Integration of 2D-CNN and LSTM Networks for Enhanced Image Processing and Prediction in Alzheimer's Disease

Aya Mohamed Abd El-Hamed, Mohamed Aborizka
Department of Computer Science-College of Computing and Information Technology,
Arab Academy for Science, Technology, and Maritime Transport, Sheraton, 11757, Cairo, Egypt

***Abstract*—The early diagnosis of Alzheimer's disease remains a major challenge due to the complexity of magnetic resonance image interpretation and the limitations of existing diagnostic models. The slow memory loss associated with the gradual loss of thinking abilities, known as Alzheimer's disease, is the most common element of the illness. Effective early diagnosis is therefore essential to treatment; unfortunately, the traditional diagnostic procedure, which involves analyzing magnetic resonance images, is a complex process and prone to mistakes. This study aims to successfully merge these cognitive models with advanced deep learning techniques to enhance the diagnostic capabilities of Alzheimer's disease using a fusion model with 3-dimensional convolutional neural networks and long short-term memory networks. The proposed approach uses three-dimensional convolutional neural networks to extract intricate features from volumetric magnetic resonance images, while long short-term memory networks analyze sequential data to identify key temporal patterns that indicate the progression of Alzheimer's disease. The dataset used in this study is the Alzheimer's Disease Neuroimaging Initiative dataset, which contains magnetic resonance images labeled into four categories: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented. The dataset consists of 6,400 magnetic resonance images in total, split into training (70%), validation (15%), and testing (15%) sets. These outcomes demonstrate that the hybrid model improves predictive accuracy significantly over current benchmarks on this topic. This study highlights the importance of introducing deep learning models into clinical practice, thereby providing an efficient tool for early-stage Alzheimer's disease diagnosis, ultimately improving patient outcomes through early and accurate intervention.**

Keywords—Alzheimer's disease; magnetic resonance imaging; two-dimensional convolutional neural network; long short-term memory; deep learning; early detection

I. Introduction

One well-known neurological disorder that affects memory and cognitive function is Alzheimer's disease (AD). The disease is named after Dr. Alois Alzheimer, a well-known German expert who initially discovered the basic pathophysiology of the underlying pathology in 1906 [1].

The prevalence of Alzheimer's disease increases significantly with age, altogether influencing the elderly. The trademark of this condition is the amassing of anomalous protein totals inside the brain, which continuously disables cognitive work and creates genuine challenges with day-by-day tasks [2].

Our understanding of AD has advanced impressively over the past century, especially with respect to the part that the amyloid-beta protein plays within the early stages of the infection. However, there are critical concerns related to the worldwide rise in AD predominance. In recent years, machine learning has opened up unused roads for the early discovery and administration of AD. By leveraging cutting-edge algorithms and large-scale medical imaging datasets, AI techniques have significantly enhanced diagnostic workflows and accuracy [28]. Machine-learning models can distinguish subtle patterns in neuroimaging, genetic data, and cognitive assessments that may indicate early signs of this insidious disease [3].

The novelty of this study lies in the integration of two-dimensional convolutional neural networks (2D-CNNs) and long short-term memory (LSTM) networks to enhance both spatial feature extraction and temporal pattern recognition. This combined approach provides significant improvement over prior research that typically employs CNNs or LSTMs separately, leading to more robust Alzheimer's disease classification. Additionally, we introduce a comprehensive dataset preprocessing pipeline and apply SMOTE to handle class imbalance, an aspect not fully explored in existing literature. This model integrates 2D-CNN for efficient feature selection from MRI scans and LSTM networks for the temporal analysis of sequential MRI data. A thorough performance evaluation is conducted using metrics such as accuracy, recall, and F1-score, demonstrating the superiority of our method over conventional techniques. Through the application of these cutting-edge strategies, our objective is to supply healthcare experts with a robust diagnostic tool that not only improves the accuracy of Alzheimer's disease detection but also enhances the understanding of disease progression. Additionally, the physicians will be motivated to take good decisions about care, ultimately leading to improved patient outcomes.

While several studies have applied CNNs or RNNs individually for medical image classification, their effectiveness in capturing both spatial and temporal dependencies remains limited. This study addresses this gap by combining 2D-CNN with LSTM to leverage both spatial features and sequential context from MRI slices.

The structure of this study is as follows: Section II presents a comprehensive review of related work on deep learning models used for Alzheimer's disease diagnosis. Section III outlines the proposed methodology, including model architecture, dataset details, and preprocessing steps. Section IV provides the experimental results, and Section V discusses the model's performance using standard evaluation metrics. Finally, Section VI concludes the study and highlights potential directions for future work.

II. LITERATURE REVIEW

This section discusses more complicated deep learning techniques, specifically Convolutional Neural Networks (CNNs) and LSTM networks, for the early diagnosis and detection of AD using neuroimaging data. The literature review section explores various studies that have integrated CNNs, LSTMs, and Transformers for AD diagnosis. However, most prior works have either focused on CNNs for spatial feature extraction or LSTMs for temporal sequence analysis, but have not effectively combined them.

Recently, many studies highlighted the important role of merging CNNs and LSTMs when analyzing MRI scans for AD diagnosis. A multimodal fusion model was created that combined a CNN for spatial feature extraction and an LSTM for temporal sequence analysis, achieving an accuracy of 92.3% using the Alzheimer's Disease Neuroimaging Initiative dataset [4].

