

A Comparative Evaluation of Machine Learning, Deep Learning, and Vision Transformer Models for Alzheimer's Disease Classification

Noushath Shaffi¹  , Viswan Vimbi² , Abdelhamid Abdesselam¹ , and Rachid Hedjam³

¹ Department of Computer Science, College of Science, Sultan Qaboos University, Muscat, Sultanate of Oman

n.shaffi@squ.edu.om

² Department of Information Technology, University of Technology and Applied Sciences-Sohar, Sultanate of Oman

vimbi.viswanutas.edu.om

³ Department of Computer Science, Bishop's University, Sherbrooke, QC, Canada

Abstract. Alzheimer's Disease (AD) is an area profusely investigated by Artificial Intelligence (AI) research. The complexity of AD feature characterization necessitates the application of AI models for an accurate diagnosis. Several AI techniques have been used in the detection and diagnosis of AD. Despite AI advancements for AD detection, the complexity of brain imaging, especially in MRI scans, poses unique challenges. This complexity has led to exploring newer architectures like vision transformers (VT). This has brought a paradigm shift in computer vision architecture, providing a unique and innovative alternative to traditional convolutional neural networks (CNN). This paper presents a comparative evaluation of AD classification using traditional and deep machine learning classifiers including VTs. The results indicate an increase in accuracy of around 5% for VT architectures compared to traditional AI models.

Keywords: Machine Learning · Deep Learning · Vision Transformers · Alzheimer's Disease · CNN

1 Introduction

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder that primarily affects memory, cognition, reasoning, and functional independence, eventually limiting a person's ability to perform routine daily activities [45,21,4]. According to recent estimates, more than 50 million people are affected by dementia worldwide, with Alzheimer's accounting for approximately 60–70% of these cases [10]. The burden extends beyond affected individuals to families, caregivers, healthcare systems, and society, producing substantial emotional, social, and economic consequences [27]. As disease progression is generally irreversible,

early identification of pathological changes is particularly important for facilitating clinical assessment, planning supportive interventions, and potentially improving long-term disease management [27,21].

Neuroimaging plays an important role in investigating structural and functional changes associated with AD. In particular, Magnetic Resonance Imaging (MRI) provides non-invasive information about brain morphology and regional atrophy without exposing patients to ionising radiation. However, the structural alterations associated with early AD can be subtle, spatially distributed, and heterogeneous across individuals. Visual assessment may therefore be affected by disease stage, image quality, acquisition protocol, and inter-observer variability. Automated computational analysis offers an opportunity to identify complex imaging patterns that may be difficult to characterise consistently through conventional visual interpretation alone [35,36].

Artificial Intelligence (AI) consequently offers powerful computational tools for supporting AD detection, classification, and disease-stage characterisation. AI techniques can assist in identifying early signs of cognitive deterioration and neurodegeneration, sometimes before severe clinical manifestations become evident [46,50,9,3]. Machine Learning (ML) methods such as Support Vector Machines, k-nearest Neighbours, Logistic Regression, Random Forest, and gradient boosting have demonstrated that carefully constructed imaging and clinical features can distinguish different stages of dementia [45,55,34,8]. Deep Learning (DL) subsequently reduced dependence on manually engineered features by allowing neural networks to learn hierarchical representations directly from medical images. CNN architectures have therefore become widely used for automated MRI classification and related neuroimaging tasks [40,43,60].

Despite their success, CNNs primarily construct representations through local convolutional operations. Capturing dependencies between anatomically distant brain regions may consequently require increasingly deep architectures. Vision Transformers (VTs) provide an alternative representation paradigm by using self-attention to model relationships among image patches across the entire spatial domain. This ability to capture both local information and long-range dependencies has motivated their increasing use in medical imaging and AD classification [49,48,47,64,24]. Transformer-based approaches are especially interesting for neuroimaging because disease-related structural changes may involve distributed anatomical regions rather than isolated local abnormalities.

While numerous studies compare conventional ML and CNN-based DL methods for AD diagnosis [61,55,60,58], comparatively fewer studies evaluate ML, CNNs, and multiple VT architectures within the same experimental framework. Such comparisons are important because high performance reported for a single architecture does not establish whether its additional computational complexity provides meaningful improvement over simpler alternatives. Benchmarking these model families using common datasets and evaluation measures can therefore provide clearer evidence regarding their relative capabilities, generalisation behaviour, and suitability for data-constrained medical imaging.

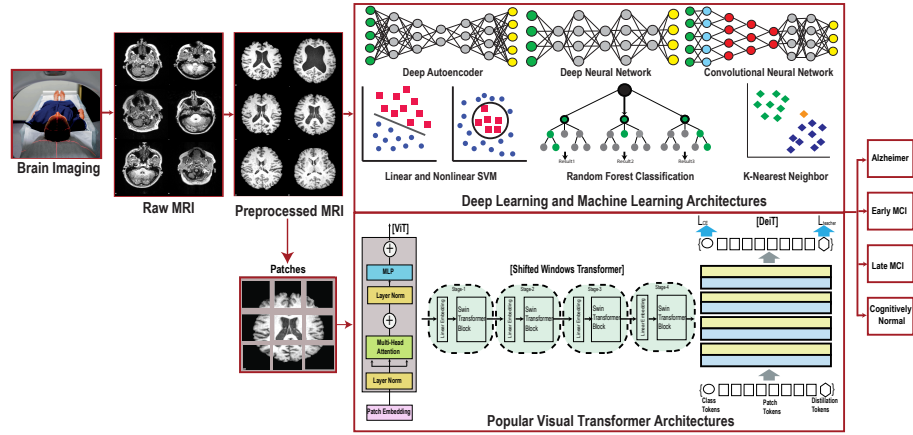


Fig. 1. Standard AI pipeline for AD classification.

