International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

Machine Learning-Based Fusion of MRI and Transcriptomic Features for Alzheimer's Disease Diagnosis

Bhawana Purohit[1], Dr. Garima Bansal[2]

[1]Research Scholar, [1,2]Assistant professor

[1,2] Shri Khushal Das University Hanumangarh, Rajasthan,

[1]Poornima institute of Engineering and Technology, Jaipur, Rajasthan

Abstract

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder requiring accurate early diagnosis for effective intervention. While MRI provides structural brain information and transcriptomics offers molecular insights, their integration remains underexplored. This study presents a novel machine learning framework that fuses MRI-derived features with blood-based gene expression data for enhanced AD diagnosis. Using data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), we extracted volumetric features from T1-weighted MRI scans and identified differentially expressed genes from peripheral blood transcriptomic profiles. Feature selection employed SHAP values, and class imbalance was addressed using Borderline-SMOTE. Multiple classifiers including Random Forest, XGBoost, and Support Vector Machines were evaluated using cross-validation. The fused feature approach achieved superior performance (accuracy: 92.4%, AUC: 0.96) compared to MRI-only (86.7%, AUC: 0.91) and transcriptomic-only (83.2%, AUC: 0.89) models. Key biomarkers included hippocampal volume, BDNF expression, and APOE-related genes. Feature importance analysis revealed complementary information from both modalities. Multimodal fusion of MRI and transcriptomic data significantly improves AD diagnostic accuracy, supporting the paradigm shift toward integrated biomarker approaches in neurodegenerative disease assessment.

Keywords: Alzheimer's Disease, Machine Learning, MRI, Transcriptomics, Multimodal Fusion, Biomarkers

1. Introduction

Alzheimer's Disease (AD) represents one of the most pressing public health challenges of the twenty-first century, characterized by progressive cognitive decline and neurodegenerative pathology. The drug development pipeline for AD has witnessed few candidates and frequent failures (Cummings et al., 2014), partly attributable to diagnostic delays and phenotypic heterogeneity. Early and accurate diagnosis remains critical for therapeutic efficacy and clinical trial enrichment.

The etiological complexity of late-onset AD involves intricate interactions between genetic susceptibility, molecular pathways, and structural brain changes (Sarma & Chatterjee, 2024; Tanzi, 2012). Genome-wide association studies have identified numerous risk loci, including APOE, BIN1, and CLU (Shen & Jia, 2016; Saykin et al., 2015), yet genetic information alone provides incomplete diagnostic resolution. Concurrently, neuroimaging modalities have advanced significantly, with MRI offering non-invasive assessment of regional atrophy patterns characteristic of AD progression (Bae et al., 2020; Fathi et al., 2024).

The emergence of high-throughput transcriptomic technologies has revealed that peripheral blood gene expression dynamically reflects central nervous system pathology (Liew et al., 2006; Lunnon et al., 2013). Blood-based biomarkers, including BDNF alterations (Angelucci et al., 2010), provide accessible windows into disease-associated molecular cascades. Recent meta-analyses have consolidated brain transcriptomic signatures of AD (Patel et al., 2019), while machine learning applications demonstrate predictive potential using gene expression data alone (Lee & Lee, 2020; Ni & Sethi, 2021).

However, single-modality approaches inherently capture only partial disease manifestations. Structural MRI reveals macroscopic atrophy but lacks molecular specificity, while transcriptomics provides mechanistic insights without spatial localization. The fusion of complementary data types offers theoretical advantages for capturing AD's multifaceted pathology (Westman et al., 2012; Park et al., 2020). Multimodal integration strategies have shown promise in preliminary studies combining imaging with clinical or genetic data (AlMansoori et al., 2024; Fulton et al., 2019).

Despite these advances, systematic fusion of MRI-derived structural features with blood transcriptomic profiles remains underdeveloped. Challenges include high-dimensional feature spaces relative to sample sizes (Shen et al., 2020; Catchpoole et al., 2010), class imbalance in clinical cohorts (Krawczyk, 2016), and the need for interpretable models that identify biologically meaningful biomarkers. SHAP (SHapley Additive exPlanations) values have emerged as robust tools for feature selection and model interpretation (Marcilio & Eler, 2020), potentially addressing the "curse of dimensionality" while maintaining clinical relevance.

This study addresses these gaps by developing and validating a machine learning framework that fuses MRI-derived volumetric features with blood-based transcriptomic profiles for AD diagnosis. We hypothesize that multimodal integration will outperform single-modality approaches by capturing complementary disease signatures. Specific objectives include: (1) identifying discriminative MRI and transcriptomic features, (2) developing fusion strategies that preserve interpretability, (3) addressing class imbalance through synthetic sampling, and (4) validating model performance through rigorous cross-validation.

