Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Enhancing Alzheimer's Disease Diagnosis using Transparent AI Models: A Survey

Madhvi Garg[1], Akshay Juneja[2], Deepak Painuli[3], Priyanka Kumari[4]

[1,2,3]Department of Computer Science and Engineering, College of Smart Computing, COER University, Roorkee, Uttarakhand, 247667 , India
[4]ECE Department, School of Engineering and Technology, CGC University, Mohali, Punjab, India
***Corresponding Author.:** akshay.j1207@gmail.com

Abstract

The integration of artificial intelligence (AI) in Alzheimer's disease detection has improved diagnostic capabilities. It enables accurate and early identification of the disease through advanced analysis of neuroimaging data, biomarkers, and cognitive assessments. The complexity of modern AI models is high and lack clinical adoption due to lack of transparency and interpretability. It creates a barrier for healthcare professionals to understand and trust the AI's decision-making process. This paper presents role of explainable artificial intelligence (XAI) to highlight the importance of high-performing AI models. The various XAI techniques are categorized based on their scope (global vs. local), implementation timing (ante-hoc vs. post-hoc), and applications (model-specific vs. model-agnostic). The global explanations provide an overall understanding of the model's behavior while local explanations focus on individual predictions. The ante-hoc models are interpretable from the start, and post-hoc models require additional techniques to explain their outputs. Also, model-specific approaches are used for particular algorithms, model-agnostic approaches are applied across different model types. The different XAI frameworks such as SHAP, LIME, Grad-CAM, and LRP are studied for AD classification to obtain model decisions without compromising accuracy. This paper also presents different challenges such as data imbalance and over fitting that provides a balance between model complexity and interpretability. The availability of small dataset of different stages of Alzheimer introduces scenes in the model's performance. Also, model performs well during training and generates poor results during testing. It is observed that XAI algorithms have improved the transparency and trustworthiness in AI-assisted diagnostic systems for effective Alzheimer's disease detection.

Keywords – Alzheimer, Grad CAM, LIME, LRP, SHAP, XAI

1. Introduction

Alzheimer's disease (AD) is a progressive and incurable neurological disorder that mostly affects the people above 45 years of age. It deteriorates their quality of life and imposes various challenges on patients and their caretakers [1, 2]. As per the World Alzheimer report [3], it is

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

observed that nearly 55 million individuals have been clinically diagnosed with AD across the globe. The predicted number of cases up to 2050 is 139 million [3]. About 75% of cases remain undiagnosed. The different symptoms of AD are memory impairment, behavioral changes, and deficits in vision and mobility that reduces the efficiency of the person to complete daily tasks [1, 4, and 5]. With the advancement of the diseases, patients are unable to live independently [2, 6].

In recent years, Artificial Intelligence (AI), machine learning (ML), and deep learning (DL) techniques have shown advancements in anomaly detection, bio signal and medical image analysis, assessment and classification of neurodevelopment disorders such as autism, detection and management of neurological conditions, COVID-19 pandemic response, elderly care and monitoring, cyber security, disease diagnosis, and smart healthcare services.

AI has enhanced the clinical diagnosis of AD as it leverages multimodal medical imaging data (for instance, 2D/3D MRI, PET, CT) with high accuracy. The automated systems identify dementia stages by learning patterns from input data, thereby facilitating decision-making with minimal human intervention. The traditional ML and DL models have achieved exceptional performance in AD detection in terms of different evaluation metrics.

However, AI models are considered as "black boxes" among medical professionals due to their lack of interpretability [7]. Moreover, the stakeholders prioritize trustworthiness and reliability over accuracy. It is found disadvantageous to utilize AI-driven computer-aided diagnosis (CAD) systems in clinical practice [8, 9].

In the last decade, ML and DL algorithms have revolutionized decision-making in biomedicine [10], biology [10], and healthcare [11]. However, the working of neurons in the hidden layers is still unknown [12]. In contrast, explainable AI (XAI) provides interpretable outcomes, such that users understand model behavior, rectify errors, and place greater confidence in predictions.

XAI is a sub-domain of AI that makes the systems transparent, understandable, and trustworthy for end-users [13]. It comprises of the techniques that include model architecture, prediction mechanisms, and decision logic, fostering trust and facilitating human-AI collaboration [14, 15]. In healthcare, XAI is essential to justify model outputs, enhance transparency, extract actionable insights, and improve clinical adoption. For instance, implementation of XAI in AD diagnosis using MRI highlights the key imaging features to perform classification. It enables healthcare practitioners to validate results and identify potential errors [16, 17]. It is essential to ensure fidelity, accountability, and trust in AI-assisted medical decision-making.

