MiGAD: Gut Microbiota-informed Genetic Surrogates for Multimodal Alzheimer's Disease Diagnosis

Zhichao Liang[1,2], Shilun Zhao[2], Fan Li[2], Zifeng Lian[2], Weilin Zhou[2], Wei Yan[2], Chi Kin Lam[1], Tao Tan[1] (✉), and Dinggang Shen[2,3,4] (✉)

[1] Faculty of Applied Sciences, Macao Polytechnic University, Macao, 999078, China
[2] School of Biomedical Engineering & State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai 201210, China
[3] Shanghai Clinical Research and Trial Center, Shanghai, 201210, China
[4] Shanghai United Imaging Intelligence Co., Ltd., Shanghai 200230, China
taotanjs@gmail.com, Dinggang.Shen@gmail.com

Abstract. Alzheimer's disease (AD) pathology begins years before clinical symptoms, making early and reliable diagnosis from a single modality inherently uncertain. Growing evidence links the gut microbiota to early AD through the gut-brain axis, but direct microbiome profiling is highly heterogeneous and temporally unstable, limiting its utility as a deployable screening biomarker. To bridge this gap, we propose a gut microbiota-informed genetic surrogate-driven multimodal framework for AD diagnosis (MiGAD), which incorporates gut-brain axis related genetic evidence and distills multimodal knowledge into an MRI-only model for deployable screening. Our framework comprises three stages. First, we perform microbiota-informed candidate variant selection by integrating multi-cohort microbiome genetic evidence and derive transferable subject-level genetic representations. Second, we perform aligned cross-modal token fusion with a hierarchical Transformer to integrate MRI, clinical tabular features, and genetic tokens for AD classification. Third, to handle missing peripheral modalities in real-world screening, we distill the multimodal teacher into an MRI-only student model. We evaluate our approach on 6,500 participants from UK Biobank and ADNI. The multimodal model achieves ACC=0.878 and AUC=0.963, while the distilled MRI-only student improves AUC from 0.825 to 0.910 over an MRI-only baseline. These results demonstrate that microbiota-informed genetic surrogates provide stable complementary signals for AD staging while preserving real-world usability.

Keywords: Alzheimer's disease · Microbiota-informed genetics · Multimodal learning · Knowledge distillation.

1 Introduction

Alzheimer's disease (AD) is a prevalent neurodegenerative disorder affecting millions worldwide [16]. As populations age, its clinical and societal burden continues to grow, while effective management remains highly dependent on timely

identification along the pre-dementia continuum. Because core AD pathology may emerge years before overt symptoms, symptom-based subtyping is often uncertain and too late for optimal intervention planning [6]. This creates an urgent clinical need for early, reliable, and scalable screening technologies for risk stratification and diagnostic staging in heterogeneous real-world populations. Although neuroimaging is central to AD assessment and imaging-based models have advanced substantially, AD is intrinsically multifactorial and influenced by peripheral organ function and systemic homeostasis [1, 11, 26]. Consequently, imaging-only models may miss biologically relevant extra-cerebral signals and generalize suboptimally across cohorts, motivating the integration of transferable peripheral mechanistic information for robust early AD screening.

Among brain-organ interaction hypotheses in AD, the gut-brain axis has emerged as a relevant peripheral pathway [19, 24]. Previous studies suggest that gut microbiota may be associated with AD-related processes through inflammatory signaling, metabolic dysregulation, and pathological molecular pathways [8, 20, 7]. However, direct microbiome profiles are strongly shaped by diet, medication, lifestyle, and environment, leading to substantial inter-individual heterogeneity and temporal variability that limit their reliability as stable screening biomarkers [13]. In contrast, host genetic variation is stable across the life course, and microbiome-related genetic evidence can provide an indirect, transferable source of gut-brain axis-related information without requiring direct microbiome profiling at deployment.

Deep learning for AD diagnosis has evolved from CNN-based unimodal models to Transformer-based multimodal fusion [14, 4, 21, 27, 25, 15, 5]. Prior work has improved robustness to incomplete inputs via modality synthesis or missing-modality modeling, but these approaches largely remain centered on brain-derived signals [21]. Other studies fuse imaging with genetics and clinical tabular data to improve discrimination [5, 14, 27, 25, 15], and recent work has begun to incorporate multi-organ evidence into AD modeling [11, 10]. Nevertheless, existing genetic inputs are often limited to generic AD-risk cues and rarely incorporate microbiota-informed genetic evidence in a transferable subject-level representation. Moreover, genetics and complete clinical tabular variables are often unavailable in frontline screening workflows, limiting the deployability of models that require full multimodal inputs.