Clinical translation also requires considerations beyond predictive accuracy. Explainability, uncertainty, privacy, personalisation, fairness, and meaningful human oversight are increasingly recognised as requirements for trustworthy health-care AI [62,57,31,33,32,29]. Consequently, identifying the highest-performing architecture represents only one component of developing clinically useful AD decision-support systems.

Against this background, this paper comparatively evaluates traditional ML, CNN-based DL, and VT architectures for multiclass AD classification using the OASIS and ADNI datasets. Our contributions are threefold: (1) a systematic comparison of conventional ML, CNN, and transformer architectures under common AD classification settings; (2) an analysis of their relative performance across datasets with different sample characteristics; and (3) identification of methodological and clinical considerations that should guide future development of explainable, multimodal, and trustworthy AD classification systems. Fig. 1 illustrates the standard AI pipeline considered in this study, including preprocessing, representation learning, classification, and evaluation.

The remainder of the manuscript is structured as follows: Section 2 reviews related research, Section 3 presents the experimental comparison, Section 4 discusses the principal findings and their implications, Section 5 outlines future research directions, and Section 6 concludes the paper.

2 Related Study

Research on computational AD diagnosis has evolved from conventional feature-based ML towards increasingly complex DL, multimodal, and transformer-based architectures. This progression reflects both improvements in computational capability and the growing availability of neuroimaging datasets such as ADNI

and OASIS. Early AI approaches generally relied on predefined imaging, demographic, cognitive, or biomarker features followed by conventional classifiers. More recent DL methods instead learn representations directly from MRI, PET, and other clinical data, while transformer-based architectures seek to capture global spatial relationships that may be difficult to represent using convolution alone [36,35,21].

Recent advances in ML and DL have therefore made these techniques important tools for AD detection and staging. ML enables predictive models to identify relationships in structured biomedical information using supervised, unsupervised, and semi-supervised learning. DL extends this capability by learning hierarchical representations from high-dimensional and unstructured data, particularly medical images [30]. This distinction is important in AD research because conventional ML can remain highly competitive when sample sizes are limited and informative features are available, whereas DL is particularly attractive when complex spatial information must be learned directly from neuroimaging data [34,44].

A major development has been the transition from purely local convolutional representations towards attention-based architectures. Vision Transformers divide an image into patches and model relationships between them through self-attention. Unlike conventional CNNs, where the receptive field expands progressively through multiple layers, transformer architectures can explicitly model interactions between spatially distant image regions. In brain MRI, this property may be advantageous because AD-related structural alterations are not restricted to a single anatomical region but may involve distributed changes across hippocampal, temporal, cortical, and other structures. The growing performance of ViT, Swin, DeiT, and hybrid CNN–Transformer architectures has therefore stimulated considerable interest in their use for neurodegenerative disease classification [13,48,47,64,24].

Another important evolution concerns the information used for diagnosis. Neurodegenerative disease is intrinsically multimodal: MRI describes structural characteristics, PET captures metabolic or molecular information, while clinical assessments, demographics, genetics, physiological measurements, and cognitive scores provide complementary evidence. Recent research has therefore increasingly investigated multimodal fusion rather than relying on one imaging source alone [19,16,63,39]. Such approaches may improve robustness because different modalities capture different manifestations of disease progression. However, they also introduce challenges involving missing modalities, feature alignment, sample-size requirements, interpretability, and privacy.

Explainability has consequently emerged as a parallel research direction. Although improved classification accuracy is valuable, clinical users must be able to assess whether predictions are based on medically plausible evidence. Feature attribution, SHAP, LIME, saliency analysis, and attention-based explanations have increasingly been explored for dementia and AD models [62,57,18,14]. This requirement is especially relevant as architectures become more complex. Transformer attention does not itself guarantee clinical explainability, and high-

performing models still require systematic validation of their reasoning and behaviour.

The literature therefore reveals three interconnected trends: movement from manually engineered ML towards representation-learning architectures, increasing adoption of transformers and hybrid architectures, and growing attention to multimodal and explainable healthcare AI. The present study is positioned primarily within the first two trends by providing a controlled comparison of conventional ML, CNN, and VT approaches, while the latter trends inform the discussion of clinical translation and future work.

Machine Learning In healthcare, ML is gaining attention for disease diagnosis, especially in conditions like Alzheimer’s Disease. Several conventional machine learning methods, such as XGBoost, Support Vector Machines (SVM), k-nearest Neighbors (KNN), and Logistic Regression (LR), have demonstrated promising performance in AD classification [55,34,51,44]. XGBoost, a powerful gradient-boosting algorithm, has been widely used due to its efficiency and accuracy. Fuliang et al. [67] used XGBoost-SHAP to address class imbalance in diagnosing AD across different disease stages. The framework outperformed random forest and AdaBoost models and provided interpretable information about influential AD risk features, consistent with the wider emphasis on explainable dementia modelling [3,8,14].

Support Vector Machines have also been extensively applied in AD diagnosis. Khedher et al. [23] developed a fully automated computer-aided diagnosis system that combines MRI feature extraction with SVM classification. Acharya et al. [2] developed a computer-aided diagnostic system for detecting AD from brain MRI using the Shearlet Transform with KNN. Logistic Regression has also shown effectiveness in classification tasks at different AD stages, as demonstrated by Sheng et al. [52]. Similar comparisons across open dementia datasets and multimodal MRI settings have been reported in Mahmud co-authored studies [34,51].