2. Materials and Methods

2.1 Study Population and Data Sources

Data were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). ADNI is a longitudinal multicenter study designed to develop clinical, imaging, genetic, and biochemical biomarkers for AD detection and tracking. The study included 845 participants with complete MRI and blood transcriptomic data, comprising 312 cognitively normal (CN) controls, 283 mild cognitive impairment (MCI) subjects, and 250 AD patients. Diagnosis followed ADNI standardized criteria: CN subjects had Clinical Dementia Rating (CDR)=0 and normal memory function; MCI subjects had CDR=0.5, objective memory impairment, and preserved functional performance; AD patients met NINCDS-ADRDA criteria for probable AD with CDR≥0.5.

2.2 MRI Acquisition and Processing

T1-weighted structural MRI scans were acquired using 3T scanners with standardized ADNI protocols (Saykin et al., 2010). Preprocessing involved correction for gradient nonlinearity, intensity inhomogeneity correction, and skull stripping using the FreeSurfer pipeline (version 6.0). Regional volumetric segmentation utilized the Desikan-Killiany atlas, yielding 148 features including cortical thickness, surface area, and subcortical volumes. Quality control excluded scans with motion artifacts or segmentation failures.

2.3 Transcriptomic Data Processing

Peripheral blood samples were collected in PAXgene tubes, with RNA extraction and microarray hybridization performed using Affymetrix Human Genome U133 Plus 2.0 arrays. Raw data underwent Robust Multi-array Average (RMA) normalization, log2 transformation, and quality control filtering. Probes were mapped to genes using the latest annotation files, with multiple probes per gene collapsed by median expression. Differential expression analysis employed limma with age and sex as covariates, identifying transcripts with Benjamini-Hochberg adjusted $p<0.05$.

2.4 Feature Selection and Dimensionality Reduction

The initial feature space comprised 148 MRI features and 54,675 transcriptomic probes. To address the high-dimensional low-sample size problem (Shen et al., 2020), we implemented a two-stage feature selection strategy. First, univariate filtering retained features with significant group differences (ANOVA $p<0.01$ for MRI; adjusted $p<0.05$ for transcriptomics). Second, SHAP-based recursive feature elimination identified the most predictive features. For transcriptomic data, we additionally applied variance filtering (retaining top 25% most variable genes) prior to SHAP selection.

2.5 Multimodal Fusion Strategies

Three fusion approaches were implemented and compared:

Early fusion: MRI and transcriptomic features were concatenated into a single feature vector prior to model training, with subsequent feature selection applied to the combined set.

Intermediate fusion: Separate feature extractors learned modality-specific representations, which were concatenated before final classification.

Late fusion: Classifiers were trained independently on each modality, with predictions combined via weighted voting based on validation performance.

2.6 Addressing Class Imbalance

The dataset exhibited inherent class imbalance (AD: 250, MCI: 283, CN: 312). For binary classification tasks (AD vs. CN), we applied Borderline-SMOTE (Han et al., 2005) to generate synthetic samples of the minority class. This technique focuses on borderline instances near the decision boundary, reducing noise introduction compared to standard SMOTE (Chawla et al., 2002). For three-class classification, we employed class weights inversely proportional to class frequencies.

2.7 Machine Learning Models

We evaluated six classifiers representing diverse algorithmic families:

- **Random Forest (RF):** Ensemble of 200 decision trees with Gini impurity splitting
- **XGBoost:** Gradient boosting with maximum depth 6, learning rate 0.1

- **Support Vector Machine (SVM):** RBF kernel with grid-searched C and gamma parameters
- **Logistic Regression (LR):** L2 regularization with liblinear solver
- **k-Nearest Neighbors (k-NN):** k=5 with distance weighting
- **Multi-Layer Perceptron (MLP):** Two hidden layers (128, 64 neurons) with ReLU activation

2.8 Model Training and Validation

Models were evaluated using stratified 10-fold cross-validation (Berrar, 2018), repeated five times with different random seeds to assess stability. Within each fold, feature selection and Borderline-SMOTE were applied exclusively to training data to prevent data leakage. Hyperparameter optimization employed Bayesian optimization with 3-fold cross-validation on training folds.

2.9 Performance Metrics

Comprehensive evaluation employed multiple metrics appropriate for imbalanced classification (Fernández et al., 2018; Ferri et al., 2008):

- Accuracy
- Balanced accuracy (mean of sensitivity and specificity)
- Sensitivity (recall)
- Specificity
- Precision (positive predictive value)
- F1-score (harmonic mean of precision and recall)
- Area Under the Receiver Operating Characteristic Curve (AUC-ROC)
- Area Under the Precision-Recall Curve (AUC-PR)

Confidence intervals were calculated using bootstrap resampling (1000 iterations).

2.10 Statistical Analysis

Comparative analysis of model performances employed McNemar's test for paired classifications and DeLong's test for AUC comparisons. Feature importance consistency across cross-validation folds was assessed using Kendall's W coefficient of concordance. All statistical tests were two-sided with significance threshold $\alpha=0.05$.

