Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Feature-Level Fusion of T2- and PD-Weighted MRI Using Vision Transformer for Five-Stage Alzheimer's Disease Classification

Zoulikha Bebboukha[1], Athmane Zitouni[2]
1 zoulikha.bebboukha@univ-biskra.dz, 2 athmane.zitouni@univ-biskra.dz
1,2 *Vision and Communication Systems Laboratory (VCS), Department of Electrical Engineering, University of Biskra, Algeria.*

ARTICLE INFO

Received: 30 Dec 2025

Revised: 04 Feb 2026

Accepted: 28 Feb 2026

ABSTRACT

Alzheimer's disease (AD) is a progressive neurodegenerative disorder in which early and accurate diagnosis remains challenging, particularly for clinically overlapping mild cognitive impairment (MCI) stages. Although magnetic resonance imaging (MRI) is widely used for AD assessment, reliance on a single imaging sequence often limits discrimination between disease stages. Feature fusion has recently emerged as an effective strategy to enhance classification performance while maintaining clinical feasibility. In this study, an efficient feature-level fusion framework is proposed that integrates complementary information from T2-weighted and proton density (PD-weighted) MRI. Deep features are extracted exclusively using a Vision Transformer (ViT) model and combined through a weighted fusion strategy, followed by k-nearest neighbors classification. Experiments conducted on the ADNI dataset demonstrate that the proposed approach achieves high discriminative performance across five classes (AD, CN, EMCI, LMCI, and MCI), reaching an overall accuracy of approximately 99.35%. The fusion framework consistently outperforms single-modality models, particularly in distinguishing transitional stages. These results indicate that ViT-based feature fusion of T2 and PD MRI provides a robust and computationally efficient solution for multi-stage Alzheimer's disease classification, with strong potential for practical clinical application.

Keywords: Alzheimer's disease, MRI images, Vision Transformer,

INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and a leading cause of dementia, characterized by gradual cognitive decline and memory impairment. With the rapid aging of the global population, early and accurate diagnosis has become essential for timely intervention and improved patient outcomes. Traditional diagnostic methods based on clinical assessment and neuropsychological tests are often subjective and limited, while neuroimaging techniques such as MRI and PET provide valuable insights into structural and functional brain changes. However, the high dimensionality of imaging data and the strong similarity between disease stages pose significant challenges for reliable classification.

Recent deep learning approaches have improved diagnostic accuracy through automated feature extraction and attention mechanisms, but many rely on complex architectures or costly PET imaging. To address these limitations, this study proposes an efficient feature-level fusion framework that integrates complementary T2-weighted and PD-weighted MRI features using a Vision Transformer and weighted fusion for multi-stage AD classification.

The remainder of this paper is organized as follows: Section II reviews related work, Section III presents the proposed methodology, Section IV reports experimental results, Section V discusses the findings, and Section VI concludes the paper.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

RELATED WORK

Recent studies have demonstrated significant progress in Alzheimer's disease (AD) classification using deep learning applied to MRI and multimodal data, with a focus on enhanced feature extraction and improved stage discrimination. Metric-learning and attention-based CNN models, such as the Conditional Deep Triplet Network (CDTN) and EnCNN, have reported accuracies of approximately 95%, highlighting the effectiveness of discriminative learning and attention mechanisms [8;4]. Other works [1] have emphasized the importance of high-quality preprocessing, particularly skull stripping, with CNN-based approaches achieving accuracies exceeding 98% on skull-stripped MRI data [5]. Hybrid architectures combining spatial and sequential modeling, such as ResNet–BiLSTM, have further improved classification performance but at the cost of increased architectural complexity [10].

In [9] proposed a novel multimodal fusion strategy, termed MRI-p value, which integrates genetic information as prior knowledge without disrupting the three-dimensional MRI structure, using a dual-path learning framework where one branch processes MRI data and the other handles genetic features

Multimodal fusion frameworks integrating MRI and PET, including ADMV-Net, have also shown strong performance, achieving accuracies above 94% through attention-based fusion mechanisms [3]. In contrast, simpler early-fusion strategies based on direct modality concatenation have yielded comparatively lower accuracy, particularly for adjacent stages such as EMCI and LMCI [6]. Overall, existing studies highlight a clear trend toward advanced feature extraction, attention mechanisms, and multimodal integration; however, many approaches rely on heterogeneous modalities or complex architectures that limit clinical feasibility. To address these limitations, the present study introduces a streamlined MRI-only framework based on Vision Transformer–driven weighted feature-level fusion of T2-weighted and PD-weighted MRI, enabling accurate five-stage AD classification with reduced computational burden.