The trade-off between model complexity and explain ability is a major issue in AI development. Although, convolution neural networks (CNNs) have achieved high accuracy, their interpretability remains limited [18]. There are simpler models with higher explain ability and less accuracy, whereas complex models yield higher accuracy along with transparency. In healthcare, predictive accuracy is most important parameter for clinical validation, yet patients and practitioners prefer human expertise over opaque machine decisions. Figure 1 depicts different stages of AD classification.

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

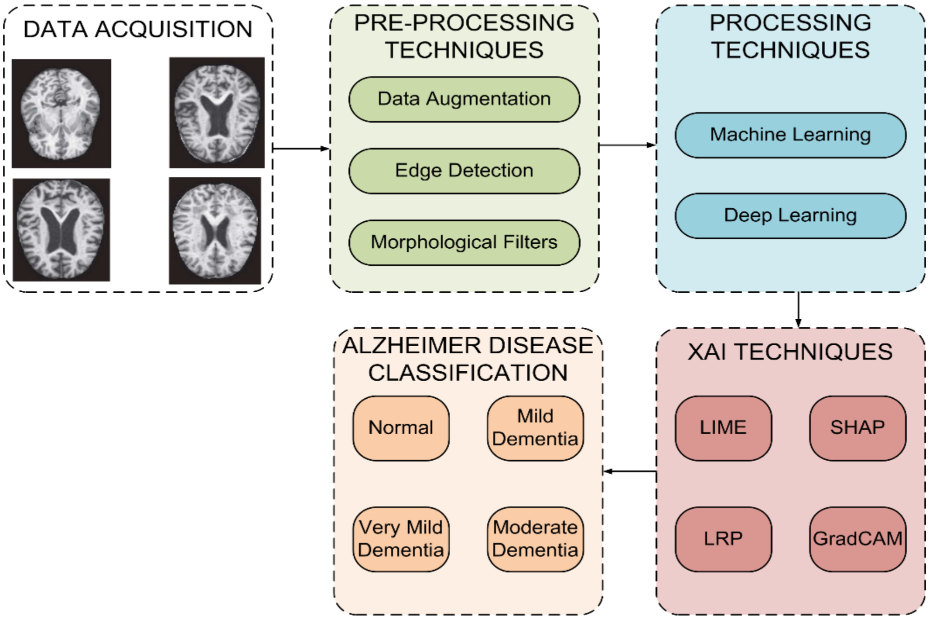

Figure 1. Different stages of Alzheimer Disease classification.

In recent years, XAI has gained popularity due to the frequent demands for accountability and the need for trustworthy AI in different domains. This paper presents a thorough review of XAI for AD classification. This study systematically analyzes the different XAI techniques used by researchers in AD detection so far.

The key contributions include:

- Analysis of XAI methods for AD diagnosis over the past decade, along with their strengths and limitations.
- Identification of trustworthy XAI techniques to promote clinical adoption.
- Discussion of challenges, advantages, and future directions in utilizing XAI in AD.

This paper comprises as follows: Section II presents the scope and importance of XAI techniques in ML and DL models. Section III discusses different XAI frameworks used for AD classification. The challenges and limitations are discussed in Section IV. Finally, the paper is concluded in Section V.

2. Background

This section presents the importance of explainable AI (XAI) methodologies, along with their applications in AD classification [13].

- Scope of explanations (i.e., local versus global interpretability)
- Implementation chronology (i.e., ante-hoc versus post-hoc approaches)
- Model applicability (i.e., model-specific versus model-agnostic techniques)

2.1. Scope-based XAI Techniques –

The explanatory scope of XAI techniques decides whether it provides interpretation at global or local level.

Global scale: The model's behavior is analyzed over the complete dataset. The model's efficiency is enhanced during decision-making. For instance, decision tree (DT) algorithms enhance the

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

intrinsic global explain ability, as their complete decision logic is explicitly represented through hierarchical tree visualization.

Local scale: The model's performance is analyzed over few data samples only. It shows the computation of individual decisions, and enhances the user confidence in model outputs. A local explanation of a DT model as a singular decision path corresponds to a specific input. A systematic accumulation of multiple local explanations yields global model.

2.2. Implementation-based XAI Techniques –

XAI methods are categorized as ante-hoc and post-hoc based on interpretability related to model training.