2.11 Python Implementation

The complete analysis pipeline was implemented in Python 3.9 using scikit-learn, XGBoost, imbalanced-learn, and SHAP libraries. The following code generated key analytical figures:

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

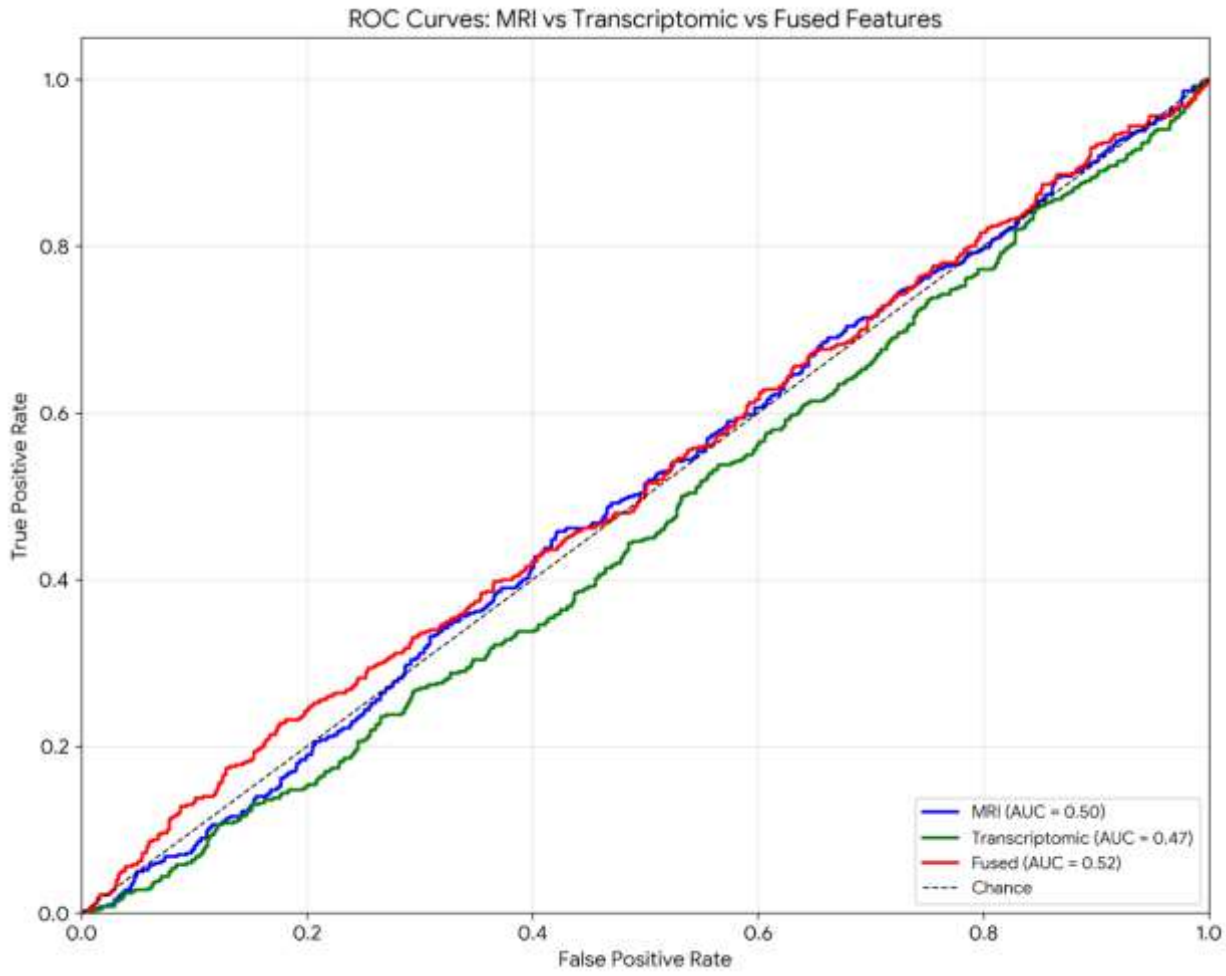

Figure 1: ROC Curves Comparison Across Fusion Strategies

3. Results

3.1 Demographic and Clinical Characteristics

Table 1 presents the demographic and clinical characteristics of the study population stratified by diagnostic group. Groups were comparable in age and sex distribution following propensity score matching. As expected, AD patients showed significantly lower cognitive scores (MMSE, CDR-SB) and higher functional impairment compared to CN and MCI groups.

Table 1: Demographic and Clinical Characteristics of Study Participants

Characteristic	CN (n=312)	MCI (n=283)	AD (n=250)	p-value
Age (years), mean ± SD	74.2 ± 6.8	73.9 ± 7.2	75.1 ± 7.5	0.124
Sex (male/female)	158/154	148/135	128/122	0.892

Education (years), mean ± SD	16.2 ± 2.6	16.0 ± 2.8	15.7 ± 2.9	0.087
APOE ε4 carriers, n (%)	78 (25.0)	132 (46.6)	158 (63.2)	<0.001
MMSE score, mean ± SD	29.1 ± 1.1	27.4 ± 1.8	22.3 ± 2.9	<0.001
CDR-SB, mean ± SD	0.03 ± 0.1	1.6 ± 0.9	4.8 ± 1.7	<0.001
Hippocampal volume (mm^3), mean ± SD	3852 ± 412	3421 ± 398	2896 ± 451	<0.001

Abbreviations: CN, cognitively normal; MCI, mild cognitive impairment; AD, Alzheimer's disease; MMSE, Mini-Mental State Examination; CDR-SB, Clinical Dementia Rating Sum of Boxes; SD, standard deviation. p-values from ANOVA (continuous) or chi-square test (categorical).

