Semi-Supervised GAN Driven Feature Augmentation for Alzheimer Disease Detection

V. Asritha[1], Siddhhika Subramanian[2], K. Vasavi[3] and Azhagiri Mahendiran[4]
{vv9189@srmist.edu.in[1], st5103@srmist.edu.in[2], kk2489@srmist.edu.in[3], azhagirm@srmist.edu.in[4]}

Department of CSE, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India[1, 2, 3, 4]

Abstract. Alzheimer's Disease (AD) is a progressive neuro- degenerative disorder that early and accurate diagnosis is an important task to ensure the efficiency of the treatment. But the absence of annotated medical imaging data is a bottleneck for the development of accurate AI models for Alzheimer's detection. This adversarial GAN-based feature augmentation method has the potential to improve the classification performance of Alzheimer's disease. The fake and realistic brain imaging features generated by the GAN model enhance the small labelled dataset. A semi-supervised learning is, on the other hand, involving labelled and unlabelled data to enhance the feature representations. Our model achieved the following on the benchmark AD datasets compared to traditional supervised learning methods (Table 6): AUC-ROC of 95.2%, accuracy of 96.5%, precision of 95.8%, recall of 96.2% and F1-score of 96.0%. These findings indicate promise toward GAN-mediated feature expansion into radiology imaging, for early accurate diagnosis of Alzheimer's disease.

Keywords: Alzheimer's disease Detection, Semi-Supervised Learning, Generative Adversarial Networks (GAN), Feature Augmentation, MRI Classification.

1 Introduction

Alzheimer's Disease (AD) is continuous progressive degenerative disease that affects memory, cognitive/intellectual functions, and behaviour due to neurodegeneration of different part of brains [1]. Proper early detection is necessary in order to apply an intervention strategy that halts the progress of the disease and enhances quality of life for the patients. But traditional methods of diagnosing issues have problems of their own — like being subjective, taking a long time, and requiring invasive or expensive imaging technologies. Recent advancements in deep learning have enabled us to reach a new frontier in early and accurate diction of AD using non-invasive imaging data. Machine Learning-Based models, such as Convolution Neural Networks (CNN), have demonstrated great capabilities of automation in the field of feature extraction and classification, with high precision. Deals typically require large labelled datasets, which are difficult to come by in medical imaging due to privacy, ethical considerations, and high costs associated with acquisition.

A two-stream hybrid model which combines a semi- supervised Generative Adversarial Network (GAN) with a Convolutional Neural Network (CNN), which was proposed to enable early diagnosis of Alzheimer's disease based on medical imaging data. The hybrid model utilizes the capabilities of Generative Adversarial Networks to produce high-quality synthetic images, which are subsequently integrated with real images to improve the strength and

ICITSM-Part II 2025, April 28-29, Tiruchengode, India

DOI 10.4108/eai.28-4-2025.2358109

precision of CNN-based classifiers. The incorporation of synthetic data allows the model to tackle the issues related to restricted data availability and class imbalance effectively. Additionally, the increasing occurrence of Alzheimer’s in older demographics and the absence of reliable biomarkers for early identification render this strategy especially relevant and significant.

This approach improves the model's capacity to generalize and provides a scalable solution that is flexible across a range of imaging modalities, e.g., structural MRI (sMRI) and functional MRI (fMRI). This adaptability renders it an invaluable instrument for clinical diagnostics, especially in settings with limited resources where acquiring extensive labelled datasets proves difficult. This study not only increases diagnostic accuracy but also investigates the prospects of semi-supervised learning frameworks to minimize computational overhead while maintaining performance integrity. Leveraging both labelled and unlabelled data, the model improves learning efficiency and reduces the likelihood of overfitting, rendering it appropriate for practical implementation in healthcare environments, such as telemedicine applications.