METHODS

This section presents the proposed feature-level fusion framework for multi-stage Alzheimer's disease classification, comprising MRI preprocessing, Vision Transformer–based feature extraction, weighted feature-level fusion, and final classification. An overview of the framework is shown in [Figure 1].

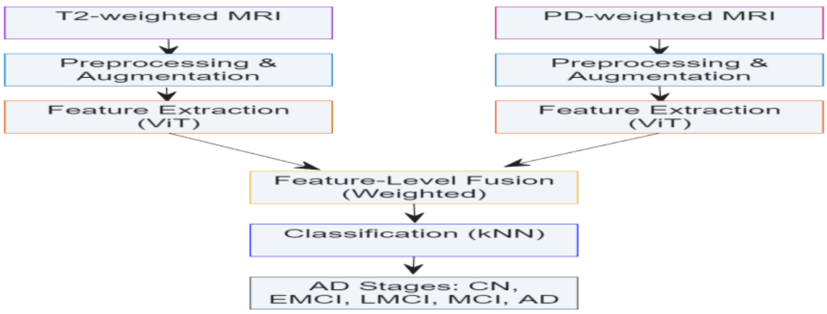

Figure 1 Block diagram of the proposed feature-level fusion approach based on Vision Transformer (ViT).

Data Source and Preprocessing: MRI data were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database [2] and consist of paired T2-weighted and proton density (PD-weighted) MRI scans acquired from subjects spanning five diagnostic categories: AD, CN, EMCI, LMCI, and MCI. All images were resized to 368 × 368 pixels and underwent data augmentation (rotation, flipping, and scaling) to prevent overfitting.

Feature Extraction: Deep features from T2-weighted and PD-weighted MRI were extracted using a pre-trained ViT. Images were divided into non-overlapping patches and encoded via self-attention to model long-range spatial relationships. After removing the classification head, the CLS token embeddings were used as modality-specific feature vectors, denoted as F_{T2} and F_{PD} , which serve as inputs for subsequent weighted feature-level fusion.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

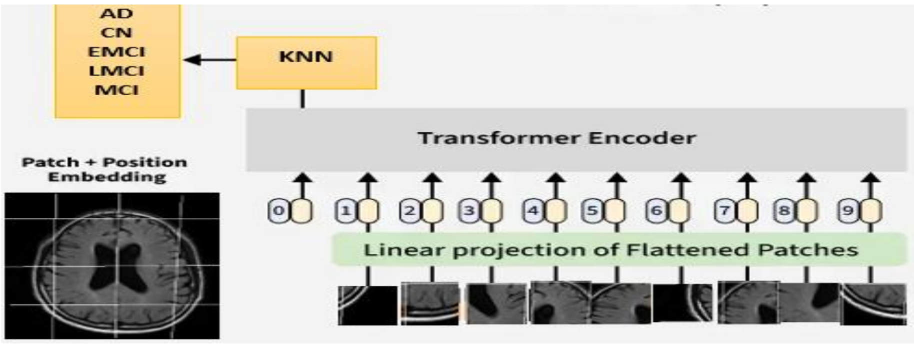

Figure 2. Vision Transformer (ViT) architecture with kNN classifier for five-class Alzheimer's disease classification.

Feature-Level Fusion and Classification: The fused feature vector F_{fused} was computed as a weighted combination of the two modality-specific features:

$$F_{fused} = \alpha * F_{T2\text{-}w} + (1-\alpha) * F_{PD\text{-}w} \quad (1)$$

where: $\alpha \in [0,1]$ is a weighting parameter that controls the relative contribution of each modality. It $\alpha=0.3$ was empirically tuned using the validation set and fixed during testing to ensure unbiased performance evaluation. This strategy preserves feature dimensionality while enhancing discriminative capability. The fused feature vectors was classified using a k-nearest neighbors (KNN) classifier with k=3, selected based on experimental optimization.

RESULTS

This section presents the experimental results of the proposed feature-level fusion framework and compares its performance with single-modality MRI models using quantitative metrics and visual analyses. To ensure robust and unbiased evaluation, 5-fold cross-validation was employed, and performance was assessed using accuracy, precision, recall, F1-score, and receiver operating characteristic (ROC) curves with area under the curve (AUC).

