Classifying Alzheimer's Disease Stages Using Convolutional Neural Networks on the OASIS MRI Dataset

Aditya Kumar
Chitkara University Institute of Engineering & Technology
Chitkara University
Rajpur, Punjab, India
adityakd.1996@gmail.com

Leema Nelson
Chitkara University Institute of Engineering & Technology
Chitkara University
Rajpur, Punjab, India
leema.nelson@chitkara.edu.in

2024 International Conference on Innovation and Novelty in Engineering and Technology (INNOVA) | 979-8-3315-0513-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/INNOVA63080.2024.10847023

***Abstract*—Alzheimer's Disease (AD) is a progressive neurodegenerative disorder, where early detection is crucial for improving patient outcomes and slowing disease progression. This study presents a Convolutional Neural Network (CNN)-based model to classify AD stages: Mild Dementia, Moderate Dementia, Very Mild Dementia, and Non-Dementia using the OASIS MRI dataset. The dataset comprises 86,437 MRI scans, with significant class imbalance addressed through weighted sampling and data augmentation. Training (60,506 images), validation (12,965 images), and testing (12,967 images) subsets were used to ensure balanced learning and robust evaluation. The model effectively captured structural patterns unique to each AD stage, achieving a high test accuracy of 98.80%, along with excellent precision, recall, and F1 scores across all classes. These findings highlight the potential of deep learning as a reliable tool for early detection and precise staging of Alzheimer's Disease, especially in distinguishing challenging cases like Non-Demented and Moderate Dementia.**

***Keywords*—Alzheimer's Disease, MRI, Convolutional Neural Network, CNN, Deep Learning, Early Detection, OASIS Dataset, Brain Imaging, Healthcare, Dementia Classification.**

I. Introduction

A. Background and Context

Alzheimer's Disease (AD) is one of the most common neurodegenerative disorders, primarily affecting older adults. Characterized by a progressive decline in cognitive function, it leads to memory loss, impaired reasoning, and behavioral changes. As the disease proceeds, patients' quality of life falls significantly since they are unable to perform regular activities. According to the World Health Organisation, Alzheimer's is the leading cause of dementia, which affects up to 60–70% of cases worldwide. Early diagnosis and intervention become even more crucial given that an aging population is expected to significantly increase the global incidence of Alzheimer's [1].

Alzheimer's is connected to pathogenic changes in the brain, including amyloid-beta plaques, neurofibrillary tangles, and substantial brain atrophy, especially in the hippocampus region, which regulates memory and learning. Usually, before the start of symptoms and early on in the disease, these structural changes are abundantly evident. This has stimulated the investigation of numerous neuroimaging techniques for Alzheimer's disease diagnosis; MRI is one of the most intriguing tools for early identification. MRI can show high-resolution images of the brain's architecture and structural defects connected with AD, including hippocampal atrophy, cortical thinning, and enlarged ventricles. MRI has shown to be a valuable tool. However, hand interpretation of these images can be labor-intensive and prone to human error. Furthermore, early-stage AD is distinguished by minute structural abnormalities that, with merely visual inspection, are sometimes difficult to detect. This has led to the merging of machine learning and deep learning algorithms to automatically recognize MRI images, therefore considerably improving diagnosis accuracy and allowing early intervention [2].

B. Problem Statement

Even with advances in diagnostic tools, Alzheimer's remains famously challenging early on in the illness. Most traditional diagnostic methods—including cognitive tests and patient history reviews—often overlook the disease in its earliest phases when care would be most helpful. By the time Alzheimer's is typically identified, patients have frequently already exhibited notable cognitive decline, thereby limiting the options for treatment. Although MRI and other neuroimaging methods identify structural markers of AD, manual analysis is insufficient for exact early diagnosis. Particularly in cases with minor structural changes, the manual review of MRI images requires substantial knowledge and may nonetheless lead to diagnostic uncertainty. Moreover, the significant volume of data generated in MRI scans makes it practically impossible for clinicians to review each image closely. In particular, convolutional neural networks and deep learning techniques have shown considerable potential here. CNNs have been particularly adopted for image classification tasks in the medical field, especially for diseases with clear pathogenic indications, including Alzheimer's [3]. They can automatically learn complex patterns from images.

