Performance Evaluation of Deep Learning Models for Alzheimer's Disease Detection

Abdul Talha
Department of ECE
B. V. Raju Institute of Technology
Narsapur, Medak, Telangana, India
20211a0405@bvrit.ac.in

Ch. Dhanasree
Department of ECE
B. V. Raju Institute of Technology
Narsapur, Medak, Telangana, India
dhanach9999@gmail.com

E. Divya
Department of ECE
B. V. Raju Institute of Technology
Narsapur, Medak, Telangana, India
divyareddyedula07@gmail.com

Kandagatla Srikar Prabhas
Department of Data Science and Artificial Intelligence
Faculty of Science and Technology (IcfaiTech), IFHE
Hyderabad, Telangana, India
srikarprabhaskandagatla@gmail.com

Syed Abudhagir, U.
Department of ECE
B. V. Raju Institute of Technology
Narsapur, Medak, Telangana, India
syedabudhagir.u@bvrit.ac.in

***Abstract*—Alzheimer's disease (AD) presents as a neurodegenerative condition characterized by dementia, impacting cognitive function, memory, and behavior. Swift identification is crucial for effective intervention and better patient outcomes. This research aims to tackle this urgent need by harnessing deep learning techniques to advance early AD detection. By merging magnetic resonance imaging (MRI) with advanced deep learning algorithms, the objective is to create a precise, non-invasive approach for early AD diagnosis. The primary goal is to uncover new features and significantly boost the accuracy of AD detection. The method involves examining convolutional neural networks alongside MRI scans to reveal patterns indicative of AD progression. Notably, the proposed model achieves impressive evaluation metrics: 99.19% accuracy, 0.023% loss, 99.08% f1-score, and 99.11% precision. This endeavor seeks to demonstrate substantial advancements in AD detection accuracy compared to existing methods. Through this interdisciplinary approach, the aim is to drive progress in early AD diagnosis, ultimately leading to more effective interventions and enhanced quality of life for those impacted by this challenging condition accuracy compared to existing methods.**

***Index Terms*—Deep learning, image classification, Alzheimer's Disease, A Convolutional Neural Network, Magnetic Resonance Imaging**

I. Introduction

Alzheimer's disease (AD), which is a progressive neurodegenerative one, takes away the biggest treasure one cannot imagine one's life without - memories and cognitive skills. First detected by Dr. Alois Alzheimer in 1906, AD remains the major etiology of dementia among the seven types that affect millions across the globe. AD is frequently tied with aging; however, it doesn't mean that aging is inevitable. Unlike occasional forgetfulness one encounters, AD has specific symptoms such as forgetting things, and the decline of one's reasoning and language, which are the features that significantly influence one's day-to-day life [1].

Scientists have been shedding light on the core pathology of AD since decades ago. The amyloid plaques as well as tau tangles are the major molecular signs of the disease. In effect, these compounds impair brain function and interactions, leading to the manifestations of the disease. The existing diagnostic tools apply for the most part cognitive screens and structural and functional brain imaging techniques such as MRI and PET. However, these strategies are highly invasive, costly, and can only determinatively diagnose AD at advanced stages [2].

The early diagnosis of AD, is still, a great concern. The incapacity of present methods to capture the disease at its early stages usually hinders the efficiency of the cure as such intervention hinders the opportunity of benefiting when intercession could be most effective. In this line, it is worth noting that the uncertainty of cognitive evaluations and the high price of brain imaging hinder the scale effect to an extent that wider screenings become extremely impossible. This shows that this area requires more objective and cost-effective tools with quick and accurate AD detection which are also accessible to everyone [3].

The deep learning's role as one of the sub-spheres of Artificial Intelligence (AI) for Alzheimer's disease (AD) diagnosis will be analyzed in this study. Some deep learning algorithms can examine so intricate data, for example, brain scans or even speech patterns, to identify the slightest bits of AD influence on those data. We, as a culminating objective, aim at a deep learning model that is capable of identifying the presence of AD through data that is easily accessible and non-invasive at early stages. Then as a result it could be access to the earlier stages of intervention, thereby getting the best patient outcomes and a promising future for Alzheimer's people and their relatives [4].