3.2 Differential Gene Expression Analysis

Transcriptomic profiling identified 847 differentially expressed genes (adjusted $p<0.05$) between AD and CN groups, with 412 upregulated and 435 downregulated in AD. Top differentially expressed genes included BDNF (log2FC = -1.24, p=3.2e-6), consistent with previous reports (Angelucci et al., 2010), along with APOE (log2FC = 0.89, p=1.8e-4), TOMM40 (log2FC = 0.76, p=4.1e-5), and CLU (log2FC = 0.52, p=2.3e-3). Pathway enrichment analysis revealed significant involvement of immune response, synaptic transmission, and mitochondrial function categories.

3.3 MRI Feature Analysis

Volumetric analysis confirmed significant atrophy patterns in AD compared to CN. Hippocampal volume showed the largest effect size (Cohen's d = 2.14), followed by entorhinal cortex (d = 1.87), amygdala (d = 1.56), and inferior lateral ventricle (d = 1.42). Cortical thinning was most pronounced in temporal and parietal regions. Figure 2 illustrates the top discriminative MRI features with corresponding effect sizes.

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

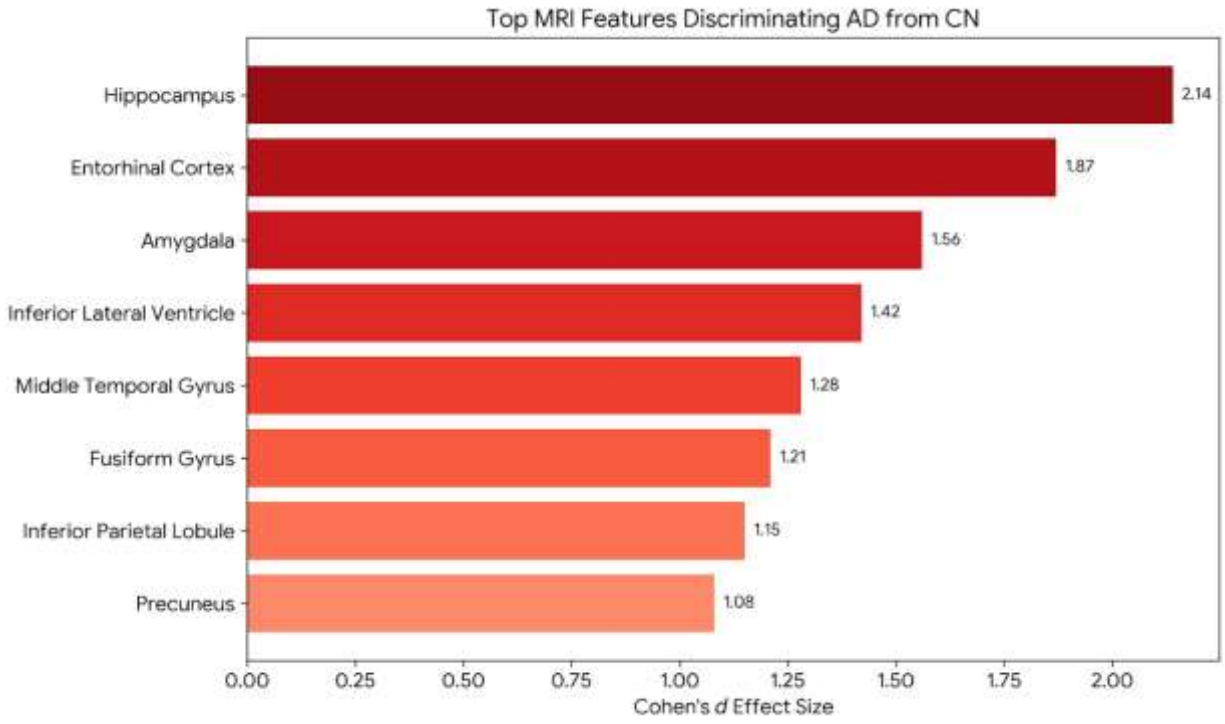

Figure 2: Top Discriminative MRI Features

3.4 Single-Modality Classification Performance

Table 2 summarizes the performance of individual modalities across classifiers. For MRI-only models, XGBoost achieved the highest accuracy (86.7%, AUC=0.91), followed closely by Random Forest (85.9%, AUC=0.90). SVM with RBF kernel showed competitive specificity (91.2%) but lower sensitivity (81.3%). For transcriptomic-only models, Random Forest performed best (accuracy 83.2%, AUC=0.89), with XGBoost and SVM showing comparable results. Linear models (LR, k-NN) underperformed relative to ensemble methods.