Still, current approaches for MRI-based Alzheimer's classification have flaws. Most models find it challenging to generalize from limited datasets, class imbalance, and overfitting. Moreover, essential for early diagnosis and treatment is the capacity of models to distinguish AD from mild cognitive impairment (MCI). More improved models can accurately detect Alzheimer's at multiple periods, including early MCI detection using MRI scans. Moreover, even though many studies focus on maximizing classification accuracy in artificial intelligence-driven medical solutions, interpretability and openness in these solutions are growing crucial. In healthcare applications, explainable artificial intelligence techniques are particularly crucial since doctors are more likely to embrace AI models if they can understand and trust the process of decision-making [4].

C. Relevance of the Topic

The importance of early detection of Alzheimer's Disease cannot be overstated. Early intervention promises both better patient outcomes and a slowing down of the course of the illness. While Alzheimer's cannot be cured early on in the

disease, treatments can help reduce symptoms and improve quality of life. Thus, addressing the growing global Alzheimer's burden calls for the creation of exact, efficient, scalable diagnostic methods. Given the possibility of MRI indicating early indications of Alzheimer's disease, evaluating MRI images using deep learning approaches is a plausible decision for early diagnosis augmentation. Alzheimer's could be found utilizing automated, non-invasive diagnostic tools integrating neuroimaging with artificial intelligence, therefore enabling doctors to make more accurate diagnoses. Having already proven success in several medical imaging uses, including cancer, retinal diseases, and heart difficulties, CNNs offer promise to assist with the diagnosis of Alzheimer's Disease [5].

The rapid advancements in deep learning technologies draw attention to the importance of this effort even more. Big datasets and increasing computing resource availability offer a credible possibility of training robust CNN models capable of achieving great diagnostic accuracy. Furthermore, the use of innovative techniques, including data augmentation, transfer learning, and model interpretability, will enable the model to generalize over many datasets, hence reducing the risk of overfitting and guaranteeing clinical relevance. From the perspective of Alzheimer's research, this finding closes a significant gap in the field. Many contemporary models have been limited by a lack of explainability and scalability. By focusing on developing a CNN-based model for Alzheimer's classification that prioritizes both accuracy and interpretability, this work intends to allow doctors to trust and include tools in their diagnostic processes [6].

D. Objectives of the Research

The main goal of this work is to build a strong deep-learning model that is able to correctly identify Alzheimer's disease using MRI data. Especially differentiating between classes of normal, mild cognitive impairment and Alzheimer's Disease, the model will be improved to identify structural brain abnormalities linked with Alzheimer's. MCI is a transitory phase in which early intervention can assist in stopping the spread of diseases, so this classification is crucial. Focussing on these variations, the method aims to assist physicians with early diagnosis and treatment planning. Further crucial objectives are to generalize the model over several sets. This effort will artificially enlarge the size of the dataset using rotation, zooming, and flipping as data augmentation methods, thereby enabling the model to acquire more full features and avoid overfitting. The goal is to establish a model whose performance ensures not only on the training set but also on unseen data, therefore guaranteeing its universal relevance in clinical settings.

Furthermore, this work tries to maximize hyperparameters like the number of layers, batch size, and learning rate, guaranteeing the performance of the model. In order to prevent overfitting among other regularisation techniques, early halting, dropout, and batch normalization will be used. Finally, the performance of the suggested model will be assessed against current models to guarantee that it provides dependability and accuracy for Alzheimer's disease diagnosis. This work intends to produce a powerful, scalable, interpretable deep learning model for Alzheimer's Disease categorization by addressing these targets. Both the academic community and clinical practice depend much on this finding since the integration of deep learning with MRI imaging offers significant opportunities to improve patient outcomes and revolutionize early diagnosis.

II. LITERATURE REVIEW

Raza et al. (2023) explored the segmentation and classification of AD using MRI images through customized CNN architectures. The work focussed primarily on segmented images of the Grey Matter (GM) in the brain in order to raise detection accuracy. Rather than starting from nothing in training the model, the researchers used transfer learning to enhance a pre-trained deep learning model for Alzheimer's classification. Over 10, 25, and 50 epochs, the accuracy of the model was evaluated to produce an overall accuracy of 97.84% [7].