This research paper is structured into various sections, with Section 2 encompassing the Literature Survey, Section 3 elucidating the proposed Model, Section 4 presenting the results, and Section 5 providing the conclusion.

II. Literature Survey

Owing to the significance of early AD diagnosis, researchers are interested in this area and are working to find a solution. Thus, the most significant studies in this area will be presented in this section. With extensive predictive model capacity, Faturrahman *et al.* [5] suggested a Deep Belief Network (DBN)-based structural methods approach for the categorization of AD. A support vector machine (SVM) has been employed as a classification model with DBN. By dividing brain pictures into categories based on how they appear in the brains of healthy individuals and those with AD.

Zaabi *et al.* [6] presented a deep way to diagnose AD. The approach consists of two main phases. Taking a block of the picture and isolating only the portion that shows the hippocampal area will be the initial step in identifying the region of interest. Classifying pictures using CNN and Transfer Learning.

Using a variety of image processing methods for AD approaches, a method for brain MR image identification and classification feature extraction techniques, and machine learning classifiers is presented by Aruchamy *et al.* [7]. Improved performance and classification accuracy of the suggested technique are achieved by modeling a classification algorithm from the traditional k-nearest-neighbor algorithm.

With a reasonable amount of computing environment resources, Majdah Alshammari *et al.* [8] provides an evolutionary deployment of CNN-based AD detection and stage categorization using the Python programming language. Model loss and accuracy are both very well achieved by the expected model's deployment.

A single 2D brain MRI slice was used by Ekin *et al.* [9] to develop a (CAE) strategy in deep learning strategy for categorizing Alzheimer's Disease + MCI vs. (HC) participants. The test set experimental outcomes showed the suggested method's efficiency in classifying AD+MCI against HC. This technique exceeds the traditional method wherein latent representations are replaced with complete MRI slices as input by deep CNNs, in terms of classification accuracy.

A unique CNN for identifying AD with comparative parameters was suggested by Mian Mohammad *et al.* [10] This particular methodology is ideally suited for the purpose of instructing a reduced collection of information. In order to accurately classify the stages of Alzheimer's disease (AD) by cutting computing costs and parameters the proposed Alzheimer's Disease Detection Network (ADD-Net) by cutting is constructed from the ground up.

Z. Xia *et al.* [11] introduced an innovative deep-learning framework for Alzheimer's disease detection, integrating both 3D CNN and 3D CLSTM with structural MRI data. Experimental results show that the model surpasses state-of-the-art methods in terms of detection performance.

3D brain MR image slices to diagnose Alzheimer's disease, utilizing Axial, Sagittal, and Coronal views of gray and white matter. Slice 51 is selected based on observations, and first-order statistical features are extracted. Classification accuracy is analyzed by Rao, K. N., Gandhi *et al.* [12] using logistic regression, Naive Bayes, SVM, and AdaBoost, with white matter slices in the coronal view showing the highest accuracy.

Utilizing ADNI data, Manzak, D *et al.* [14] developed a model to detect Alzheimer's disease without MRI measurements, achieving 67% accuracy using only 8 key features from baseline and 24 months' measurements. By simplifying the model and reducing data costs, this approach holds the potential for fast, accurate, and cost-effective early-stage Alzheimer's detection, applicable to neural network processors and potentially other diseases.

The study highlights the potential of Convolutional Neural Networks in achieving 98.05% accuracy by Mandal, P. K., & Mahto *et al.* [15], suggesting avenues for improvement, such as refining pre-processing methods and exploring different MRI layers for enhanced disease progression insights.

Alqahtani, N., Alam S *et al.* [16] utilized an extensive array of cost-effective time-series attributes such as patients' comorbidities, cognitive scores, medication histories, and demographics, elucidating the superiority of DBN-MOA models over current methods. The random forest model demonstrates higher accuracy, outperforming SVM, RF, KNN, LR, and DT in testing.

Jabason, E., Ahmad, M. O *et al.* [17] introduced a novel stacked sparse autoencoder-based imputation method for inferring missing values in multiclass Alzheimer's disease diagnosis using MR and clinical features. The unsupervised learning algorithm optimizes sparsity and weight decay regularization to learn significant features, improving classification performance significantly on ADNI data. Future work will focus on predicting additional clinical features using deep learning techniques for enhanced AD diagnosis.