Ante-hoc Techniques: The term "ante-hoc" means "before this." It incorporates explain ability directly into the model before its training, such as linear regression, decision trees, and Bayesian models with their interpretable coefficients, visualizable rule hierarchies, and probabilistic reasoning frameworks, respectively.

Post-hoc Techniques: Further, post-hoc methods, known as "after this" implement interpretability techniques to already-trained models. These approaches are used for complex models such as support vector machines and CNNs.

2.3. Applications-based XAI Techniques –

Another categorization is model-specific and model-agnostic approaches.

Model-specific methods: These models are integrated with specific architectures and are not transferred to other model types. They work by reverse-engineering the data-process path to identify the factors that influence their decisions.

Model-agnostic methods: These models are implemented with machine learning model irrespective of their architecture. These methods treat the model as a black box that needs no knowledge of its internal parameters. These models are important in clinical AD diagnosis in which different institutions employed diverse model architectures. They have a consistent framework for model interpretation across various implementations.

3. XAI Frameworks

This section presents different XAI frameworks that exist in the literature.

3.1. Local Interpretable Model Agnostic Explanations (LIME)

It is a versatile solution to obtain instance-specific explanations of black-box model. It is a model-agnostic approach that uses input data samples and observes corresponding output variations to construct local trustworthy surrogate models [15]. This perturbation-based methodology enables LIME to identify and rank the most essential features that contributes to individual predictions, irrespective of the model design.

The local explanation is preferred over global model behavior. The computational efficiency and interpretability are enhanced. The framework has agnostic nature that makes it efficient to explain complex classifiers in production environments.

LIME generates visual explanations for image classification tasks that highlight salient regions. The text-based models receive either textual or numerical feature importance scores, while tabular data is explained using various interpretable representations. Thus, LIME is widely adopted across domains that require transparent AI decision-making, such as AD detection. Table 1 presents the summary of existing AD classification techniques that utilized LIME for model prediction.

3.2. Shapley Additive explanations (SHAP)

It produces interpretable explanations for machine learning models by calculating Shapley values. These values represent the contribution of each feature in model's prediction. It studies all possible combinations of features and their weighted impacts. It follows the four key principles: efficiency, symmetry, null player, and additively. The efficiency refers to the total attribution is

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306
Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

equal to the prediction output, while symmetry states that equal features receive equal attribution. Further, the null player is the features with no impact and receives zero attribution.
The additively defines the consistent aggregation across feature subsets. SHAP presents local explanations that offer consistent and precise interpretations of individual predictions. It is used to analyze the numerical data that generates intuitive visualizations with feature importance

Table -1 Summary of LIME-based technique for AD classification

Ref.	Classes	Dataset	Technique	Accuracy (in %)
[19]	4	Image, Numeric	CNN	97.3
[20]	2	Image	Resent	86.86
			VGG16	87.44
[21]	2	Numeric	ANN	79.96
			CNN	83.96
[22]	2	Categorical	BERT-DE	86.25

Patterns. Table 2 presents the summary of existing AD classification techniques that utilized SHAP for model prediction.

3.3. Layer wise Relevance Propagation (LRP)

LRP is designed to interpret CNNs. It propagates relevance scores backward through the network, i.e., it begins analysis from the output layer and move towards the input layer. The contribution of each neuron is calculated to predict the output class and visualize a heat map. It highlights the most influential regions of the input image or video. LRP provide detailed, pixel-level explanations that help to understand the visual data processing using CNNs. The heat maps are used to identify different parts of an image that focuses on the model during predictions. It shows transparency in the model's decision-making process. Table 3 presents the summary of existing AD classification techniques that utilized LRP for model prediction.

Table -2 Summary of SHAP-based techniques for AD classification

Ref.	No. of Classes	Technique	Accuracy (in %)
[23]	4	RF	First Layer – 93.95; Second Layer – 87.08
[24]	3	RF, XGBoost	–
[25]	3	RF	75
[26]	3	XGBoost	–
[27]	2	SVM	80.8
[28]	4	XGBoost	84.2
[29]	2	RF, SVM, XGBoost	–
[30]	3	SVM-SMOTE	80.7

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306
Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

[31]	3	LGBM	–
[32]	4	XGBoost, RF, SVM	–
[33]	2	XGBoost	92.6
[34]	3	RF	–

Table 3. Summary of LRP-based techniques for AD classification.

