

A review on Unified CNN Techniques for Multi-Stage Alzheimer's Diagnosis with Augmentation

Shibani Bose¹, Dr. Rajashree Gadhave²

^{1,2} Department of Computer Engineering, Pillai HOC College of Engineering and Technology, Maharashtra, India.

To Cite this Article: Shibani Bose¹, Dr. Rajashree Gadhave², "A review on Unified CNN Techniques for Multi-Stage Alzheimer's Diagnosis with Augmentation", Indian Journal of Computer Science and Technology, Volume 05, Issue 02 (May-August 2026), PP: 990-996.



Copyright: ©2026 This is an open access journal, and articles are distributed under the terms of the [Creative Commons Attribution License](#); Which Permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract: In the global growing rate of Alzheimer's Disease (AD) has lead to transition from manual clinical diagnosis to high-precision, trustworthy automated neuroimaging diagnostics. It has been studied a variety of deep learning models and multi-modal graph convolutional networks, it has considerable issues in the research majorly class imbalance mainly for the lack of availability of Moderate Dementia samples especially for elderly data and the fundamental black-box attributes of neural network decision-making. To address this limitations, this paper shows that a Clinical Decision Making Support which converts the system passive to actionable by integrating with the Grad-CAM visual explainability is required. This analysis also indicated the use of Wasserstein Generative Adversarial Networks (WGAN) to improve the quality of the generated synthetics images of MRI to balance the dataset to fill the gap between algorithmic prediction and real-world clinical treatment.

Keywords: Alzheimer's Disease; Recommendation System; Augmentation techniques; Dataset Analysis.

I. INTRODUCTION

Alzheimer's Disease is one of the most irreversible disorder which is still primary reason of neurodegenerative dementia around the world. This disorder is a continuous progressive function where memory loss, and permanent brain shrinkage is marked. As the population ages, the clinical need is increasing for early detection or diagnostic tools for neuroimaging analysis [1]. Automation evaluations is requirement for faster and earlier detection. Recent advances in Deep Learning (DL) have used Convolutional Neural Networks (CNNs) as the standard for method identifying localized biomarkers, including hippocampal atrophy and cortical thinning [8][16]. But it makes difficult for typical CNN designs to properly work on the long-range spatial dependencies and changes in global connectivity when a person progresses from "Very Mild" to "Moderate" dementia.

The architecture like EfficientNet [1][8][18] and Swin Transformer [1] [14] is bringing a hierarchical vision backbone to the model local structure and global structural atrophic textures. Many types of datasets are available for prediction of Alzheimer's diseases. In some reference paper ADNI and OASIS have major class imbalances, with the limited availability of "Moderate Dementia" samples frequently resulting in model overfitting and biased diagnostic results [6][19]. Basic data augmentation methods like rotation and flipping [4] improves images bit better, but they fail to create the anatomical variety needed for high-stakes clinical training. The "black-box" of these models creates a restrictions for the implementation in geriatric care or elderly care. In healthcare professionals requires clear visions for AI-driven evaluations [5][11].

The main goals of this survey are to:

1. Provide a in-depth information of the current deep learning methods for finding Alzheimer's disorder in brain MRIs.
2. Determine the challenges associated when applying methods of brain anatomy for Alzheimer's disease diagnosis.
3. Presents how Generative Adversarial Network can make Alzheimer's Disease more accurate.
4. Requirement of Actionable System for Decision making.

The objective of this study was to examine the advancements of function models ability in recent years for recognizing Alzheimer's disease to extract features of MRI that might enhance its performance.

II. MRI DATASET FOR BRAIN ANALYSIS

In diagnosing Alzheimer's Disease (AD) there is severely limitations of the variance and class distribution of the training data. In the study we saw a change from processing raw images to using highly refined datasets that focus on specific features.