[Figure 3] presents the confusion matrices obtained using T2-weighted MRI features, PD-weighted MRI features, and the proposed feature-level fusion model. When used individually, the T2-weighted and PD-weighted models achieved accuracies of 96.68% and 98.57%, respectively, and exhibited noticeable confusion in early-stage categories such as EMCI. In contrast, the proposed Vision Transformer–based feature-level fusion model achieved an overall classification accuracy of 99.35%, showing a clear reduction in misclassification across all classes, particularly for clinically adjacent stages (AD, CN, EMCI, LMCI, and MCI).

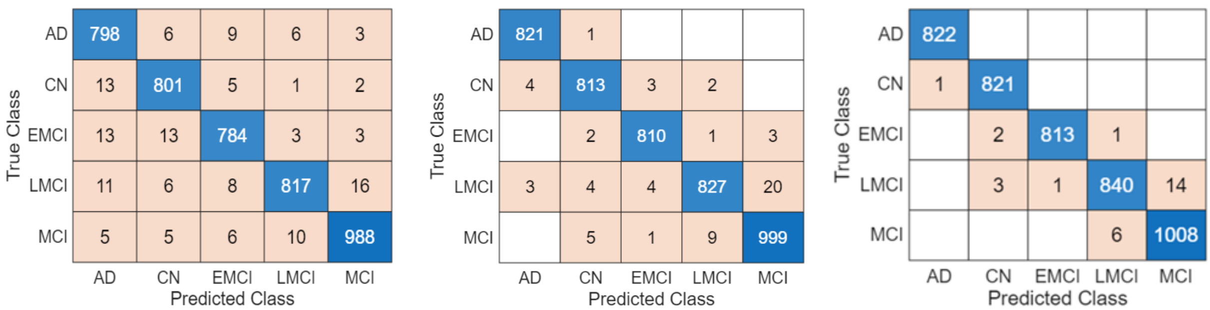

Figure 3. Aggregated confusion matrices across five cross-validation folds for T2-weighted MRI (left), PD-weighted MRI (middle), and feature-level fusion (right).

These results demonstrate the superior discriminative capability and stability of the fusion-based approach compared to single-modality MRI models.

As summarized in [Table 1], the proposed feature-level fusion approach consistently outperforms single-modality models across all evaluation metrics, achieving the highest overall accuracy. The confusion matrix in [Figure 3]

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

confirms improved class separability and reduced misclassification for all five categories (AD, CN, EMCI, LMCI, and MCI).

Table 1 Performance comparison of single-modality and feature-level fusion models

Imaging Modality	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)
T2-weighted MRI	96.68	96.68	96.71	96.69
PD-weighted MRI	98.57	98.57	98.52	98.54
Feature-Level Fusion	99.35	99.35	99.36	99.35

The ROC curves in [Figure 4] demonstrate near-perfect performance, with AUC values of 1.000 for AD, CN, and EMCI, and 0.997 and 0.998 for the more challenging LMCI and MCI stages, respectively, clearly surpassing the single-modality results . These results confirm that Vision Transformer–based fusion of T2-weighted and PD-weighted MRI significantly enhances discriminative capability for multi-stage Alzheimer's disease classification.

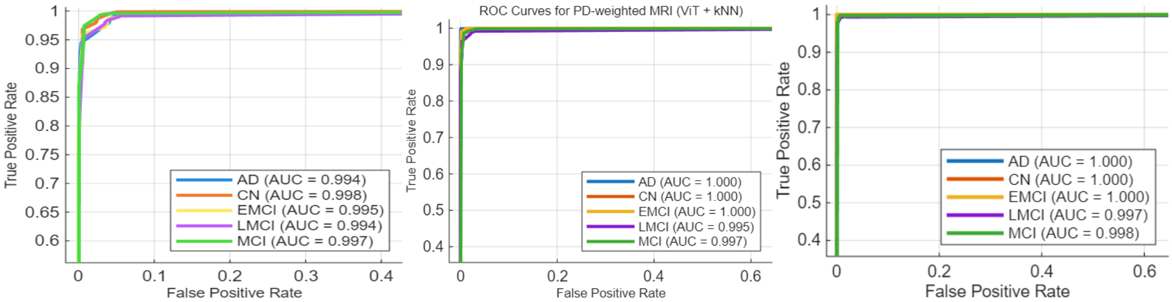

Figure 4. Receiver Operating Characteristic (ROC) curves of T2-weighted MRI, PD-weighted MRI, and feature-level fused show from left to right

Table 2. Comparison of the proposed method with representative reference studies for Alzheimer's disease classification