Deep Learning EfficientNet-B0 optimizes accuracy and efficiency in brain MRI classification while reducing computational costs. InceptionResNet v2 enhances training speed for multi-class AD classification and Inception v3 captures features at various scales, yielding competitive accuracy. The ResNet family (50, 101, 152) employs deep residual learning to model complex features effectively. These architectures have been extensively investigated for multiclass AD classification [17,11,19,18,40,42,43,12,60].

Kundaram et al. [25] developed a CNN-based classifier using ADNI, and Balsaia et al. [7] employed a 3D CNN on T1-weighted images. Transfer Learning has emerged as a valuable approach in AD classification. Jain et al. [20] leveraged VGG-16, while Wu et al. [66] compared AlexNet and GoogleNet. Related Mahmud co-authored work has evaluated transfer-learning ensembles and CNN-based MRI classifiers for AD and related neurodegenerative disorders [58,60,64].

One can also refer to Shaffi et al. [45,44] for a comprehensive examination of ML and DL algorithms for AD classification.

Vision Transformers Vision Transformers (ViTs) have emerged as powerful architectures for computer vision applications, particularly image classification and segmentation [22]. The original ViT demonstrated that transformer-based models could surpass CNNs in image classification on large datasets. The Swin Transformer [26] introduces a hierarchical design and shifting windows, and DeiT [56] reduces data requirements while maintaining competitive performance. In AD research, Mahmud and collaborators have investigated a dedicated ViT pipeline, ensembles of transformer architectures, CNN–ViT hybrid models, and hybrid ViT–graph models for multiclass MRI classification [48,47,64,24].

Maram et al. [5] applied ViT to OASIS MRI scans. Sherwani et al. [53] compared CNN against Deep ViT and CaIT. Odusami et al. [37] combined MRI and PET data from ADNI, while Lyu et al. [28] investigated cross-domain transfer learning. Saman et al. [41] incorporated re-attention and CaIT mechanisms, Huang et al. [15] introduced the Resize Swin Transformer, and Shin et al. [54] demonstrated ViT for AD classification using PET imaging. More recent Mahmud co-authored studies extend this direction to multimodal AD detection, brain structure–function modelling, and generative representations [16,63,38,68].

3 Experimental Results

3.1 Experimental Setup

MRI Data The experiments were conducted using two widely used neuroimaging datasets: the Open Access Series of Imaging Studies (OASIS, www.oasis-brains.org) and the Alzheimer’s Disease Neuroimaging Initiative (ADNI, <https://adni.loni.usc.edu/>).

OASIS Dataset The OASIS dataset used in this study contains 6,400 images distributed across four categories according to the Clinical Dementia Rating scale. These include non-demented (3,200), very mild demented (2,240), mild demented (896), and moderately demented (64). A 70:30 train-test partition was adopted, yielding 5,120 images for training and 1,280 for testing. The preprocessed images were obtained from the Hugging Face platform [1].

ADNI Dataset T1-weighted MRI images were downloaded for subjects aged 50 to 65 from both genders. A total of 1,056 raw axial-plane MRI images were grouped into four classes: AD (223), EMCI (475), LMCI (262), and CN (96). The dataset was split 90:10 into 950 training and 106 test samples. The images were sourced from ADNI-1, ADNI-2, and ADNI-GO.

MRI Preprocessing The preprocessing pipeline comprised reorientation, registration, removal of non-brain tissues, and histogram equalization. These procedures were carried out using the FMRIB Software Library, with BET used for skull stripping. A detailed description is provided by Viswan et al. [65,59]; related work has also highlighted the importance of MRI preprocessing and representation choices for AI-based AD classification [36,35].

Implementation Previously trained VT weights were obtained through Hugging Face. Conventional machine learning algorithms were developed with Scikit-learn, while pretrained CNN architectures were accessed through TensorFlow. All models were fine-tuned and evaluated using a Windows system equipped with an NVIDIA RTX 3060 graphics card and a CPU operating at 3.2 GHz. The implementation is available at <https://github.com/snoushath/ViT-CNN-AD.git>.

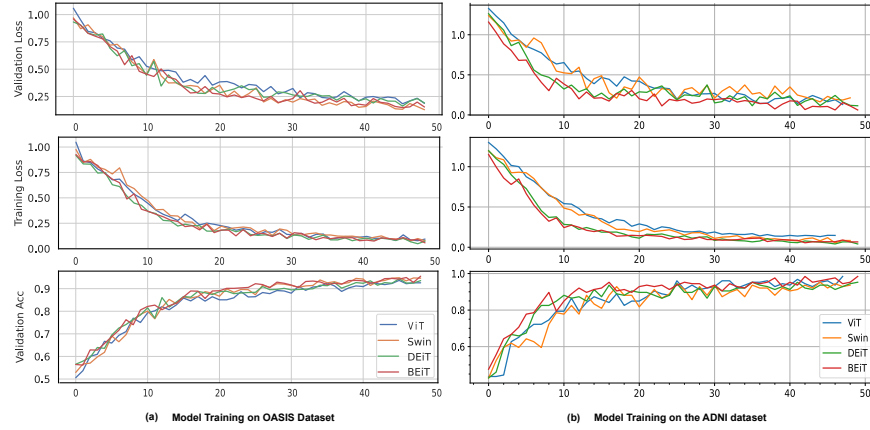


Fig. 2. Model training of vanilla VT models: (a) OASIS and (b) ADNI datasets.

3.2 VT Training and Testing

This experiment investigates the training behavior of four VT architectures on OASIS and ADNI. While OASIS provides a relatively larger number of training samples, ADNI represents a more data-constrained setting. Figure 2 presents training loss, validation loss, and validation accuracy over 50 epochs. Similar performance trends were observed across both datasets, supporting previous observations that transformer-based AD classifiers can be effective under varying data regimes [48,47,64,24].