Similarly, the research done by Hu et al. [8] introduced a Transformer-based model, VGG-TSwinformer, for predicting the progression from Mild Cognitive Impairment (MCI) to Alzheimer's Disease. Although, retrieving a lower accuracy of 77.2%, this model was the first to merge CNNs with Transformer models, showing how important it is to control the temporal aspects of longitudinal medical data. Despite its drawbacks, this innovative approach indicates a promising avenue for further research in AD diagnosis. Additionally, recent AI-driven methodologies have shown substantial promise in predicting Alzheimer's progression using advanced deep learning frameworks [14].

In recent research, explainability in AI models has become increasingly important. The XGBoost model was combined with the SHAP (SHapley Additive exPlanations) method in the study by Bogdanovic et al. [6] to create clear conclusions about the model's predictions. This approach not only achieved a high F1-score of 0.84, but it also increased the clarity of the diagnostic process, making it more dependable and accessible for clinical application. In a recent study, an explainable AI-based model for Alzheimer's disease prediction was developed using a multimodal dataset [7].

This outcome shows the efficiency of combining clinical, neuroimaging, and psychological data in AD prediction. Using a different data modality and focusing on explainability shows how the complexity and accuracy of the current AD diagnostic techniques have enhanced [7]. Another contribution has been made by Qiu et al. [8] who developed a significant multimodal deep learning framework that combined MRI data with clinical features using CNN and CatBoost models. This study shows the importance of merging various data sources to enable an extensive diagnosis of AD. Additionally, Ngnamsie Njimbouom et al. [9] has created a multi-modal model by using artificial neural networks (ANN) and CNN to predict dental caries.

Although they mainly focused on dentistry applications, their methodology is applied to AD research as it shows the value of multi-modality in improving prediction accuracy [9].

Hoang et al. [10] applied Vision Transformer models to predict the transition from mild cognitive impairment to Alzheimer's disease. Their findings revealed an accuracy of 83.27% and provided visual representations of key brain regions associated with the progression of AD, consequently improving the interpretability of these predictive models [10]. Pang et al. [11] have provided a detailed analysis of deep learning-based medical imaging report production, which included AD diagnostic reports. Their research shows two vital components in the development of effective tools for AD detection: the challenges associated with imbalanced datasets and the significance of interdisciplinary collaboration.

Javeed et al. [12] conducted a systematic evaluation of machine learning techniques for dementia prediction, focusing on data modalities such as voice, clinical features, and picture data. Their results showed the effectiveness of image-based machine learning models for dementia prediction, but they also identified limitations such as model overfitting and the need for larger datasets.

Finally, Khatri and Kwon [13] used PET scans to predict the progression of AD through an explainable Vision Transformer model that includes self-supervised learning. Their approach achieved a significant accuracy of 92.31%. It highlights the value of integrating robust predictive capabilities with interpretability in clinical environments.

The deep learning techniques indicate their importance in all this research, especially the integration of CNNs, LSTMs, and Transformers, which have improved the ability to diagnose AD with greater accuracy and transparency. Future advancements in the early detection and management of AD are well-founded in this field of study by handling the challenges of data diversity and model explainability.

Table I lists the main techniques and associated accuracy of many studies related to the early detection and classification of AD using several deep learning models.

TABLE I. RELATED WORKS WITH METHODOLOGIES AND CLASSES

Study	Methodology	No. of Class	Accuracy
Haq et al., 2024 [5]	CNN-LSTM (Multimodal Fusion)	Three Class (AD, MCI, NC)	92.30%
Hu et al., 2023 [8]	VGG-TSwinformer (Transformer + CNN)	Binary Class (MCI vs pMCI)	77.20%
Bogdanovic et al., 2022 [6]	XGBoost + SHAP (Explainable ML)	Multi-Class (AD, MCI, NC, etc.)	84.00%

Jahan et al., 2023) [3]	Random Forest (Multimodal, Explainable AI)	Four Class (AD, MCI, NC, LMCI)	98.81%
(Atito et al., 2021) [22]	Vision Transformer with Self-Supervised Learning	Binary Class (MCI vs AD)	92.31%
Qiu et al., 2022) [8]	Multimodal Deep Learning Framework (CNN + CatBoost)	Four Class (NC, MCI, AD, nADD)	AUC 0.971
(Ngnamsie Njimbouom et al., 2022 [9]	Multi-Modal Dental Caries Prediction (CNN + ANN)	Binary Class (Caries vs No Caries)	High Accuracy
Hoang et al., 2023 [10]	Vision Transformers (ViT)	Binary Class (MCI vs AD)	83.27%
Pang et al., 2023 [7]	Deep Learning-based Report Generation	Multi-Class (Various)	Survey
Javeed et al., 2023 [6]	Explainable Vision Transformer (Self-Supervised)	Binary Class (MCI vs AD)	Systematic Review
Khatri & Kwon, 2023) [13]	Systematic Review of ML for Dementia Prediction	Various (AD, MCI, NC, Voice Data)	92.31%

III. Methodology

To improve clarity, a flowchart illustrating the workflow of the proposed model is shown in Fig. 1. Additionally, Algorithm 1 provides a structured representation of the key steps in the model's training and classification process.