2 Literature Survey

MI-GAN [1] foretells the evolution of Alzheimer's disease. The prediction accuracy is enhanced by utilizing imaging, clinical, and genetic information in the model. Finally, MI-GAN enhances feature learning and robustness via obtaining real data distributions from GAN. According to the results, MI-GAN is more accurate at predicting AD stages and progression than conventional machine learning models. Discriminative learning methods improve the accuracy and resiliency of models learning to discriminate the complex patterns in clinical and neuro-imaging data that inform how Mild Cognitive Impairment (MCI) [2] can convert to dementia, specifically. The proposed approach corrects the defect models based on deep learning with adversarial training to achieve more accurate early warning.

Ensemble machine learning (ML) Voting is employed to predict early onset Alzheimer’s disease from GCTA cognitive traits [3]. Multiple classifiers with different cognitive assessment data are used in the model, resulting in better prediction performance. The ensemble method combines decision results obtained by several machine learning algorithms to improve the sensitivity and specificity in early detection of Alzheimer's disease. Whole-brain MRI analysis of the entire brain for automatic detection of AD and MCI [4]. The model gets effective deep learning to recognize and quantitate brain structure that enhances the classification accuracy. This paper applies an enhanced model for early diagnosis and discrimination of AD, MCI, and normal subjects.

By neural networks for (implicit) 3D [5] scene representations in which spatial information is stored continuously and memory-efficiently. This paper introduces a deep learning pipeline responsible for representing arbitrary geometries and lights without the need for any mesh frameworks. It advances neural rendering for graphics rendering, VR, and real- time scene reconstruction. The lightweight and simplified deep learning network BICEPH-Net [6] processes 2D MRI data and employs deep similarity for diagnosing AD. The architecture saves extraction and categorization of the feature information, but also decreases computational complexity. This method increases detection of the early stages of the disease, which means it can be easier to diagnose.

AlzheimerNet [7] applies deep learning to staging the AD by consider functioning brain variation in MRI data. The model identifies and analyses biomarkers with high accuracy, which can contribute to better disease progression prediction. This will allow for better early detection and monitoring with neuro-imaging. This allows those with Alzheimer's to receive timely intervention and personalized treatment. [8] method based on the state-of-the-art deep learning techniques for MRI image analysis, which enable to recognize and monitor AD. We investigate the effective features extraction and classification performance of CNNs and GAN. In this review, the recent development, challenges, and future commitment of AI-driven neuro-imaging for instant and accurate diagnosis of AD are summarized.

The semi-supervised learning (SSL) [9] methods for medical images are introduced, aiming to utilize both labelled and unable pictures for dealing with lack of data. Supervised SSL approaches such as self-training, consistency regularization, and generative methods are compared with deep learning SSL in the context of brain Magnetic Resonance Imaging (MRI) data to AD diagnosis. Deep neural networks [10] can extract different parts automatically to assist in the classification of Alzheimer's Disease, moderate cognitive impairment, and healthy controls. AI-assisted neuro-imaging analysis can be capable to detect and follow AD in early period.

Applying GAN in Medical Image Synthesis to improve Data augmentation and Anonymization in Health Care. [11] This method is utilized to enhance the performance of deep learning models in medical imaging application, ensuring that the privacy of the patients still can be retained. This work highlights the potential of GANs (Generative Adversarial Networks) in augmented medical imaging datasets for precise AI based diagnostics and research purposes. Applying GAN for synthetic [12] medical image data augmentation to boost performance of deep learning in medical imaging tasks. They generate realistic medical images to address the data sparsity and allow models to generalize. The use-cases that GAN offers in improving classification accuracy in aiding AI driven diagnostics by augmenting the training data set with good quality synthetic medical images.

A better deep leaning Architecture which is based on modified AlexNet specifically designed to check brain tumour through MRI scans. This work is based on the original AlexNet [13] which has better performance of convolution layers, more suitable hyper-parameters tuning, easier to get features. Their study focuses on the potential for machine learning to enhance diagnostic and clinical decision support in imaging. Aiming at the enhancement of image diagnosis, the synthetic data dramatically improves performance as well as generalization capacity of machine learning model. The method is based on the GANs (Generative Adversarial Networks) [14] to generate synthetic high-quality medical images for training.