To address these challenges, we propose a gut microbiota-informed genetic surrogate model for multimodal Alzheimer's disease diagnosis (MiGAD). Our model leverages multi-cohort microbiome genetics to derive transferable microbiota-linked candidate variants and subject-level genetic representations, fuses these surrogates with MRI and clinical features through a cross-modally aligned hierarchical Transformer, and distills multimodal knowledge into an MRI-only model for deployable screening when peripheral modalities are missing. Our contributions are summarized as follows: (1) We develop a microbiota-informed genetic surrogate construction pipeline that enables stable and transferable incorporation of gut-brain axis related genetic evidence into subject-level AD prediction. (2) We introduce an explicit alignment-and-fusion Transformer to improve

cross-modal consistency and robustness for multimodal AD classification. (3) We present a multimodal-to-MRI knowledge transfer strategy via distillation that reduces reliance on non-imaging inputs while retaining competitive performance for MRI-only inference.

2 Methodology

An overview of the MiGAD framework is shown in Fig. 1. It consists of three stages: (I) microbiota-informed candidate variant selection by integrating multi-cohort microbiome and AD genetic evidence through Mendelian randomization and colocalization; (II) aligned cross-modal token fusion for AD diagnosis, where MRI, clinical variables, and selected genotypes are encoded as modality tokens and fused for three-class classification of normal cognition (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD); and (III) multimodal-to-MRI knowledge distillation for deployable screening, where a teacher transfers knowledge to an MRI-only student under missing peripheral modalities.

2.1 Microbiota-informed candidate variant selection

In Stage I, we derive a microbiota-informed candidate variant set by integrating multi-cohort gut microbiome genome-wide association studies (GWAS) with AD-related GWAS summary statistics. Using two-sample Mendelian randomization with sensitivity analyses, we select microbiome traits showing robust associations with AD-related outcomes from the candidate trait set $\mathcal{M}$, yielding $\mathcal{M}^\star \subseteq \mathcal{M}$ [12]. For each trait $m \in \mathcal{M}^\star$, we apply Linkage Disequilibrium (LD) clumping to microbiome-associated instruments ($r^2 < r_0^2$) and obtain an LD-independent lead-variant set $\tilde{\mathcal{Z}}_m$.

For each lead variant $\tilde{z} \in \tilde{\mathcal{Z}}_m$, we perform colocalization between microbiome GWAS and AD-related GWAS within a local window and collect variants from the credible set under the shared-causal hypothesis (e.g., 95% posterior coverage) [12]. The final candidate set is

$$\mathcal{S} = \Big(\bigcup_{m \in \mathcal{M}^\star} \bigcup_{\tilde{z} \in \tilde{\mathcal{Z}}_m} \mathcal{C}(\tilde{z}) \Big) \cup \Big(\bigcup_{m \in \mathcal{M}^\star} \tilde{\mathcal{Z}}_m \Big), \tag{1}$$

where $\mathcal{C}(\tilde{z})$ denotes the locus-specific credible set from colocalization. If colocalization evidence is insufficient, we retain $\tilde{\mathcal{Z}}_m$ to preserve signal coverage. We then order $\mathcal{S} = \{s_k\}_{k=1}^K$ and extract additive allele dosages for subject i:

$$g_{i,k} = \mathrm{dosage}_i(s_k) \in \{0, 1, 2, \mathrm{NA}\}, \qquad k = 1, \ldots, K, \tag{2}$$

with allele harmonization across cohorts and explicit missingness preservation. This yields the genotype matrix $\mathbf{G} = [g_{i,k}]$, which is used as the genetic input to the multimodal model.