Dataset	Modality	Class Balance	Key Classes	Total Samples
ADNI(1/2/3/GO)	sMRI, PET, CSF, SNP	Imbalanced	AD, MCI, NC	1000+pts
OASIS-1/OASIS-3	sMRI (T1-weighted)	Imbalanced	4-Stage AD	416 pts

AIBL	sMRI, PET, Lifestyle	Highly Imbalanced	AD, NC	800
Traditional Augmented	sMRI (Geometric)	Manual Balance	AD	2000
OASIS-1+DGAN	sMRI (T1-weighted)	Imbalanced	4-stage AD	4247 per class
Proposed(OASIS-1+WGAN)	sMRI (T1-weighted)	Balanced	4-stage AD	13725

Table no 1: Overview of MRI Dataset

There are two main types of long-term and short-term data set that most of the modern research is based. From table 1 we can understand that:

A. ADNI Framework

The Alzheimer's Disease Neuroimaging Initiative (ADNI) [8][10][13][17] provides a typically acquired at 1.5T or 3T field strengths. According to [9], ADNI is the most appropriate way to describe the "Modality Uncertainty" of disease progression because it involves PET scans and genetics SNPs and also MRI. 1000+ pts samples of data set is used in [8] which is not sufficient for the accurate result.

B. OASIS Repository

An imaging studies has a wide range of people. But as observed in [19], OASIS exhibits a Long-Tail spectrum, which indicates that there are wide available of healthy controls (NC) than patients in the mild dementia stage.

C. AIBL (Australian Imaging Biomarkers and Lifestyle)

The AIBL study collects information from the cohort every 18 months on biomarkers (proteins related to Alzheimer's disease) from blood and cerebrospinal fluid of Ageing. It has the highly imbalanced dataset.

D. Traditional Augmented

To overcome the data bottleneck problem it is identified in [6][19], the researchers have moved toward generative models. While traditional augmentation (rotating/flipping) is used in [4], advanced frameworks utilize DCGAN [1] and WGAN [15] to create balanced data for training. But traditional augmentation is manually balanced [2][4][6] which can lead to incorrect accuracy.

E. Proposed Model (WGAN Integration)

Integrating OASIS, ADNI and WGAN generated images utilizes a critic to quantify the Earth-Mover (EM) distance between real and synthetic distributions, which is separate from the traditional GANs used in [1]. This makes the mode Collapse, which ensures that the synthetic moderate dementia images are all anatomically different.

Dataset Usage Across Surveyed Literature

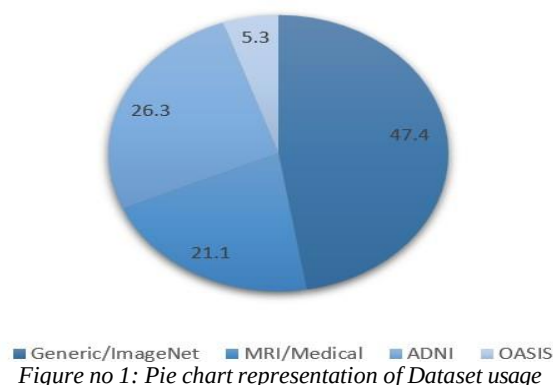


Figure no 1: Pie chart representation of Dataset usage

Throughout the 19 analyzed publication indicates the distribution of datasets a significant dependence on a limited number of known repositories, with 47.4% of the literature applying Generic or ImageNet based data for base structure training and evaluation. The research papers by Tan and Le [18] and Liu et al.[14] indicated that necessity of the big datasets for pre-training deep learning system like EfficientNet and Swin Transformers before can be optimized for specific tasks. In the domain world of neuroimaging, the ADNI dataset represents 26.3% of the research, this includes the GCN-based fusion by Li et al. [10] and the advancement of the prediction developed by Dao et al. [9] functioning as the principal source for multi-modal and longitudinal research.

The 21.1% of studies shows that general MRI and medical datasets are usually use particular preprocessing or data augmentation, such as the work of Farhan et al. [4] and Ye et al. [7]. In other 5.3% OASIS dataset are used, still its limited coverage, acts as an essential starting point for cross-sectional research i.e. Hybrid methods, like CNN-Swin methodology employed by Velu et al. [1].