The findings reported in Table 1 indicate comparable test results among the evaluated VT models. On OASIS, ViT, Swin Transformer, DEiT, and BEiT achieve high accuracy, specificity and sensitivity. A similar trend is observed

Table 1. Performance of the evaluated VT architectures across different classification metrics on OASIS and ADNI.

Models	OASIS Dataset					ADNI Dataset				
	Acc	Spe	Sen	FNR	FPR	Acc	Spe	Sen	FNR	FPR
ViT	97.2	98.9	92.0	8.0	1.1	99.0	99.6	99.2	8.0	3.6
Swin	98.0	99.1	94.1	5.8	0.9	99.5	99.6	98.9	1.0	0.4
DeiT	98.4	99.3	97.0	3.0	0.6	99.0	99.6	98.9	1.1	0.3
BEiT	97.7	99.0	95.3	4.0	0.9	99.0	99.7	98.9	1.0	0.3

on ADNI. These findings are consistent with Mahmud co-authored AD studies reporting strong performance from transformer and hybrid architectures [48,47,64,24].

3.3 Comparison with CNN Models

Although VTs have progressed considerably, CNNs continue to play an important role in medical computer vision. We compare VT- and CNN-based models using ADNI and OASIS. CNN architectures include EfficientNetB0, Inception-ResNetV2, InceptionV3, ResNet50, ResNet101 and ResNet152. For fine-tuning, the original classification layer was replaced with the layer sequence adopted in [45,44]. Previous Mahmud co-authored comparisons of pretrained deep networks, DenseNet ensembles and Siamese CNNs provide additional context for these architectures [60,40,43,12].

Using OASIS Dataset The four vanilla VT models achieved strong performance on OASIS. In comparison, EfficientNetB0 produced relatively high training and validation errors and achieved a test accuracy of 43.80%. InceptionResNetV2 and InceptionV3 also exhibited differences between training and validation behavior.

Table 2. Performance comparison of VT and CNN architectures on OASIS.

Architectures	Test Acc	Val Acc	Train Loss	Val Loss
EfficientNetB0	43.80	92.59	74.12	68.25
InceptionResNetV2	52.10	81.44	10.29	97.48
InceptionV3	73.12	93.29	57.48	68.60
ResNet50	84.92	96.59	30.30	44.74
ResNet101	82.96	96.40	30.74	45.10
ResNet152	85.54	97.11	29.19	40.24
DeiT	98.43	95.50	5.70	14.83

Using the ADNI Dataset The results obtained on the smaller ADNI dataset (Table 3) further demonstrate strong VT performance relative to CNNs. BEiT,

ViT, and Swin adapted effectively to the limited training data. This finding accords with recent Mahmud co-authored studies investigating transformer-based and hybrid AD classification under medical-imaging constraints [48,47,64,24].

Table 3. Performance comparison of VT and CNN architectures on ADNI.

Architectures	Test Acc	Val Acc	Train Loss	Val Loss
EfficientNetB0	88.03	99.00	39.32	26.06
InceptionResNetV2	59.80	83.66	1.0259	1.0107
InceptionV3	85.64	96.62	35.24	53.60
ResNet50	94.73	99.46	6.27	18.74
ResNet101	94.25	99.14	3.34	32.52
ResNet152	87.55	99.15	9.90	23.49
BEiT	100.0	98.41	5.0	6.0

3.4 Comparison with ML Models

DL algorithms have achieved considerable success across computer vision applications. Nevertheless, their use may not always be necessary when data are limited. Shaffi et al. [45,44] showed that diverse ML classifiers can achieve competitive performance in data-constrained medical applications. Accordingly, we consider XGBoost, SVM, KNN and LR and compare them with VT architectures. Related Mahmud co-authored studies have also compared conventional ML for dementia and multi-planar MRI-based AD classification [34,51,3].

Table 4. Performance comparison of ML algorithms on OASIS.

Method	Accuracy	Specificity	Sensitivity	FNR	FPR
XGBoost	0.8902	0.9633	0.8905	0.1094	0.0366
SVM	0.9132	0.9107	0.9135	0.0864	0.0893
KNN	0.9371	0.9790	0.9377	0.0622	0.0209
LR	0.907	0.969	0.9074	0.0925	0.0309
BEiT [6]	0.9843	0.9931	0.9696	0.0303	0.0068

The ML results are presented in Tables 4 and 5. KNN achieved the highest OASIS accuracy among the ML models, while SVM was particularly strong on ADNI. The VT architectures consistently achieved higher values across the evaluated metrics. This broader pattern is compatible with recent Mahmud co-authored comparisons of conventional, deep and transformer-based AD classifiers [44,60,48,47].

The performance evaluation of all AI model families is presented in Table 6. The results indicate that the evaluated VT models substantially outperform both CNNs and traditional ML algorithms in this experimental setting.

Table 5. Performance comparison of ML algorithms on ADNI.

Method	Accuracy	Specificity	Sensitivity	FNR	FPR
XGBoost	0.9809	0.9925	0.9803	0.0196	0.0074
SVM	0.9904	0.9967	0.9895	0.0104	0.0032
KNN	0.9809	0.9931	0.9803	0.0196	0.0068
LR	0.9809	0.9925	0.9803	0.0196	0.0074
Swin[26]	0.9952	0.9963	0.9887	0.0112	0.0036

Table 6. Overall performance comparison of ML, CNN, and VT models on OASIS.