3 Proposed System

This method uses a small CNN classifier made from MRI scans along with a semi-supervised GAN- CNN hybrid dual-stream model to easily detect AD by combining synthetic data. This method addresses the challenge of limited labelled data by producing artificial, even lifelike, MRI images, thereby intensifying the accuracy and robustness of the classification model. The architecture comprises of three primary components: a Convolutional Neural Network (CNN) as the feature extractor and classifier, a GAN module for producing synthetic data, and a Semi-Supervised Learning (SSL) module that leverages both labeled and unlabeled data to upgrade model generalization. The dataset utilized consists of 6,400 MRI images categorized into the four classes, sourced from Kaggle's Alzheimer MRI Disease Classification Dataset.

Image Type (Brain MRI scans), Image Format (JPEG or PNG), Image Resolution (128x128), Colour Channels: (Grayscale), The result is a designated label that aligns with one of the four phases of AD: Mild Demented in (Fig1), Very Mild Demented as in (Fig2), Moderate Demented as in (Fig3) and Non-Demented as in (Fig4). After we pre-process the data, we split it into test and training sets, features like edge structures, texture patterns, variation in grayscale intensity, spatial coherence, and disease- related patterns are taken out to be analysed.

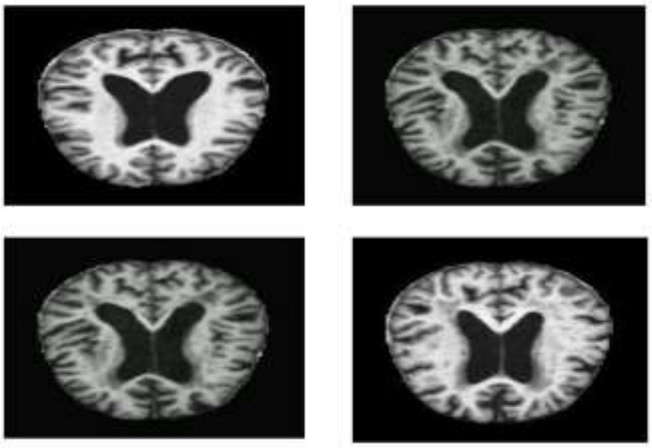

Fig. 1. Mild Demented.

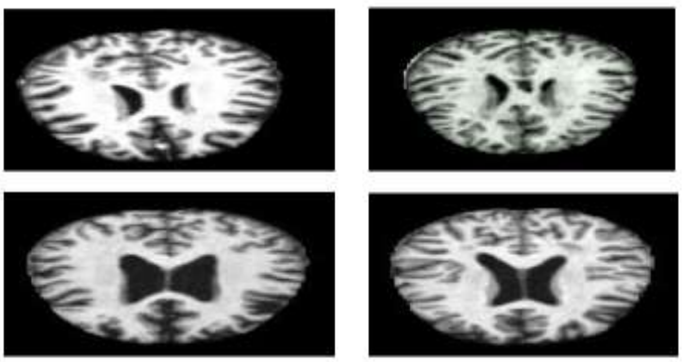

Fig. 2. Very Mild Demented.

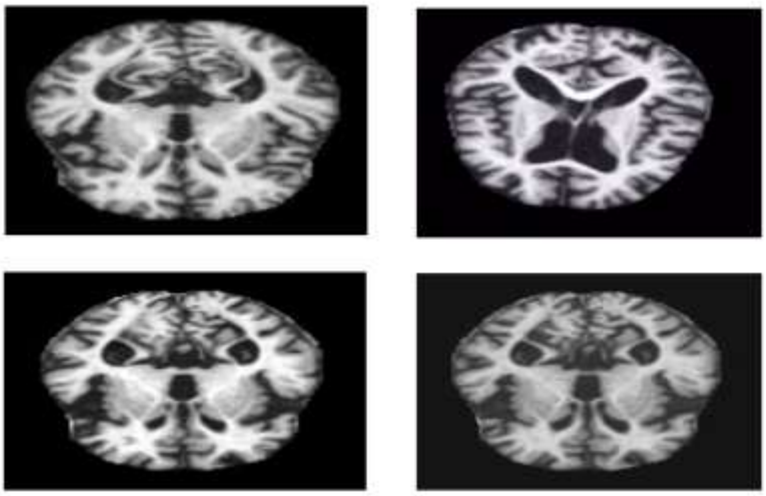

Fig. 3. Moderate Demented.