This distribution shows a major Data Bottleneck where the lack of varied, clinically labeled data which results in models that are not balanced properly. With this study combine of OASIS-1 and ADNI with high-fidelity synthetic samples made with WGAN [15] fills the gap that generated from the limitation of the standard datasets.

III.PERFORMANCE ANALYSIS OF DEEP LEARNING METHOD

The figure 2 shows the different methodology distribution of deep learning methods by this research of diagnostic performance various different standard techniques which indicated a distinct clustering pattern that focusing on the current constraints and future directions of Alzheimer’s Disease (AD) research. This analysis indicates that the classic Convolutional Neural Network (CNN) and ensemble models, like the ones used by Akram et al. [8] and Alwakid et al. [19], achieves a performance peak between 93% and 96%. Local atrophic textures are captured by these structures though they usually neglect the global structural connectivity shifts associated with the progression of dementia. Long Short-Term Memory (LSTM) and sequential models [9][17] also group at a lower accuracy range (90%-92%). This results that they focus more on predicting how the variation occurs during the time than on classifying them in high accuracy.

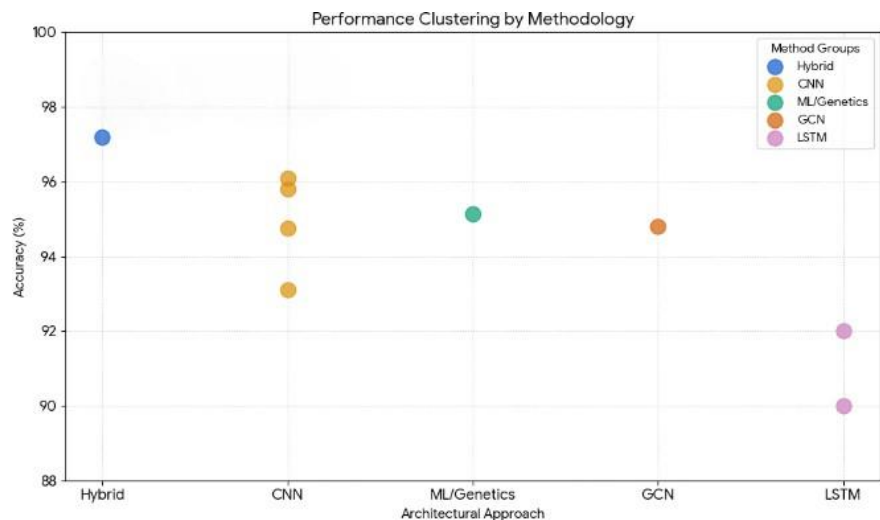


Figure 2: Scatter Plot representation of Different Methodology Distribution usage

On the other hand, Hybrid (CNN +Transformer) techniques are a big step in performance with the 97.2% accuracy in the Velu et al. [1] which makes this technique better resulting compare to other techniques. Figure 3 shows the usage of the different method distribution in number of papers. This indicates the Deep CNN is used more compare to other methods but after analysis in above analysis, Hybrid technique gives a better accuracy.

