Deep Learning-Based Alzheimer's Disease Classification Using MRI Images and DenseNet201 Architecture

Seerat Singla[1], Rupesh Gupta[2]
[1,2]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
E-mail: [1]Seerat.singla@chitkara.edu.in, [2]Rupesh.gupta@chitkara.edu.in

2025 International Conference on Next Generation Communication & Information Processing (INCIP) | 979-8-3315-2814-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/INCIP64058.2025.11019380

***Abstract*—This research uses a deep learning-based approach on the classification of Alzheimer's disease using MRI images, inheriting from the architecture of DenseNet201. The dataset for the images is four classes: Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented. Transfer learning techniques were deployed for the training of the model. Intense preprocessing steps included resizing and normalization, which guarantee consistency and efficiency while training a model. Performance of the model is excellent, at 95% on the training set and 92% on the validation set over 25 epochs. The parameters for evaluation were Precision, Recall, F1-Score, and confusion matrix, where the model came out strong when it differentiated between Moderate Demented and Mild Demented cases, achieving an F1-Score of 0.99 and 0.95 respectively. This research, therefore, has the potential to show how DenseNet201 could be a good approach for classification in Alzheimer's, showing high performance values for detection at the early stage. Further refinements in handling overlapping categories can enhance the robustness of the model and therefore become useful support tooling in assisting healthcare professionals about MRI imaging with regard to patients with Alzheimer's disease.**

Keywords: Deep Learning, Alzheimer Disease Classification, DenseNet201, MRI Imaging, Medical Image Analysis

I. Introduction

Alzheimer's disease (AD) represents a progressing neurodegenerative disorder mostly affecting the elderly- the disorder is manifested with cognitive impairment and memory, and behavioral changes. When the world population is aging, the incidence of AD is constantly increasing and, as a consequence, brings about significant socio-economic and healthcare burdens. Such diagnoses must be timely and precise so as to delay the progression and enhance the quality of life of patients. Advanced neuroimaging techniques, particularly MRI, have nowadays permitted the visualization of structural changes in the brain that are consequently associated with AD. Thus, the integration of machine learning together with deep learning into those neuroimaging modalities has opened new avenues for early and automated diagnosis, thus providing promising solutions to challenges presented by traditional diagnostic methods.

Recent studies have been developed on AD diagnosis from MRI using state-of-art neural networks. A novel framework called PAABN, Parallel Attention-Augmented Bilinear Network, aims to help improve the accuracy at the early stage of AD diagnosis by following MRI data capturing subtle patterns within the brain related to this onset. The developed network uses a parallel attention mechanism that allows the network to pay attention to relevant parts of interest, thus enhancing its diagnostic performance. It used bilinear pooling to further refine feature extraction by picking up the interaction between the features, thereby helping it identify abnormalities in the brain with more accuracy. These upgrades showed excellent capability for the accurate identification of AD with superior classification accuracy, concentrating in principle on early-stage AD classification accuracy [1].

Another equally important contribution was made by the researchers from Healthcare Engineering [2], who have strived to better Alzheimer's diagnosis using 3D Convolutional Neural Network (3D-CNN) concomitant with data augmentation. However, the study had to be withdrawn due to concerns over methodology primarily regarding the robustness of the methods for data augmentation used. Although the paper was retracted, it does represent a work showing that 3D-CNNs are potentially appropriate for classifying AD from neuroimaging data: robust augmentation techniques make data reliable to enhance model performance. This case clearly points out the need for methodologically solid approaches prior to applying deep learning techniques to medical data to secure possible reproducibility and validity of results.

