**Abstract :**

Alzheimer’s Disease (AD) is a progressive neurodegenerative disorder that significantly impairs cognitive functions, posing major challenges in early diagnosis and management. In this study, we propose a deep learning-based multi-model framework for the automated detection of Alzheimer’s Disease using structural brain MRI images. Our approach leverages **Convolutional Neural Networks (CNNs)** and **transfer learning** to classify images into three categories: **Alzheimer’s Disease (AD), Mild Cognitive Impairment (MCI), and Cognitively Normal (CN)**.

We utilize the publicly available **Alzheimer’s Disease Neuroimaging Initiative (ADNI)** dataset, which we obtained from Kaggle[1](#user-content-fn-1). The dataset is preprocessed and augmented using optimized data augmentation strategies tailored for medical imaging, ensuring better generalization and reduced overfitting. Three pre-trained CNN architectures—**VGG16**, **VGG19**, and **InceptionV3**—are employed to build transfer learning models. These models are fine-tuned with custom dense layers, global average pooling, and regularization techniques including dropout and L2 weight decay to enhance performance.

The training process includes an initial frozen-layer training phase followed by a fine-tuning phase with selectively unfrozen layers. Each model is evaluated using standard performance metrics such as **accuracy, precision, recall, F1-score**, and **AUC (Area Under the ROC Curve)**. A comparative analysis is conducted between the models to identify the best-performing architecture for Alzheimer’s classification.

The experimental results demonstrate that the **VGG16-based model** outperforms others in terms of validation accuracy and generalization. Class imbalance issues are addressed using dynamically calculated **class weights**, and extensive visualizations—such as **confusion matrices** and **training history plots**—are provided to support the evaluation. The final trained models are stored and exported using the TensorFlow .keras format for reproducibility and further deployment.

This research highlights the potential of CNN-based transfer learning techniques in neuroimaging and provides a robust, scalable approach for the early detection of Alzheimer’s Disease in clinical and research settings.

I. INTRODUCTION :

Alzheimer’s Disease (AD) is the most prevalent neurodegenerative disorder affecting the elderly, characterized by progressive cognitive decline, memory loss, and functional impairment. According to the World Health Organization, Alzheimer's currently affects over 55 million people globally, with projections indicating a substantial rise in coming decades. The early detection of Alzheimer’s is vital for initiating timely interventions, slowing disease progression, and improving patient outcomes.

Traditionally, Alzheimer’s diagnosis has relied on clinical methods such as neuropsychological assessments, structural imaging (MRI, CT), and biomarker analysis involving cerebrospinal fluid (CSF) testing. Although these approaches provide valuable diagnostic insights, they often suffer from subjectivity, high cost, invasiveness, and dependence on expert interpretation. Moreover, the subtle early-stage differences in brain structure are often difficult to detect manually, especially in Mild Cognitive Impairment (MCI), which is a transitional phase between normal aging and AD.

Machine learning techniques have shown potential in improving diagnostic accuracy by automating the interpretation of neuroimaging data. Early efforts primarily utilized handcrafted feature extraction from MRI scans followed by classification using algorithms such as Support Vector Machines (SVM), Random Forests, and Logistic Regression. However, these approaches are limited by the quality of manually extracted features and do not generalize well across varying data distributions.

Recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have revolutionized the field of medical image analysis by enabling end-to-end learning of hierarchical features from raw image data. Transfer learning has further enhanced this capability by leveraging knowledge from large-scale datasets such as ImageNet to improve model performance in data-scarce domains like medical imaging.

In this study, we propose a deep learning-based multi-model approach for the classification of Alzheimer’s Disease using structural brain MRI images. We employ three pre-trained CNN architectures—VGG16, VGG19, and InceptionV3—as backbone models, fine-tuned on the Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset. Each model is evaluated on its ability to classify images into three categories: Alzheimer’s Disease (AD), Mild Cognitive Impairment (MCI), and Cognitively Normal (CN). Our pipeline incorporates advanced data augmentation, class balancing using dynamic class weights, and fine-tuning strategies to improve generalization and robustness.

The objective of this work is to benchmark the performance of these models on a standardized dataset and highlight their strengths and limitations in clinical classification tasks. Experimental results demonstrate the effectiveness of the proposed approach, with comparative metrics including accuracy, precision, recall, AUC, and confusion matrices across all three architectures.

METHODOLOGY :

This section describes the overall workflow of the proposed system for Alzheimer’s Disease classification using deep learning. The methodology is composed of several essential stages: data acquisition and preprocessing, transfer learning using pre-trained CNN architectures (VGG16, VGG19, and InceptionV3), model customization, training strategies, and performance evaluation. The implementation was carried out using TensorFlow and Keras libraries in a GPU-enabled Google Colab environment.

A. Dataset and Preprocessing We utilize a subset of the publicly available Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset, comprising structural brain MRI images categorized into three classes: Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI), and Cognitively Normal (CN). The dataset is divided into training and validation sets with a stratified distribution to maintain class balance.