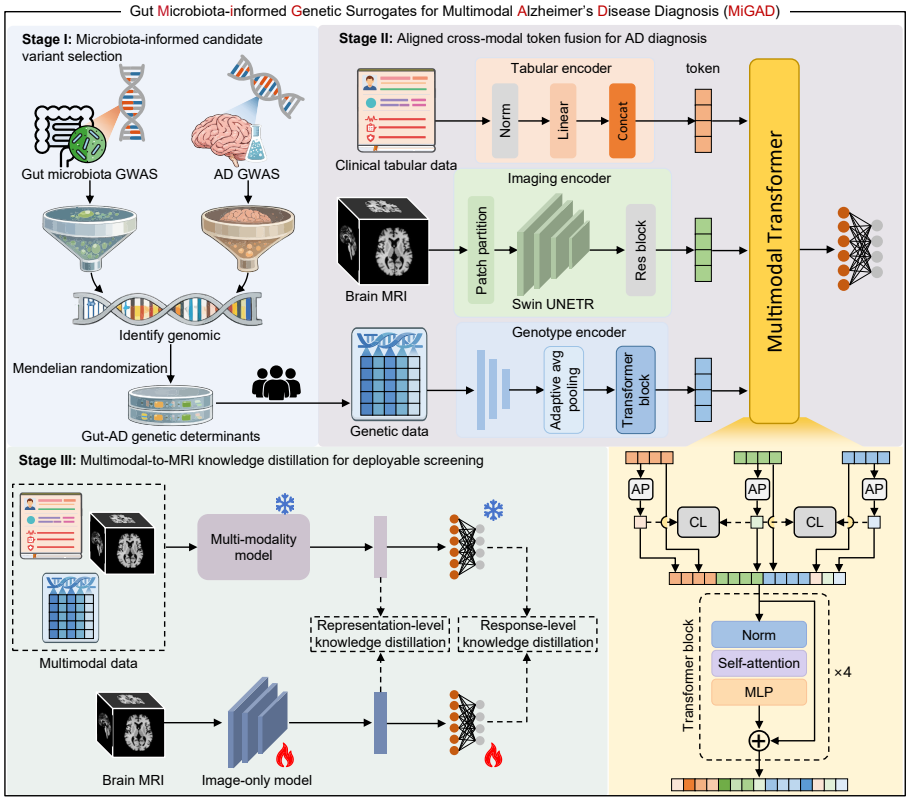

Fig. 1: Overview of the proposed gut microbiota-informed multimodal AD diagnosis framework. AP: attention pooling; CL: contrastive learning.

2.2 Aligned cross-modal token fusion for AD diagnosis

For each subject i, we use three inputs: a T1-weighted MRI volume, a clinical tabular feature vector, and genotype dosages at the microbiota-informed candidate variants $\mathcal{S} = \{s_k\}_{k=1}^{K}$. Each modality is encoded into a token sequence with a shared embedding dimension d, while missingness is explicitly represented.

Tabular tokens. Let $\mathbf{t}_i$ denote the tabular vector. Each tabular variable is embedded as a d-dimensional token (linear embedding for continuous features and learnable embeddings for categorical features), yielding $\mathbf{X}_i^{(\mathrm{tab})} \in R^{L_{\mathrm{tab}} \times d}$.

Imaging tokens. A pretrained Swin UNETR encoder extracts imaging features, which are flattened and projected to form image tokens $\mathbf{X}_i^{(\mathrm{img})} \in R^{L_{\mathrm{img}} \times d}$.

Genetic tokens. Each dosage $g_{i,k}$ is mapped to a value embedding and combined with a learnable variant identity embedding for s_k, followed by lightweight sequence compression (1D convolution + pooling) and a small Transformer to obtain a fixed-length token sequence $\mathbf{X}_i^{(\mathrm{geno})} \in R^{L_{\mathrm{geno}} \times d}$.

Modality pooling. Since token lengths differ across modalities and some modalities may be missing, we summarize each modality $m \in \{\text{img}, \text{tab}, \text{geno}\}$ using attention pooling with a learnable query $\mathbf{q}^{(m)}$:

$$\mathbf{z}_i^{(m)} = \sum_{\ell=1}^{L_m} a_{i,\ell}^{(m)} \mathbf{x}_{i,\ell}^{(m)}, \qquad a_{i,\ell}^{(m)} = \mathrm{softmax}_\ell\left(\frac{\langle \mathbf{x}_{i,\ell}^{(m)}, \mathbf{q}^{(m)} \rangle}{\sqrt{d}}\right), \tag{3}$$

where $\mathbf{X}_i^{(m)} = [\mathbf{x}_{i,1}^{(m)}, \ldots, \mathbf{x}_{i,L_m}^{(m)}]$ and L_m is the token length of modality m.