Models	Methods	Accuracy	Specificity	Sensitivity	FNR	FPR
ML	XGBoost	89.02	96.33	89.05	10.94	3.66
	SVM	91.32	91.07	91.35	8.64	8.93
	KNN	93.71	97.90	93.77	6.22	2.09
	LR	90.7	96.90	90.74	9.25	3.09
CNN	EfficientNetB0	43.80	52.81	41.21	76.23	0.331
	InceptionResNetV2	52.10	76.55	27.14	72.85	23.44
	ResNet50	84.92	93.33	83.01	16.98	6.66
	ResNet101	82.96	92.60	80.13	19.86	7.39
	ResNet152	85.54	93.82	82.33	17.66	6.17
VT	ViT	97.18	98.82	91.95	8.08	1.18
	Swin	98.04	99.13	94.15	5.84	0.86
	DeiT	98.43	99.31	96.96	3.03	0.68
	BEiT	97.73	99.05	95.33	4.66	0.94

4 Discussion

The experimental results demonstrate a consistent performance advantage for the evaluated Vision Transformer architectures over both conventional ML and CNN models. On OASIS, the strongest conventional ML model reaches an accuracy of 93.71%, whereas DeiT achieves 98.43%. A similar pattern occurs on ADNI, where SVM reaches 99.04% and Swin reaches 99.52%. Although these differences are numerically modest on ADNI because all models perform strongly, the results indicate that transformer-based representations can remain effective even in a comparatively data-constrained setting. This observation is consistent with recent Mahmud co-authored studies investigating ViT-based, ensemble, CNN–Transformer, and graph-enhanced transformer architectures for AD classification [48,47,64,24].

The comparison with CNN architectures is particularly noteworthy. Deep CNNs such as ResNet perform substantially better than EfficientNetB0 and Inception-based models in several experiments, illustrating that classification performance cannot be attributed simply to the use of DL. Architecture, pre-training, optimisation, dataset composition, and fine-tuning strategy all affect generalisation. The strong VT results may partly reflect the ability of self-attention to represent relationships among spatially separated MRI regions. However, the present experiments do not directly establish that attention itself

causes the improvement. Controlled ablation studies comparing representation mechanisms, parameter counts, pretraining datasets, and optimisation procedures would be required to establish such a causal interpretation.

The very high accuracies observed on ADNI also warrant careful interpretation. Performance approaching 100% may indicate strong discriminative information, but medical imaging studies are particularly susceptible to dataset-specific signals, preprocessing effects, scanner characteristics, class composition, and potential correlations between samples from the same individual. External validation and strictly subject-independent evaluation are therefore essential before interpreting high benchmark accuracy as evidence of clinical generalisability. Future comparisons should report uncertainty intervals and include independent datasets or cross-dataset testing wherever possible.

From a clinical perspective, predictive performance is also insufficient on its own. An AD classification model should provide understandable evidence for its predictions and communicate uncertainty when imaging findings are ambiguous. Existing work on explainable AD classification and broader XAI highlights the importance of identifying clinically meaningful imaging regions and avoiding opaque decisions [62,57,18,14]. Likewise, emerging healthcare AI frameworks emphasise privacy, human oversight, personalisation, and inclusion [31,33,32,29]. The results presented here should therefore be viewed as a comparative assessment of model discrimination rather than evidence that any individual architecture is ready for autonomous clinical deployment.

Finally, multimodal modelling represents an important progression beyond the single-modality MRI experiments considered here. Combining structural imaging with PET, cognitive assessments, demographic information, biomarkers, or longitudinal observations could provide a more complete representation of AD progression [19,16,63]. Future transformer architectures may be particularly well suited to this setting because attention mechanisms can be extended to interactions between heterogeneous modalities. The principal challenge will be ensuring that increased architectural complexity produces clinically meaningful and externally reproducible improvements rather than better fitting of individual benchmark datasets.

5 Scope for Future Work

Building on the comparative findings and limitations discussed above, several directions warrant further investigation. First, the findings are based on OASIS and ADNI and may not fully represent diverse real-world clinical populations. Second, explainability is essential when AI models are deployed in medical practice, while this study focused primarily on predictive performance. A future comparative analysis should therefore examine model explanations using established XAI methods and AD-specific evidence [62,57,18,14]. Third, future work should evaluate multimodal learning that combines MRI with clinical, physiological, functional, or other complementary information [19,16,63,38,68,39]. Finally,

translation to clinical environments should consider privacy, uncertainty, human oversight, accessibility and inclusion alongside accuracy [31,33,32,29,4].

6 Conclusion

In this study, we compared ML, DL, and VT models for AD classification using MRI data from OASIS and ADNI. The results demonstrate the effectiveness of DL models and show that, in the evaluated experiments, vision transformer architectures consistently outperform the CNN and conventional ML alternatives. These findings are consistent with a growing body of Mahmud co-authored work on transformer, hybrid and multimodal AI for AD classification [48,47,64,24,16].

Despite the computational efficiency of traditional ML algorithms such as KNN, SVM, and XGBoost, DL and VT models achieved stronger classification metrics in this study. Future work should extend the analysis to multimodal data, explainability, externally validated cohorts, and trustworthy deployment. In this context, explainable and privacy-preserving personalised AI, preventive and inclusive healthcare AI, human-centred multi-agent healthcare systems, and inclusive multimodal digital-human modelling provide complementary directions for responsible clinical translation [31,33,32,29].