Subetha, T. *et al.* [18] conducts a systematic review of dementia leading to Alzheimer's disease, examining diverse approaches for AD diagnosis. It evaluates recent work on early AD detection using machine learning, IoT, and AI techniques, while also discussing future research directions and challenges in handling Alzheimer's data. However, evaluations are limited to pathologically unproven datasets, and comparing different imaging modalities remains challenging.

III. Methodology

A. Description of Dataset

The dataset, uploaded on Kaggle, has gone through thorough proofreading by its uploader which consequently makes it trustable. The dataset, consisting of 6400 samples, is sufficient for conducting research. Before the data release, the dataset underwent exhaustive cleaning, including resizing and organization of individual three-channel RGB images with the size of 128 x 128. These images are categorized into four distinct classes: NOD (No Dementia), VMD (Very Mild Dementia), MD (Mild Dementia), and MOD (Moderate Dementia). The class distribution within the dataset is as follows: NOD consists of 3200 samples, VMD has 2240 samples, MD has 896 samples, and MOD has 64 samples.

B. Model Architecture

The Convolutional Neural Network (CNN) architecture with a thick block was a novel approach to Alzheimer's disease detection that we presented in this research proposal. Computer vision tasks including image classification, picture segmentation, and object recognition are the main applications of CNN architecture, which is taken by the anatomical composition of the human brain. According to Figure 1, the suggested design is made up of four convolution blocks, each of which has a ReLU activation function and a convolution filter. Three dense blocks with dropout and a Softmax classification layer are then displayed.

C. Model parameters

This neural network design utilizes a sequence of convolutional layers with max-pooling layers for extracting features from input images. The initial convolutional layer, ConvLayer1, comprises 32 filters with a (3,3) kernel size, generating 128x128 feature maps. Subsequent max-pooling layers halve the spatial dimensions, down-sampling the feature maps effectively. As the network deepens, the filter count progressively increases; ConvLayer2 employs 64 filters, and ConvLayer3 and ConvLayer4 use 64 and 128 filters, respectively. After each convolutional layer, max-pooling reduces spatial dimensions. The flattened output from

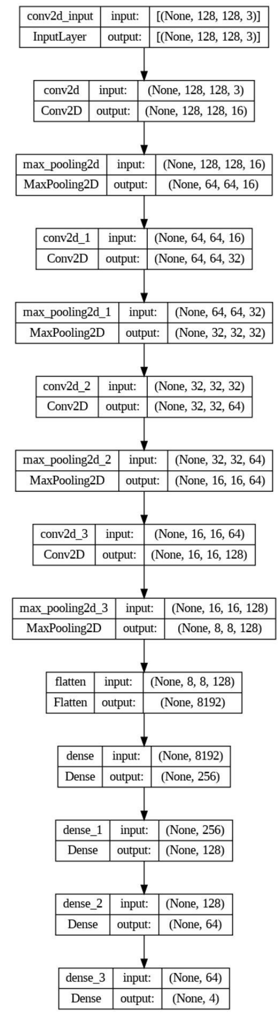

Fig. 1. Architecture of Proposed Network of AD

the final max-pooling layer feeds into a fully connected neural network with dense layers. These layers gradually decrease dimensionality, starting from 256 neurons, then 128, ultimately producing classification via a SoftMax activation function with 4 neurons in the last layer. Dropout layers are introduced post-dense layers to mitigate overfitting by randomly deactivating neurons during training. In total, the network comprises 2,268,996 trainable parameters and is tailored for a classification task with four output classes.

Various methods have been explored for classification, yielding different levels of success. Faturrahman (2017) used Deep Belief Networks (DBN) achieving 91.76% accuracy. Zaabi (2020) employed CNN and Transfer Learning for 92.68% accuracy. Aruchamy (2020) proposed a modified kNN algorithm with 93.18% accuracy. Alshammari (2021) introduced a modified CNN for 97% accuracy. Yagis (2021) explored