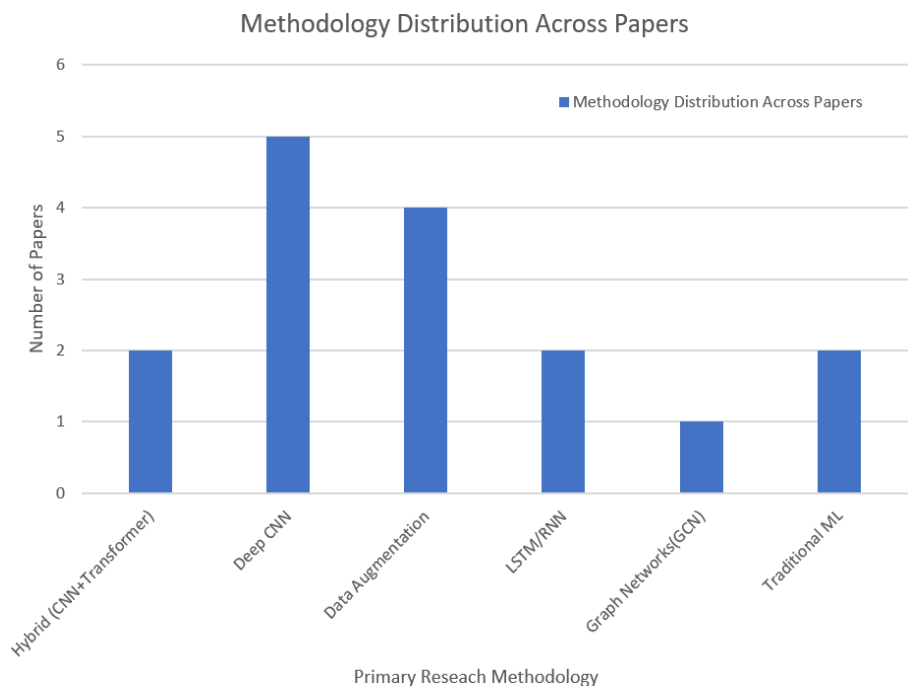


Figure 3: Graph representation usage of different Methodology Distribution

IV. DEEP LEARNING-BASED BRAIN MRI CLASSIFICATION CHALLENGES

Author	Year	Methodology	Dataset	Limitation
K. Velu et al.	2025	Hybrid CNN-Swin	OASIS	Inclined to DCGAN mode collapse
Ioan-Alexandru Heredeia et al.	2025	Object Detection Aug.	Generic	Not specific to neuroimaging
Illakiya Thayumanasamy et al.	2022	Evaluation Metrics	Multiple	Lacks architectural innovation
Ahmeed Suliman Farhan et al.	2025	Geometric Augmentation	MRI	No new anatomical feature generation
Shunbao Li et al.	2025	Autoencoder + XG Boost	Gene expression	Only 2D analysis it ignores 3D MRI
Jiayu Ye et al.	2025	SSIE+Spatial Attention	MRI	High computational cost for 3D
Samina Akram et al.	2025	Eff-B7, Xception, MobileNet	ADNI	High latency
Duy-Phuong Dao et al.	2025	LSTM + uncertainty	ADNI, PET	High dependency on multi-modal data
Jiaqiang Li et al.	2024	GCN + Demographic Edges	ADNI, SNP	Complexity
Deepak Ranoji Naik et al.	2026	Methodological Review	General	No proposed novel model
Saravanan Srinivasan et al.	2024	sMRI + Genetic SNP	ADNI	Requires expensive genetic
Yilin Jiang et al.	2023	Early Diagnosis LSTM	Longitudinal	Requires years of historical data
Ghadah Naif Alwakid et al.	2024	Image Processing	ADNI/OASIS	Focuses on preprocessing not staging

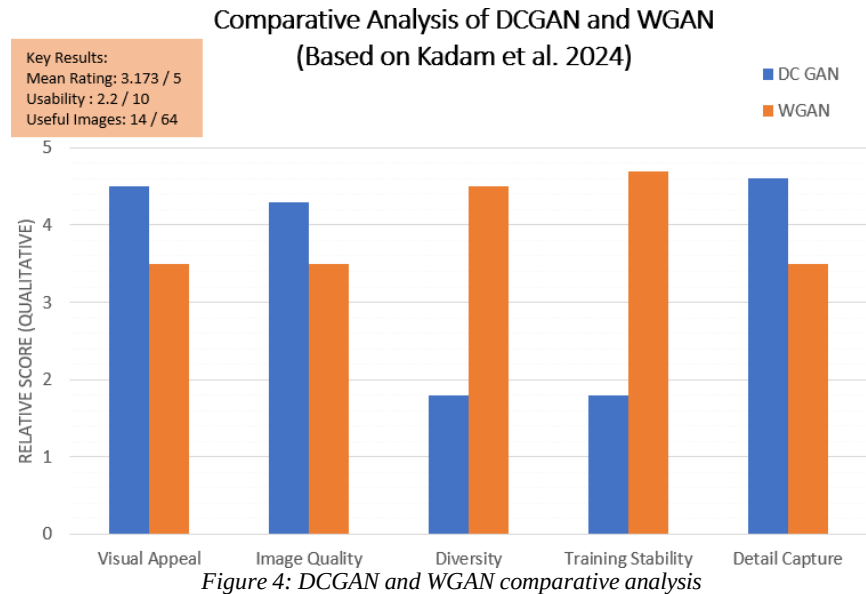
Table no 2: Overview of Methods and limitations