References

1. Hugging face datasets. Retrieved from <https://huggingface.co/datasets> (2022), accessed on 09-19-2024
2. Acharya, U.R., Fernandes, S.L., WeiKoh, J.E., Ciaccio, E.J., Fabell, M.K.M., Tanik, U.J., Rajinikanth, V., Yeong, C.H.: Automated detection of alzheimer's disease using brain mri images—a study with various feature extraction techniques. *Journal of medical systems* **43**, 1–14 (2019)
3. Ahmed, R., Fahad, N., Miah, M.S.U., Hossen, M.J., Morol, M.K., Mahmud, M., Mostafizur Rahman, M.: A novel integrated logistic regression model enhanced with recursive feature elimination and explainable artificial intelligence for dementia prediction. *Healthcare Analytics* **6** (2024). <https://doi.org/10.1016/j.health.2024.100362>
4. Al Mamun, S., Kaiser, M.S., Mahmud, M.: An artificial intelligence based approach towards inclusive healthcare provisioning in society 5.0: A perspective on brain disorder. *Lecture Notes in Computer Science* **12960 LNAI**, 157 – 169 (2021)
5. Almufareh, M.F., Tehsin, S., Humayun, M., Kausar, S.: Artificial cognition for detection of mental disability: a vision transformer approach for alzheimer's disease. In: *Healthcare*. vol. 11, p. 2763. MDPI (2023)
6. Bao, H., Dong, L., Piao, S., Wei, F.: BEit: BERT pre-training of image transformers. *arXiv preprint arXiv:2106.08254* (2021)
7. Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., Filippi, M., Initiative, A.D.N., et al.: Automated classification of alzheimer's disease and mild cognitive impairment using a single mri and deep neural networks. *NeuroImage: Clinical* **21**, 101645 (2019)
8. Chen, T., Su, P., Shen, Y., Chen, L., Mahmud, M., Zhao, Y., Antoniou, G.: A dominant set-informed interpretable fuzzy system for automated diagnosis of dementia. *Frontiers in Neuroscience* **16** (2022). <https://doi.org/10.3389/fnins.2022.867664>

9. Fabietti, M., Mahmud, M., Lotfi, A., Leparulo, A., Fontana, R., Vassanelli, S., Fasolato, C.: Early detection of alzheimer's disease from cortical and hippocampal local field potentials using an ensembled machine learning model. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **31**, 2839 – 2848 (2023)
10. Gauthier, S., Webster, C., Sarvaes, S., Morais, J., Rosa-Neto, P.: *World Alzheimer Report 2022: Life after diagnosis - navigating treatment, care and support* (2022)
11. Hajamohideen, F., Shaffi, N., Mahmud, M., Subramanian, K., Al Sariri, A., Vimbi, V., Abdesselam, A.: Four-way classification of alzheimer's disease using deep siamese convolutional neural network with triplet-loss function. *Brain Informatics* **10**(1), 1–13 (2023)
12. Hajamohideen, F., et al.: Four-way classification of alzheimer's disease using deep siamese convolutional neural network with triplet-loss function. *Brain Informatics* **10**(1) (2023). <https://doi.org/10.1186/s40708-023-00184-w>
13. Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., et al.: A survey on vision transformer. *IEEE transactions on pattern analysis and machine intelligence* **45**(1), 87–110 (2022)
14. Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., Hussain, A.: Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation* **16**(1), 45 – 74 (2024)
15. Huang, Y., Li, W.: Resizer swin transformer-based classification using smri for alzheimer's disease. *Appl. Sci.* **13**(16), 9310 (2023)
16. Islam, M.S., Rahman, M.M., Mahmud, M.: A multimodal deep learning system for alzheimer's disease detection with automated clinical feature selection. *Lecture Notes in Computer Science* **16347 LNAI**, 283 – 295 (2026)
17. Jahan, S., Abu Taher, K., Kaiser, M.S., Mahmud, M., Rahman, M.S., Hosen, A.S., Ra, I.H.: Explainable ai-based alzheimer's prediction and management using multimodal data. *Plos one* **18**(11), e0294253 (2023)
18. Jahan, S., Saif Adib, M.R., Mahmud, M., Kaiser, M.S.: Comparison between explainable ai algorithms for alzheimer's disease prediction using efficientnet models. *Lecture Notes in Computer Science* **13974 LNAI**, 357 – 368 (2023). https://doi.org/10.1007/978-3-031-43075-6_31
19. Jahan, S., Taher, K.A., Kaiser, M.S., Mahmud, M., Rahman, M.S., Hosen, A.S., Ra, I.H.: Explainable ai-based alzheimer's prediction and management using multimodal data. *PLoS ONE* **18**(11 November) (2023). <https://doi.org/10.1371/journal.pone.0294253>
20. Jain, R., Jain, N., Aggarwal, A., et al.: Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cogn. Syst. Res.* **57**, 147–159 (2019)
21. Javed, A.R., Saadia, A., Mughal, H., Gadekallu, T.R., Rizwan, M., Maddikunta, P.K.R., Mahmud, M., Liyanage, M., Hussain, A.: Artificial intelligence for cognitive health assessment: State-of-the-art, open challenges and future directions. *Cognitive Computation* **15**(6), 1767 – 1812 (2023). <https://doi.org/10.1007/s12559-023-10153-4>
22. Khan, S., Naseer, M., Hayat, M., Zamir, S.W., Khan, F.S., Shah, M.: Transformers in vision: A survey. *ACM computing surveys (CSUR)* **54**(10s), 1–41 (2022)
23. Khedher, L., et al.: Independent component analysis-support vector machine-based computer-aided diagnosis system for Alzheimer's with visual support. *Int. j. neural syst.* **27**(03), 1650050 (2017)