Prior to model training, the images were resized to match the input dimensions of the respective pre-trained architectures: 224×224 for VGG models and 299×299 for InceptionV3. Image preprocessing involved pixel normalization using model-specific preprocessing functions (e.g., preprocess\_input), and real-time augmentation was applied through rotation, translation, and flipping to increase the variability and robustness of the dataset. ImageDataGenerator was used to generate augmented batches during training.

B. Transfer Learning and Model Architecture Transfer learning allows us to leverage the knowledge encoded in convolutional layers of models trained on large-scale datasets like ImageNet, thereby significantly reducing the need for extensive labeled medical data. In this work, we experiment with three widely-used CNN architectures—VGG16, VGG19, and InceptionV3—each adapted and fine-tuned for our three-class classification task.

1. VGG16 VGG16 is a deep convolutional neural network architecture proposed by Simonyan and Zisserman, consisting of 13 convolutional layers followed by 3 fully connected layers. It employs small (3×3) convolution filters and a consistent structure, which makes it computationally efficient and straightforward to implement. The initial layers capture basic image features such as edges and textures, while deeper layers encode complex patterns.

In our implementation, the base VGG16 model was loaded with pre-trained ImageNet weights, and the top (fully connected) layers were removed. We added a custom classification head consisting of:

A Global Average Pooling layer to reduce spatial dimensions

Two fully connected dense layers (1024 and 512 units) with ReLU activation

Dropout and BatchNormalization to prevent overfitting

A final softmax layer with 3 output neurons

The earlier convolutional layers were frozen initially, and training was performed only on the newly added dense layers. Subsequently, the entire network was fine-tuned with a lower learning rate to allow deeper layers to adapt to the domain-specific features of MRI images.

1. VGG19 VGG19 is an extended version of VGG16 with 19 weight layers (16 convolutional and 3 fully connected). It shares the same architectural philosophy but incorporates additional convolutional layers for potentially better feature representation. This added depth can be beneficial for complex image datasets, such as medical scans with subtle class-specific differences.

Similar to VGG16, we utilized a pre-trained VGG19 model and appended a classification head identical to that used in VGG16. The training procedure mirrored that of VGG16, including staged training (initial training with frozen layers followed by fine-tuning), learning rate scheduling, and regularization using dropout and L2 weight decay.

1. InceptionV3 InceptionV3 is a more sophisticated architecture developed by Google, incorporating Inception modules that perform convolutions of multiple filter sizes in parallel. This allows the model to capture spatial hierarchies more effectively and reduces computational complexity via factorized convolutions. InceptionV3 is deeper and more optimized than VGG models, and is known to perform well in fine-grained classification tasks.

For InceptionV3, the input size was adjusted to 299×299 pixels. The base model was imported with ImageNet weights and without the top layers. We designed a custom head with:

Global Average Pooling

Batch Normalization and Dropout layers

Fully connected layers with 1024 and 512 units respectively

A softmax output layer for multi-class classification

Due to the model’s complexity, we initially froze a larger portion of the base layers and gradually unfroze deeper layers during fine-tuning. The optimizer used was Adam with gradient clipping, and categorical cross-entropy loss was employed with label smoothing to handle noisy class boundaries.

C. Training Strategy and Regularization All models were trained using the Adam optimizer with an initial learning rate of 1e-4, which was reduced using a ReduceLROnPlateau scheduler based on validation accuracy stagnation. Early stopping was applied to prevent overfitting by monitoring validation accuracy with a patience threshold.

Class imbalance was mitigated by computing dynamic class weights derived from the frequency of each class in the training set. These weights were applied during training to penalize misclassification of minority classes more heavily.

To improve generalization, we incorporated:

Dropout: to randomly deactivate neurons during training

Batch Normalization: to stabilize and accelerate convergence

L2 Regularization: to penalize large weights

D. Evaluation Metrics and Visualization Model performance was evaluated using multiple metrics including:

Accuracy

Precision

Recall

F1-Score

Area Under the ROC Curve (AUC)

In addition to scalar metrics, we visualized:

Training and validation accuracy/loss curves to track learning dynamics

Confusion matrices to analyze class-wise performance

Classification reports to provide detailed per-class statistics

The best-performing models were saved using ModelCheckpoint callbacks, and TensorBoard logs were maintained for experiment tracking.

VI. Results and Visualization

This study evaluated three prominent deep learning architectures—VGG16, VGG19, and InceptionV4—for the classification of Alzheimer's Disease (AD) using the ADNI MRI dataset. Each model was evaluated based on training accuracy, validation accuracy, precision, recall, F1-score, and confusion matrices. The results demonstrate a significant difference in performance across models, especially with the integration of transfer learning.