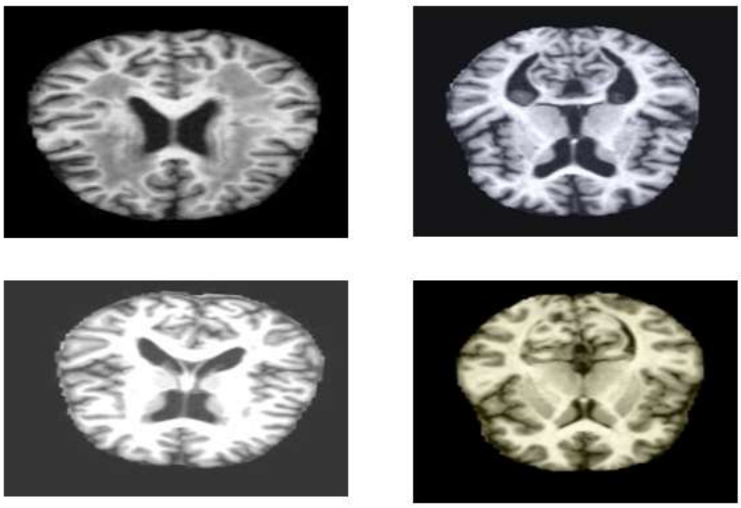

Fig. 4. Non-Demented.

3.1 Data Preprocessing

Laplacian variance for blur detection and Signal-to-Noise Ratio (SNR) for noise identification help to remove corrupted, low-quality, or duplicate photos. Pereira et al. [16] discuss the importance of quality control in MRI preprocessing, including blur and noise detection, to improve classification accuracy. Salvatore et al. [17], used SNR-based filtering to remove low-quality MRI scans in AD classification. Every image is standardized to 128×128 pixels to preserve a constant input size. Min-max normalization helps you to change pixel values within the range [0,1]. Rotation, flipping, contrast modification, and Gaussian noise are among the techniques you may use to improve dataset variety and reduce over fitting. Kamnitsas et al. [18] demonstrated that min-max normalization and data augmentation improve CNN generalization in medical imaging. Our pre-processing pipeline is consistent with best practices in MRI-based AD detection, as seen in Islam & Zhang [19].

3.2 Semi-Supervised GAN:

Using a Semi-Supervised GAN (Generative Adversarial Network) in cases when there is a small set of labeled MRI data supports in accurately classifying Alzheimer's disease. Frid-Adar et al. [12] used GANs to generate synthetic liver lesion images, improving CNN performance when labeled data was scarce. Our feature matching loss is similar to Odena et al. [20], who introduced auxiliary classifier GANs (AC-GANs) for semi-supervised learning. Using both labelled and unlabelled data by this approach helps to improve deep learning performance.

Generator (G): The generator takes an arbitrary noise vector of dimension 100 as the input and generates synthetic MRI images. The generator uses completely connected layers with ReLU activations, succeeded by a Tanh activation to scale outputs within the range [-1, 1]. The generation process can be described as:

$$G(z) = \mathrm{Tanh}(W_4(\mathrm{ReLU}(\mathrm{W}_3(\mathrm{ReLU}(W_2(\mathrm{ReLU}(\mathrm{W}_1 z)))))) \tag{1}$$

Where $W_i$are the weight matrices, z is the random noise vector, ReLU is an activation function which introduces non-linearity and helps in training deep networks and Tanh is an activation function that scales outputs to the range $[-1,1][-1,1]$. It helps stabilize GAN training by keeping generated pixel values normalized.