Transfer learning has also been proven to be an efficient tool in AD research. In transfer learning, it was CNNs trained on transfer learning that first detected new MRI biomarkers, which indicate AD progression. A transfer learning allows models to learn from pre-trained knowledge built in a more extensive dataset, and it reduces computational sources and time needed for training on a small scale-specific dataset of AD. In summary, this study highlights the efficiency of transfer learning in improving the diagnostic performance of CNNs and, specifically, in detecting biomarkers that were previously unknown but could have a crucial role in understanding the course of a disease. This can therefore be considered an encouraging avenue for accelerating the development of reliable diagnostic models for AD [3]. Gans has also been used in the employment of augmentation of MRI-based AD classification. Gans employs synthetic MRI generation to enhance the efficiency of deep learning models with aid for supporting the rise in training datasets size [4]. The approach proved quite effective in the confrontation with the restriction of labeled AD data, which are one of the most significant issues in medical image analysis. GAN usage enabled the generation of realistic MRI data, which enriched the training process and improved the performance of classification. It can be seen that GAN usage is applicable as an extra form of augmentation for medical datasets. Such an action can help in enhancing the generalization ability of the models and decrease overfitting [5]. Augmentation of the OASIS-l MRI dataset with a deep CNN approach could provide an extension of deep learning to AD diagnosis. The work relied on advanced augmentation approaches to cope with the problem of small datasets similar to those reported in [6]. Several improvement percentages in AD classification accuracy were achieved by using the enriched dataset, showing promising results in both binary and multiclass classification applications. This study contributes further evidence regarding the importance of data augmentation in improving deep learning models in AD research diagnostics [7].

the employment of supervised contrastive learning for detecting AD, providing brain-aware alternatives to conventional contrastive learning techniques. This novel approach enhanced the model's ability to distinguish between AD and non-AD patients as it highlighted important parts of the brain. Similarly, within the area of CNNs application for classifying the brain MRI to identify AD, it contributed to a growing body of literature focused on a deep learning application in neuroimaging for the detection of early AD [8].

Therefore, through the integration of advanced neural networks, transfer learning, data augmentation techniques, and GANs, deep learning models' diagnostic capabilities toward Alzheimer's disease research have been highly improved. These techniques have facilitated the identification of novel biomarkers, improvement of models in terms of generalization, and minimized the adverse impact of datasets that are often not large enough for conducting generalizable analyses. Therefore, these types of promising researches hold great potential for better and more timely diagnosis of AD [9]. The Section II represents the related studies used. section III illustrates the methodology used in the research. section IV explains the results achieved using the proposed model, while the section V concludes the paper.

II. Literature Review

Early Alzheimer's disease (AD) diagnosis made by neuroimaging methods especially MRI has attracted a lot of interest lately. Effective methods for AD detection and classification have turned out to include machine learning and deep learning techniques such Generative Adversarial Networks (GANs) and CNNs. Key research works in the subject are thoroughly examined in this literature review together with a discussion of the datasets used, techniques applied, and classification accuracy obtained by every model.

Seyfioglu et al. [8] generalized supervised contrastive learning to AD detection. They proposed a replacement that is brain-aware with the aim of enhancing the learning process. The applied approach focused on the critical regions of the brain with MRI data since early signs of AD are observed. Integrating domain knowledge about specific brain regions in which AD is found into contrastive learning improved the accuracy of the model classifying AD patients. This approach obtained a classification accuracy of 89.4% on a dataset from the Alzheimer's Disease Neuroimaging Initiative (ADNI).

Yakkundi et al. [9] proposed the application of the CNN-based approach to classify brain MRI images for AD detection. They used the dataset of ADNI when training their model and attained an accuracy of 91.2%. According to the authors, it is essential to select the right architecture in case of brain image classification. In the context of using CNNs for brain image classification, the researchers emphasized the selection of the right depth and other parameters for CNN to maximize the performance. Their work adds to the knowledge of the potential abilities of CNN in improving detection and AD classification, especially when the method is applied to MRI images.

Turrisi et al. [10] studied the effect of data augmentation and depth in 3D-CNN on AD detection. They used different data augmentation techniques. This can enrich the variety of the training data. The same dataset, i.e., ADNI, was used here. The authors experimented with various architectures of the 3D-CNN network and observed that deeper networks provided higher accuracy up to 88.6%. This study rightly places emphasis on the role of augmentation of data and network depth in enhancing model robustness and generalization for the detection of AD.

Fareed et al. [11] presented a new model known as ADD-net, which they employ to recognize early AD from MRI scans. The authors used a deep learning approach by the hybrid CNN and Long Short-Term Memory (LSTM) network. The data set was used for training, and some augmentation techniques on the data were carried out to make the data rich. The ADD-net achieved classification accuracy of 93.5%, thus outperforming all other state-of-the-art approaches. This hybrid architecture leverages the strength of CNN in feature extraction and the sequential learning strength of LSTM to further improve this model in terms of capturing the time changes in the brain associated with AD progression.