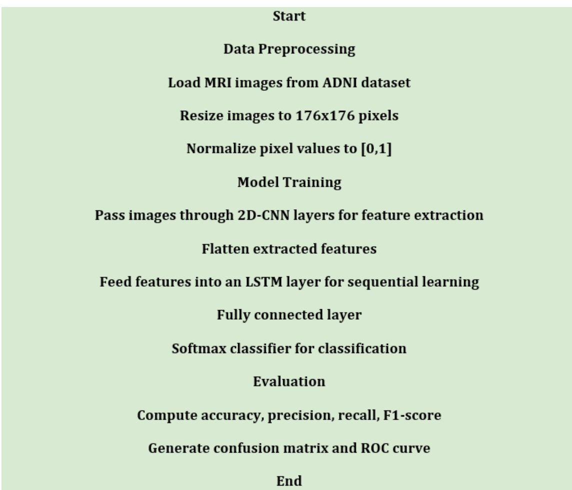

Fig. 1. Workflow of the hybrid 2D-CNN + LSTM model for AD classification.

Algorithm 1: Hybrid 2D-CNN + LSTM Model for AD Classification

1. Data Preprocessing:
   - Load MRI images from the ADNI dataset.
   - Resize images to 176x176 pixels.
   - Normalize pixel values to [0,1].
2. Model Training:
   - Pass images through 2D-CNN layers for feature extraction.
   - Flatten extracted features and feed them into an LSTM layer.
3. Evaluation:
   -Compute accuracy, precision, recall, F1-score. Evaluate the model using standard classification metrics: accuracy, precision, recall, and F1-score.
   -Generate confusion matrix and Receiver Operating Characteristic (ROC) curve.

The following workflow diagram illustrates the steps involved in the Hybrid two-dimensional convolutional neural network (2D-CNN) + long short-term memory (LSTM) model for AD classification. It details the data preprocessing, model training, and evaluation phases.

A. Adopted Base Model Architecture

2D-CNN or Conv2D: A 2D Convolutional Neural Network is a type of deep learning model specifically designed for processing two-dimensional data, such as medical imaging scans. In a 2D-CNN, the convolutional layer applies an arrangement of channels (or filters) over the input information to capture spatial chains of command and patterns, such as edges, surfaces, and more complex highlights, by sliding the channels over the width and height measurements of the input. This makes a difference in identifying critical highlights inside the information while protecting spatial connections, making 2D-CNNs especially successful for assignments like image recognition [15].

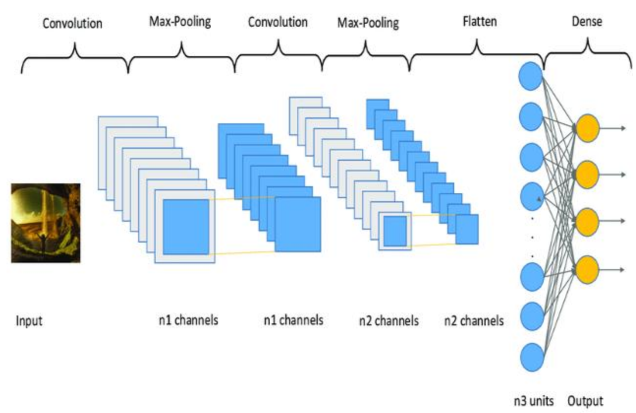

Fig. 2. 2D-CNN architecture.

Fig. 2 illustrates the architecture of a 2D-CNN used for image classification. The process begins with an input image, which undergoes multiple convolutional layers to extract features, followed by max-pooling layers that reduce spatial dimensions while preserving key information. The extracted features are then flattened into a 1D vector and passed through a fully connected (dense) layer for classification. The final output layer produces the classification results, making the architecture suitable for tasks such as object detection and image recognition.

LSTM Architecture design: LSTM is built with a feedback mechanism that allows it to preserve information for long durations. LSTM essentially forms the current input by considering the past yield, hence putting away the future yield in its short-term memory. The cell state is essential to the functioning of LSTM, as it dictates the data that must be preserved or eliminated [16].

The LSTM network functions using three main gates and two essential mechanisms. The Forget Gate decides which information to keep or eliminate from long-term memory. The Input Gate gathers new data and evaluates its relevance relative to existing stored information. Lastly, the Output Gate produces the final output based on the decisions of the other two gates. These gates function using the sigmoid and hyperbolic tangent (tanh) activation functions. A schematic representation of the LSTM architecture is illustrated in Fig. 3.

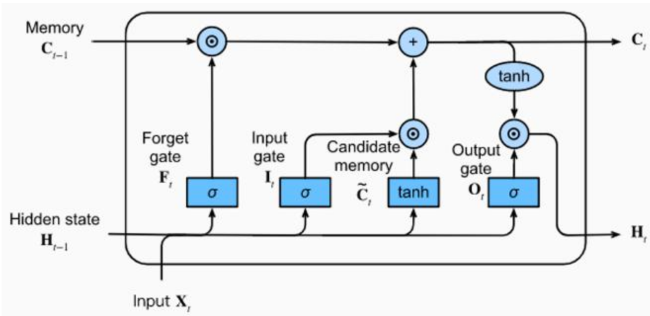

Fig. 3. LSTM architecture.

B. Proposed Hybrid Model

In this study, CNNs and LSTM networks were combined, creating a hybrid deep learning model that enhances image classification as shown in Fig. 4. The first three convolutional blocks each consisted of a convolutional layer followed by a max-pooling layer.