Cross-modal alignment. We enforce CLIP-style alignment on pooled representations by mapping each modality representation $\mathbf{z}_i^{(m)}$ into a shared alignment space using a modality-specific projection head $f_m(\cdot)$:

$$\mathbf{u}_i^{(m)} = \frac{f_m(\mathbf{z}_i^{(m)})}{\|f_m(\mathbf{z}_i^{(m)})\|_2}. \tag{4}$$

We then apply an InfoNCE objective using imaging as the anchor:

$$\mathcal{L}_{align} = \sum_{m \in \{\text{tab},\text{geno}\}} -\frac{1}{|\mathcal{B}_m|} \sum_{i \in \mathcal{B}_m} \log \frac{\exp(\mathbf{u}_i^{(\text{img})} \cdot \mathbf{u}_i^{(m)}/\tau)}{\sum_{j \in \mathcal{B}_m} \exp(\mathbf{u}_i^{(\text{img})} \cdot \mathbf{u}_j^{(m)}/\tau)}, \tag{5}$$

where τ is a temperature parameter and $\mathcal{B}_m$ is the set of samples in the batch for which modality m is available.

Multimodal Transformer. We concatenate a task token, the pooled modality tokens $\{\mathbf{z}_i^{(m)}\}$, and the modality token sequences $\{\mathbf{X}_i^{(m)}\}$, and feed them into a backbone Transformer to obtain a fused representation $\mathbf{h}_i$, followed by a linear classifier. The training objective is

$$\mathcal{L} = \mathcal{L}_{cls} + \lambda \mathcal{L}_{align}, \tag{6}$$

where $\mathcal{L}_{cls} = \mathrm{CE}(y_i, \hat{\mathbf{y}}_i)$ is the supervised cross-entropy loss and λ controls the alignment strength.

2.3 Multimodal-to-MRI knowledge distillation for deployable screening

In practical screening, genetic and complete tabular information are often unavailable. To enable deployable inference, we distill the trained multimodal model (teacher) into an MRI-only model (student). The teacher maps multimodal inputs to logits $\mathbf{o}^T$ and a fused representation $\mathbf{h}^T$ from the final fusion layer, while the student takes MRI only and outputs $\mathbf{o}^S$ and $\mathbf{h}^S$. During distillation, the teacher is frozen and only the student is optimized.

We train the student with a combination of (i) hard-label supervision, (ii) response-level distillation that matches the teacher's softened class distribution, and (iii) representation-level distillation that aligns student and teacher embeddings:

$$\mathcal{L}_{KD} = \alpha\, \mathrm{CE}(y, \mathbf{o}^S) + (1-\alpha)\, T^2\, \mathrm{KL}\big(\mathrm{softmax}(\mathbf{o}^T/T) \parallel \mathrm{softmax}(\mathbf{o}^S/T)\big) + \beta \|\mathbf{h}^S - \mathbf{h}^T\|_2^2, \tag{7}$$

where T is the distillation temperature, and α and β balance hard-label learning and embedding alignment.

3 Experimental

3.1 Experimental settings

Cohorts and data preprocessing. We evaluate MiGAD on UK Biobank (UKB) and the Alzheimer's Disease Neuroimaging Initiative (ADNI). After cohort-specific quality control and cross-modality matching, 6,500 participants were included (UKB: 4,141; ADNI: 2,359). All have T1-weighted MRI and harmonized clinical tabular features, and 4,794 additionally have genotypes at the selected variant set. Diagnostic labels were harmonized into three classes, with ADNI MCI defined by standard criteria and UKB labels derived via a reproducible adjudication pipeline [17, 18]. To avoid cohort imbalance caused by random pooling, UKB and ADNI were split separately into training, validation, and test sets with an 8:1:1 ratio, and the corresponding subsets were then combined for model training and evaluation. Clinical tabular features included demographic variables, cognitive and functional assessment scores, and routine clinical variables available in both cohorts, with continuous variables standardized using training-split statistics and categorical variables one-hot encoded. MRI volumes were standardized to a 128^3 space and encoded by a pretrained Swin UNETR, and genotypes were extracted as additive dosages with allele harmonization and missingness preservation. Cohort characteristics are summarized in Table 1.

Table 1: Study population and characteristics stratified by diagnosis.