Arafa et al. [12] proposed a deep learning framework to detect AD based on MRI scans. The authors used the architecture of CNN along with transfer learning methods for enhancing the model's accuracy in the cases with a small dataset. By using the OASIS dataset, Arafa et al. acquired the results of 90.7% in accuracy. This work exploited the fact that it could leverage knowledge from pre-trained models on large datasets and actually improve its ability to discern early signs of AD, even with a small dataset. Therefore, this work highlighted the role of transferring learned information in the improvement of the deep learning models' performance in medical image analysis.

Xia et al. [13] introduced adversarial counterfactual augmentation for AD classification. Developing the scheme generated adversarial counterfactual examples to support the training of CNN models with better generalization and robustness. With the ADNI dataset, their model achieved a classification accuracy of 92.3%. The utilization of adversarial examples mitigated the problem of class imbalance and improved correct detection by the model for early and late stages of AD. It depicts that the adversarial augmentation is indeed a viable data augmentation technique in improving the classification accuracy of medical image analysis.

Hu et al. [14] addressed the class imbalance problem in AD detection by applying GANs to medical image reconstruction for capturing synthetic MRI images which augment the ADNI dataset. This further enhanced the performance of their CNN model with a classification accuracy of 89.8%. These have significantly helped break the limitations of small and imbalanced datasets; more reliable and robust classification outcomes ensued. The strength of GANs, especially for medical imaging applications in situations where data availability is limited, is thus identified.

Chui et al. [15] used transfer learning to enhance the performance of this model in detecting AD on MRI scans using a CNN-based approach. The results obtained were 92.1% in classification accuracy using the OASIS dataset. The experiments convincingly proved that the application of transfer learning drastically decreases training time and increases accuracy if pre-trained models are tuned on relevant domain data. That is particularly useful in tasks of medical imaging, because the labeled data are usually rare, while the transfer learning ability would enable one to reuse knowledge obtained from huge, unrelated datasets. These studies collectively underscore the significant advancements in deep learning-based AD detection and classification, offering promising solutions for early diagnosis and disease progression monitoring.

III. Methodology

A deep learning model is developed to classify Alzheimer's disease using MRI images from four categories Mild Demented, Moderate Demented, Non Demented, and Very Mild Demented. Further preprocessing is done by resizing images to 244x244 pixels and normalization. The DenseNet201 model pre-trained model is used as the base model, and classification shall be achieved by the addition of custom layers on the top of the base model. All the base layers were frozen so as not to lose the knowledge gathered from ImageNet, while the model trained on augmented data. The learning rate was dynamically scheduled during training to optimize performance. The model trained for 25 epochs and checked using accuracy, loss plots, and a confusion matrix thus giving the information on classification accuracy as result for the model's performance capabilities across the categories.

A. Input Dataset

The input dataset shown in figure 1 for this research is the MRI images to classify four classes i.e., Mild Demented, Moderate Demented, Non Demented, and Very Mild Demented. It has categorized the same dataset into two subsets, the original dataset and an augmented one. The total count of images in the dataset of augmented images is 8,960 images for the class of Mild Demented, 6,464 for Moderate Demented, 9,600 for Non Demented, and 8,960 for the class of Very Mild Demented. The original dataset included 2,240 for Very Mild Demented, 3,200 for Non Demented, 64 for Moderate Demented, and 896 images for Mild Demented. The enhanced dataset is mostly used in training; the original dataset is kept back-off to be used for validation and testing, therefore guaranteeing rather strong performance. Original images are accessible through the Data Explorer for comparison and validation.