Shamrat et al. (2023) proposed AlzheimerNet, a fine-tuned CNN classifier aimed to identify five phases of AD along with the Normal Control (NC) classifier. The researchers applied CLAHE image enhancement and data augmentation using the ADNI MRI dataset to fix the class imbalance, therefore generating a 60,000-image dataset. Five previously published models were tested: InceptionV3 scored 96.31% accuracy. Built from fine-tuning InceptionV3 using an RMSprop optimizer and a learning rate of 0.00001, AlzheimerNet obtained a test accuracy of 98.67%). A two-tailed Wilcoxon signed-rank test validated AlzheimerNet's improved performance, $p < 0.05$ [8].

Murugan et al. (2021) proposed a DL approach for AD stage detection utilizing MRI scans by use of CNNs. Working with constraints in traditional diagnosis methods, the authors developed DEMNET, a structure designed to precisely classify four phases of dementia. Using the AD Neuroimaging Initiative, the model corrected the Kaggle dataset class imbalance issues. With an AUC of 97%, a Cohen's Kappa value of 0.93, and an accuracy of 95.23%, DEMNET exceeded past methods and provided accurate visualizations of Alzheimer's risk [9].

Al Shehri et al. (2022) proposed a DL-based model for AD classification using ResNet-50 and DenseNet-169 architectures. The model distinguishes Alzheimer's from other disorders as moderate, very mild, mild, non-dementia, or moderate dementia. DenseNet-169 outperformed ResNet-50 by 0.8870 and 0.8192, respectively, with training and testing accuracy of 0.977 and 0.843, respectively [10].

Helaly et al. (2022) developed an end-to-end framework for the classification of AD using CNNs on the ADNI dataset. The work used two approaches: VGG19 model-based transfer learning and basic CNN architectures for 2D and 3D picture classification. While the CNN models obtained accuracies of 93.61% and 95.17% for 2D and 3D classifications, respectively, the fine-tuned VGG19 model attained 97% accuracy [11].

Janghel et al. (2020) investigated DL techniques for Alzheimer's Disease diagnosis using CNNs on MRI scans. Examining four architectures— Lenet-5, AlexNet, ZFNet, and R-CNN—seeking to exceed traditional MMSE score and manual MRI examination. ZFNet reached the best accuracy with 97.68% for 75–25 cross-valuation and 98.75% for 90–10 cross-valuation [12].

III. METHODOLOGY

This section provides a detailed description of the dataset, data augmentation techniques, model architecture, and the

proposed work to develop and evaluate the CNN model for Alzheimer's classification using MRI images, as shown in Fig. 1.