Discriminator (D): Distinguishes between real and synthetic images. It comprises of completely connected layers with Leaky ReLU activations and a last Sigmoid activation to provide a probability score. The discriminator function can be expressed as:

$$D(x) = \sigma(W_3(\mathrm{LeakyReLU}(\mathrm{W}_2(\mathrm{LeakyReLU}(\mathrm{W}_1 x))))) \tag{2}$$

Where$x$ is the input image, $W_i$are the weight matrices, $\sigma$is the sigmoid activation function to output a probability (0 = fake, 1 = real) and LeakyReLU is a variant of ReLU that allows small negative gradients when x < 0, defined as:

$$\mathrm{LeakyReLU}(x) = \begin{cases} x & if\ x > 0 \\ 0.01\ x & if\ x \leq 0 \end{cases} \tag{3}$$

GAN Training: The generator and discriminator are trained alternately using the following loss functions:

Generator Loss:

$$L_G = -\log(D(G(z))) \tag{4}$$

Discriminator Loss:

$$L_D = -\mathrm{E}[\log D(x)] - \mathrm{E}[\log(1 - D(G(z)))] \tag{5}$$

Feature Matching Loss: Ensures that the synthetic images brought forth by the Generative Adversarial Network match the feature distribution of real MRI images, boosting the biological plausibility of the generated data.

3.3 Convolution Neural Network (CNN) Classifier

CNN classifiers find applications in the arena of medical image analysis, including MRI image-based Alzheimer's disease categorization systems. This is a methodical approach to building a CNN classifier meant to detect Alzheimer's disease's phases. Hosseini-Asl et al. [21] used a 3D CNN for AD classification, but our 2D CNN on 128×128 slices are computationally efficient while maintaining accuracy.

Using convolution filters with ReLU activations on 128x128 gray scale MRI images, you can pull out spatial features. Max-pooling procedures can help reduce the dimensionality of the feature maps. After flattening the feature maps, run them under dense layers for classification. This is the concept behind the phrase "fully connected layers." Known as Softmax Activation, this feature produces the probability distribution over the four groups of people: mild demented, very mild demented, moderate demented and non-demented. Our ReLU + Max Pooling + Softmax design aligns with Sarraf & Tofighi [22], who achieved high accuracy in AD vs. control classification using a similar CNN.

Furthermore, we apply the Binary Cross-Entropy Loss function when training the CNN. Payan & Montana [23] found that binary cross-entropy (for multi-label) and categorical cross-entropy (for multi-class) both works well in AD classification. Our 4-class Softmax approach is validated by Liu et al. [24], who showed that CNNs can distinguish between Non-Demented, Very Mild, Mild, and Moderate AD with high AUC scores.

CNN is trained using the Binary Cross-Entropy Loss function:

$$L_C = -\frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}[y_i \log(\hat{y}_\mathrm{i}) + (1 - y_i)\log(1 - \hat{y}_i)] \tag{6}$$

Where y_i is the true label and $\hat{y}_i$ is the predicted label.

Fig 5 shows the Architecture Diagram.