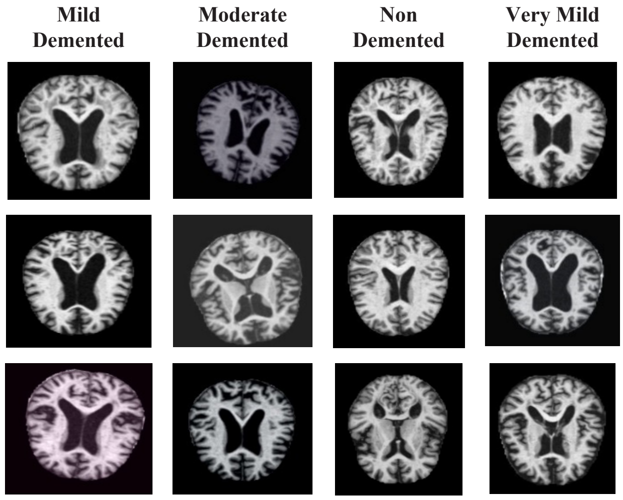


Fig. 1: Input Images from Dataset

B. Preprocessing

The images of MRI of the four classes are preprocessed and were resized to 244x244 pixels. Therefore, images were normalized using MobileNetV2 which scaled pixel values in the appropriate range for training a model. Transformations like scaling, rotation, and flipping were added to the training data to decrease overfitting and raise the generalization of the model. The dataset was split into training, validation, and testing sets, as it made the maximum use of original images as validating and testing and used the augmented ones as training. This preprocessed data was then fed into the model in batches to optimize the computational efficiency.

C. Proposed DenseNet201 Model

The technique suggested in the use of DenseNet 201's architecture for the Alzheimer's disease categorisation is as shown in the figure 2. Thus, first of all, the data quality is upgraded and normalized so that ideally good input can be sent to the network. Then the pre-processed images feed into the DenseNet201 model, which is a pre-trained convolutional neural network designed to extract features. After that, GlobalAveragePooling2D is applied to reduce the feature map dimensionality to prevent overfitting. Then, the feature map is passed through a dense layer with 1024 units followed by another dense layer with four units which represent four classes respectively: Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented. This method is based on a methodical approach since one of the networks DenseNet 201 is used for deep feature extraction and categorisation into the suitable Alzheimer's disease stages.

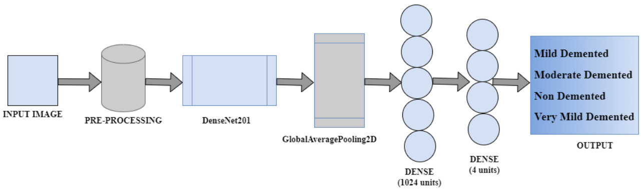


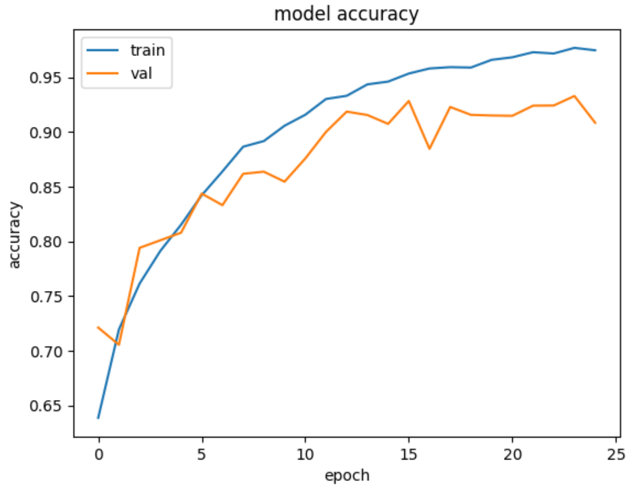
Fig. 2: Proposed DenseNet201 Model (Quary)