24. Kumari, A., Basha, N.K., Shaffi, N., Viswan, V., Mahmud, M.: Alzvitgnet: A hybrid vision transformer and graph convolutional network framework for multiclass alzheimer's disease classification from mri images. *Lecture Notes in Computer Science* **16347 LNAI**, 259 – 270 (2026). https://doi.org/10.1007/978-981-95-9575-4_20
25. Kundaram, S.S., Pathak, K.C.: Deep learning-based alzheimer disease detection. In: *Proc. MCCS 2019*. pp. 587–597. Springer (2021)
26. Liu, Z., Lin, Y., Cao, Y., et al.: Swin transformer: Hierarchical vision transformer using shifted windows. In: *Proc. ICCV*. pp. 10012–10022 (2021)
27. Long, S., Benoist, C., Weidner, W.: World Alzheimer Report 2023: reducing dementia risk: never too early, never too late. *Alzheimer's Dis. Int.* (2023)
28. Lyu, Y., Yu, X., Zhu, D., et al.: Classification of alzheimer's disease via vision transformer: Classification of alzheimer's disease via vision transformer. In: *Proc. ICPETRA*. pp. 463–468 (2022)
29. Mahmud, M., Baslyman, M., et al.: Inclusive ai for digital human modelling: A perspective on explainable multimodal digital twins for accessibility and inclusion. In: *International Conference on Human-Computer Interaction*. pp. 1–16. Springer (2026)
30. Mahmud, M., Kaiser, M.S., McGinnity, T.M., Hussain, A.: Deep learning in mining biological data. *Cogn. comput.* **13**(1), 1–33 (2021)
31. Mahmud, M., Kaiser, M.S., et al.: Towards explainable and privacy-preserving artificial intelligence for personalisation in autism spectrum disorder. *Lecture Notes in Computer Science* **13309 LNCS**, 356 – 370 (2022)
32. Mahmud, M., Shaffi, N., Kaiser, M.S., et al.: Privacy-preserving multi-agent systems for human-centred inclusive healthcare ai. *AI Magazine* **47**(3), e70087 (2026). <https://doi.org/10.1002/aaai.70087>
33. Mahmud, M., et al.: Preventive healthcare through privacy-preserving, explainable and inclusive artificial intelligence. *Lecture Notes in Computer Science* **16340 LNCS**, 359 – 374 (2026)
34. Miah, Y., Prima, C.N.E., Seema, S.J., Mahmud, M., Shamim Kaiser, M.: Performance comparison of machine learning techniques in identifying dementia from open access clinical datasets. *Advances in Intelligent Systems and Computing* **1188**, 79 – 89 (2021)
35. Noor, M.B.T., Zenia, N.Z., Kaiser, M.S., Mahmud, M., Al Mamun, S.: Detecting neurodegenerative disease from mri: A brief review on a deep learning perspective. *Lecture Notes in Computer Science* **11976 LNAI**, 115 – 125 (2019). https://doi.org/10.1007/978-3-030-37078-7_12
36. Noor, M.B.T., Zenia, N.Z., Kaiser, M.S., Mamun, S.A., Mahmud, M.: Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of alzheimer's disease, parkinson's disease and schizophrenia. *Brain Informatics* **7**(1) (2020). <https://doi.org/10.1186/s40708-020-00112-2>
37. Odusami, M., Maskeliūnas, R., Damaševičius, R.: Pixel-level fusion approach with vision transformer for early detection of alzheimer's disease. *Electronics* **12**(5), 1218 (2023)
38. Pan, J., et al.: Generative ai empowers estimation of brain structure-function connectivity for alzheimer's disease. *IEEE Transactions on Consumer Electronics* **72**(2), 3781 – 3788 (2026)
39. Rahman, M.A., Brown, D.J., Shopland, N., Burton, A., Mahmud, M.: Explainable multimodal machine learning for engagement analysis by continuous performance test. *Lecture Notes in Computer Science* **13309 LNCS**, 386 – 399 (2022)

40. Ruiz, J., Mahmud, M., Modasshir, M., Shamim Kaiser, M.: 3d densenet ensemble in 4-way classification of alzheimer's disease. *Lecture Notes in Computer Science* **12241 LNAI**, 85 – 96 (2020). https://doi.org/10.1007/978-3-030-59277-6_8
41. Sarraf, S., Sarraf, A., DeSouza, D.D., Anderson, J.A., Kabia, M., Initiative, A.D.N.: Ovitad: Optimized vision transformer to predict various stages of alzheimer's disease using resting-state fmri and structural mri data. *Brain Sciences* **13**(2), 260 (2023)
42. Shaffi, N., Hajamohideen, F., Abdesselam, A., Mahmud, M., Subramanian, K.: Ensemble classifiers for a 4-way classification of alzheimer's disease. *Communications in Computer and Information Science* **1724 CCIS**, 219 – 230 (2022). https://doi.org/10.1007/978-3-031-24801-6_16
43. Shaffi, N., Hajamohideen, F., Mahmud, M., Abdesselam, A., Subramanian, K., Sariri, A.A.: Triplet-loss based siamese convolutional neural network for 4-way classification of alzheimer's disease. *Lecture Notes in Computer Science* **13406 LNAI**, 277 – 287 (2022). https://doi.org/10.1007/978-3-031-15037-1_23
44. Shaffi, N., Subramanian, K., Vimbi, V., Hajamohideen, F., Abdesselam, A., Mahmud, M.: Performance evaluation of deep, shallow and ensemble machine learning methods for the automated classification of alzheimer's disease. *International Journal of Neural Systems* **34**(7) (2024). <https://doi.org/10.1142/S0129065724500291>
45. Shaffi, N., Subramanian, K., Vimbi, V., Hajamohideen, F., Abdesselam, A., Mahmud, M.: Performance evaluation of deep, shallow and ensemble machine learning methods for the automated classification of Alzheimer's disease. *International Journal of Neural Systems* **2450029** (2024)
46. Shaffi, N., Vimbi, V., Mahmud, M., Subramanian, K., Hajamohideen, F.: Bagging the best: A hybrid svm-knn ensemble for accurate and early detection of Alzheimer's and parkinson's diseases. In: *International Conference on Brain Informatics*. pp. 443–455 (2023)
47. Shaffi, N., Viswan, V., Mahmud, M.: Ensemble of vision transformer architectures for efficient alzheimer's disease classification. *Brain Informatics* **11**(1) (2024). <https://doi.org/10.1186/s40708-024-00238-7>
48. Shaffi, N., Viswan, V., Mahmud, M.: Vistad: A vision transformer pipeline for the classification of alzheimer's disease (2024). <https://doi.org/10.1109/IJCNN60899.2024.10650975>
49. Shaffi, N., Viswan, V., Mahmud, M.: VisTAD: A vision transformer pipeline for the classification of Alzheimer's disease. In: *2024 International Joint Conference on Neural Networks (IJCNN)*. pp. 1–8. IEEE (2024)
50. Shaffi, N., Viswan, V., Mahmud, M., Hajamohideen, F., Subramanian, K.: Multi-planar mri-based classification of alzheimer's disease using tree-based machine learning algorithms. In: *2023 IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*. pp. 496–502. IEEE (2023)
51. Shaffi, N., Viswan, V., Mahmud, M., Hajamohideen, F., Subramanian, K.: Multi-planar mri-based classification of alzheimer's disease using tree-based machine learning algorithms. p. 496 – 502 (2023)
52. Sheng, J., Wu, S., Zhang, Q., Li, Z., Huang, H.: A binary classification study of alzheimer's disease based on a novel subclass weighted logistic regression method. *IEEE Access* **10**, 68846–68856 (2022). <https://doi.org/10.1109/ACCESS.2022.3186888>
53. Sherwani, P., Nandhakumar, P., Srivastava, P., et al.: Comparative analysis of alzheimer's disease detection via mri scans using convolutional neural network and vision transformer. In: *Proc. ICECONF*. pp. 1–9 (2023)