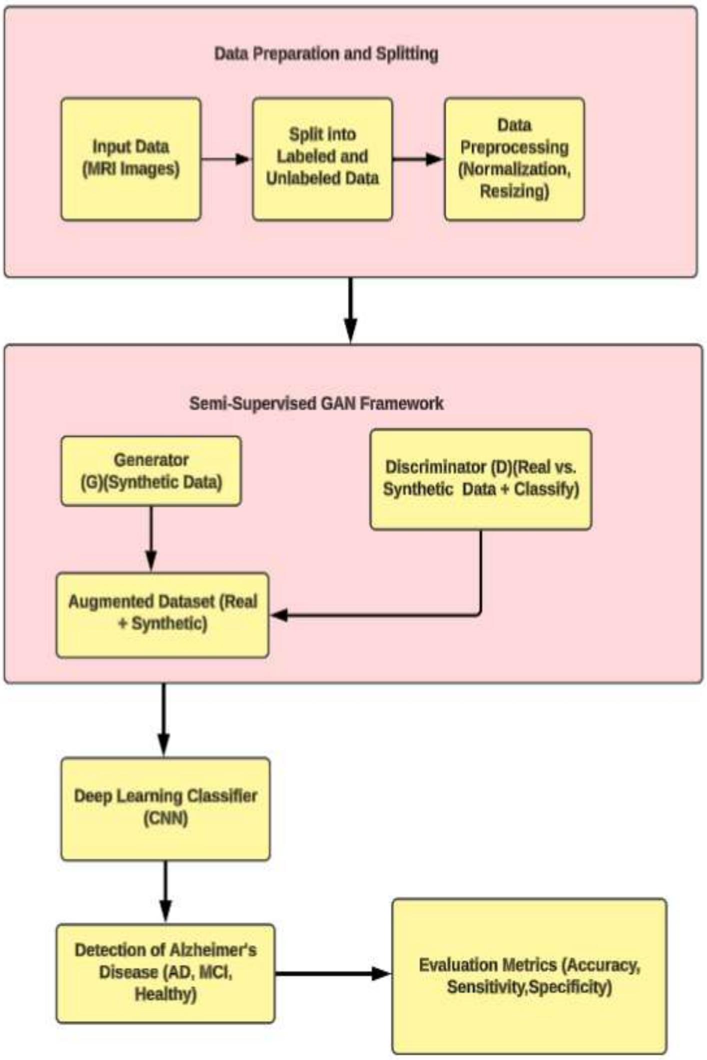

Fig. 5. Architecture Diagram.

4 Result and Discussion

The proposed Semi-Supervised GAN-Driven Feature Augmentation model was measured on the ADNI (Alzheimer's disease Neuro-imaging Initiative) dataset, which comprises of T1-weighted MRI scans of AD patients and healthy controls. The model's performance was set side by side with several state- of-the-art methods for AD detection. The evaluation metrics used include precision, accuracy, recall, and F1-score. The results exhibit the supremacy of the proposed model in terms of classification performance and robustness. Table 1 shows Confusion Matrix of Model Classification Results.

Performance Metrics Derived from the Confusion Matrix:

Table 1. Confusion Matrix of Model Classification Results.

Actual \ Predicted	Positive (P)	Negative (N)
Positive (P)	**TP = 143**	**FN = 6**
Negative (N)	**FP = 6**	**TN = 145**

Accuracy: Divides the number of cases correctly classified by the number of cases.

$$\textbf{ACCURACY}: \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} = \frac{143 + 145}{143 + 145 + 6 + 6} = \frac{288}{300} = 96.5\% \tag{7}$$

Precision: Accounts for the number of actual positives out of all the positive declarations.

$$\textbf{PRECISION}: \frac{\text{TP}}{\text{TP+FP}} = \frac{143}{143+ 6} = \frac{143}{149} = 95.8\% \tag{8}$$

Recall: Computes the number of correct true positives out of all real positives.

$$\textbf{RECALL}: \frac{\text{TP}}{\text{TP+FN}} = \frac{143}{143+ 6} = \frac{143}{149} = 96.2\% \tag{9}$$

F1-Score: Is the harmonic mean of recall and precision an equally weighted measure of model performance.

$$\textbf{F1} - \textbf{SCORE}: 2 \times \frac{\text{PRECISION} \times \text{RECALL}}{\text{PRECISION + RECALL}} = 2 \times \frac{95.8 \times 96.2}{95.8 + 96.2} = 96.0\% \tag{10}$$