Dimension	**RBHML [7]**	**ADMV-Net [3]**	**Proposed Study**
Core methodology	Hybrid fusion (ML/CNN)	Multi-view framework with attention modules	Feature-level fusion (ViT)
Imaging modalities	MRI + non-imaging data	MRI + PET	MRI (T2-w + PD-w)
Number of classes	Binary	3 classes	5 classes
Accuracy	96.3%	94.83%	**99.35%**

The comparative analysis presented in [Table 2] between the proposed study and the reference studies highlights a tangible advantage in terms of accuracy and comprehensiveness. While RBHML relies on hybrid, heterogeneous data fusion (images, structured data, and text) for binary or coarse classification with an accuracy of 96.3%, and ADMV-Net uses a framework that fuses MRI and the costly PET images for three classes with an accuracy of 94.83%, the proposed study offers a more focused and efficient MRI-only framework (T2-weighted and PD-weighted images). The essential contribution of this work lies in its ability to provide a detailed five-class classification (AD, CN, EMCI, LMCI, MCI) with an accuracy of 99.35%, demonstrating that fusion of available MRI sequence features is sufficient to surpass more complex and costly MRI–PET models, especially in distinguishing subtle transitional stages.

CONCLUSION

This study demonstrates that feature-level fusion of T2-weighted and proton density (PD-weighted) MRI significantly enhances multi-stage Alzheimer's disease classification. By leveraging Vision Transformer–based deep feature extraction and an empirically weighted fusion strategy, the proposed framework achieves consistently high performance across all evaluation metrics, reaching an overall accuracy of 99.35% for five-class classification. Quantitative results and ROC/AUC analysis confirm improved class separability, particularly for clinically overlapping stages such as EMCI and LMCI. Importantly, the proposed MRI-only approach avoids reliance on costly or less accessible modalities while maintaining strong discriminative capability and computational efficiency, highlighting its potential applicability as a practical and scalable decision-support tool for fine-grained Alzheimer's disease staging in clinical settings

REFRENCES

[1] Alhagi, A. M. R., & Ata, O. (2025). Hybrid Ensemble Deep Learning Framework with Snake and EVO Optimization for Multiclass Classification of Alzheimer's Disease Using MRI Neuroimaging. Electronics, 14(21), 4328. https://doi.org/10.3390/electronics14214328

[2] Alzheimer's Disease Neuroimaging Initiative. (n.d.). ADNI. Retrieved 27 January 2026, from https://adni-lde.loni.usc.edu/

[3] Feng, J., Zhao, X., Liu, Z., Ding, Y., & Wang, F. (2025). A multi-view multimodal deep learning framework for Alzheimer's disease diagnosis. Frontiers in Neuroscience, 19, 1658776. https://doi.org/10.3389/fnins.2025.1658776

[4] J. Arumugam. (2025). An Introspective Study on Attention-Based Transfer Learning in CNNs for Alzheimer's Disease Detection. Journal of Information Systems Engineering and Management, 10(7s), 344–352. https://doi.org/10.52783/jisem.v10i7s.866

[5] Minal A. Zope. (2025). Developing Advanced Deep Learning Layers for Enhanced Automatic Feature Extraction and Higher Accuracy in Alzheimer's Disease Classification. Journal of Information Systems Engineering and Management, 10(10s), 303–319. https://doi.org/10.52783/jisem.v10i10s.1380

[6] Odusami, M., Maskeliūnas, R., Damaševičius, R., & Misra, S. (2023). Explainable Deep-Learning-Based Diagnosis of Alzheimer's Disease Using Multimodal Input Fusion of PET and MRI Images. Journal of Medical and Biological Engineering, 43(3), 291–302. https://doi.org/10.1007/s40846-023-00801-3

[7] Rajakumari, J. G. A., & Balajiraja, D. N. (2025). Advanced Disease Prediction using Multimodal Features and Hybrid Learning Models.

[8] Sachan, S. K., Wagh, H. M., Shaikh, A. M., Raval, D. C., Banu, E. A., & Vishwakarma, P. (2025). Alzheimer's Disease Detection using CDTN over.

[9] Wang, J., Wen, S., Liu, W., Meng, X., & Jiao, Z. (2024). Deep joint learning diagnosis of Alzheimer's disease based on multimodal feature fusion. BioData Mining, 17(1), 48. https://doi.org/10.1186/s13040-024-00395-9

[10]Zs. Khaleel. (2025). Impact of Skull Segmentation in MRI images for Alzheimer's Diagnosis based on Transfer Learning Techniques. Journal of Information Systems Engineering and Management, 10(12s), 490–504. https://doi.org/10.52783/jisem.v10i12s.1858