Diagnosis	Age (y), mean±sd	Male, n(%)	Education (y), mean±sd
UKB			
NC [n=3790]	60.78±6.90	2467 (65.09%)	17.52±2.20
MCI [n=288]	68.39±7.27	111 (38.54%)	16.44±2.23
AD [n=63]	70.89±6.39	32 (50.79%)	16.49±1.94
ADNI			
NC [n=868]	72.51±6.60	382 (44.01%)	16.53±2.53
MCI [n=1081]	72.83±7.64	631 (58.37%)	15.98±2.76
AD [n=410]	74.94±7.77	230 (56.10%)	15.23±2.91

Microbiota-informed candidate variants. To capture gut-brain axis genetic signals without requiring direct microbiome profiling, we curate a microbiota-informed candidate variant set by integrating multi-cohort gut microbiome GWAS and AD-related GWAS summary statistics via the Stage I pipeline. This yields 12,634 candidate variants for subject-level genotype extraction. Fig. 2 contrasts representative top association signals from microbiome GWAS and AD GWAS.

Implementation and training settings. LD threshold r_0^2 is set to 0.1 [9]. All models are implemented in PyTorch and trained on a single NVIDIA A40

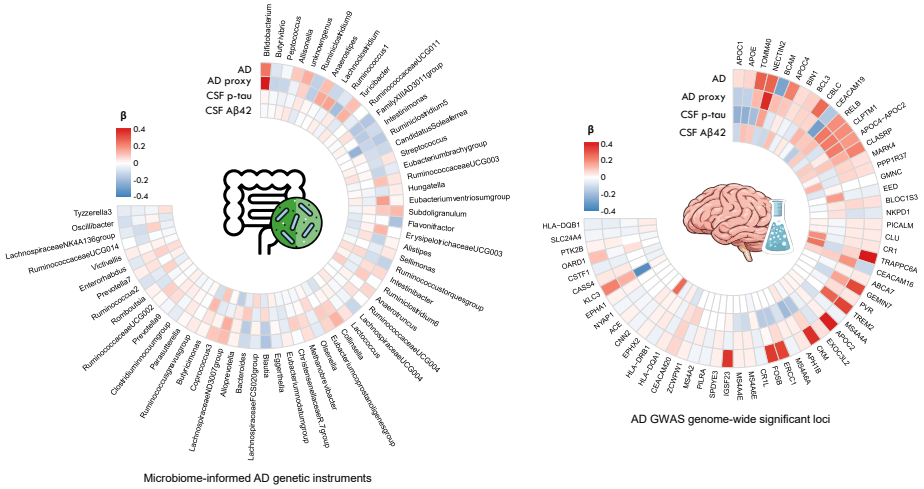

Fig. 2: Population-level genetic signals in gut microbiota GWAS (left) and AD GWAS (right), visualized by top 60 lead associations (colors denote signed standardized effect sizes).

GPU (46 GB). The multimodal teacher is optimized with AdamW (using initial learning rate 10^{-3}, batch size 64) for 200 epochs. For knowledge distillation, we use temperature $T = 2.0$ and weights $\alpha = 0.3$ and $\beta = 1.0$ to balance hard-label supervision, response-level distillation, and representation-level matching.

3.2 Experimental results

Performance of the Proposed Method. Fig. 3(a) summarizes three-class classification performance under different modality configurations. Among unimodal baselines, clinical tabular features are the most informative (AUC=0.935), whereas imaging-only and genetics-only models are less discriminative (AUC=0.825 and 0.789). Multimodal fusion consistently improves performance, and MiGAD achieves the best overall results (ACC=0.878, AUC=0.963), outperforming both bimodal variants. For deployable screening, the distilled MRI-only student substantially improves over the imaging-only baseline (AUC increases from 0.825 to 0.910). Beyond aggregate performance, Fig. 3(b-d) shows that MiGAD reduces clinically important stage-mixing between MCI and AD. Predictions for MCI and AD cases concentrate closer to their correct categories with less spillover to the neighboring stage, and our model lies nearer the desirable lower-left region in the clinical pareto view.