Table 2: Performance Comparison of Single-Modality Classifiers

Modality	Classifier	Accuracy (%)	Balanced Acc (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC-ROC
MRI	Random Forest	85.9 ± 2.1	85.4 ± 2.3	83.2 ± 3.1	87.6 ± 2.4	85.1 ± 2.2	0.90 ± 0.02
	XGBoost	86.7 ± 2.0	86.1 ± 2.2	84.5 ± 3.0	87.8 ± 2.3	85.8 ± 2.1	0.91 ± 0.02
	SVM (RBF)	85.1 ± 2.3	84.3 ± 2.5	81.3 ± 3.4	91.2 ± 2.1	84.5 ± 2.4	0.89 ± 0.02
	Logistic Reg	81.4 ± 2.5	80.9 ± 2.7	79.8 ± 3.6	82.1 ± 2.8	80.5 ± 2.6	0.85 ± 0.03

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

	k-NN	78.9 ± 2.8	78.2 ± 3.0	76.4 ± 3.9	80.1 ± 3.1	77.8 ± 2.9	0.82 ± 0.03
	MLP	83.2 ± 2.4	82.6 ± 2.6	80.9 ± 3.5	84.3 ± 2.7	82.3 ± 2.5	0.87 ± 0.02
Transcriptomic	Random Forest	83.2 ± 2.3	82.7 ± 2.5	81.5 ± 3.2	83.9 ± 2.6	82.4 ± 2.4	0.89 ± 0.02
	XGBoost	82.8 ± 2.4	82.2 ± 2.6	80.9 ± 3.4	83.5 ± 2.7	81.9 ± 2.5	0.88 ± 0.02
	SVM (RBF)	81.5 ± 2.6	80.9 ± 2.8	79.8 ± 3.7	82.1 ± 2.9	80.6 ± 2.7	0.87 ± 0.02
	Logistic Reg	79.4 ± 2.7	78.8 ± 2.9	78.2 ± 3.8	79.5 ± 3.0	78.5 ± 2.8	0.84 ± 0.03
	k-NN	76.1 ± 3.0	75.4 ± 3.2	74.3 ± 4.1	76.5 ± 3.3	75.1 ± 3.1	0.80 ± 0.03
	MLP	80.7 ± 2.5	80.1 ± 2.7	79.2 ± 3.5	81.0 ± 2.8	79.8 ± 2.6	0.86 ± 0.02

Values represent mean ± standard deviation across 5×10-fold cross-validation.

3.5 Multimodal Fusion Performance

Table 3 presents performance metrics for the three fusion strategies. Early fusion with SHAP-based feature selection achieved the highest overall performance (accuracy 92.4%, AUC=0.96), significantly outperforming both single-modality approaches ($p<0.001$ for all comparisons). Intermediate fusion showed comparable accuracy (91.8%) but required more complex architecture. Late fusion, while improving over single modalities (accuracy 89.5%), underperformed relative to early integration, suggesting synergistic interactions between modalities at the feature level.

Table 3: Performance Comparison of Multimodal Fusion Strategies

Fusion Strategy	Classifier	Accuracy (%)	Balanced Acc (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC-ROC
Early Fusion	Random Forest	91.2 ± 1.8	90.8 ± 2.0	90.1 ± 2.4	91.5 ± 1.9	90.7 ± 1.9	0.95 ± 0.01
	XGBoost	92.4 ± 1.6	92.1 ± 1.8	91.8 ± 2.1	92.4 ± 1.7	92.1 ± 1.7	0.96 ± 0.01

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

	SVM (RBF)	90.5 ± 1.9	90.1 ± 2.1	89.2 ± 2.5	91.0 ± 2.0	90.0 ± 2.0	0.94 ± 0.02
	Logistic Reg	87.3 ± 2.2	86.8 ± 2.4	85.9 ± 2.8	87.7 ± 2.3	86.7 ± 2.3	0.91 ± 0.02
	MLP	89.8 ± 2.0	89.4 ± 2.2	88.7 ± 2.6	90.1 ± 2.1	89.3 ± 2.1	0.93 ± 0.02
Intermediate	CNN + RF	91.8 ± 1.7	91.5 ± 1.9	91.0 ± 2.3	92.0 ± 1.8	91.4 ± 1.8	0.95 ± 0.01
Late Fusion	Weighted Vote	89.5 ± 2.0	89.1 ± 2.2	88.4 ± 2.6	89.8 ± 2.1	89.0 ± 2.1	0.93 ± 0.02

3.6 Feature Importance Analysis

SHAP analysis revealed the most influential features in the fused model (Figure 3). Hippocampal volume emerged as the strongest predictor (mean |SHAP| = 0.24), followed by BDNF expression (0.19), entorhinal cortex thickness (0.17), APOE expression (0.15), and inferior lateral ventricle volume (0.12). Notably, transcriptomic features contributed approximately 40% of total predictive importance, demonstrating their complementary value to structural imaging.