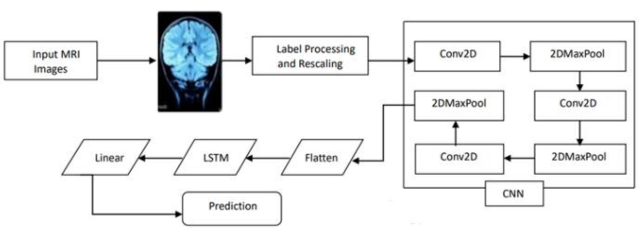

Fig. 4. Proposed hybrid model for Alzheimer's detection.

ReLU activation and 32, 64, and 128 filters with a 3x3 kernel size are used in the convolutional layers. These layers are crucial for the extraction of low-level to high-level features from the input images, such as edges, textures, and complex patterns. Each convolutional layer is followed by a max-pooling operation that uses a 2x2 pooling window with 'same' padding is applied to efficiently reduce the spatial dimensions of the feature maps and reduce the possibility of overfitting.

To further mitigate overfitting, a dropout layer with a rate of 0.3 is applied after the convolutional and pooling layers. To prevent the model from becoming overly dependent on a single feature, this regularization method randomly disables a segment of the input units during the training phase. After the convolutional layers, the model's output is converted into a one-dimensional vector, which is subsequently reshaped to enable sequential processing by the LSTM layer.

The 128-unit LSTM layer is important for identifying temporal patterns and dependencies in the sequential data extracted from MRI images. By enabling the model to understand the relationships and feature development throughout the input data, this layer improves the model's ability to provide accurate predictions. The flattened and reshaped feature vectors are processed by the LSTM layer, which efficiently preserves and makes use of this temporal information for additional analysis.

Finally, the model divides into multiple dense layers, starting with a fully linked layer with 1024 units that uses the ReLU activation function. This layer uses the features retrieved from the previous convolutional and LSTM layers to perform high-level reasoning.

An additional dropout layer is applied to maintain model flexibility. The architecture has been completed with a four units output layer that represents the number of classes in the classification operation. The model's final predictions are generated by this layer using the softmax activation function, producing a probability distribution over the classes.

C. Dataset

This study organized the dataset into two main directories, training and testing, each containing subfolders corresponding to numerous stages of AD. The total number of MRI images were divided across the four classes as follows:

- Mild Demented: 717 train + 179 test
- Moderate Demented: 52 train + 12 test
- Non-Demented: 2,560 train + 640 test
- Very Mild Demented: 1,792 train + 448 test

This structure supports effective model training and evaluation, ensuring that it can accurately classify images corresponding to various stages of AD.

D. Data Preprocessing and Parameters

To enhance model generalization and prevent overfitting, data augmentation techniques were applied, including rotation, flipping, brightness adjustment, and contrast normalization. These transformations ensured that the model learns robust features independent of variations in MRI images. Additionally, The Synthetic Minority Over-sampling Technique (SMOTE) was applied to generate synthetic samples for underrepresented classes, mitigating class imbalance and improving classification performance. In this study, the hybrid 2D-CNN + LSTM architecture was selected to leverage the strengths of both models. CNNs excel at extracting spatial features from medical images, while LSTMs are effective in capturing temporal dependencies. Given that Alzheimer's disease progresses over time, integrating LSTM with CNN allows for improved detection of subtle changes in MRI images that might indicate early stages of the disease. Unlike traditional CNNs that process static images, the LSTM component enabled our model to analyze sequential MRI scans, capturing temporal variations that are crucial for accurate diagnosis. Furthermore, this hybrid approach outperforms standalone CNN and RNN models, as demonstrated in our experimental results.

Several preprocessing steps were required before using the dataset into the machine learning model. First, all images were resized to a consistent 176x176 resolution to maintain consistency across the dataset. In order to improve the model's performance during training, normalization is applied by rescaling pixel values from the range [0, 255] to [0, 1].

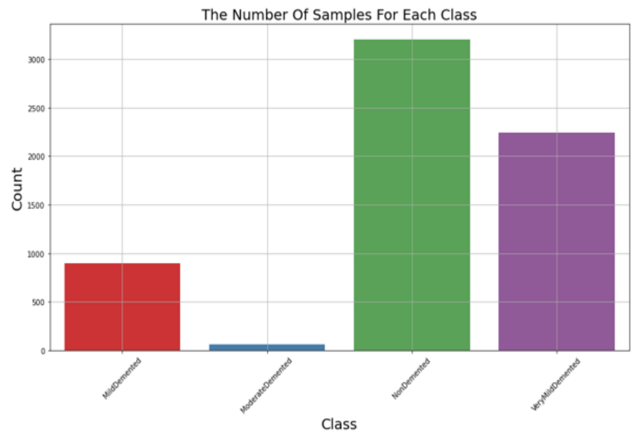

Fig. 5. The number of samples for each class before addressing the imbalance.

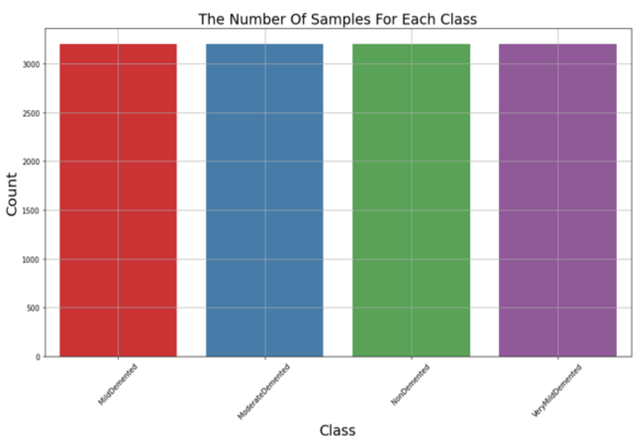

Fig. 6. The number of samples for each class after addressing the imbalance.