Ref.	No. of Classes	Technique	Performance Metric
[35]	2	CNN	Specificity – 94%
[36]	4	CNN	Accuracy – 78.12%
[37]	2	CNN	–
[38]	2	3D CNN	–

3.4. Gradient-weighted Class Activation Mapping (GradCAM)

Grad CAM is used to visualize important regions in input images that influence CNN predictions. The gradient information is obtained from the final convolution layer. Grad CAM shows a wide overview of significant regions, that makes it effective for quick visual assessments. It is easy to implement and effective for various CNN architectures. The heat maps are used to study complex model internals and human-interpretable explanations. Table 4 presents the summary of existing AD classification techniques that utilized Grad CAM for model prediction.

3.5. Combined Techniques

In recent works, the researchers have combined different XAI techniques to perform AD classification (see Table 5).

4. Challenges and Limitations

This section presents the various challenges present during implementation of XAI techniques in AD classification.

4.1. Data Imbalance

The majority of medical datasets suffers from class imbalance, such that the healthy samples are greater than samples with AD. The models show biasness towards the majority class that reduces their sensitivity to AD classes.

Therefore, the researchers employ various data balancing techniques such as oversampling, under sampling of the majority class, and SMOTE. These techniques generate synthetic AD cases to balance the healthy and diseased classes during training.

Table 4. Summary of Grad CAM-based techniques for AD classification.

Ref.	No. of classes	Model	Accuracy (in %)
[39]	2	CNN with self-attention	81
[40]	4	CNN	82
		VGG16	87

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306
Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

[41]	2	VGG16	80.7
	2	Residual Network	85.1
	2	Attenuation Network	86
[42]	3	CNN	–
[43]	2	CNN and BRNN	–
[44]	2	CNN	–

Table 5. Summary of combined XAI techniques for AD classification.

Ref.	No. of classes	Classifier	Model	Accuracy (in %)
[45]	2	LIME, SHAP	SVM	85.9
			ken	87.27
			MLP	91.94
[46]	2	HAM, PCR	CNN	95.4%
[47]	3	GNNExplainer	GNN	53.5 ± 4.5%
[48]	2	Occlusion Sensitivity Mapping, Saliency Map, LRP	EfficientNet-B0	80
[16]	3	–	3D CNN	–
[49]	2	DT	Boost	91.0
[24]	2	3D Ultra metric Contour Map, CAM, Grad CAM	3D CNN	76.6%
[50]	3	Occlusion	3D CNN	77.0%

4.2. Over fitting

Another challenge in the implementation of DL for AD classification is over fitting. It occurs when models learn training data patterns that include noise and irrelevant variations, rather than learning essential features. It results in high training accuracy and low testing accuracy. The risk of over fitting increases with small or imbalanced datasets, due to data privacy regulations in medical domain. It is difficult to obtain the annotated neuroimaging dataset.

There are various regularization techniques that are used to prevent overfitting. For instance, dropout regularization randomly eliminates neurons during training. It forces the network to develop redundant representations and reduces the dependency on specific nodes. The data augmentation techniques such as rotation, flipping, and scaling improves the robustness of the architecture during training.

5. Conclusion

In this paper, the need and importance of XAI in the field of Alzheimer Disease detection is discussed. It minimizes the complexity of AI models and improves the clinical interpretability using different frameworks that ensure rationality and transparency in decision-making. The XAI

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

JRICST A Next Generation Research

techniques are categorized as global-level and local-level based on their scope. They are also categorized as ante-hoc and post-hoc techniques based on their implementation. Moreover, these techniques are also classified as either model-specific or model-agnostic based on their applications. Furthermore, the different XAI frameworks such as SHAP, LIME, GradCAM, and LRP have been integrated with ML and DL models to enhance trust in AI predictions, without affecting diagnostic performance. It is followed by various challenges in real-world clinical adoption such as data imbalance and over fitting. It is observed that the balance between model complexity and explain ability is needed for clinical acceptance to design a transparent and reliable AI-assisted diagnostic systems.

References

[1] Kanpur, S., Bharadwaj, C. B., Hardwar, R., & Kayak, J. S. (2024). An approach for XAI visualizations for explains ability of Alzheimer's detection. In 2024 2nd International Conference on Networking, Embedded and Wireless Systems (ICNEWS) (pp. 1-6). IEEE.

[2] Visa, V., Shaffer, N., Mahmud, M., Subramanian, K., & Hajamohideen, F. (2024). Explainable artificial intelligence in Alzheimer's disease classification: A systematic review. Cognitive Computation, 16(1), 1-44.