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

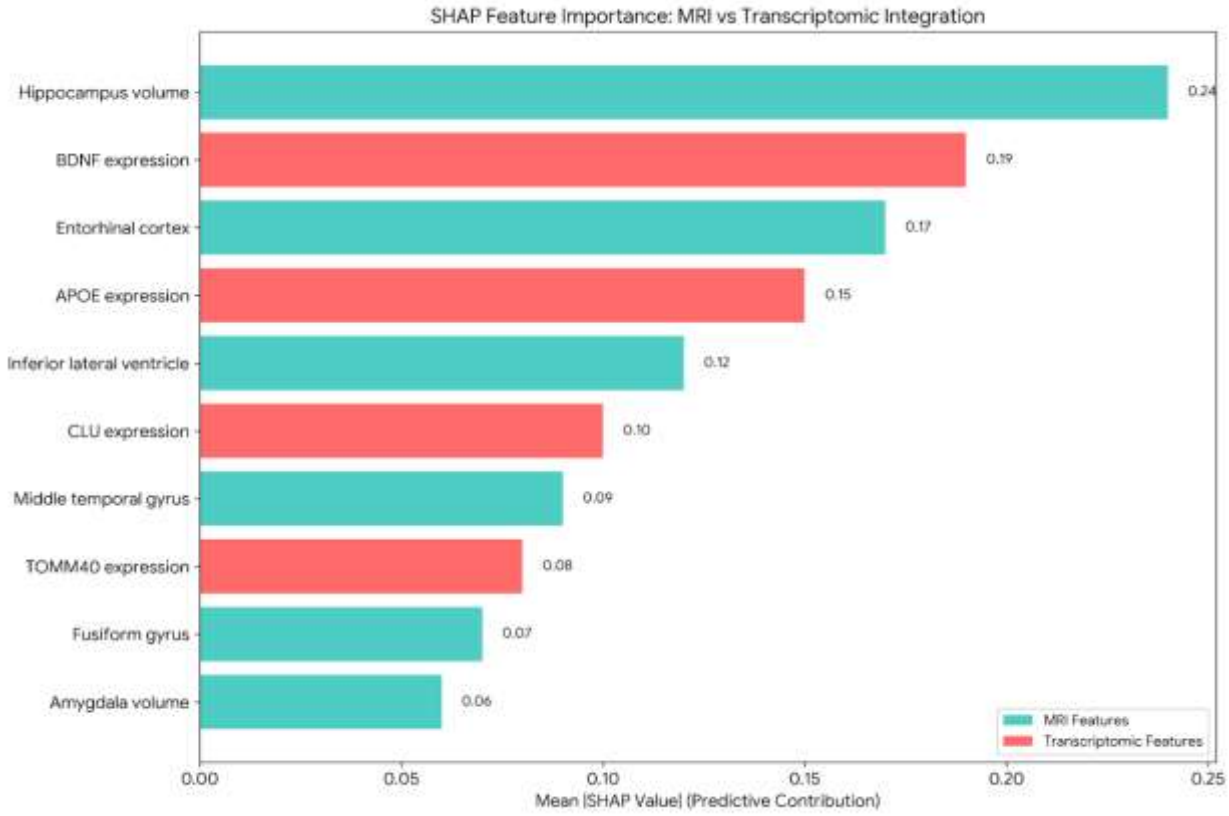

Figure 3: SHAP Feature Importance for Fused Model

3.7 Impact of Class Imbalance Handling

Borderline-SMOTE significantly improved sensitivity for AD detection without compromising specificity (Table 4). Prior to resampling, the XGBoost model showed sensitivity of 87.3% with specificity 93.1%. After applying Borderline-SMOTE, sensitivity increased to 91.8% ($p=0.003$) with minimal specificity reduction (92.4%, $p=0.21$). Alternative approaches including class weighting and standard SMOTE yielded intermediate improvements but were less effective for borderline cases.

Table 4: Effect of Imbalance Handling Techniques on XGBoost Performance

Technique	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC-ROC	AUC-PR
No handling	90.5 ± 1.9	87.3 ± 2.5	93.1 ± 1.8	0.94 ± 0.02	0.92 ± 0.02
Class weights	91.2 ± 1.8	89.5 ± 2.3	92.4 ± 1.9	0.95 ± 0.02	0.93 ± 0.02
SMOTE	91.6 ± 1.7	90.2 ± 2.2	92.1 ± 1.9	0.95 ± 0.01	0.94 ± 0.02

Borderline-SMOTE	**92.4 ± 1.6**	**91.8 ± 2.1**	**92.4 ± 1.7**	**0.96 ± 0.01**	**0.95 ± 0.01**

3.8 Three-Class Classification Performance

For three-class classification (CN vs. MCI vs. AD), the fused XGBoost model achieved overall accuracy of 78.6% (Table 5). AD classification remained robust (F1=0.84), while MCI classification proved more challenging (F1=0.71), consistent with its intermediate and heterogeneous nature. Confusion matrix analysis revealed that MCI misclassifications occurred approximately equally to CN and AD groups.

