S^{2}GAIN: Spatio-Spectral Guided Affinity Imputation Network for Alzheimer's Disease Diagnosis with Incomplete Modality

Yanan Wu[1], Yihong Dong[1]*, and Yuehan Wu[1]

Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China
dongyihong@nbu.edu.cn

Abstract. Multimodal neuroimaging provides complementary information for early computer-aided diagnosis of Alzheimer's disease (AD). However, in clinical practice, modalities such as PET are often missing due to cost and accessibility constraints, hindering stable cross-modal inference. Existing approaches either synthesize the missing modality in the image domain, which may introduce artifacts and noise that interfere with classification, or perform feature-domain alignment and completion while emphasizing spatial consistency, making it difficult to handle MRI-PET heterogeneity in texture patterns and frequency components. Moreover, forcing unpredictable cross-modal components to fit can inject noise and degrade diagnosis. To address these issues, this paper proposes the Spatio-Spectral Guided Affinity Imputation Network (S^{2}GAIN), which employs a cascaded dual-stream cross-modal interaction backbone and multi-stage lightweight imputers to infer missing PET features layer by layer. A spatio-spectral joint alignment strategy is introduced to apply contrastive constraints in both the spatial domain and the frequency-domain magnitude spectrum, improving heterogeneous representation alignment. An affinity-guided reliability-weighting mechanism is further designed to suppress noise propagation from uninferable components. Finally, image-domain reconstruction regularization is employed to improve consistency between imputed metabolic representations and PET. Experimental results on ADNI demonstrate that S^{2}GAIN achieves superior and consistent performance under missing-modality settings.

Keywords: Incomplete Multimodal Learning · Missing Modality · Spatio-Spectral Alignment · Affinity-guided Reliability · Alzheimer's Disease.

1 Introduction

Alzheimer's disease (AD) is a major cause of dementia and is typically characterized by progressive memory decline and cognitive impairment [1]. As effective curative treatments remain unavailable, early identification of cognitive status

* Corresponding author.

is crucial for timely intervention and delayed disease progression [17, 3]. Neuroimaging modalities such as MRI and FDG-PET, which noninvasively reflect structural atrophy and metabolic abnormalities in the brain, have become key evidence for computer-aided diagnosis [7, 18]. Prior studies have shown that multimodal imaging fusion can exploit complementary pathological information and improve diagnostic performance compared with a single modality [14].

However, integrating multimodal neuroimaging data for early AD detection remains highly challenging in clinical practice. Although PET imaging can provide valuable information for AD diagnosis, its high cost and the need for specialized facilities and personnel mean that many patients cannot undergo PET scans [13]. In addition, the use of radioactive tracers in PET introduces safety considerations, which further limits its broad applicability.

Traditional case analyses often discard samples with missing modalities, causing sample waste, reducing information utilization, and increasing the risk of overfitting. To leverage incomplete data, one line of work generates the target modality from available modalities [20]. In recent years, GANs have been widely used to learn such nonlinear mappings [2, 12, 10]. For example, Gao et al. [2] proposed TPA-GAN for missing-modality imputation, while Xu et al. [23] introduced DSIP, which preserves domain-specific information via style transfer and dual reconstruction pathways, and incorporates EAP and CAA to capture key anatomical details. However, because MRI primarily emphasizes anatomical structure whereas PET reflects metabolic function, the two modalities differ fundamentally in imaging mechanisms, making cross-modal synthesis prone to artifacts and fine-detail loss. Such synthesis-induced noise can interfere with downstream classifiers and limit diagnostic performance.

Recently, some studies have shifted toward latent subspace learning approaches [24, 19, 8]. For example, AGDiC [8] performs distribution transformation using normalizing flows, and Xiong et al. [22] proposed DCFMnet, which combines feature disentanglement and codebook learning to align latent representations of incomplete data in a shared space via feature-similarity matching. Such methods can avoid noise introduced by pixel-space generation; however, existing feature-level approaches still have two limitations. First, alignment largely relies on spatial-domain constraints, yet MRI and PET also exhibit substantial differences in texture patterns and frequency components; spatial alignment alone is insufficient to bridge the semantic gap between modalities. Second, feature completion is often forced to use observed modalities to fit all information of the missing modality, overlooking that some functional PET information has no corresponding representation in MRI. Forcing such predictions can inject noise into the latent space and interfere with the classification decision boundary.