TABLE I
THE NUMBER OF PARAMETERS IN THE MODEL

Layer	Output Shape	Parameters
ConvLayer1	(None, 128, 128, 32)	896
Pool1	(None, 64, 64, 32)	0
ConvLayer2	(None, 64, 64, 64)	18496
Pool2	(None, 32, 32, 64)	0
ConvLayer3	(None, 32, 32, 64)	36928
Pool3	(None, 16, 16, 64)	0
ConvLayer4	(None, 16, 16, 128)	73856
Pool4	(None, 8, 8, 128)	0
flatten1	(None, 8192)	0
dense1	(None, 256)	2097408
dropout1	(None, 256)	0
dense2	(None, 128)	32896
dropout2	(None, 128)	0
dense3	(None, 64)	8256
dense4	(None, 4)	260
Total parameters:	2268996	(8.66 MB)
Trainable parameters:	2268996	(8.66 MB)
Non-trainable parameters:	0	0 MB

Convolutional Autoencoders achieving 74.66% accuracy. Fareed (2022) introduced ADD-Net with 98.5% accuracy and high precision, recall, and F1-score. Z. Xia (2020) used 3D CNN and CLSTM for 94.19% accuracy. Finally, our proposed network achieved 99.19% accuracy, low loss (0.023), and impressive precision, recall, and F1-score, demonstrating its effectiveness.

IV. RESULTS

Environmental Setup The trials were conducted on a system that had Nvidia RTX 3090 Ti – 24 GB GPU RAM, Intel core i7-12700K (5.0GHz). The evaluation of the model was executed by employing the test set, which was generated through the division of the dataset prior to the model's training.

The hybrid Alzheimer's disease detection model, using a basic CNN demonstrated promising results in rigorous evaluations using diverse neuroimaging datasets. Using a variety of performance measures, such as F1 score, recall, accuracy, and precision, the model consistently outperformed in distinguishing Alzheimer's patterns.

The collaborative strengths of the basic CNN in capturing spatial information lead to enhanced diagnostic accuracy. Validation on separate datasets underscored the model's robustness and generalization capabilities, while interpretability was ensured through visualization techniques. Figure 2 explains the accuracy of the model,with X represents the epoch and Y representing the accuracy.

Datasets underscored the model's robustness and generalization capabilities, while interpretability was ensured through visualization techniques. Figure 2 explains the accuracy of the model, the x-axis depicts the

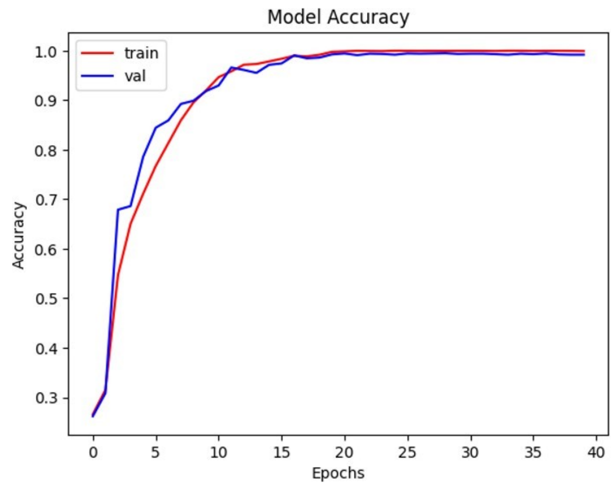

Fig. 2. Accuracy Plot of Proposed Model

epochs, and the Y-axis depicts accuracy. After training the model for a few epochs the loss remains negligible which is represented in Figure 3. Figures 4 and 5 represent the plots of F1-Score and precision of the proposed model respectively.

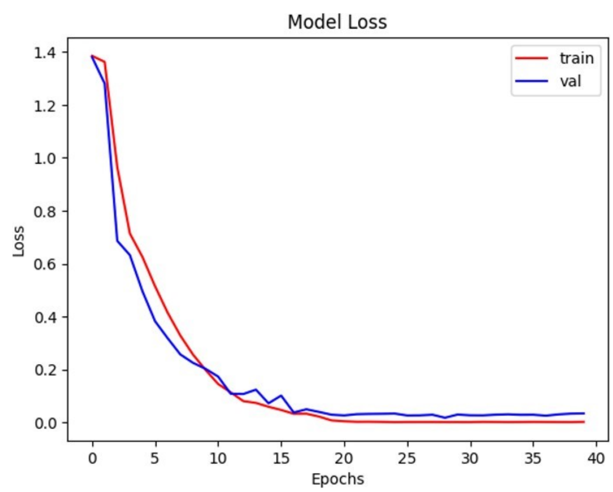

Fig. 3. Loss function of Proposed Model

In the realms of machine learning and statistics, a confusion matrix serves as a tabular assessment tool, critically analyzing the effectiveness of a classification model by providing a comprehensive overview of its performance. Figure 6 shows the prediction between the actual outcomes and the predicted outcome. Most of the samples are predicted accurately but some of the samples are predicted incorrectly.