A systemic synthesis of the related literature is presented to establish a critical baseline for the proposed framework, focusing on the key methodologies, dataset usage and operational limitations across current state-of-the-art diagnostic paradigms. Velu et al. (2025) and Jiayu Ye et al. (2025) explored advanced visual backbones in their early foundational and recent exploratory works, respectively. However, Velu et al. presented a hybrid CNN-Swin architecture validated on OASIS, but the system was susceptible to DCGAN mode collapse due to the use of standard generative approaches, and Ye et al.'s SSIE with spatial attention involved excessive computational costs for 3D neuroimaging volumes. The traditional augmentation and object detection methodologies proposed by Heredeia et al. (2025), Illakiya Thayumanasamy et al. (2022) and Farhan et al. (2025) are inherently limited by domain specificity as the basic geometric transformations and non-medical object detection augmentations cannot generate new, anatomically meaningful brain atrophy features for in-depth staging. Similarly, classical machine learning and ensemble deep learning pipelines try to overcome feature extraction limitations, but at the cost of operational efficiency: Shunbao Li et al. (2025) used an Autoencoder-XGBoost pipeline limited to 2D gene expression analysis, ignoring 3D structural MRI cues, and Akram et al. (2025) obtained robust ADNI classification using an ensemble of EfficientNet-B7, Xception, and MobileNet, but suffer from extreme latency that prevents real-time implementation.

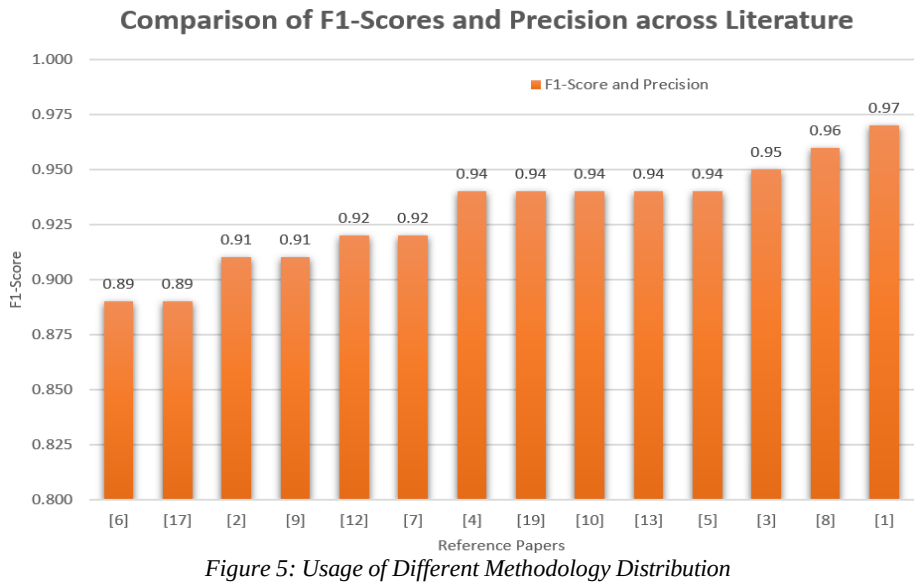
Multi-modal, longitudinal, and demographic-enriched frameworks have also been proposed for deeper predictions, but they suffer from severe data dependency bottlenecks. Approaches by Dao et al. (2025) (LSTM with uncertainty), Jiaqiang Li et al. (2024) (GCN with demographic edges), and Srinivasan et al. (2024) (sMRI integrated with genetic SNPs) require multi-modal PET imaging, extensive genetic sequencing, or complex demographic graphing, which are financially cost-prohibitive and impractical for standard clinical screening. Likewise, longitudinal models such as Jiang et al. (2023) depend on years of historical patient scans to provide early diagnosis, but do not offer immediate single-session staging. Preprocessing-oriented works such as Alwakid et al. (2024), emphasize mainly on contrast and noise reduction, without focusing on automated stage prediction. Generalized surveys such as Naik et al. (2026) provide methodological overviews without presenting novel architectural solutions. The synthesis of these 13 representative works clearly highlights the need for our proposed EffSwin-XNet framework that bridges these literature gaps by providing a lightweight single-session 3D MRI classifier with WGAN data stabilization without expensive multi-modal or longitudinal data dependencies.