To address the above challenges, we propose a Spatio-Spectral Guided Affinity Imputation Network (S^2GAIN), which aims to mitigate modality alignment bias and noise in existing methods through spatio-spectral joint alignment strategy and an affinity-guided reliability-weighting mechanism, thereby enabling computer-aided diagnosis of brain diseases under missing-modality settings. The main contributions can be summarized as follows:

1. A spatio-spectral joint alignment strategy is developed to overcome the limitation of conventional alignment that is confined to the spatial domain. Specifically, a fast Fourier transform (FFT)-based frequency-domain constraint is introduced. By simultaneously aligning the spatial locations and frequency domain magnitude spectrum, precise heterogeneous in both domains features is achieved.
2. An affinity-guided reliability-weighting mechanism is proposed to mitigate noise injection arising from forcing full information fitting in prior approaches. During feature learning, an affinity matrix is used as a reliability measure of cross-modal predictability, and an affinity-guided adaptive weighting scheme is designed to encourage the model to emphasize reliable cross-modal associations while suppressing non-inferable noise, thereby improving the purity and robustness of the imputed features.
3. Extensive experiments on the ADNI dataset demonstrate that S^{2}GAIN outperforms state-of-the-art methods in handling missing modalities and improving diagnostic performance.

2 Proposed Model

To exploit incomplete multimodal data for AD diagnosis, S^{2}GAIN is proposed (Fig. 1). The framework consists of three modules. First, a cascaded cross-modal interaction network based on multi-stage Dense Blocks is constructed to progressively extract cross-modal features from shallow details to deep semantics while preserving cross-level complementary information across modalities. Second, a feature completion and refinement module is introduced, in which a cross-modal imputer is deployed at each stage to estimate missing PET information from MRI features. To improve imputed feature quality, a spatio-spectral joint alignment strategy (S^{2}JA) enforces feature consistency, and an affinity-guided reliability-weighting mechanism (AGRW) suppresses noise. To preserve spatial structure in imputed representations, a PET reconstruction decoder is introduced at the deepest layer to map completed deep metabolic features back to the image domain, and image-domain reconstruction regularization (R^2) is imposed to preserve metabolic texture and spatial consistency to improve imputed feature quality. Finally, during training, full-modality supervision is used to optimize mapping relationships; during inference, data streams are automatically switched ensuring diagnosis when PET is missing via imputed features.

2.1 Cascaded Cross-Modal Interaction Backbone

As shown in Fig. 1, S^{2}GAIN adopts an interactive dual-stream network constructed using Dense Blocks as basic feature-extraction units, and information from potentially missing modalities is progressively recovered in the feature space through stage-wise interactions. The original high-dimensional MRI and PET images are first processed by modality-specific encoders E_m to extract shallow features. These features are then fed into a cascaded backbone with L stages. At

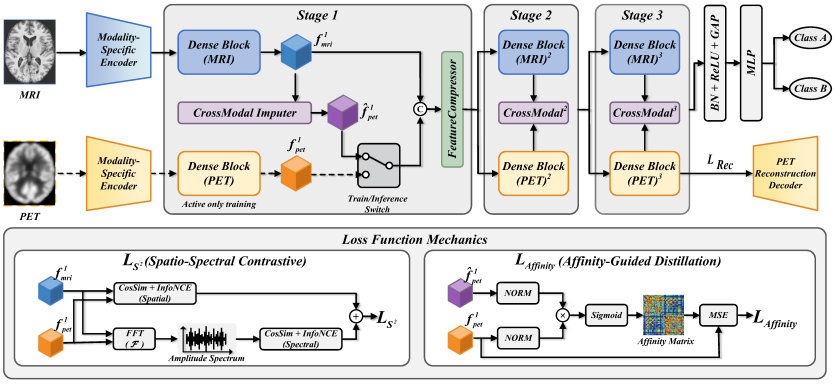

Fig. 1: Illustration of the S^{2}GAIN framework. Dashed lines indicate data streams that may be absent during inference.

the l-th stage ($l \in \{1, \ldots, L\}$), the previous-stage output f_{in}^{l-1} is taken as input to a 3D Dense Block $\mathcal{F}_{\text{dense}}^{l}$. For the MRI branch, $f_{\text{mri}}^{l} = \mathcal{F}_{\text{dense}}^{l}(f_{in}^{l-1})$. During training, the PET branch follows the same procedure to extract f_{pet}^{l}.

To compensate for PET absence during inference, a lightweight cross-modal imputer, G^l, is integrated at each stage to learn a nonlinear mapping from the anatomical feature domain to the metabolic feature domain. Accordingly, even when PET is missing, the metabolic feature $\hat{f}\text{pet}^l$ can still be inferred from the current MRI feature as $\hat{f}\text{pet}^l = G^l(f_{\text{mri}}^{l}) = \sigma!\left(W_g^l * f_{\text{mri}}^{l} + b_g^l\right)$.