Overall, findings highlight the model's efficiency as a valuable tool for early and accurate Alzheimer's detection, emphasizing the success of the hybrid architecture

TABLE II
THE PERFORMANCE METHOD IS COMPARED WITH OTHER STATE-OF-THE-ART METHODS

Approach	Year	Method	Accuracy	Loss	Precision	Recall	F1-Score
Faturrahman [5]	2017	DBN (Deep Belief Network)	91.76%	-	-	-	-
Zaabi [6]	2020	CNN + Transfer Learning	92.68%	-	-	-	-
Aruchamy [7]	2020	Modified kNN Algorithm	93.18%	-	-	-	-
Alshammari [8]	2021	Modified CNN	97%	-	-	-	-
Yagis [9]	2021	Convolutional Autoencoder	74.66%	-	-	-	-
Fareed [10]	2022	ADD-Net	98.5%	0.0549	98.61%	98.61%	99.76%
Z. Xia [11]	2020	3D CNN + 3D CLSTM	94.19%	-	-	-	-
Our Proposed method			99.19%	0.023	99.11%	99.11%	99.05%

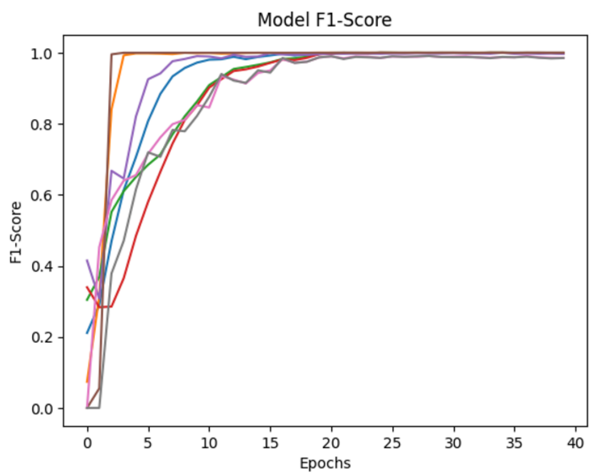

Fig. 4. F1-Score of Proposed Model

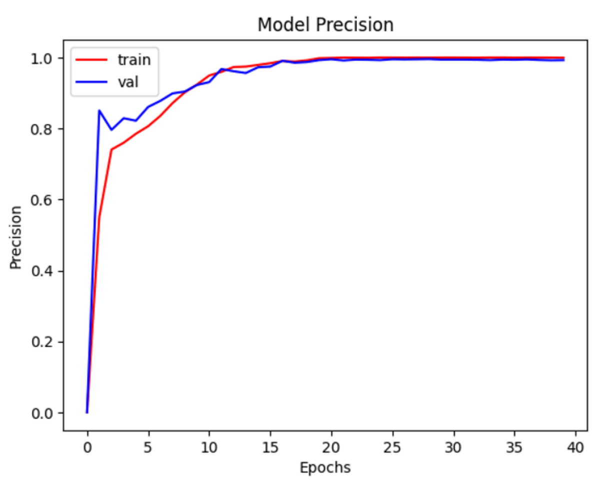

Fig. 5. Precision of Proposed Model

in improving diagnostic precision in complex medical applications.

Table III presents various evaluation metrics for a particular model or system. Firstly, the accuracy stands at an impressive 99.14%, indicating the proportion of correctly classified instances out of the total. A low loss value of 0.023 suggests the model's efficiency in minimizing errors during training. The F1-score, a measure of a model's accuracy that considers both precision and recall, is reported at 99.08%, highlighting the balance between precision and recall. Precision, which denotes the accuracy of positive predictions, is noted at 99.11%, further emphasizing the model's ability to identify relevant instances correctly. These metrics collectively demonstrate the high performance and reliability of the evaluated model in its classification task.