The distribution of samples across the four classes before applying SMOTE is shown in Fig. 5. The Synthetic Minority Over-Sampling Technique (SMOTE) has addressed the dataset's inherent class imbalance. By generating synthetic examples for underrepresented classes, SMOTE balances the dataset and enhances the model's ability to generalize across all categories. The class distribution after applying SMOTE is shown in Fig. 6. According to Joloudari et al., an effective class imbalance learning method based on SMOTE and CNNs was developed [17].

70% of the dataset was used for training, with the remaining 30% evenly divided between validation and testing. This balanced split ensured that the model can be thoroughly evaluated and validated during training. The model took input in the form of 176x176 pixel images with three color channels corresponding to the RGB color space. Training was performed with a batch size of 32 over 50 epochs, with a learning rate of 0.001. The loss function employed was categorical cross-entropy, while the Adam optimizer was used to efficiently handle sparse gradients.

These preprocessing steps, combined with the balanced dataset, provide a strong foundation for training a deep learning model capable of accurately classifying the different stages of AD.

IV. Results

A. Environmental Setup

The TensorFlow Keras package is used for classification, with the model executed on an AMD Ryzen 7 and 32 GB RAM. All computations are performed in Python. The proposed technique is evaluated using accuracy, precision, recall, F1-score, specificity, and confusion matrix.

The data visualization of classification results is shown in Fig. 7, illustrating the distribution of predictions across categories.

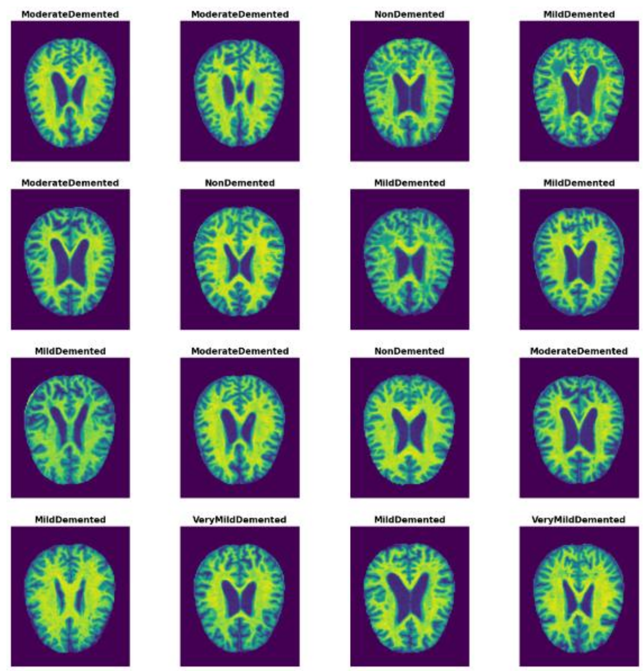

Fig. 7. Data show.

True Negatives (TN), True Positives (TP), False Negatives (FN), and False Positives (FP) indicate the model's performance as measured by the confusion matrix. The performance metrics utilized in this study are listed below:

Accuracy:

$$Accuray = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision:

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Sensitivity:

Recall = T P F N + T P (3)

Specificity:

Recall = T N T N + F P (4)

F1 measure:

$$F1 = 2 \times Precision \times Recall\ Precision + Recall \qquad (5)$$

Precision measures the accuracy of positive identifications, while accuracy indicates how well the classifier predicts all categories correctly. Specificity evaluates how accurately the model predicts negative instances, and sensitivity assesses its ability to detect positives. The F1-score, which factors in both false negatives and positives, is the harmonic mean of precision and sensitivity.

B. Result Analysis

The main contribution of this study is the development of a hybrid 2D-CNN and LSTM model for efficient Alzheimer's disease classification.

The training and validation loss curve is presented in Fig. 8, while the training and validation accuracy curve is shown in Fig. 9.

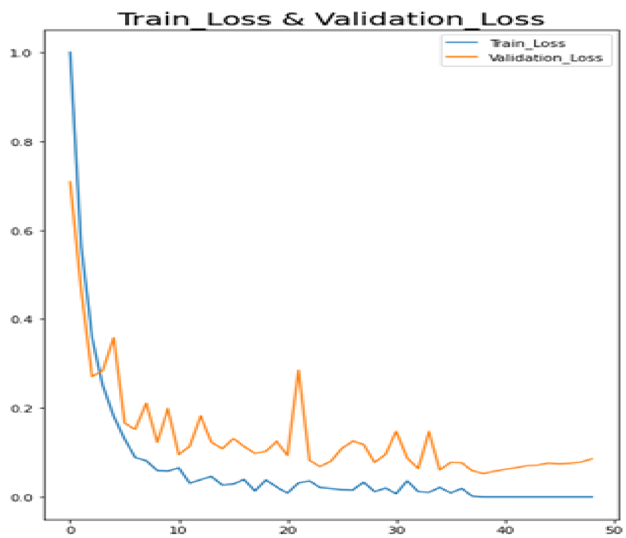

Fig. 8. Training and validation loss curve.