At each stage, MRI and PET features are concatenated along the channel dimension, and the fused representation is reduced in dimensionality and spatially downsampled as input to the next stage, ensuring complete cross-modal contextual information.

2.2 Spatio-Spectral Joint Alignment Strategy

To infer latent PET features from MRI in feature space, S^{2}JA is introduced. With a frequency-domain aware mechanism, f_{mri}^{l} and f_{pet}^{l} are encouraged to become closer in spatial distributions and frequency-domain patterns, providing a consistent foundation for feature imputation.

First, a spatial-domain distribution alignment constraint is imposed to capture anatomical correspondences and cross-modal structural consistency. For features extracted at stage l, features are flattened into vectors and spatial cosine similarity is computed. The spatial alignment objective $\mathcal{L}_{\text{spa}}^{l}$ is defined as:

$$\mathcal{L}_{\text{spa}}^{l} = -\mathbb{E}\left[\log\left(\sigma\left(\frac{\text{CosSim}\left(f_{\text{mri}}^{l}, f_{\text{pet}}^{l}\right)}{\tau}\right)\right)\right] \tag{1}$$

where τ is the temperature parameter. To compensate for the inability of spatial constraints alone to recover complex metabolic texture details, a 3D fast

Fourier transform (FFT) is introduced to project features into the frequency domain. The magnitude spectrum, $\mathcal{A}(F) = |\mathcal{F}(F)|$, is extracted to characterize the frequency-domain energy distribution, and a contrastive loss is applied to enforce consistency between the magnitude spectra of MRI and PET features:

$$\mathcal{L}^l_{\text{spe}} = -\mathbb{E}\left[\log\left(\sigma\left(\frac{\text{CosSim}\big(\mathcal{A}(f^l_{\text{mri}}), \mathcal{A}(f^l_{\text{pet}})\big)}{\tau}\right)\right)\right] \tag{2}$$

The spatio-spectral joint alignment loss at stage l is defined as

$$\mathcal{L}^l_{S^2} = \alpha\mathcal{L}^l_{\text{spa}} + \beta\mathcal{L}^l_{\text{spe}} \tag{3}$$

With this mechanism, MRI-branch features are aligned with PET-branch features in spatial structure and frequency-domain texture before imputation, thereby reducing cross-modal imputation uncertainty.

2.3 Affinity-guided Optimization and Image-domain Reconstruction

After distribution-consistent feature representations are obtained, a cross-modal imputer G^l is integrated at each stage to impute hierarchical latent features of the missing modality. Although inter-domain discrepancies are reduced by alignment, heterogeneity between MRI and PET leaves some anatomical information in MRI unrelated to metabolic activity. Therefore, AGRW is adopted to assess imputation reliability and optimize generation.

First, the imputed features $\hat{f}^l_{\text{pet}}$ and the corresponding ground-truth features f^l_{pet} extracted at the same stage are normalized, after which a stage-wise affinity matrix A^l is computed as $A^l = \sum_{c,d,h,w}\left(\frac{\hat{f}^l_{\text{pet}}}{\|\hat{f}^l_{\text{pet}}\|_2} \odot \frac{f^l_{\text{pet}}}{\|f^l_{\text{pet}}\|_2}\right)$ to reflect the reliability of synthesized features under the current anatomical context.

Based on the computed affinity, an adaptive weight W^l is generated using a sigmoid function. To ensure that the affinity serves only as a reliability criterion and does not participate in backpropagation, gradient stopping is applied, i.e., $W^l = \text{Sigmoid}(A^l).\text{detach}()$.

The resulting weight defines a weighted mean squared error between imputed and ground-truth features, yielding the affinity-guided distillation loss $\mathcal{L}^l_{\text{aff}}$:

$$\mathcal{L}^l_{\text{aff}} = mean\left(W^l \cdot (\hat{f}^l_{\text{pet}} - f^l_{\text{pet}})^2\right) \tag{4}$$

By treating affinity as a reliability filter, high affinity yields W^l close to 1 and encourages learning of reliable cross-modal associations; for non-inferable content arising from modality heterogeneity, the weight is reduced to suppress noise propagation, thereby improving the purity and robustness of imputed features.

Although S^{2}JA and AGRW reduce feature-space discrepancy between MRI and PET and suppress noise propagation from modality heterogeneity, feature-level constraints alone yields imputed metabolic representations that are discriminative yet fail to preserve interpretable metabolic patterns in the PET image

domain. To further enforce structural consistency, a lightweight PET reconstruction decoder is introduced at the deepest stage to map deep metabolic features back to the PET image domain, and image-domain supervision is applied to encourage imputed features to retain reconstructable spatial information.