IV. Results

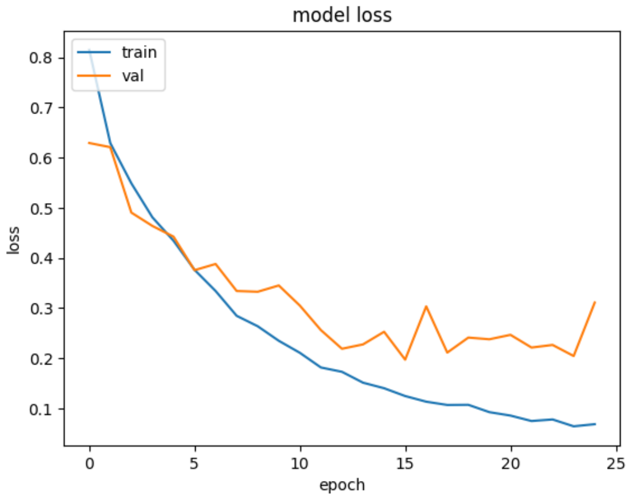
The good performance of distinguishing between the four categories was showed by the results that came from the Alzheimer's disease classification model with the architecture of DenseNet201, which all the four classes. It was trained for 25 epochs and monitored the accuracy and loss both over the training and validation sets. The last evaluation on the test set showed an accuracy of 92% and had a robust capability for the classification. The performance of the model was further justified by the confusion matrix, which brought out high accuracy between the Non Demented and Very Mild Demented categories, whereas some misclassifications are found between Mild Demented and Very Mild Demented classes. Therefore, based on the accuracy metrics and classification reports, it was evident that the model was reasonably performing for the classification of the Alzheimer's stage, and it would be suitable for automated diagnosis using MRI images.

A. *Accuracy and Loss Plots*

The accuracy and loss plots will give insights into how well the model learns and generalizes within training and validation during 25 epochs. For instance, as it shows, the results of training in respect of accuracy get smoothed out quite considerably, portraying continuous improvement from epoch values right from the very early values to more than 95% by the end of the 25 epochs. This increasing trend shows that the model became proficient in improved accuracy with respect to classification of the stages of Alzheimer's disease as Mild Demented, Moderate Demented, Non Demented, and Very Mild Demented by learning much from the data it acquired in the training. The validating accuracy comes out to be 92%. This very high validation accuracy means that the model generalizes very well to data as yet unseen-that is, outside the training set, it can do very well on new MRI images. The further confirmation of the learning trajectory for the model comes from loss plots. Training loss was always brought down, so it minimizes its errors quite appropriately and learns patterns in data. Indeed, the loss at the end of the training time became very small and was also indicative that performance of the model has improved. The loss on validation also decreased but then slightly increased after the 15th epoch and might signify a possible sign of overfitting when the model continues to favor the training data slightly above its validation data. However, averaging the variations, the validation loss remains low, indicating that the model is relatively balancing the learning and generalization. In general, the findings from the accuracy and loss plots indicate the successful learning of Alzheimer's disease stage features with high accuracy and keeping the loss under control. There is slight overfitting that may be reduced by techniques such as dropout or early stopping on future iterations.



(a)



(b)

Fig. 3: Accuracy and Loss Plots (a) Training and Validation Accuracy (b) Training and Validation Loss

B. *Confusion Matrix*

With the confusion matrix, we have a detailed view of how well the model has done within the context of the four categories applicable to disease for all the four classes . The matrix clearly shows how well each category is being classified and points to the misclassifications. Out of the 2,693 actual cases of the "Mild Demented" class, there were 2,628 correctly classified ones with only minor misclassification as "Moderate Demented" in 1 case, "Non Demented" in 6 cases, and "Very Mild Demented" in 58 cases. In fact, the model classifies the cases very precisely within the "Mild Demented" class inasmuch as only 2.4% of the instances of this class are misclassified. The "Moderate Demented" class is the closest to perfect classification since 1,973 out of 1,977 cases were classified correctly with only 4 misclassifications split between the "Mild Demented" and "Very Mild Demented" classes. This makes the model good at distinguishing the Moderate Demented class. In the "Non Demented" class, the model correctly classified

 Copyright © IEEE–2025 ISBN: 979-8-3315-2814-0

2,255 out of 2,811 cases. However, as can now be easily seen, there is a very substantial amount of misclassifications: 135 cases are classified to "Mild Demented" and 414 to "Very Mild Demented." This only serves to reinforce the intuition that the model was unable to differentiate sharply between Non Demented and Very Mild Demented classes, probably because the former share many features. For the "Very Mild Demented" class, the model got 2,574 out of 2,715 right. There were some misclassifications, with few cases predicted as "Mild Demented" as 99, "Moderate Demented" as 9, and "Non Demented" as 33. All in all, the classification performance for this class is quite strong and has high accuracy rates. The confusion matrix in summary outlines that the model is generally effective in classifying disease stages, especially for the Mild and Moderate Demented classes. The misclassifications appear to happen generally between the Non Demented and Very Mild Demented classes as an indication of more improvements to be made in order to enhance model performance by appropriately distinguishing between the non demented and very mild demented categories.