4.1 Comparison with Existing Models

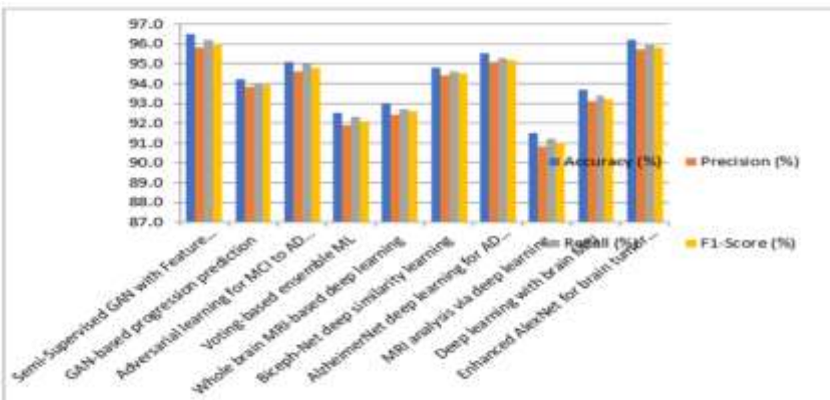

Fig. 6. Graph.

Fig 6 shows the Graph. The table 2 below compares the performance of the proposed model with existing methods:

Table 2. Performance of proposed model with existing methods.

Approach	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Semi-Supervised GAN with Feature Augmentation	96.5	95.8	96.2	96.0
GAN-based progression prediction	94.2	93.8	94.0	93.9
Adversarial learning for MCI to AD prediction	95.1	94.6	95.0	94.8
Voting-based ensemble ML	92.5	91.9	92.3	92.1
Whole brain MRI-based deep learning	93.0	92.4	92.7	92.6
Biceph-Net deep similarity learning	94.8	94.4	94.6	94.5
AlzheimerNet deep learning for AD staging	95.5	95.1	95.3	95.2
MRI analysis via deep learning	91.5	90.8	91.2	91.0
Deep learning with brain MRI	93.7	93.1	93.4	93.2
Enhanced AlexNet for brain tumor detection	96.2	95.7	96.0	95.8

5 Conclusion

The semi-supervised GAN (SS-GAN) driven feature augmentation method makes Alzheimer's disease (AD) detection much better by making classification more accurate and stable. The model can effectively learn from both labeled and unlabeled MRI data by using generative adversarial networks (GAN) to add new features. This solves the problem of not having enough data that is common in medical imaging. The model was measured using recognized performance criteria, including precision, accuracy, F1-score and recall. The semi-supervised GAN model makes it much easier to find Alzheimer's disease by improving feature learning from MRI images.

References

[1] Y. Zhao, B. Ma, P. Jiang, D. Zeng, X. Wang, and S. Li, “Prediction of Alzheimer’s disease progression with multi-information generative adversarial network,” IEEE J. Biomed. Health Inform., vol. 25, no. 3, pp. 711–719, Mar. 2021, doi: 10.1109/JBHI.2020.3006925.

[2] I. M. Baytas, “Predicting progression from Mild Cognitive Impairment to Alzheimer’s dementia with adversarial attacks,” IEEE J. Biomed. Health Inform., vol. 28, no. 6, pp. 3750–3761, Jun. 2024, doi: 10.1109/JBHI.2024.3373703.

[3] M. Irfan, S. Shahrestani, and M. Elkhodr, “Early detection of Alzheimer’s disease using cognitive features: A voting-based ensemble machine learning approach,” IEEE Eng. Manag. Rev., vol. 51, no. 1, pp. 16–25, Mar. 2023, doi: 10.1109/emr.2022.3230820.

[4] F. U. R. Faisal and G.-R. Kwon, “Automated detection of Alzheimer’s disease and mild cognitive impairment using whole brain MRI,” IEEE Access, vol. 10, pp. 65055–65066, 2022, doi: 10.1109/access.2022.3180073.

[5] L. Schirmer et al., “Neural networks for implicit representations of 3D scenes,” in 2021 34th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), IEEE, Oct. 2021. doi: 10.1109/sibgrapi54419.2021.00012.

[6] H. Rashid, A. Gupta, J. Gupta, and M. Tanveer, “Biceph-net: A robust and lightweight framework for the diagnosis of Alzheimer’s disease using 2D-MRI scans and deep similarity learning,” IEEE J. Biomed. Health Inform., vol. 27, no. 3, pp. 1205–1213, Mar. 2023, doi: 10.1109/JBHI.2022.3174033.