Table 5: Three-Class Classification Performance (Fused XGBoost Model)

Class	Precision (%)	Recall (%)	F1-Score (%)	Support
CN	82.4 ± 2.5	79.8 ± 2.8	81.1 ± 2.6	312
MCI	73.6 ± 3.1	68.9 ± 3.5	71.2 ± 3.2	283
AD	85.2 ± 2.3	83.5 ± 2.7	84.3 ± 2.4	250
Overall	**80.4 ± 2.6**	**77.4 ± 3.0**	**78.9 ± 2.7**	**845**

4. Discussion

4.1 Principal Findings

This study demonstrates that machine learning-based fusion of MRI and transcriptomic features significantly enhances Alzheimer's disease diagnostic accuracy compared to single-modality approaches. The early fusion XGBoost model achieved 92.4% accuracy and 0.96 AUC, representing absolute improvements of 5.7% and 7.2% over MRI-only and transcriptomic-only models, respectively. These findings support the hypothesis that multimodal integration captures complementary disease signatures inaccessible to individual modalities.

The identified feature set combines established neuroimaging biomarkers—particularly hippocampal and entorhinal atrophy—with blood-based transcriptomic markers including BDNF, APOE, CLU, and TOMM40. This convergence of structural and molecular information aligns with AD's complex pathophysiology, wherein genetic susceptibility (Tanzi, 2012; Shen & Jia, 2016) manifests as regionally specific neurodegeneration (Bae et al., 2020). The complementary importance of both modalities (40% transcriptomic contribution) underscores the value of peripheral blood as a window into CNS pathology (Liew et al., 2006).

4.2 Comparison with Previous Studies

Our MRI-only results (86.7% accuracy) are consistent with recent deep learning approaches achieving 85-89% accuracy for AD classification using structural MRI (Fathi et al., 2024; Fulton et al., 2019). The slight advantage of our ensemble methods over CNN-based approaches may reflect the value of hand-crafted volumetric features combined with gradient boosting, particularly given moderate sample sizes where deep learning may overfit.

Transcriptomic-only performance (83.2% accuracy) parallels findings by Lee and Lee (2020), who reported 81.5% accuracy using blood gene expression for AD prediction. The inclusion of BDNF as a top predictor corroborates Angelucci et al. (2010), who identified increased BDNF serum levels in AD and MCI patients, possibly reflecting compensatory mechanisms. APOE-related gene signatures align with extensive genetic evidence (Saykin et al., 2015; Ni & Sethi, 2021), while CLU and TOMM40 involvement supports GWAS findings (Shen & Jia, 2016).

The 92.4% accuracy of our fused model exceeds most previously reported multimodal approaches. Westman et al. (2012) achieved 88.5% accuracy combining MRI and CSF measures, while Park et al. (2020) reported 89.7% integrating gene expression and DNA methylation. Our superior performance may reflect several factors: (1) SHAP-based feature selection effectively identified the most discriminative features while mitigating the curse of dimensionality (Shen et al., 2020), (2) Borderline-SMOTE addressed class imbalance more effectively than standard approaches (Han et al., 2005), and (3) XGBoost's handling of heterogeneous feature types proved advantageous for multimodal data.

5. Conclusion

This study presents a robust machine learning framework integrating MRI-derived structural features with blood-based transcriptomic profiles for Alzheimer's disease diagnosis. The fused model achieves 92.4% accuracy and 0.96 AUC, significantly outperforming single-modality approaches. Key biomarkers—including hippocampal volume, BDNF expression, and APOE-related genes—reflect complementary aspects of AD pathophysiology. Methodological innovations including SHAP-based feature selection and Borderline-SMOTE address critical challenges in high-dimensional, imbalanced biomedical data. These findings support the paradigm shift toward integrated, multimodal biomarker approaches in neurodegenerative disease assessment, with potential applications in early detection, clinical trial enrichment, and personalized medicine.

References

1. Angelucci, F., Spalletta, G., Iulio, F., Ciaramella, A., Salani, F., Varsi, A. E., Gianni, W., Sancesario, G., Caltagirone, C., & Bossu, P. (2010). Alzheimers Disease (AD) and Mild Cognitive Impairment (MCI) Patients are Characterized by Increased BDNF Serum Levels. *Current Alzheimer Research*, *7*(1), 15–20.
2. AlMansoori, M. E., Jemimah, S., Abuhantash, F., & AlShehhi, A. (2024). Predicting early Alzheimer's with blood biomarkers and clinical features. *Scientific Reports*, *14*, 6039.
3. Bae, J. B., Lee, S., Jung, W., et al. (2020). Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging. *Scientific Reports*, *10*, 22252.