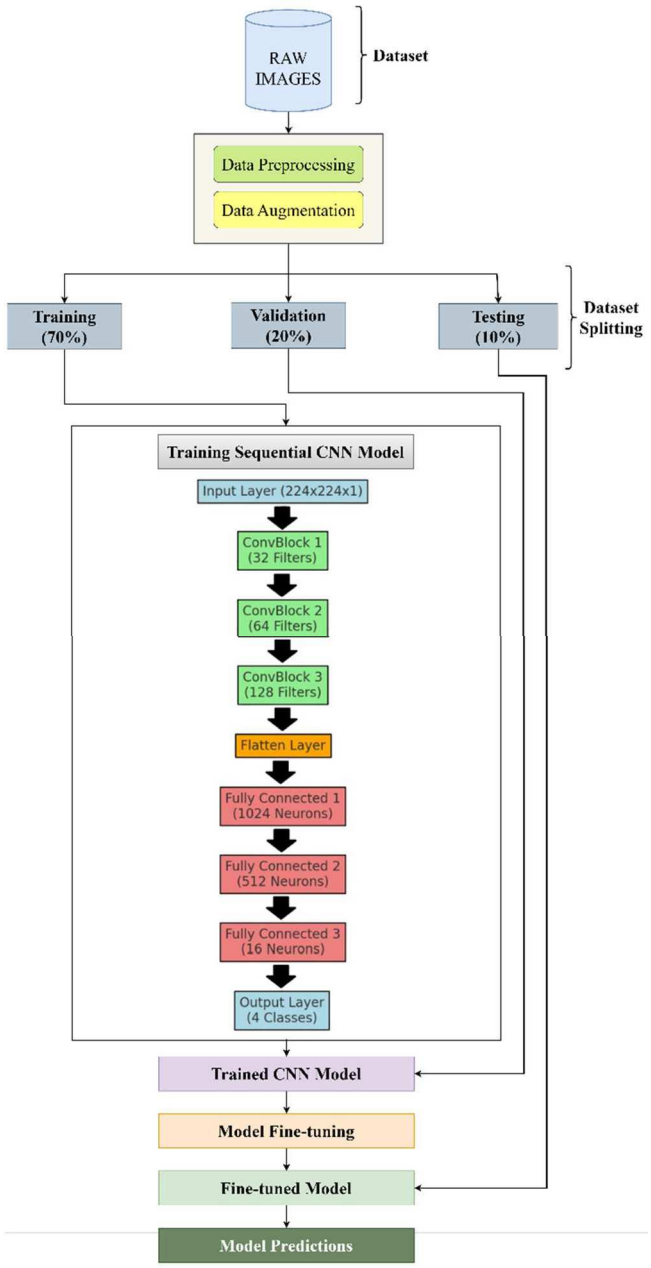

Fig. 1. Proposed Methodology

A. Dataset Description

Source of the OASIS MRI dataset used for this study: Kaggle repository [13]. There are 86,437 brain MRI images in all, divided into four classifications according to varying phases of Alzheimer's disease progression: Moderate Dementia, Very Mild Dementia, Non-Demented, and Mild Dementia, as shown in Fig. 2. Deep learning algorithms aiming at early detection of Alzheimer's Disease by studying structural changes in the brain depend critically on this dataset. With most of the photos categorized as Non-Demented (67,222 images), the class distribution of the dataset is clearly skewed. Under-represented are the other groups: Mild Dementia (5002 images), Very Mild Dementia (13,725 images), and Moderate Dementia (488 images). Table I depicts a thorough analysis of the class distribution. Data augmentation and weighted sampling methods are used to guarantee that the deep learning model can efficiently learn patterns across all classes without bias, hence addressing this disparity.

The dataset representation emphasizes the several phases of dementia, as shown in the MRI images. As Fig. 3 shows, the dataset is also separated into three subsets for testing, validation, and training. Training uses 60,506 images specifically; validation uses 12,965 images; and testing uses 12,967 images. This all-encompassing division guarantees that the model is rigorously tested at every level of development, therefore enabling the development of a strong and generalizable model that is able to correct Alzheimer's categorization.

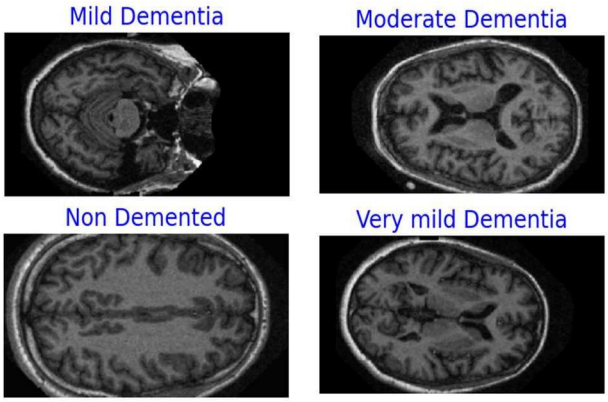

Fig. 2. Dataset Representation.

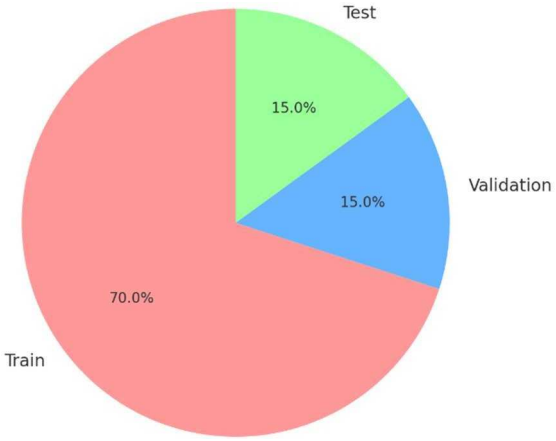

Fig. 3. Dataset Distribution.