[3] Nimbi, V., Shaffer, N., & Mahmud, M. (2024). Interpreting artificial intelligence models: A systematic review on the application of LIME and SHAP in Alzheimer's disease detection. Brain Informatics, 11(1), 10.

[4] Taine Khosroshahi, M., et al. (2025). Explainable artificial intelligence in neuroimaging of Alzheimer's disease. Diagnostics, 15(5), 612.

[5] Booting, D., et al. (2025). ADAMAEX—Alzheimer's disease classification via attention-enhanced auto encoders and XAI. Egyptian Informatics Journal, 30, 100688.

[6] Bib, N., Courtney, J., & McGuiness, K. (2025). Enhancing brain disease diagnosis with XAI: A review of recent studies. ACM Transactions on Computing for Healthcare, 6(2), 1-35.

[7] Ray, A. (2020). Explainable AI: From black box to glass box. Journal of the Academy of Marketing Science, 48(1), 137-141.

[8] Kauri, D., et al. (2022). Trustworthy artificial intelligence: A review. ACM Computing Surveys, 55(2), 1-38.

[9] Near, M., et al. (2021). A systematic review of human-computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques. IEEE Access, 9, 153316-153348.

[10] Mamoshina, P., et al. (2016). Applications of deep learning in biomedicine. Molecular Pharmaceutics, 13(5), 1445-1454.

[11] Juneau, A., et al. (2024). XcepCovidNet: Deep neural networks-based COVID-19 diagnosis. Multimedia Tools and Applications, 1-31.

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

[12] Yang, G., et al. (2022). Inbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. Information Fusion, 77, 29-52.

[13] Kumar, D., & Mehta, M. A. (2023). An overview of explainable AI methods, forms and frameworks. In Springer International Publishing (pp. 43-59).

[14] Loch, H. W., et al. (2022). Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011-2022). Computer Methods and Programs in Biomedicine, 221, 107161.

[15] Ribera, M. T., Singh, S., & Guest in, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135-1144).

[16] Descant, L. A., et al. (2022). An explainable convolution neural network for the early diagnosis of Alzheimer's disease from 18F-FDG PET. Journal of Digital Imaging, 1-15.

[17] Ghassemi, M., et al. (2021). The false hope of current approaches to explainable artificial intelligence in health care. The Lancet Digital Health, 3(11), e745-e750.

[18] Jung, Y. J., et al. (2021). Explaining CNN and RNN using selective layer-wise relevance propagation. IEEE Access, 9, 18670-18681.

[19] Kamala, M. S., et al. (2021). Alzheimer's patient analysis using image and gene expression data and explainable-AI to present associated genes. IEEE Transactions on Instrumentation and Measurement, 70, 1-7.

[20] Shad, H. A., et al. (2021). Exploring Alzheimer's disease prediction with XAI in various neural network models. In TENCON 2021-2021 IEEE Region 10 Conference (pp. 720-725).

[21] Sidulova, M., et al. (2021). Analysis explainable image for Alzheimer's disease and mild cognitive impairment diagnosis. In IEEE Applied Imagery Pattern Recognition Workshop (AIPR) (pp. 1-6).

[22] Iliac, L., & Astounds, D. (2022). Explainable identification of dementia from transcripts using transformer networks. IEEE Journal of Biomedical and Health Informatics, 26(8), 4153-4164.

[23] El-Sapping, S., et al. (2021). A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer's disease. Scientific Reports, 11(1), 2660.

[24] Yang, C., et al. (2018). Visual explanations from deep 3D convolution neural networks for Alzheimer's disease classification. In AMIA Annual Symposium Proceedings (p. 1571).

[25] Lombardi, A., et al. (2022). A robust framework to investigate the reliability and stability of explainable artificial intelligence markers of mild cognitive impairment and Alzheimer's disease. Brain Informatics, 9(1), 1-17.

[26] Hernandez, M., et al. (2022). Explainable AI toward understanding the performance of the top three TADPOLE Challenge methods in the forecast of Alzheimer's disease diagnosis. Plops One, 17(5), e0264695.

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

[27] Lai, Y., et al. (2022). Identification of endoplasmic reticulum stress-associated genes and subtypes for prediction of Alzheimer's disease based on interpretable machine learning. Frontiers in Pharmacology, 13.

