:1702.00187.
27. ^Fuse, H., et al. (2018). Detection of Alzheimer's disease with shape analysis of MRI images. In 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symp

osium on Advanced Intelligent Systems (ISIS). IEEE.

28. ^*Huang, G., et al. (2017). Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.*

29. ^*Howard, A. G., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applicati ons. arXiv preprint. arXiv:1704.04861.*

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.