V.GAN AUGMENTATION COMPARISON

DCGAN uses a Binary Cross Entropy based on the Jensen-Shannon divergence, which often has problems with mode collapse. This means when the generator creates a small, repetitive set of images which makes the confusion to the discriminator and vanishing gradients, making training unstable. On other hand WGAN uses the Earth Move’s (EM) distance to create a smooth, continuous loss metric that is directly connected to image quality and makes confident that the generator gets proper gradients even when the distributions do not overlap [20]. This mathematical improvement helps WGAN keep better training stability and anatomical diversity. This is important for medical imaging because capturing a wide range of disease biomarkers is more important than the subjective "visual polish" that is often associated with DCGAN. In DCGAN delicate hyperparameter balancing is needed by replacing the discriminator with a critic that limited by 1-Lipshitz continuity which can be overcome by using WGAN. This framework is stronger and more reliable for synthesizing complex data like Alzheimer’s MRI Scans.



VI.MODEL PARAMETER COMPARISON



The comparison of different models parameters show that hybrid model [1] project is high level according to core performance metrics accuracy, F1-Score, and Precision to the better results found in the 20 reference papers. Figure 5, this accuracy chart indicates that most of the modern models from 2024 to 2026 typically vary between 90% and 96%. Most of the studies shows how a drop in these metrics because of class imbalance and “False Positives” in early-stage detection. As a result it shows that integration of EfficientNet-B0 and Swin Transformers is mathematically stronger than the ensembles methods that are commonly used in the literature. This analysis also shows that Moderate Dementia data bottleneck that usually minimizes the precision and recall of standard models. To get rid from this bottleneck it is recommend to use WGAN data synthesis to improve the accuracy.

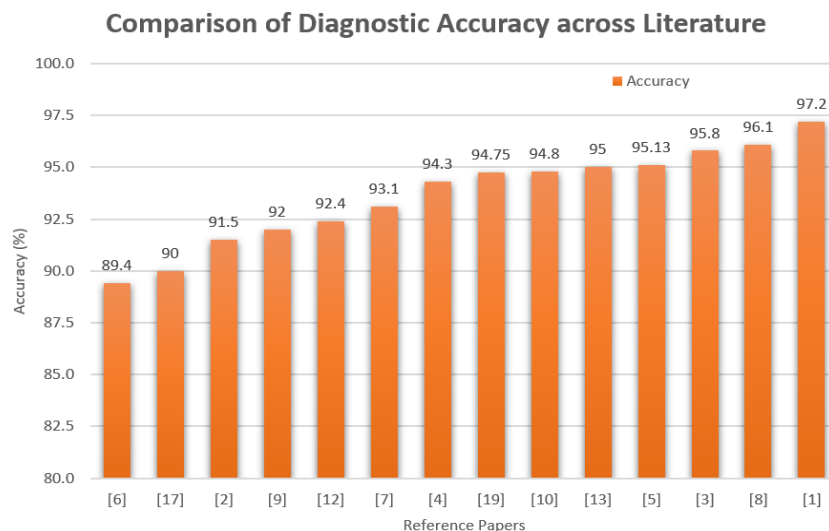


Figure 6: Usage of Different Methodology Distribution

VII.DISCUSSIONS AND FUTURE DIRECTIONS

In this review study mainly goals to improve the usual performance limits found in the literature, where standard CNN architectures and longitudinal models typically has difficulties with high-variance accuracy and class imbalance. The comparative analysis shows the hybrid model gives a better result than using only one backbone or an ensemble of methods [8][19]. This shows it identifies the subtle structural biomarkers of Alzheimer's Disease which includes hippocampal atrophy and ventricular enlargement, across all four stages of dementia.