Ablation Study. Table 2 quantifies the contribution of each component. Replacing microbiota-informed variants with AD-risk variants consistently degrades performance, suggesting that microbiota-linked genetic signals provide additional gut-brain axis cues beyond canonical AD loci. Substituting the genotype token encoder with an MLP and removing cross-modal alignment further

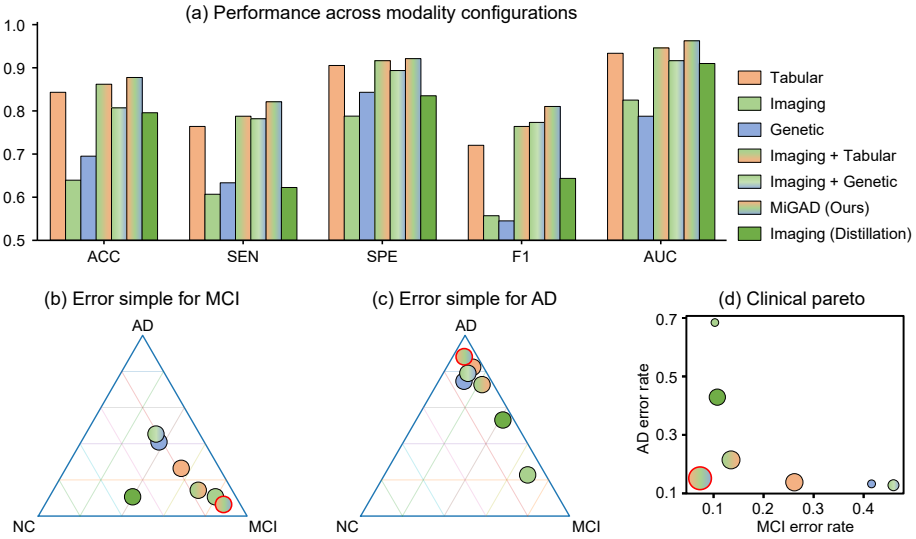

Fig. 3: Performance and clinically oriented error analysis. (a) Overall performance across modality settings. (b-c) Predicted-class compositions for MCI and AD cases. (d) Trade-off between MCI-to-AD and AD-to-MCI confusions. Marker area is proportional to AUC.

reduces accuracy, supporting the value of structure-aware tokenization and explicit feature alignment. Replacing the Transformer fusion backbone with a 3D ResNet yields the largest drop, highlighting the importance of global token interactions for multimodal integration.

Comparison with Other Multimodal Methods. We further compare MiGAD with classical baselines (SVM [2] and DL+RF [22]) and representative multimodal fusion methods (HyperFusion [3], DAFT [23], and MADDi [5]) under identical data splits and modality availability. As shown in Table 3, MiGAD achieves the best overall performance (ACC=0.878, AUC=0.963), outperforming HyperFusion (AUC=0.918), DAFT (AUC=0.935), MADDi (AUC=0.945), as well as SVM (AUC=0.755) and DL+RF (AUC=0.892). We attribute these gains to our alignment-regularized Transformer fusion, which can better integrate complementary signals and yield more robust prediction.

4 Conclusion

In this work, we propose a microbiota-informed multimodal framework for AD diagnosis by fusing brain MRI, clinical phenotypes, and microbiota-related host genetic surrogate signals with cross-modal alignment. On 6,500 UKB and ADNI participants, our model achieves the best overall performance and consistently outperforms unimodal and representative multimodal baselines. We further distill the multimodal teacher into an MRI-only student, retaining strong accuracy

Table 2: Ablation study on NC *vs.* MCI *vs.* AD classification. MV: microbiota-informed variants; GE: genotype token encoder; AL: contrastive alignment; TB: Transformer backbone.

Components				Performance	
MV	**GE**	**AL**	**TB**	**ACC**	**AUC**
✗	✓	✓	✓	0.862	0.940
✓	✗	✓	✓	0.865	0.937
✓	✓	✗	✓	0.875	0.947
✓	✓	✓	✗	0.842	0.916
✓	✓	✓	✓	**0.878**	**0.963**

Table 3: Comparison with representative multimodal methods on NC *vs.* MCI *vs.* AD classification under identical data splits and missing-modality handling.

Method	**ACC**	**F1**	**AUC**
SVM	0.674	0.704	0.755
DL+RF	0.816	0.710	0.892
Hyperfusion	0.839	0.706	0.918
DAFT	0.842	0.750	0.935
MADDi	0.866	0.767	0.945
MiGAD	**0.878**	**0.812**	**0.963**

when peripheral modalities are unavailable, highlighting a practical route toward scalable and deployable screening.