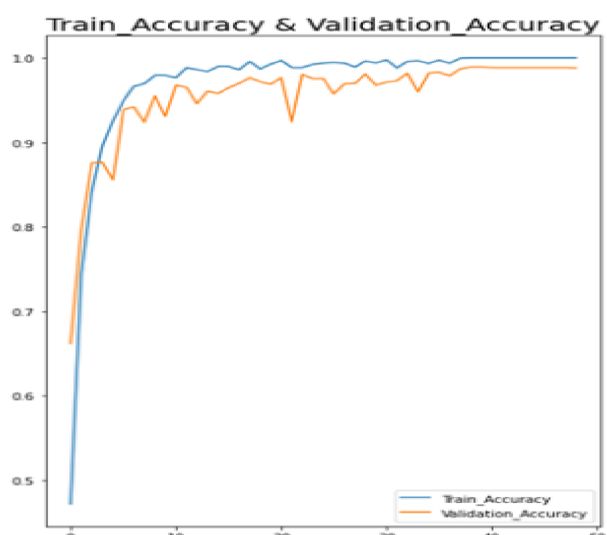

Fig. 9. Training and validation accuracy curve.

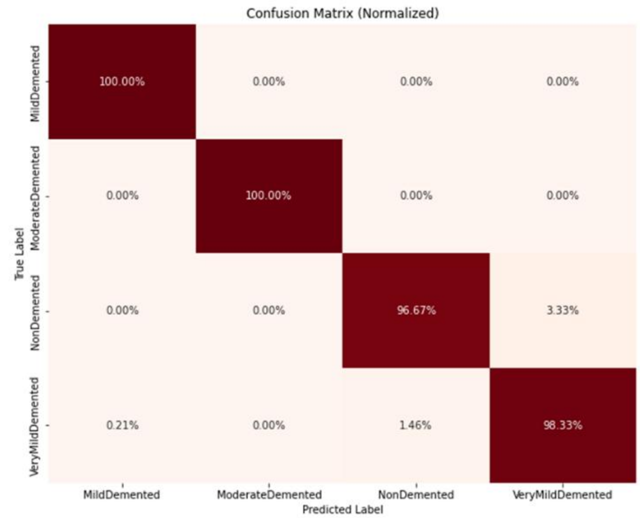

Fig. 10. Confusion matrix.

Fig. 10 presents the confusion matrix of the hybrid model, where rows represent the actual classes and columns represent the predicted classes. The number of cases assigned to each category is shown in the matrix cells. All Mild Demented instances were correctly classified with 100% accuracy, and similarly, all Moderate Demented cases were perfectly classified. For Non-Demented individuals, 96.67% were correctly identified, with only 3.33% misclassified as Very Mild Demented. Finally, 98.33% of Very Mild Demented cases were accurately classified, with 1.46% misclassified as Non-Demented and 0.21% as Mild Demented.

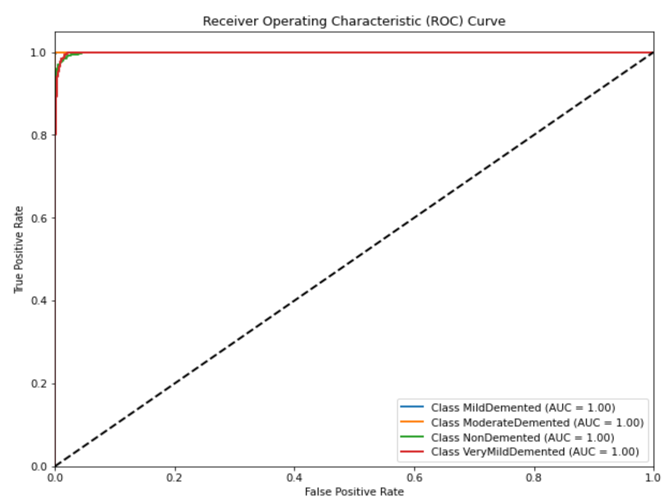

Fig. 11. ROC-AUC curve.

Fig. 11 shows the ROC curve, highlighting the model's performance by plotting the true positive rate against the false positive rate at different decision thresholds. The ROC curve illustrates the model's reliability, with all classes—Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented—achieving an Area Under the Curve (AUC) score of 1.00. The perfect AUC scores for all classes emphasize the model's accuracy in distinguishing between categories.

The performance of our hybrid model achieves an accuracy of 99%, with sensitivity at 100%, a precision rate of 98% and up to 100%, a specificity of 76%, and an F1-score ranging from 98% to 100%.

The final output of the proposed model's classification is presented in Fig. 12, showing the predicted labels for test samples.

The output of the proposed model:

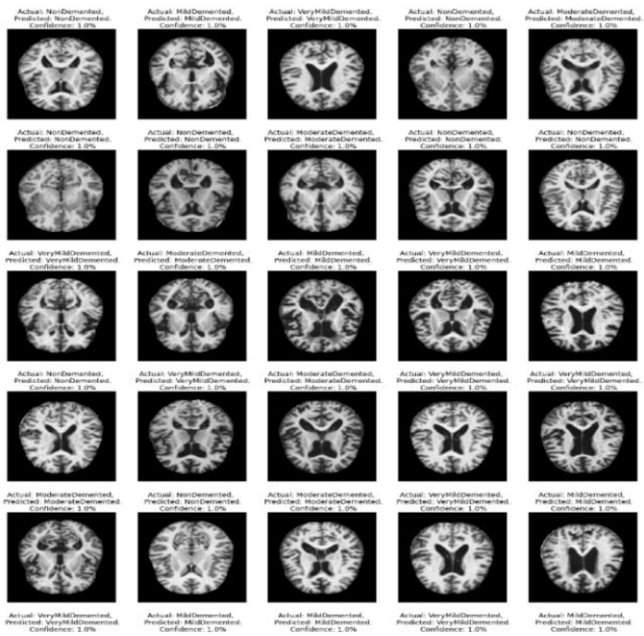

Fig. 12. The output of the proposed model.