[28] Bogdanovic, B., et al. (2022). In-depth insights into Alzheimer's disease by using explainable machine learning approach. Scientific Reports, 12(1), 1-26.

[29] Chun, M. Y., et al. (2022). Prediction of conversion to dementia using interpretable machine learning in patients with amnesic mild cognitive impairment. Frontiers in Aging Neuroscience, 14.

[30] CSU, X., & Yan, X. (2022). A convenient and reliable multi-class classification model based on explainable artificial intelligence for Alzheimer's disease. In IEEE AEECA (pp. 671-675).

[31] Salish, A., et al. (2022). Investigating explainable artificial intelligence for MRI-based classification of dementia: A new stability criterion for explainable methods. In IEEE ICIP (pp. 4003-4007).

[32] Bloch, L., & Friedrich, C. M. (2022). Machine learning workflow to explain black-box models for early Alzheimer's disease classification evaluated for multiple datasets. ArXiv: 2205.05907.

[33] Dansk, S. O., et al. (2021). Developing an explainable machine learning-based personalised dementia risk prediction model: A transfer learning approach with ensemble learning algorithms. Frontiers in Big Data, 4, 21.

[34] Hammond, T. C., et al. (2020). B-amyloidal and tau drive early Alzheimer's disease decline while glucose hypo metabolism drives late decline. Communications Biology, 3(1), 1-13.

[35] Bohlen, M., et al. (2019). Layer-wise relevance propagation for explaining deep neural network decisions in MRI-based Alzheimer's disease classification. Frontiers in Aging Neuroscience, 11, 194.

[36] Sudan, K. M., et al. (2022). Alzheimer's disease analysis using explainable artificial intelligence (XAI). In IEEE ICSCDS (pp. 419-423).

[37] Wang, D., et al. (2023). Deep neural network heat maps capture Alzheimer's disease patterns reported in a large meta-analysis of neuroimaging studies. NeuroImage, 269.

[38] Pohl, T., et al. (2022). Interpretability of deep neural networks used for the diagnosis of Alzheimer's disease. International Journal of Imaging Systems and Technology, 32(2), 673-686.

[39] Ruengchaijatuporn, N., et al. (2022). An explainable self-attention deep neural network for detecting mild cognitive impairment using multi-input digital drawing tasks. Alzheimer's Research & Therapy, 14(1), 1-11.

[40] Jain, V., et al. (2021). A novel AI-based system for detection and severity prediction of dementia using MRI. IEEE Access, 9, 154324-154346.

Received: 2026-04-04
Accepted: 2026-07-14
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030306

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

[41] Zhang, X., et al. (2021). An explainable 3D residual self-attention deep neural network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. IEEE Journal of Biomedical and Health Informatics.

[42] Oh, K., et al. (2023). Learn-explain-reinforce: Counterfactual reasoning and its guidance to reinforce an Alzheimer's disease diagnosis model. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(4), 4843-4857.

[43] Graham, N., et al. (2023). Prediction of Alzheimer's progression based on multimodal deep-learning-based fusion and visual explain ability of time-series data. Information Fusion, 92, 363-388.

[44] Shojaei, S., et al. (2023). An evolutionary explainable deep learning approach for Alzheimer's MRI classification. Expert Systems with Applications, 220.

[45] Lovelies, G., et al. (2022). Explanation-driven HCI model to examine the mini-mental state for Alzheimer's disease. ACM Transactions on Multimedia Computing, Communications, and Applications.

[46] Yu, L., et al. (2022). A novel explainable neural network for Alzheimer's disease diagnosis. Pattern Recognition, 131.

[47] Kim, M., et al. (2021). Interpretable temporal graph neural network for prognostic prediction of Alzheimer's disease using longitudinal neuroimaging data. In IEEE BIBM (pp. 1381-1384).

[48] Borden, V., et al. (2022). Explainable AI points to white matter hyper intensities for Alzheimer's disease identification: A preliminary study. In IEEE EMBC (pp. 484-487).

[49] Garcia-Gutierrez, F., et al. (2022). GA-MADRID: Design and validation of a machine learning tool for the diagnosis of Alzheimer's disease and front temporal dementia using genetic algorithms. Medical & Biological Engineering & Computing, 60(9), 2737-2756.

[50] Wreaked, J., et al. (2018). Visualizing convolution networks for MRI-based diagnosis of Alzheimer's disease. In Understanding and Interpreting Machine Learning in Medical Image Computing Applications (pp. 24-31). Springer.