TABLE I. CLASS DISTRIBUTION OF IMAGES IN THE DATASET.

Class	*Total Images*	*Training Set*	*Validation Set*	*Testing Set*
Mild Dementia	5002	3438	765	799
Moderate Dementia	488	333	78	77
Non Demented	67222	47111	10088	10023
Very mild Dementia	13725	9623	2034	2068

B. Data Augmentation and preprocessing

Based on MRI scans, the model intended for Alzheimer's classification consists of convolutional neural networks (CNN) with many convolutional and fully linked layers. The

architecture consists of four conv blocks with two convolutional layers, batch normalization, and ReLU activation functions. These layers, from the MRI scans, first extract low-level and then high-level traits. Max pooling downsamples the feature maps after every ConvBlock, reducing spatial dimensions and maintaining prominent features. As the number of filters increases with each ConvBlock, the model captures progressively more complex information essential for separating between the phases of Alzheimer's Disease from 32 filters up to 512.

After the convolutional layers, the output is flattened and passed through three fully linked layers (linear blocks). The first layer comprises 1024 neurons; the second layer has 512 neurons; the last fully connected layer reduces the size to 16 neurons. Globally, statistical features, mean, and standard deviation generated from the input images are mixed with the output of the fully linked layers to enhance performance. Using a softmax layer, the four Alzheimer's categories, Moderate Dementia, Very Mild Dementia, Non-Demented, and Mild Dementia, have class probabilities predicted.

The Adam optimizer is used to optimize variable learning rates. Early halting and dropout regularisation help to avoid overfitting; learning rate scheduling provides seamless convergence and fine-tunes the learning process. These techniques assure accuracy as well as generalism.

C. Model Architecture

The proposed model architecture for Alzheimer's classification is based on a CNN designed to capture structural patterns in brain MRI data effectively. The design ensures uniformity in data dimensions, starting with an input layer that manages MRI images shrunk to 224x224 pixels. Using filters to extract low-level properties such as edges and textures, several convolutional layers, which progressively build up to higher-level representations of brain areas, use ReLU activation functions to complement every convolutional layer to offer non-linearity and max pooling layers to downsample the feature maps, hence reducing dimensionality while keeping important information. Batch normalization helps to increase generalization and straightens the learning process. The flattened retrieved features pass via fully linked layers to further refine the representations. From the last softmax layer, class probabilities for the four Alzheimer's types, Non-Demented, Very Mild Dementia, Mild Dementia, and Moderate Dementia, outputs overfitting is avoided by dropout regularisation; for efficient and flexible learning, the Adam optimizer initial learning rate was 0.001 with learning rate schedule is utilized. This architecture preserves computational economy by balancing performance with complexity, therefore guaranteeing strong classification.

D. Experimental Setup

The experiments were conducted using Python and TensorFlow software environment. Training and evaluation were done on an NVIDIA V100 GPU with 32 GB of VRAM to address the computational demands required by the OASIS MRI dataset. The training involved ten epochs with a batch size of 64. Adam optimizer with an initial learning rate 0. 001 was applied with a ReduceLROnPlateau scheduler. The software environments used in the current study were Python 3.8, TensorFlow 2.9, and CUDA 11.2 for GPU processing. For the purpose of reproducing experiments, random number generators for TensorFlow and NumPy libraries were locked.

E. Proposed Work

The proposed work aims to develop a DL-based system for classifying AD stages using MRI images, focusing on improving diagnostic accuracy and generalization. The model will use CNN architecture, which is especially meant to simulate complicated brain areas related to different phases of dementia. CNN will classify MRI images into four: mild, moderate dementia, very mild, and non-demented. The proposed model is novel in that it integrates taught convolutional features with global statistical properties, such as mean and standard deviation, therefore improving the model's capacity to capture localized and global patterns in the brain.