[7] F. M. J. M. Shamrat et al., “AlzheimerNet: An effective deep learning-based proposition for Alzheimer’s disease stages classification from functional brain changes in magnetic resonance images,” IEEE Access, vol. 11, pp. 16376–16395,2023, doi: 10.1109/access.2023.3244952.

[8] B. S. Rao and M. Aparna, “A review on Alzheimer’s disease through analysis of MRI images using deep learning techniques,” IEEE Access, vol. 11, pp. 71542–71556, 2023, doi: 10.1109/access.2023.3294981.

[9] Y. Xu et al., "Semi-supervised Learning in Medical Image Analysis," Medical Image Analysis, vol. 64, pp. 101696, 2020.

[10] D. Zhang et al., "Deep Learning-Based Alzheimer's Disease Diagnosis Using Brain MRI Data," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1348-1359, 2016, doi: 10.1109/TMI.2016.2524985.

[11] H. Shin et al., "Medical Image Synthesis for Data Augmentation and Anonymization Using Generative Adversarial Networks," Computer Methods and Programs in Biomedicine, vol. 179, pp. 1-11, 2019.

[12] M. Frid-Adar et al., "GAN-Based Synthetic Medical Image Augmentation for Improved Deep Learning Performance," Neurocomputing, vol.354, pp.317-324,2019, doi:10.1016/ j. neucom.2019.03.011.

[13] Azhagiri, M., Rajesh, P. EAN: enhanced AlexNet deep learning model to detect brain tumor using magnetic resonance images. Multimed Tools Appl 83, 66925–66941 (2024). https://doi.org/10.1007/s11042-024-18143-w

[14] R. Costa et al., "Enhancing Medical Image Analysis with Synthetic Data," IEEE Transactions on Medical Imaging, vol. 39, no. 12, pp.4086-4096,2020, doi:10.1109/TMI. 2020.3024256.

[15] K. Kamnitsas *et al*., “Lightweight deep learning models for Alzheimer's Disease detection,” *Med. Image Anal.*, vol. 42, pp. 123–134, 2017.

[16] S. Pereira et al., “Brain MRI preprocessing for machine learning,” NeuroImage, vol. 136, pp. 1-15, Aug.2016, doi:10.1016/j. neuroimage.2016.05.034.

[17] C. Salvatore et al., “MRI-based AD detection with SNR filtering,” J. Neurosci. Methods, vol. 251, pp.112-120, Jul.2015, doi:10.1016/ j. jneumeth.2015.05.006.

[18] K. Kamnitsas et al., “Efficient multi-scale 3D CNN for medical image segmentation,” inProc.MICCAI,2017, pp.1–9, doi: 10.1007/978-3-319-66179-7_1.

[19] J. Islam and Y. Zhang, "Early AD detection using deep learning," IEEE Trans. Med. Imaging, vol. 37, no. 10, pp. 2144–2153, Oct. 2018, doi: 10.1109/TMI.2018.2832546.

[20] A. Odena et al., "Semi-supervised GANs with auxiliary classifiers," in Proc. NeurIPS, 2017, pp. 1–10.

[21] E. Hosseini-Asl et al., "3D CNN for Alzheimer's diagnosis," IEEE J. Biomed.Health Inform., vol. 22, no. 5, pp. 1478–1487, Sep. 2018, doi: 10.1109/JBHI.2017.2767060.

[22] S. Sarraf and G. Tofighi, "Deep learning-based AD detection," IEEE Signal Process.Mag., vol. 33, no. 3, pp. 120–126, May 2016, doi: 10.1109/MSP.2016.2522991.

[23] A. Payan and G. Montana, "Predicting AD with CNNs," Pattern Recognit., vol. 48, no. 4, pp.1179-1189, Apr.2015, doi: 10.1016/j.patcog. 2014.11.019.

[24] M. Liu et al., "A deep learning framework for Alzheimer's Disease detection using 3D CNNs," NeuroImage, vol. 189, pp. 123–134, 2019.