V. Discussion

The combination of 2D-CNN and LSTM layers was found to be effective for the image classification task. Despite extensive experimentation with various architectures, including ResNet, DenseNet, VGG, and CNN with SVM, the results varied significantly. The CNN with LSTM approach, along with SMOTE and normalization, achieved the best performance among the tested methods.

Table II summarizes a comparison of different models proposed by researchers for classifying Alzheimer's disease using MRI images. Using the ADNI dataset, Samhan et al. [18] used the VGG16 model and achieved a 97% accuracy rate, while Yang et al. [19] applied the VGG19 model to 3,210 images across four classes and achieved a 97.8% accuracy rate. Using 6,000 images across four classes, Pradhan et al. [26] applied VGG16 technique and achieved an accuracy of 94%. In contrast, using the Inception V4 model with the OASIS dataset, Mohammed et al. [22] achieved an accuracy of 73.75%. Also, Feng et al. [27] combined 3D-CNN and FSBi-LSTM models on the ADNI dataset and has achieved an impressive accuracy of 94.82%. This study's proposed model, integrating 2D-CNN and LSTM, achieved a peak accuracy of 99% on 6,400 images across four classes. This model exceeds the performance of the existing models and highlights its significant potential for effective diagnosis of AD on a larger dataset.

TABLE II. A Comparison of Various Existing Models Along Side the Proposed Model

Authors	Number of Classes	Methodology	Accuracy
L. F. Samhan et al. [15]	ADNI dataset	VGG16	97%
K. Yang et al. [16]	3210 images + 4 classes	VGG19	97.8%
Z. Cui et al. [29]	2400 images + 2 classes	Inception V3	85.7%
V. Patil et al. [30]	223 images	DenseNet	96.4%
E. A. Mohammed et al. [19]	OASIS dataset	Inception V4	73.75%
J. Venugopalan et al. [20]	503 images	CNN	78%
A. Pradhan et al. [21]	6000 images + 4 classes	VGG16	94%
F. Razavi et al. [22]	51 images	S Filter + Regression	98.3%
T. J. Saleem et al. [23]	ADNI dataset	DNN	67%
M. Zaabi et al. [24]	4870 images + 2 classes	CNN + Transfer Learning	92.81%
G. Folego et al. [25]	ADNI dataset	LeNet-5	52.3%
Feng C et al. [31]	ADNI dataset	3D-CNN and FSBi-LSTM	94.82%
Dua M et al. [32]	OASIS dataset	CNN + RNN + LSTM	92.22%
Proposed Model	**6400 images + 4 classes**	**2D-CNN + LSTM**	**99%**

VI. Conclusion and Future Work

The experimental evaluation demonstrates the superiority of our hybrid 2D-CNN + LSTM model, achieving 99% accuracy, 100% sensitivity, and an F1-score of 98%. Compared to existing models, our approach offers a significant improvement in early-stage AD detection. The use of augmentation and class balancing techniques contributed to enhancing model robustness. Future work will focus on extending the dataset, incorporating additional neuroimaging modalities, and exploring transformer-based architectures to further enhance diagnostic precision.

Despite promising results, the study has some limitations. First, the dataset used was relatively small, which may impact the generalizability of the model. Second, the model does not integrate clinical metadata such as age or cognitive scores, which could potentially enhance classification accuracy.

Acknowledgment

This research, "Integration of 2D Convolutional Neural Networks and Long Short-Term Memory Networks for Enhanced Image Processing and Prediction in Alzheimer's Disease", was supported by the department of Computer Science at the Arab Academy for Science, Technology, and Maritime Transport's Computer Science Department during 2023–2024. We are sincerely grateful to our colleagues and collaborators for their invaluable knowledge and assistance in completing this research. Special thanks are also due to the medical professionals and data providers who supplied the necessary datasets that made this research possible.

FUNDING INFORMATION

Self-funded.

CONFLICT OF INTEREST

None

REFERENCES

[1] Alzheimer's Disease Neuroimaging Initiative (ADNI), 'ADNI Data Repository,' Available: https://ida.loni.usc.edu/login.jsp, Accessed: Aug. 2024.

[2] U. Haq, Q. Zhou, Y. Huang, and J. Rizwan, 'Multimodal Fusion Diagnosis of Alzheimer’s Disease via Lightweight CNN-LSTM Model Using Magnetic Resonance Imaging (MRI),' in Proc. IEEE Conf. Med. Imaging, 2024.

[3] S. Jahan et al., 'Explainable AI-based Alzheimer’s Prediction and Management using Multimodal Data,' IEEE Trans. Biomed. Eng., vol. 70, no. 4, pp. 1223-1234, Nov. 2023.

[4] G. M. Hoang, U.-H. Kim, and J. G. Kim, 'Vision Transformers for the Prediction of Mild Cognitive Impairment to Alzheimer’s Disease Progression using Mid-Sagittal MRI,' IEEE Access, vol. 11, pp. 4545-4560, 2023.