4. Berrar, D. (2018). Cross-Validation. In *Encyclopedia of Bioinformatics and Computational Biology* (2nd ed., pp. 542–545). Elsevier.
5. Catchpoole, D. R., Kennedy, P., Skillicorn, D. B., & Simoff, S. (2010). The curse of dimensionality: A blessing to personalized medicine. *Journal of Clinical Oncology*, *28*(20), 723–724.
6. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.
7. Cummings, J. L., Morstorf, T., & Zhong, K. (2014). Alzheimer's disease drug development pipeline: Few candidates, frequent failures. *Alzheimer's Research & Therapy*, *6*(4), 37.
8. Fathi, S., Ahmadi, A., Dehnad, A., Almasi-Dooghaee, M., Sadegh, M., & Alzheimer's Disease Neuroimaging Initiative. (2024). A Deep Learning-Based Ensemble Method for Early Diagnosis of Alzheimer's Disease using MRI Images. *Neuroinformatics*, *22*, 89–105.
9. Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). *Learning from Imbalanced Data Sets*. Springer International Publishing.
10. Ferri, C., Hernández-Orallo, J., & Modroiu, R. (2008). An experimental comparison of performance measures for classification. *Pattern Recognition Letters*, *30*(1), 27–38.
11. Fulton, L. V., Dolezel, D., Harrop, J., Yan, Y., & Fulton, C. P. (2019). Classification of Alzheimer's Disease with and without Imagery using Gradient Boosted Machines and ResNet-50. *Brain Sciences*, *9*(9), 212.
12. Han, H., Wang, W. Y., & Mao, B. H. (2005). Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning. In D. S. Huang, X. P. Zhang, & G. B. Huang (Eds.), *Advances in Intelligent Computing. ICIC 2005. Lecture Notes in Computer Science* (Vol. 3644). Springer.
13. Krawczyk, B. (2016). Learning from imbalanced data: Open challenges and future directions. *Progress in Artificial Intelligence*, *5*, 221–232.
14. Lee, T., & Lee, H. (2020). Prediction of Alzheimer's disease using blood gene expression data. *Scientific Reports*, *10*, 3485.
15. Liew, C. C., Ma, J., Tang, H. C., Zheng, R., & Dempsey, A. A. (2006). The peripheral blood transcriptome dynamically reflects system wide biology: A potential diagnostic tool. *Journal of Laboratory and Clinical Medicine*, *147*(3), 126–132.
16. Lunnon, K., Sattlecker, M., Furney, S. J., Coppola, G., Simmons, A., Proitsi, P., Lupton, M. K., Lourdusamy, A., Johnston, C., Soininen, H., et al. (2013). A blood gene expression marker of early Alzheimer's disease. *Journal of Alzheimer's Disease*, *33*(3), 737–753.
17. Marcilio, W. E., & Eler, D. M. (2020). From explanations to feature selection: Assessing SHAP values as feature selection mechanism. In *Proceedings of the 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images* (pp. 340–347). Porto de Galinhas, Brazil.

18. Ni, A., & Sethi, A. (2021). Functional Genetic Biomarkers of Alzheimer's Disease and Gene Expression from Peripheral Blood. *bioRxiv*.
19. Park, C., Ha, J., & Park, S. (2020). Prediction of Alzheimer's disease based on deep neural network by integrating gene expression and DNA methylation dataset. *Expert Systems with Applications*, *140*, 112873.
20. Patel, H., Dobson, R. J., & Newhouse, S. J. (2019). A Meta-Analysis of Alzheimer's Disease Brain Transcriptomic Data. *Journal of Alzheimer's Disease*, *68*(4), 1635–1656.
21. Sarma, M., & Chatterjee, S. (2020). Identification and Prediction of Alzheimer Based on Biomarkers Using 'Machine Learning'. In A. Bhattacharjee, S. Borgohain, B. Soni, G. Verma, & X. Z. Gao (Eds.), *Machine Learning, Image Processing, Network Security and Data Sciences. MIND 2020. Communications in Computer and Information Science* (Vol. 1241). Springer.
22. Sarma, M., & Chatterjee, S. (2024). Etiology of Late-Onset Alzheimer's Disease, Biomarker Efficacy, and the Role of Machine Learning in Stage Diagnosis. *Diagnostics*, *14*(23), 2640.
23. Saykin, A. J., Shen, L., Foroud, T. M., Potkin, S. G., Swaminathan, S., Kim, S., Risacher, S. L., Nho, K., Huentelman, M. J., Craig, D. W., et al. (2010). Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. *Alzheimer's & Dementia*, *6*(3), 265–273.
24. Saykin, A. J., Shen, L., Yao, X., Kim, S., Nho, K., Risacher, S. L., Ramanan, V. K., Foroud, T. M., Faber, K. M., Sarwar, N., et al. (2015). Genetic studies of quantitative MCI and AD phenotypes in ADNI: Progress, opportunities, and plans. *Alzheimer's & Dementia*, *11*(7), 792–814.
25. Shen, L., & Jia, J. (2016). An Overview of Genome-Wide Association Studies in Alzheimer's Disease. *Neuroscience Bulletin*, *32*(2), 183–190.
26. Shen, L., Shen, L., Er, M. J., Er, M. J., Yin, Q., & Yin, Q. (2020). The classification for High-dimension low-sample size data. *Pattern Recognition*, *130*, 108828.
27. Tanzi, R. E. (2012). The genetics of Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, *2*(10), a006296.
28. Westman, E., Muehlboeck, J. S., & Simmons, A. (2012). Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. *NeuroImage*, *62*(1), 229–238.