This work solves class imbalance in the dataset using weighted random sampling and data augmentation techniques to guarantee the appropriate representation of minority classes (such as Moderate Dementia) during training. Among data augmentation techniques that will artificially expand the dataset and, therefore, raise the generalizing capability of the model are rotation, zooming, and flipping. Furthermore, advanced optimization techniques such as dropout regularisation, learning rate scheduling, and Adam Optimiser are employed to prevent overfitting and increase training effectiveness.

The model will be evaluated in part by recall, precision, F1-score, and accuracy. The proposed research provides doctors with an automated, reliable tool for early diagnosis and disease monitoring, thereby significantly improving the field of medical imaging and Alzheimer's detection.

IV. Results and Discussion

This study thoroughly evaluates the CNN model's performance in classifying Alzheimer's Disease stages from MRI images, focusing on key criteria such as accuracy and loss plots, confusion matrix, and performance metrics to assess its overall effectiveness.

A. Accuracy and Loss plots

The accuracy graph over the ten epochs indicates a continual rise in both training and validation accuracy, that is, a persistent increase, as shown in Fig. 4. The first epoch training accuracy for the model is 94.3%; validation accuracy starts at 96.9%. As training proceeds, the accuracy rises rapidly; by the tenth epoch, it is 99.7% for training and 98.8% for validation. The little variation in training and validation accuracy points to a model that fits unseen data quite nicely. The constantly high validation accuracy, which peaks at 99.5%, indicates the model is not overfitting since the validation set's performance closely follows the training set. The accuracy stabilizes after the fourth epoch, suggesting that early in the training process, the model converged and achieved optimal performance.

The loss plot over the epochs demonstrates a perfect decline in both training and validation loss, as shown in Fig. 5. Starting at the training loss of 0.0143 in the first epoch and validation loss of 0.0128, the losses keep progressively dropping. By the ninth epoch, the training loss reduces to 0.0117; the validation loss settles at about 0.0118. The very modest difference between training and validation losses supports the non-overfitting of the model even more. Though they exhibit slight variations, particularly in the validation loss between epochs 9 and 10, generally, the losses remain consistent, so demonstrating good model learning.