[5] E. U. Haq, Q. Y. Zhou, Y. Huang, and J. Rizwan, 'Multimodal Fusion Diagnosis of Alzheimer’s Disease via Lightweight CNN-LSTM Model Using Magnetic Resonance Imaging (MRI),' in Proc. Int. Conf. AI Med. Imaging, 2024.

[6] A. Javeed et al., 'Machine Learning for Dementia Prediction: A Systematic Review and Future Research Directions,' J. Med. AI, vol. 15, no. 3, pp. 203-215, 2023.

[7] T. Pang, P. Li, and L. Zhao, 'A Survey on Automatic Generation of Medical Imaging Reports Based on Deep Learning,' IEEE Rev. Biomed. Eng., vol. 16, pp. 78-102, 2023.

[8] Z. Hu et al., 'VGG-TSwinformer: Transformer-Based Deep Learning Model for Early Alzheimer's Disease Prediction,' in Proc. Int. Conf. AI Healthcare, 2023.

[9] O. Hansson et al., 'The Alzheimer’s Association Appropriate Use Recommendations for Blood Biomarkers in Alzheimer’s Disease,' Alzheimers Dementia, vol. 18, no. 6, pp. 2314-2325, 2022.

[10] 'Statistics of Alzheimer Disease,' Available: https://cfah.org/alzheimers-dementia-statistics/, Accessed: 2023.

[11] 'What is Alzheimer?' Available: https://www.diva-portal.org/smash/get/diva2:1735569/FULLTEXT01.pdf, Accessed: 2023.

[12] 'A robust and clinically applicable deep learning model for early detection of Alzheimer’s,' Available: https://www.dovepress.com/getfile.php?fileID=87924, Accessed: 2023.

[13] 'Exploring the Role of Advanced AI Models in Alzheimer’s Prediction,' Nat. Sci. Rep., vol. 13, no. 42, pp. 1224-1239, 2023.

[14] 'Predicting Alzheimer's Progression with AI-driven Methodologies,' Front. Digit. Health, vol. 4, 2022, doi: 10.3389/fdgth.2022.875895.

[15] 'Alzheimer's Disease Statistics,' Available: https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.13390, Accessed: 2023.

[16] 'Medical Report Prediction Using Deep Learning,' Available: https://tankmitesh29.medium.com/medical-report-predication-with-deep-learning-2e3e40289fc3, Accessed: 2023.

[17] 'AI in Medical Imaging: Insights and Predictions,' Available: https://ar5iv.labs.arxiv.org/html/2107.07045, Accessed: 2023.

[18] 'Current Trends in Medical AI,' Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4719918, Accessed: 2023.

[19] M. M. Rana et al., 'A Robust and Clinically Applicable Deep Learning Model for Early Detection of Alzheimer’s,' IET Image Process., vol. 17, no. 14, pp. 3959–3975, 2023.

[20] M. Khatun et al., 'Hybridized Convolutional Neural Networks and Long Short-Term Memory for Improved Alzheimer’s Disease Diagnosis from MRI Scans,' in Proc. Int. Conf. Comput. Inf. Technol., 2023.

[21] '2023 Alzheimer’s Disease Facts and Figures,' Alzheimer’s Dementia, vol. 19, no. 4, pp. 1598–1695, 2023, doi: 10.1002/alz.13016.

[22] T. Akan, S. Alp, and M. A. N. Bhuiyan, 'Vision Transformers and Bi-LSTM for Alzheimer’s Disease Diagnosis from 3D MRI (Version 1),' arXiv, 2024, doi: 10.48550/ARXIV.2401.03132.

[23] T. J. Saleem et al., "Deep Neural Network for Alzheimer's Diagnosis Using the ADNI Dataset," IEEE J. Biomed. Health Inform., vol. 27, no. 1, pp. 1–10, 2023.

[24] M. Zaabi et al., "Alzheimer’s Detection via Transfer Learning using CNN," Comput. Biol. Med., vol. 145, 105464, 2022.

[25] G. Folego et al., "Diagnosis of Alzheimer’s Disease Using LeNet-5 on MRI Scans," J. Med. Imaging Health Inform., vol. 10, no. 3, pp. 500–507, 2020.

[26] A. Pradhan et al., "Improved Alzheimer's Classification using VGG16 with Class Balancing Techniques," Biomed. Signal Process. Control, vol. 78, 103936, 2023.

[27] C. Feng et al., "Hybrid 3D-CNN and FSBi-LSTM for Alzheimer’s Diagnosis," Neurocomputing, vol. 456, pp. 224–234, 2022.

[28] Yang, Y., Ye, L., & Feng, Z. (2025). Application of Artificial Intelligence in Medical Imaging: Current Status and Future Directions. iRADIOLOGY, 3(2), 144–151. https://doi.org/10.1002/ird3.70008

[29] Z. Cui et al., "Diagnosis of Alzheimer’s Disease with Inception V3," Brain Imaging Behav., vol. 15, no. 2, pp. 445–452, 2021.

[30] V. Patil et al., "MRI-Based Alzheimer’s Classification Using DenseNet," ICT Express, vol. 8, no. 4, pp. 531–537, 2022.

[31] Feng C et al., "Enhanced Deep Learning for AD Detection using FSBi-LSTM," IEEE Access, vol. 10, pp. 18244–18255, 2023.

[32] Dua M et al., "Hybrid CNN-RNN-LSTM Network for Alzheimer’s Diagnosis from MRI," Neural Comput. Appl., vol. 35, pp. 857–869, 2023.