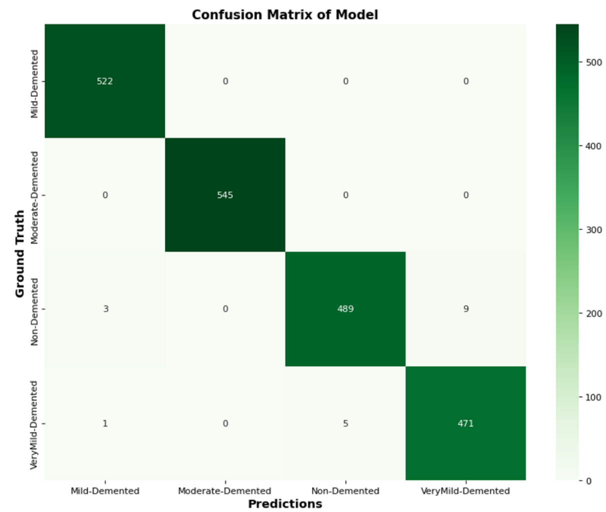

Fig. 6. Confusion Matrix for the Network

TABLE III
EVALUATION METRICS

S.no	Evaluation Metric	Percentage
1	Accuracy	99.14%
2	Loss	0.023
3	F1-Score	99.08%
4	Precision	99.11%

V. CONCLUSION AND FUTURE WORK

Alzheimer's disease progresses over time and affects patients at different times, an accurate diagnosis depends on knowing the disease's current stages, For the proposed model, the corresponding values were 99.19 percent, 0.023 percent, 99.08 percent, 99.11 percent, accuracy, loss, and f1-score. For the given dataset, the recommended model incurred the least amount of loss while performing more accurately than the other traditional approaches. The Kaggle dataset consists of four classes namely mild demented, very mild demented, moderate demented, and non–demented images. effectiveness.99.19%,0.023%, 99.08%, and 99.11% accuracy, loss, f1-score, and precision, respectively. The accuracy, loss, f1-score, precision, and accuracy of the suggested model were 99.19 percent, 0.023 percent, 99.08 percent, and 99.11 percent, respectively. The model outperformed conventional techniques in accuracy and minimal loss for the specified dataset. For prospects, the noise during the model's training can be minimized to enhance its efficiency.

One of the project imperfections is the data balancing algorithms that can least tend to high noise and overfitting of the data. To deal with this issue, the project should build an algorithm possessing the property of the highest noise reduction while being the least likely to compromise model performance. Furthermore, the data scarcity problem creates another impediment impacting the development of the data-augmentation or transfer learning techniques that can be used to increase the said dataset for training purposes. The final critical step is the provision of clinical validation to prove the reasonability of the algorithm in the real-world healthcare environment. So, the algorithm's applicability to healthcare facilities may be perfectly done.

ACKNOWLEDGMENT

We sincerely thank our management for providing a high-end NVIDIA RTX 3090 Ti - 24GB GPU Machine for our research work.

REFERENCES

[1] E. Nichols and T. Vos, "The estimation of the global prevalence of 753 dementia from 1990–2019 and forecasted prevalence through 2050: An 754 analysis for the global burden of disease (GBD) study 2019," Alzheimer's 755 Dementia, vol. 17, no. S10, pp. e105–e125, Dec. 2021. 756

[2] M. A. Ebrahimighahnavieh, S. Luo, and R. Chiong, "Deep learning 760 to detect Alzheimer's disease from neuroimaging: A systematic litera- 761 ture review," Comput. Methods Programs Biomed., vol. 187, Apr. 2020, 762 Art. no. 10

[3] L. G. Apostolova, A. E. Green, S. Babakchanian, K. S. Hwang, Y. Y. Chou, 742 A. W. Toga, and P. M. Thompson, "Hippocampal atrophy and ventricular 743 enlargement in normal aging, mild cognitive impairment and Alzheimer's 744 disease," Alzheimer Disease Associated Disorders, vol. 26, no. 1, p. 17, 745 2012.

[4] M. Tanveer, B. Richhariya, R. U. Khan, A. H. Rashid, P. Khanna, 757 M. Prasad, and T. C. Lin, "Machine learning techniques for the diagnosis 758 of Alzheimer's disease: A review," ACM Trans. Multimedia Comput. 759 Commun. Appl., vol. 16, no. 1s, pp. 1–35, Apr. 2020.