54. Shin, H., Jeon, S., Seol, Y., Kim, S., Kang, D.: Vision transformer approach for classification of alzheimer's disease using 18f-florbetaben brain images. *Applied Sciences* **13**(6), 3453 (2023)
55. Tanveer, M., Richhariya, B., Khan, R.U., Rashid, A.H., Khanna, P., Prasad, M., Lin, C.T.: Machine learning techniques for the diagnosis of alzheimer's disease: A review. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **16**(1s), 1–35 (2020)
56. Touvron, H., Cord, M., Douze, M., et al.: Training data-efficient image transformers & distillation through attention. In: *Proc. ICML*. pp. 10347–10357 (2021)
57. Vimbi, V., Shaffi, N., Mahmud, M.: Interpreting artificial intelligence models: a systematic review on the application of lime and shap in alzheimer's disease detection. *Brain Informatics* **11**(1) (2024). <https://doi.org/10.1186/s40708-024-00222-1>
58. Vimbi, V., Shaffi, N., Mahmud, M., Subramanian, K., Hajamohideen, F.: Transfer learning-based ensemble of deep neural architectures for alzheimer's and parkinson's disease classification. *Communications in Computer and Information Science* **2065 CCIS**, 186 – 204 (2024)
59. Viswan, V., Hajamohideen, F., Subramanian, K., Shaffi, N., Mahmud, M.: Enhancing insights: unravelling the potential of preprocessing MRI for artificial intelligence based Alzheimer's disease classification (2024)
60. Viswan, V., Shaffi, N., Mahmud, M., Subramanian, K., Hajamohideen, F.: A comparative study of pretrained deep neural networks for classifying alzheimer's and parkinson's disease. p. 1334 – 1339 (2023)
61. Viswan, V., Shaffi, N., Mahmud, M., Subramanian, K., Hajamohideen, F.: A comparative study of pretrained deep neural networks for classifying Alzheimer's and parkinson's disease. In: *2023 IEEE symposium series on computational intelligence (SSCI)*. pp. 1334–1339. IEEE (2023)
62. Viswan, V., Shaffi, N., Mahmud, M., Subramanian, K., Hajamohideen, F.: Explainable artificial intelligence in alzheimer's disease classification: A systematic review. *Cognitive Computation* **16**(1), 1 – 44 (2024). <https://doi.org/10.1007/s12559-023-10192-x>
63. Viswan, V., Shaffi, N., Malathy, E., Chemmalar Selvi, G., Kavitha, B., Abdessalam, A., Wang, S., Suganthan, P.N., Shezawi, I.A., Mahmud, M.: Multimodal fusion and explainability of artificial intelligence models in alzheimer's disease detection. *Brain Informatics* **13**(1) (2026). <https://doi.org/10.1186/s40708-025-00291-w>
64. Viswan, V., Shaffi, N., Sirasanagandla, S.R., Mahmud, M.: Enhancing alzheimer's disease detection: A hybrid approach with cnns and vision transformers. *Communications in Computer and Information Science* **2607 CCIS**, 264 – 276 (2025)
65. Viswan, V., Shaffi, N., Subramanian, K., Hajamohideen, F.: Optimizing medical imaging quality: An in-depth examination of preprocessing methods for brain MRIs. In: *Proc. AII*. pp. 65–81 (2023)
66. Wu, C., Guo, S., Hong, Y., et al.: Discrimination and conversion prediction of mild cognitive impairment using convolutional neural networks. *Quant. imaging med. surg.* **8**(10), 992 (2018)
67. Yi, F., Yang, H., Chen, D., Qin, Y., Han, H., Cui, J., Bai, W., Ma, Y., Zhang, R., Yu, H.: Xgboost-shap-based interpretable diagnostic framework for alzheimer's disease. *BMC medical informatics and decision making* **23**(1), 137 (2023)
68. Zhou, T., et al.: Bg-gan: Generative ai enable representing brain structure-function connections for alzheimer's disease. *IEEE Transactions on Consumer Electronics* **71**(2), 3087 – 3098 (2025)