Acknowledgments. This work was supported in part by National Natural Science Foundation of China (grant numbers U23A20295, 62131015, 82441023, 82394432), the Shenzhen Medical Research Fund (D2501013), the China Ministry of Science and Technology (STI2030-Major Projects-2022ZD0209000, STI2030-Major Projects-2022ZD0213100), The Key R&D Program of Guangdong Province, China (grant number 2023B0303040001), and HPC Platform of ShanghaiTech University. This study was also supported by Macao Polytechnic University Grant (RP/FCA-17/2025), and the Science and Technology Development Fund of Macao (0041/2023/RIB2).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Alzheimer's Association: 2013 alzheimer's disease facts and figures. Alzheimer's & Dementia **9**(2), 208–245 (Mar 2013). https://doi.org/10.1016/j.jalz.2013.02.003, https://pubmed.ncbi.nlm.nih.gov/23507120/
2. Bucholc, M., Ding, X., Wang, H., Glass, D.H., Wang, H., Prasad, G., Maguire, L.P., Bjourson, A.J., McClean, P.L., Todd, S., et al.: A practical computerized decision support system for predicting the severity of alzheimer's disease of an individual. Expert systems with applications **130**, 157–171 (2019)
3. Duenias, D., Nichyporuk, B., Arbel, T., Riklin Raviv, T.: Hyperfusion: A hypernetwork approach to multimodal integration of tabular and medical imaging data for predictive modeling. Medical Image Analysis **102**, 103503 (2025)
4. Fan, J., Cao, X., Wang, Q., Yap, P.T., Shen, D.: Adversarial learning for mono-or multi-modal registration. Medical image analysis **58**, 101545 (2019)
5. Golovanevsky, M., Eickhoff, C., Singh, R.: Multimodal attention-based deep learning for alzheimer's disease diagnosis. Journal of the American Medical Informatics Association **29**(12), 2014–2022 (2022)
6. Iddi, S., Li, D., Aisen, P.S., Rafii, M.S., Thompson, W.K., Donohue, M.C.: Predicting the course of alzheimer's progression. Brain Informatics **6**, 6 (Jun 2019)