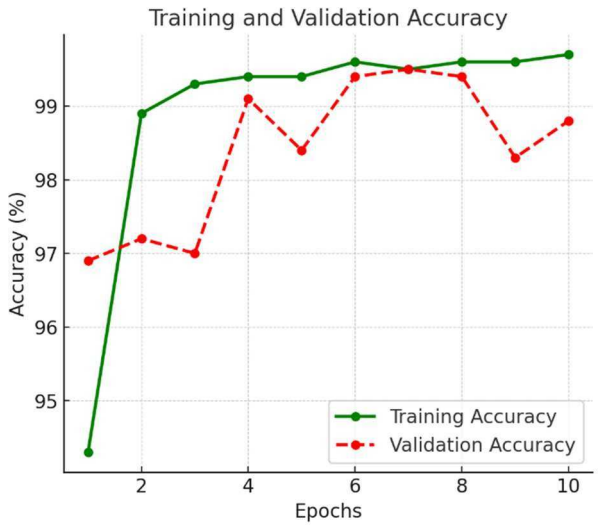

Fig. 4. Training and Validation Accuracy

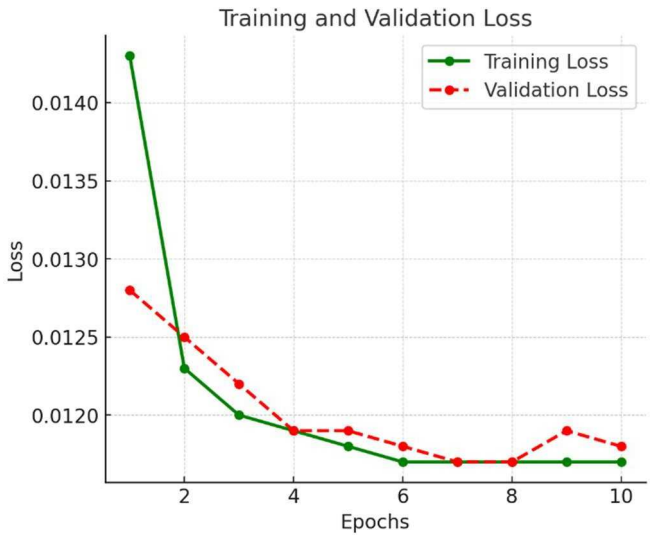

Fig. 5. Training and Validation Loss

B. *Confusion matrix:*

The confusion matrix, shown in Fig. 6, provides a detailed breakdown of the classification performance across the four classes: Moderate Dementia, Very Mild Dementia, Non-Demented, and Mild Dementia. The model performs well in spotting Non-Demented events, generating a minimal misclassification rate with 6726 properly classified cases out of 6737. Although 15 cases are misclassified as Non-Demented, the model performs well again, correctly identifying 460 cases in the Mild Dementia class.

Although rare in the sample, the model suggested remarkable accuracy and recall for Moderate Dementia since it successfully recognized 45 instances with no misclassifications to other classes. There was a little issue with Very Mild Dementia since 1310 instances were correctly detected, but 78 cases were misclassified as Non-Demented. This misclassification pattern could indicate some overlap between the criteria of Very Mild Dementia and Non-Demented categories, given the little differences in brain structure at different periods of the disease.

The model performs very well generally, notably in the Non-Demented and Moderate Dementia categories, with minor misclassifications mostly within closely related classes. The results suggest that more development should focus on separating Very Mild Dementia from Non-Demented people.

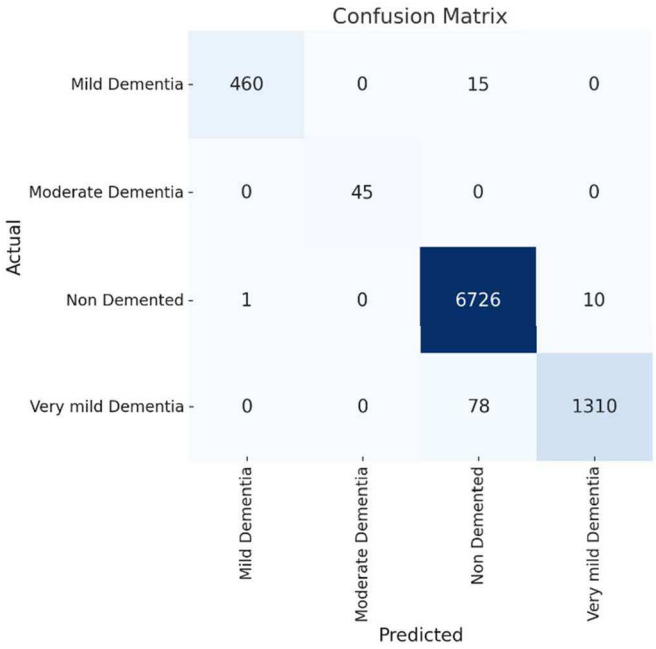

Fig. 6. Confusion matrix

C. *Performance Parameters:*

Table II indicates that the model works rather well over all classes and obtains an accuracy of 98.8%. The model received a precision of 1.00 for the Mild Dementia class, implying that all Mild Dementia predictions were accurate; the recall was 0.97, indicating that 97% of true Mild Dementia cases were suitably diagnosed. Having an F1 score of 0.98, this class exhibits uneven performance. With perfect accuracy, recall, and F1-score of 1.00, the model in the Moderate Dementia class performed very well despite their reduced presence in the sample, thereby exactly identifying all occurrences. For the majority class Non-Demented, the model did pretty well with a precision of 0.99 and a recall of 1.00, suggesting practically all Non-Demented occurrences were properly classified with rather few false positives. With 0.99 the F1-score, the model handles this class with strength. Regarding Very Mild Dementia, recall is 0.94, and precision is 0.99, meaning that some cases were misclassified as Non-Demented even if most were correctly recognized. Still performing rather well in this class; the F1 score is 0.97. With a macro and weighted average F1-score of 0.99, the model sufficiently manages imbalanced class distributions and makes correct predictions, displaying regular strong performance across all classes.