[5] Faturrahman, M., Wasito, I., Hanifah, N., & Mufidah, R. (2017, October). Structural MRI classification for Alzheimer's disease detection using deep belief network. In 2017 11th International Conference on Information & Communication Technology and System (ICTS) (pp. 37-42). IEEE.

[6] Zaabi, M., Smaoui, N., Derbel, H., & Hariri, W. (2020, July). Alzheimer's disease detection using convolutional neural networks and transfer learning Based methods. In 2020 17th International Multi-Conference on Systems, Signals & Devices (SSD) (pp. 939-943). IEEE.

[7] Aruchamy, S., Mounya, V., & Verma, A. (2020, December). Alzheimer's disease classification in brain MRI using modified kNN algorithm. In 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC) (pp. 1-6). IEEE.

[8] Alshammari, M., & Mezher, M. (2021, March). A modified Convolutional Neural Networks for MRI-based images for detection and stage classification of alzheimer disease. In 2021 National Computing Colleges Conference (NCCC) (pp. 1-7). IEEE.

[9] Yagis, E., De Herrera, A. G. S., & Citi, L. (2021, June). Convolutional Autoencoder based Deep Learning Approach for Alzheimer's Disease Diagnosis using Brain MRI. In 2021 IEEE 34th International Symposium on Computer- Based Medical Systems (CBMS) (pp. 486-491). IEEE.

[10] Fareed, M. M. S., Zikria, S., Ahmed, G., Mahmood, S., Aslam, M., Jillani, S. F., ... & Asad, M. (2022). ADD-Net: An Effective Deep Learning Model for Early Detection of Alzheimer Disease in MRI Scans. IEEE Access, 10, 96930- 96951.

[11] Z. Xia et al., "A Novel End-to-End Hybrid Network for Alzheimer's Disease Detection Using 3D CNN and 3D CLSTM," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 2020, pp. 1-4, doi: 10.1109/ISBI45749.2020.9098621.

[12] Rao, K. N., Gandhi, B. R., Rao, M. V., Javvadi, S., Vellela, S. S., & Basha, S. K. (2023, June). Prediction and Classification of Alzheimer's Disease using Machine Learning Techniques in 3D MR Images. In 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS) (pp. 85-90). IEEE.

[13] Alshammari, M., & Mezher, M. (2021, March). A modified Convolutional Neural Networks for MRI-based images for detection and stage classification of alzheimer disease. In 2021 National Computing Colleges Conference (NCCC) (pp. 1-7). IEEE.

[14] Raza, N., Naseer, A., Tamoor, M., & Zafar, K. (2023). Alzheimer disease classification through transfer learning approach. Diagnostics, 13(4), 801.

[15] Manzak, D., Çetinel, G., & Manzak, A. (2019, August). Automated classification of Alzheimer's disease using deep neural network (DNN) by random forest feature elimination. In 2019 14th International Conference on Computer Science & Education (ICCSE) (pp. 1050-1053). IEEE.

[16] Mandal, P. K., & Mahto, R. V. (2023). Deep Multi-Branch CNN Architecture for Early Alzheimer's Detection from Brain MRIs. Sensors, 23(19), 8192.

[17] Alqahtani, N., Alam, S., Aqeel, I., Shuaib, M., Mohsen Khormi, I., Khan, S. B., & Malibari, A. A. (2023). Deep belief networks (DBN) with IoT-based alzheimer's disease detection and classification. Applied Sciences, 13(13), 7833.

[18] Jabason, E., Ahmad, M. O., & Swamy, M. S. (2018, October). Missing Structural and Clinical Features Imputation for Semi-supervised Alzheimer's Disease Classification using Stacked Sparse Autoencoder. In 2018 IEEE biomedical circuits and systems conference (BioCAS) (pp. 1-4). IEEE.

[19] Subetha, T., Khilar, R., & Sahoo, S. K. (2020, July). An Early Prediction and Detection of Alzheimer's Disease: A Comparative Analysis on Various Assistive Technologies. In 2020 International Conference on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE) (pp. 1-5). IEEE.