7. Jin, J., Xu, Z., Zhang, L., Zhang, C., Zhao, X., Mao, Y., Zhang, H., Liang, X., Wu, J., Yang, Y., et al.: Gut-derived β-amyloid: Likely a centerpiece of the gut–brain axis contributing to alzheimer's pathogenesis. Gut microbes **15**(1), 2167172 (2023)
8. Kleniuk, J., Edison, P.: Association of cardiovascular risk and sex with cortical tau deposition in cognitively normal and at-risk alzheimer's subjects: Neuroimaging/new imaging methods. Alzheimer's & Dementia **16**, e046050 (2020)
9. Kurilshikov, A., Medina-Gomez, C., Bacigalupe, R., Radjabzadeh, D., Wang, J., Demirkan, A., Le Roy, C.I., Raygoza Garay, J.A., Finnicum, C.T., Liu, X., et al.: Large-scale association analyses identify host factors influencing human gut microbiome composition. Nature genetics **53**(2), 156–165 (2021)
10. Li, F., Bai, S., Liu, Y., Chen, Z., Zhao, S., Ding, Z., Xie, F., Xu, Y., Yue, L., Zhang, H., Zhang, Y., Sun, K., Shen, D.: Enhancing diagnosis of mild cognitive impairment through brain-heart-gut metabolic networks in whole-body pet imaging. Cell Reports Medicine **7**(2) (2026). https://doi.org/10.1016/j.xcrm.2026.102629, https://doi.org/10.1016/j.xcrm.2026.102629
11. Li, F., Zhao, S., Bai, S., Liu, Y., Zhang, K., Xu, Y., Zhang, Y., Sun, K., Shen, D.: Brain-heart-gut guided multi-constraint knowledge distillation for early alzheimer's disease diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 45–54. Springer (2025)
12. Li, J., Yuan, Z., Li, J., Liu, Z., Wang, Y., Cui, M., Suo, C., Jin, L., Ding, D., Chen, X., et al.: Immune, blood-brain barrier, and metabolic biomarkers mediate gut-brain axis crosstalk in alzheimer's disease. Biomarker Research **13**(1), 137 (2025)
13. Liang, Z., Li, F., Jia, D., Duan, Y., Xie, X., Sun, K., Cui, Z., Tan, T., Shen, D.: Tree-diffusion: Octree-based conditional diffusion model for small bowel skeleton generation with geometric direction modeling. In: 2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 2483–2488. IEEE (2025)
14. Liu, M., Li, F., Yan, H., Wang, K., Ma, Y., Shen, L., Xu, M., Initiative, A.D.N., et al.: A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in alzheimer's disease. Neuroimage **208**, 116459 (2020)
15. Liu, M., Zhang, D., Adeli, E., Shen, D.: Inherent structure-based multiview learning with multitemplate feature representation for alzheimer's disease diagnosis. IEEE Transactions on Biomedical Engineering **63**(7), 1473–1482 (2015)
16. Naqvi, E.: Alzheimer's disease statistics. Alzheimer's News Today (Jun 2017), https://alzheimersnewstoday.com/alzheimers-disease-statistics/, accessed: 2026-02-12
17. Petersen, R.C.: Mild cognitive impairment. CONTINUUM: lifelong Learning in Neurology **22**(2), 404–418 (2016)
18. Ritchie, C., Trepel, D., Edwards, S., Hahn-Pedersen, J.H., Chan, M.S., Bray, B.D., Clark, A., Wichmann, C.A., Evans, M.: The impact of mild cognitive impairment on healthcare utilization and costs: A uk biobank study. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring **17**(2), e70065 (2025)
19. Song, R., Pan, K.Y., Xu, H., Qi, X., Buchman, A.S., Bennett, D.A., Xu, W.: Association of cardiovascular risk burden with risk of dementia and brain pathologies: a population-based cohort study. Alzheimer's & Dementia **17**(12), 1914–1922 (2021)
20. Stakos, D.A., Stamatelopoulos, K., Bampatsias, D., Sachse, M., Zormpas, E., Vlachogiannis, N.I., Tual-Chalot, S., Stellos, K.: The alzheimer's disease amyloid-beta hypothesis in cardiovascular aging and disease: Jacc focus seminar. Journal of the American College of Cardiology **75**(8), 952–967 (2020)

21. Sun, K., Zhang, Y., Liu, J., Yu, L., Zhou, Y., Xie, F., Guo, Q., Zhang, H., Wang, Q., Shen, D.: Achieving multi-modal brain disease diagnosis performance using only single-modal images through generative ai. Communications Engineering **3**(1), 96 (2024)
22. Venugopalan, J., Tong, L., Hassanzadeh, H.R., Wang, M.D.: Multimodal deep learning models for early detection of alzheimer's disease stage. Scientific reports **11**(1), 3254 (2021)
23. Wolf, T.N., Pölsterl, S., Wachinger, C.: DAFT: A Universal Module to Interweave Tabular Data and 3D Images in CNNs. NeuroImage p. 119505 (2022). https://doi.org/10.1016/j.neuroimage.2022.119505, https://www.sciencedirect.com/science/article/pii/S1053811922006218
24. Xiang, J., Tang, J., Kang, F., Ye, J., Cui, Y., Zhang, Z., Wang, J., Wu, S., Ye, K.: Gut-induced alpha-synuclein and tau propagation initiate parkinson's and alzheimer's disease co-pathology and behavior impairments. Neuron **112**(21), 3585–3601 (2024)
25. Xue, C., Kowshik, S.S., Lteif, D., Puducheri, S., Jasodanand, V.H., Zhou, O.T., Walia, A.S., Guney, O.B., Zhang, J.D., Poésy, S., et al.: Ai-based differential diagnosis of dementia etiologies on multimodal data. Nature Medicine **30**(10), 2977–2989 (2024)
26. Zhang, J., Zhang, Y., Wang, J., Xia, Y., Zhang, J., Chen, L.: Recent advances in alzheimer's disease: Mechanisms, clinical trials and new drug development strategies. Signal transduction and targeted therapy **9**(1), 211 (2024)
27. Zhang, Y., Sun, K., Liu, Y., Xie, F., Guo, Q., Shen, D.: A modality-flexible framework for alzheimer's disease diagnosis following clinical routine. IEEE Journal of Biomedical and Health Informatics (2024)