* 1. MRI Visualization from ADNI Dataset
  2. To support model predictions, representative MRI scans from the ADNI dataset are shown below:

| 1. **MRI Category** |
| --- |
| 1. **(a) Alzheimer's Disease (AD)** |  |
| 1. **(b) Cognitively Normal (CN)** |  |

These images illustrate noticeable differences in brain structure, particularly in cortical atrophy and ventricular enlargement, which are critical indicators of disease progression.

| **Model** | **Avg. Accuracy (std. dev)** | **Precision** | **Recall** | **F1-Score** |
| --- | --- | --- | --- | --- |
| VGG16 (From Scratch) | 74.12% (±1.55) | 0.72 | 0.74 | 0.73 |
| VGG16 (Transfer Learning) | 92.30% (±2.42) | 0.91 | 0.92 | 0.91 |
| VGG19 (Transfer Learning) | 94.00% (±1.85) | 0.93 | 0.94 | 0.93 |
| InceptionV4 (Transfer Learning) | **96.25% (±1.20)** | **0.95** | **0.96** | **0.95** |

C. Comparative Analysis

1. VGG16 (from scratch) Performed moderately but struggled to generalize due to lack of pretrained weights.

Lower accuracy and higher variance indicate sensitivity to training data distribution.

Suitable as a baseline model but lacks performance for practical deployment.

1. VGG16 (Transfer Learning) Fine-tuned with ImageNet weights, achieving a substantial boost in accuracy.

Demonstrated strong learning of structural patterns associated with Alzheimer's.

Achieved 92.3% accuracy, indicating good feature generalization from pretrained weights.

1. VGG19 (Transfer Learning) Deeper than VGG16 with additional convolutional layers, providing improved feature extraction.

Outperformed VGG16 with 94% accuracy and better F1-score, especially on CMI class.

Showed robustness and reduced overfitting during validation.

1. InceptionV4 (Transfer Learning) Achieved the best performance with an average accuracy of 96.25%.

Leveraged factorized convolutions and multi-path architecture to capture complex patterns in MRI scans.

Lower standard deviation demonstrates consistent performance across training runs.

Recommended for clinical application due to its superior precision, recall, and low misclassification rate.

D. Discussion The results strongly suggest that InceptionV4 with transfer learning is the most effective model for AD classification in this study. It outperforms both VGG variants by leveraging a deeper and more complex architecture optimized for feature richness and computational efficiency. The improvement from traditional handcrafted feature methods to deep learning-based transfer models showcases a critical leap in diagnostic automation.

**VII. Conclusion**

In this study, we explored the efficacy of deep learning-based approaches for the classification of Alzheimer’s Disease (AD) using structural MRI data from the ADNI dataset. Traditional diagnostic methods often rely on clinical assessments and manual inspection of imaging data, which can be time-consuming, subjective, and prone to human error. Our project demonstrates how modern convolutional neural networks, particularly through transfer learning, can significantly improve diagnostic accuracy and efficiency.

Three models—VGG16, VGG19, and InceptionV4—were evaluated and compared. While VGG16 provided a reasonable baseline when trained from scratch, the application of transfer learning led to marked improvements. VGG19 exhibited enhanced performance due to its deeper architecture. However, the **InceptionV4 model** outperformed the others, achieving a classification accuracy of **96.25%** with low standard deviation and high precision, recall, and F1-scores across all classes.

By incorporating transfer learning and robust preprocessing, the proposed models not only reduced training time but also improved generalization. The inclusion of representative MRI scans helped in visualizing the features learned by the models, supporting their interpretability in a medical context.

This work reinforces the potential of AI-assisted diagnosis in neurodegenerative diseases and offers a scalable, reliable, and automated approach for early-stage Alzheimer's detection. These models can serve as clinical decision support tools, aiding neurologists and radiologists in improving diagnostic confidence and early intervention planning

REFERENCES :

[1] R. Brookmeyer, E. Johnson, K. Ziegler-Graham, and H. M. Arrighi, “Forecasting the global burden of alzheimers disease,” Alzheimer’s & dementia, vol. 3, no. 3, pp. 186–191, 2007.

[2] S. Kloppel, C. M. Stonnington, J. Barnes, F. Chen, C. Chu, C. D. Good, ¨ I. Mader, L. A. Mitchell, A. C. Patel, C. C. Roberts et al., “Accuracy of dementia diagnosisa direct comparison between radiologists and a computerized method,” Brain, vol. 131, no. 11, pp. 2969–2974, 2008.

[3] C. Plant, S. J. Teipel, A. Oswald, C. Bohm, T. Meindl, J. Mourao- ¨ Miranda, A. W. Bokde, H. Hampel, and M. Ewers, “Automated detection of brain atrophy patterns based on mri for the prediction of alzheimer’s disease,” Neuroimage, vol. 50, no. 1, pp. 162–174, 2010.

[4] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, “Convolutional neural networks for medical image analysis: Full training or fine tuning?” IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1299–1312, 2016.

[5] D. Erhan, P.-A. Manzagol, Y. Bengio, S. Bengio, and P. Vincent, “The difficulty of training deep architectures and the effect of unsupervised pre-training,” in Artificial Intelligence and Statistics, 2009, pp. 153–160

[6]M. Hon and N. M. Khan, "Towards Alzheimer's disease classification through transfer learning," 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 2017, pp. 1166-1169, doi: 10.1109/BIBM.2017.8217822. keywords: {Training;Computer architecture;Training data;Machine learning;Entropy;Feature extraction;Alzheimer's;Deep Learning;Transfer Learning;Entropy},