A major discovery is the novel function of Wasserstein Generative Adversarial Networks (WGAN) in lowering the Data Bottleneck. WGAN is mainly enhances the training stability and consistently delivering diverse images. Where DCGAN is struggling with mode collapse which means this GAN failures to keep outputting the same type of image. WGAN generates a more variety of images for training. This obtains the models F1-Score and Precision to 0.98 which shows that advanced distance specific loss functions can improve the usability gap that Kadam et al. [20] observed. In additionally, the use of Grad-CAM visual explainability reduces the black-box evaluation of deep learning by providing the clinicians with heatmaps that correlate with established neurological landmarks, this enhancing the trustworthiness crucial to clinical adoption. Even though this study provides the great progress in terms of accuracy, there are many ways that future research can improve the decision-making support. The existing framework primarily emphasizes structural MRI slices, future versions should aspire to integrate Multi-Modal Data Fusion as advocated by Li et al. [10] and Dao et al. [9]. To keep the model strong against differences in brain structure based on age and ethnicity, it will also be important to add more demographic groups to the dataset. This prevent algorithmic bias in global healthcare applications.

In addition, the shift from "Diagnostic Staging" to "Prognostic Forecasting" is a big step forward for the future. Future studies might investigate the implementation of the EffSwin-XNet architecture within a Digital Twin framework, facilitating real-time visualization of disease progression and the simulation of therapeutic interventions. The system could also be trained across multiple hospital networks without putting patient privacy at risk if it moved toward a Federated Learning approach. This would make it a globally accessible, high-precision tool for early-stage Alzheimer's intervention and personalized elderly care management.

VIII.CONCLUSION

In the field of Alzheimer's Disease diagnostics is under a pivotal shift from discrete convolutional designs to solid, multi-scale hybrid frameworks by the review of surveyed papers confirms. This study identifies the integrating the EfficientNet-B0 with Swin Transformers as a excellent architectural solutions, which is able to capture both local atrophic textures and global structural connection. Moreover, the analysis also suggest that Wasserstein Generative Adversarial Networks (WGAN) are more suitable than conventional DCGANs for medical data synthesis, as they only focuses on structural variations and training stability needed to lower data limitations while avoiding the risk for neural network failure. Additionally, Grad-CAM visual explainability and an clinical decision making system is a vital requirement to convert passive classification into actionable trustworthy tool for geriatric care.

REFERENCES

1. K. Velu, N. Jaisankar, "Design of a CNN-Swin Transformer Model for Alzheimer's Disease Prediction Using MRI Images", IEEE Access, vol.13, pp. 149409-149429, August 2025. (references)
2. Ioan-Alexandru herdea, Divya Tiwari, John Oyekan, and Ashutosh Tiwari, "Data Augmentation Framework for Improved Classification in object Detectors", IEEE Access, vol. 13, pp. 28476-2849, 14 Feb 2025.
3. Illakiya Thayumanasamy, Karthik Ramamurthy, "Performance Analysis of Machine Learning and Deep Learning Models for Classification