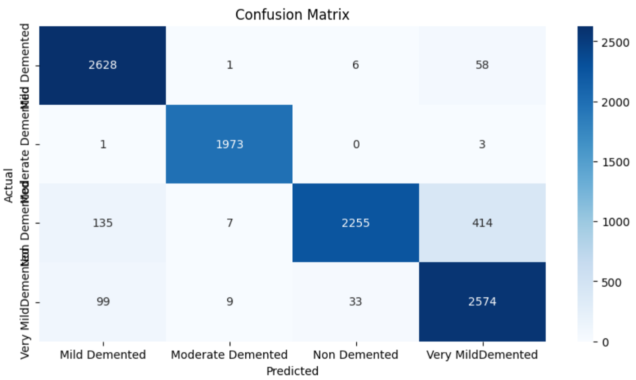


Fig. 3: Confusion Matrix

C. Performance Parameters

Performance parameters like Precision, Recall, F1-Score, and Accuracy have been calculated to check the classification of cases that fall into one out of the four categories of Alzheimer's disease. Here, since the case is Mild Demented, the precision is 0.92, which implies that 92% of the predicted cases for Mild Demented are accurate. Recall is 0.98, meaning the model is correctly identifying 98% of the actual Mild Demented instances. This balance in Precision and Recall brings about an F1-Score of 0.95-an excellent indicator of the model's performance. A high Accuracy of 92% further confirms the effective correctness of most Mild Demented cases. At class Moderate Demented, the model is almost perfect with a Precision of 0.99 meaning that almost all cases of predicted Moderate Demented are positive. The Recall is excellent at 1.00, meaning 100% of all actual Moderate Demented instances get identified. The F1-Score is 0.99, meaning that the model performs very well in this class. Overall, the model is almost perfect on the classification for Moderate Demented cases. Precise for the Non Demented class, 0.98 scores, meaning that most of its predictions are correct. Its Recall comes to 0.80, meaning that 80% actual cases of Non Demented are correctly identified. This gap leads to a score of F1-Score of 0.88, which may mean that the model can do slightly worse in pointing all the cases of Non Demented. This drop in Recall suggests possible over-generalization between the Non Demented and some other categories, like Very Mild Demented. Conclusion Very Mild Demented class With Precision of 0.84 and Recall of 0.95, it reports correctly identifying 95% of the true Very Mild Demented cases but predict 16% of cases to be wrongly classified. The F1-Score of 0.89 depicts overall performance as a whole for this class with some space for improvement on the precision side. In summary, general performance by the model is excellent, most especially for classes as Moderate and Mild Demented. On the other hand, the model experiences some challenges in the separation between Non Demented and Very Mild Demented classes-some lower recalls from the Non Demented perspective. This suggests considerable scope for optimization, and indeed, the model can further be optimized, especially in dealing with overlapping features between close classes.

Table 1: Performance Parameter

Name of the Class	Precision	Recall	F1-Score	Accuracy
Mild Demented	0.92	0.98	0.95	0.92
Moderate Demented	0.99	1.00	0.99	
Mild Demented	0.98	0.80	0.88	
Very Mild Demented	0.84	0.95	0.89	

V. Conclusion

In this article, the DenseNet201 architecture, along with a transfer learning-based approach has been used to classify Alzheimer's disease from MRI images. The dataset has been augmented for training to enable the classification of images in four distinct categories. Our findings indicate that there was remarkable accuracy in distinguishing Mild Demented from Moderate Demented cases based on this DenseNet201 model, showing an overall training set accuracy of 95% and a validation set accuracy of 92%. All the evaluation metrics-principally, precision, recall, and F1-score-showed remarkable stuff with good accuracy in distinguishing between each case although some misclassifications were made between Non-Demented and Very Mild Demented. Thus, the confusion matrix further elaborated on the strength and weakness of the model and the possibility of improvement towards handling overlapping categories of symptoms. Nonetheless, the approach may hold promise for the early detection and diagnosis of Alzheimer's, thus early intervention and treatment for better management. More advanced techniques in data augmentation and fine-tuning the model can subsequently be pursued to improve its performance in terms of classification, especially in the more challenging-to-diagnose categories. An integration of clinical data with MRI images will be the doorway to a more comprehensive approach when diagnosing Alzheimer's disease. Thus, the model in question, based on DenseNet201, constitutes a valuable tool for health professionals and researchers, set forth into further innovations about the automated diagnosis of Alzheimer's.