TABLE II. PERFORMANCE PARAMETERS

Class	*Precision*	*Recall*	*F1-Score*	*Accuracy*
Mild Dementia	1.00	0.97	0.98	0.99
Moderate Dementia	1.00	1.00	1.00	
Non Demented	0.99	1.00	0.99	
Very mild Dementia	0.99	0.94	0.97	

V. CONCLUSION

This work utilized a CNN to classify MRI scan phases of AD. Correct identification of Alzheimer's disease is quite

essential since it directly influences medication and disease management. Inspired by the OASIS MRI dataset, it improved and preprocessed to manage class imbalance. Built using several convolutional layers and dense layers, the CNN architecture was enhanced using weighted sampling and dropout regularisation. With a test loss of 0.0119 and a test accuracy of 98.80%, the model generated excellent performance. Overall classes, the confusion matrix also displayed good classification capacity; results for the Non-Demented and Moderate Dementia groups were robust. Every class's F1 score recall and precision exceeded 0.97, which underscores the model's ability to balance false positives and negatives reasonably. The accuracy, recall, and F1-scores for categories such as Moderate Dementia landed at 1.00, meaning perfect classification in this group. Finally, this work demonstrates that deep learning may generate quite accurate and dependable models for Alzheimer's classification, thereby offering considerable opportunity for clinical use in the early diagnosis and monitoring of disease evolution. Further developments in separating closely similar classes, such as Very Mild Dementia, may support performance even more.

REFERENCES

[1] P. Kalaivani, P. Iyyanar, C. Rajan, R. Harshini Priya, P. Janani, and A. S. Jayasudha, "Oral cancer detection using deep learning," in 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA), 2024.

[2] Z. Breijyeh and R. Karaman, "Comprehensive review on Alzheimer's disease: Causes and treatment," Molecules, vol. 25, no. 24, p. 5789, 2020.

[3] G. Kaur, N. Sharma, D. Upadhyay, M. Singh, and R. Gupta, "Brain tumour disease detection in MRI images using xception model," in 2024 3rd International Conference for Innovation in Technology (INOCON), 2024.

[4] M. Yildirim and A. Cinar, "Classification of Alzheimer's disease MRI images with CNN based hybrid method," *Ing. Syst. D Inf.*, vol. 25, no. 4, pp. 413–418, 2020.

[5] R. Singh, C. Prabha, H. M. Dixit, and S. Kumari, "Alzheimer Disease Detection using Deep Learning," in 2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), 2023.

[6] G. Garg, C. Prabha, B. Ahuja, R. Singh, and A. Agarwal, "An In-Depth Study of Alzheimer's Detection: Leveraging OASIS MRI with a 19-Layer CNN," in 2024 International Conference on Expert Clouds and Applications (ICOECA), IEEE, 2024, pp. 758–763.

[7] N. Raza, A. Naseer, M. Tamoor, and K. Zafar, "Alzheimer disease classification through transfer learning approach," Diagnostics (Basel), vol. 13, no. 4, p. 801, 2023.

[8] F. J. M. Shamrat, S. Akter, S. Azam, A. Karim, P. Ghosh, Z. Tasnim, ... and K. Ahmed, "AlzheimerNet: An effective deep learning based proposition for Alzheimer's disease stages classification from functional brain changes in magnetic resonance images," IEEE Access, vol. 11, pp. 16376–16395, 2023.

[9] S. Murugan, C. Venkatesan, M. G. Sumithra, X. Z. Gao, B. Elakkiya, M. Akila, and S. Manoharan, "DEMNET: A deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images," IEEE Access, vol. 9, pp. 90319–90329, 2021.

[10] A. Shehri, "Alzheimer's disease diagnosis and classification using deep learning techniques," PeerJ Computer Science, vol. 8, 2022.

[11] H. A. Helaly, M. Badawy, and A. Y. Haikal, "Deep learning approach for early detection of Alzheimer's disease," Cognit. Comput., vol. 14, no. 5, pp. 1711–1727, 2022.

[12] R. R. Janghel, "Deep-learning-based classification and diagnosis of Alzheimer's disease," in Deep Learning and Neural Networks, IGI Global, 2020, pp. 1358–1382.

[13] "OASIS Alzheimer's Detection," Kaggle, Jun. 18, 2023. https://www.kaggle.com/datasets/ninadaithal/imagesoasis