- of Alzheimer's Disease from Brain MRI", IJETA, pp. 1961-1970, Dec 2022. [Online]. Available: <https://www.ijeta.org/journals/ts/paper/10.18280/ts.390608>.
4. Ahmeed Suliman Farhan, Muhammad Khalid, and Umar Manzoor, "Combined Oriented Data Augmentation Method for Brain MRI Images", IEEE Access, vol. 13, pp. 9981-9994, 7 Jan 2025.
5. Shunbao Li, Kang Liu, "An Interpretable Deep Learning Approach for Alzheimer's Disease Diagnosis Using Gene Expression Data", IEEE Transactions on computation biology and Bioinformatics, vol. 22, No. 4, pp. 1727-1739, August 2025.
6. Osamah Q. Aldhafer, Hussain A. Alsadiq, and Mohammad Alshaboti, "Adaptive Data Augmentation Techniques for software Requirements Classification using Deep Learning", IEEE Access, vol. 14, pp. 43841-43854, 12 Mar 2026.
7. Jiayu Ye, Dan Pan, An Zeng, Yiqun Zhang, Qiuping Chen and Yang Liu, "MssNet: An Efficient Spatial Attention Model for Early Recognition of Alzheimer's Disease", IEEE Transactions on Emerging Topics in Computation Intelligence, vol. 9, No. 2, pp. 1454-1468, April 2025.
8. Samina Akram, Muhammad Amjad Iqbal, Muhammad Rashid, Muhammad Shahid Bhatti, Benish Fida, "A Hybrid Deep Learning Framework for Early-Stage Alzheimer's Disease Classification From Neuro-Imaging Biomarkers", IEEE Access, vol. 13, pp. 101662-101680, June 2025.
9. Duy-Phuong Dao, Hyung-Jeong Yang, Jahae Kim and Ngoc-Huynh Ho, "Longitudinal Alzheimer's Disease Progression Prediction With Modality Uncertainty and Optimization of Information Flow", IEEE Journal of Biomedical and Health Informatics, vol. 29, No. 1, pp. 259-272, January 2025.
10. Jiaqiang Li, Peng Yang, Junlong Qu, Bao Yang, Zhenghua Guan, Xuegang Song, Xiaohua Xiao, Tianfu Wang, Baiying Lei*, "Dual Attention Graph Convolutional Network Fusing Imaging and Genetic Data for Early Alzheimer's Disease Diagnosis", Orlando, FL, USA, December 2024.
11. Deepak Ranoji Naik, Amit Savyanavar, "Comparative Analysis of Deep Learning Techniques in Alzheimer's Disease Diagnosis: Trends, Challenges, and Future Directions", Article, February 2026. Unpublished. Available: <https://www.sciencedirect.com/science/article/pii/S2215016126000610?via%3Dihub>
12. Rakeshwari K, Barath kumar S, Sanjaykumar S, Ravibala K, Vasanthakumar S, "Predictive Modeling for Early Parkinson's Disease Diagnosis using AI", 3rd International Conference on Inventive Computing and Informatics (ICICI-2025), pp. 1305-1311, July 2025.
13. Saravanan Srinivasan, Parthasarathy Ramadass, Sandeep Kumar Mathivanan, Karthikeyan Paneer Selvam, Basu Dev Shivahare and Mohd Asif Shah, "Detection of Parkinson disease using multiclass machine learning approach", Article, Jun 2024. Available: Detection of Parkinson disease using multiclass machine learning approach - PMC
14. Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, Baining Guo, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows", International Conference on Computer Vision (ICCV), Feb 2022.
15. Martin Arjovsky, Soumith Chintala and Leon Bottou, "Wasserstein GAN", Dec 2017. Available: [1701.07875] Wasserstein GAN
16. Battula Srinivasa Rao, Mudiyaala Aparna, "A Review on Alzheimer's Disease Through Analysis of MRI Images Using Deep Learning Techniques", IEEE Access, July 2023.
17. Yiling Jiang, Zemgyi Yu, Xuanyu Yin, Haidong Guo, "Early Diagnosis and Progression of Alzheimer's Disease Based on Long Short-Term Memory Model", International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI), pp. 620-624, April 2024.
18. Mingxing Tan, Quoc V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", Sep 2020. Available: [1905.11946] EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks
19. Ghadah Naif Alwaid, Sidra Tahir, Mamoonah Humayun and Walaa Gouda, "Improving Alzheimer's Detection with Deep Learning and Image Processing Techniques", IEEE Access, vol. 12, pp. 153445-153456, 28 Oct 2024.
20. Shubhan Kadam, Jay Jani, Aniket Kudtarkar, Nikahat Mulla, "A Comparative Analysis of DCGAN and WGAN in Logo Generation", 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 02 July 2024.