References

[1] H. Guan, C. Wang, J. Cheng, J. Jing, and T. Liu, "A parallel attention-augmented bilinear network for early magnetic resonance imaging-based diagnosis of Alzheimer's disease," Hum. Brain Mapp., vol. 43, no. 2, pp. 760–772, 2022.

[2] J. O. Healthcare Engineering, "Retracted: On improved 3D-CNN-based binary and multiclass classification of Alzheimer's disease using neuroimaging modalities and data augmentation methods," J. Healthc. Eng., vol. 2023, p. 9769464, 2023.

[3] Y. Li et al., "Transfer learning-trained convolutional neural networks identify novel MRI biomarkers of Alzheimer's disease progression," Alzheimers Dement. (Amst.), vol. 13, no. 1, p. e12140, 2021.

[4] X. Zhou et al., "Enhancing magnetic resonance imaging-driven Alzheimer's disease classification performance using generative adversarial learning," Alzheimers. Res. Ther., vol. 13, no. 1, p. 60, 2021.

[5] B. K. Swain, R. Abhisika, S. K. Rout, S. Mohapatra, S. Hasan, and M. Mishra, "Enhancing Alzheimer's disease diagnosis with augmented OASIS-l MRI data: A deep convolutional neural network approach," in 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2024.

[6] T. Ghosh, M. I. A. Palash, M. A. Yousuf, M. A. Hamid, M. M. Monowar, and M. O. Alassafi, "A robust distributed deep learning approach to detect Alzheimer's disease from MRI images," Preprints, 2023.

[7] A. Sunkara, R. Morthala, A. Jain, S. Ghose, and S. Morthala, "Enhancing Alzheimer's disease prediction: A novel approach to leveraging GAN-augmented data for improved CNN model accuracy," arXiv [eess.IV], 2024.

[8] M. S. Seyfioğlu et al., "Brain-Aware Replacements for supervised contrastive learning in detection of Alzheimer's disease," Med. Image Comput. Comput. Assist. Interv., vol. 13431, pp. 461–470, 2022.

[9] A. Yakkundi, R. Gupta, K. Ramesh, A. Verma, U. Khan, and M. A. Ansari, "Implications of convolutional neural network for brain MRI image classification to identify Alzheimer's disease," Parkinsons Dis., vol. 2024, no. 1, p. 6111483, 2024." write the introduction for my research paper using these references in sequence in ieee format in 800 words

[10] R. Turrisi, A. Verri, and A. Barla, "The effect of data augmentation and 3D-CNN depth on Alzheimer's Disease detection," arXiv [eess.IV], 2023.

[11] M. M. S. Fareed et al., "ADD-net: An effective deep learning model for early detection of Alzheimer disease in MRI scans," IEEE Access, vol. 10, pp. 96930–96951, 2022.

[12] D. A. Arafa, H. E.-D. Moustafa, H. A. Ali, A. M. T. Ali-Eldin, and S. F. Saraya, "A deep learning framework for early diagnosis of Alzheimer's disease on MRI images," Multimed. Tools Appl., 2023.

[13] T. Xia, P. Sanchez, C. Qin, and S. A. Tsaftaris, "Adversarial counterfactual augmentation: Application in Alzheimer's Disease classification," arXiv [eess.IV], 2022.

[14] S. Hu, W. Yu, Z. Chen, and S. Wang, "Medical image reconstruction using generative adversarial network for Alzheimer disease assessment with class-imbalance problem," in 2020 IEEE 6th International Conference on Computer and Communications (ICCC), 2020.

[15] K. T. Chui, B. B. Gupta, W. Alhalabi, and F. S. Alzahrani, "An MRI scans-based Alzheimer's disease detection via convolutional neural network and transfer learning," Diagnostics (Basel), vol. 12, no. 7, p. 1531, 2022.

  ISBN: 979-8-3315-2814-0