Specifically, a decoder $D(\cdot)$ is constructed to reconstruct the PET image $\hat{x}_{\text{pet}}$ from f_{pet}^{L} via progressive trilinear upsampling followed by 3D convolutions, i.e., $\hat{x}_{\text{pet}} = D(f_{\text{pet}}^{L})$. During training, the ground-truth PET image x_{pet} is used as the supervision signal, and an L_1 reconstruction loss is employed:

$$\mathcal{L}_{\text{rec}} = \|\hat{x}_{\text{pet}} - x_{\text{pet}}\|_1 \tag{5}$$

This reconstruction branch is jointly optimized during training as an auxiliary regularizer, constraining structural drift in deep metabolic representations and reducing uncertainty in the imputation learning process.

Finally, the above auxiliary losses are combined with the primary classification loss of the downstream task to jointly optimize multimodal feature imputation and disease diagnosis. The overall objective $\mathcal{L}_{total}$ is defined as follows:

$$\mathcal{L}_{total} = \mathcal{L}_{cls} + \sum_{l=1}^{L} \lambda_1 \mathcal{L}_{S^2}^{l} + \sum_{l=1}^{L} \lambda_2 \mathcal{L}_{aff}^{l} + \lambda_3 \mathcal{L}_{rec} \tag{6}$$

3 Experiments

Experiments were conducted on two ADNI subsets (ADNI1 and ADNI2) [6]. Following prior studies [9, 4], each including baseline data from AD, CN, pMCI and sMCI, where pMCI denotes conversion to AD within 36 months after baseline and sMCI denotes no conversion. Table 1 reports modality counts and demographics. MRI preprocessing included AC-PC alignment, N3 correction [16], skull stripping, and registration to the Colin27 template [5]; PET preprocessing included co-registration to MRI, skull removal using the MRI-derived brain mask, and registration to Colin27 using MRI transformation parameters. All images were cropped to $160 \times 192 \times 160$ and downsampled to $80 \times 96 \times 80$. The model was optimized with Adam (learning rate 1×10^{-4}; batch size 2).

3.1 Comparison and Ablation Results

Table 2 reports cross-cohort comparisons on ADNI-1 and ADNI-2 against five representative baselines: JSRL [11] (generator-feature-based MRI-to-PET mapping), TPA-GAN [2] (GAN-based modality imputation), ShaSpec [21] (shared/modality-specific representation generation), MJL [15] (multimodal fusion encoding), and DCFMnet [22] (disentanglement- and codebook-based cross-modal alignment). Overall, S^{2}GAIN achieves the best performance across both cohorts. For AD vs CN, it ranks first on most metrics and shows better stability; although TPA-GAN and DCFMnet are competitive in some settings (e.g., DCFMnet AUC = 94.00), they remain inferior to S^{2}GAIN (best AUC = 96.12). For sMCI vs pMCI,

Table 1: Demographic and clinical characteristics of the subjects in the ADNI1 and ADNI2 cohorts. Age and Mini-Mental State Examination (MMSE) scores are reported.

Dataset	Category	MRI/PET	Gender(F/M)	Age	MMSE
ADNI-1	AD	188/93	99/89	75.3 ± 7.5	23.3 ± 2.0
	CN	229/102	119/110	75.9 ± 5.0	29.1 ± 1.0
	sMCI	204/100	134/70	74.8 ± 7.16	27.2 ± 1.6
	pMCI	175/79	106/69	74.6 ± 6.9	26.7 ± 1.7
ADNI-2	AD	150/145	87/63	74.7 ± 8.1	23.0 ± 2.0
	CN	188/186	90/98	73.5 ± 6.3	29.0 ± 1.2
	sMCI	239/228	133/105	71.6 ± 7.4	28.1 ± 1.7
	pMCI	86/84	48/38	74.1 ± 7.4	28.0 ± 1.7

all methods degrade substantially, but S^{2}GAIN still achieves the best results, including the best AUC (77.76) with a more balanced sensitivity/specificity trade-off. By contrast, JSRL tends to show lower sensitivity, ShaSpec yields higher specificity in some settings (71.57) but weaker progressive-case identification, and MJL performs worse overall. These results support the effectiveness of S^{2}JA and AGRW constraints for robust progression-related representation learning under missing PET.

Table 2: Performance comparison on ADNI-1/ADNI-2. Bold and underlined numbers indicate the best and second-best results, respectively.

Task	Methods	AD vs CN				sMCI vs pMCI			
		ACC	AUC	SEN	SPE	ACC	AUC	SEN	SPE
ADNI-1 training and ADNI-2 testing	JSRL	85.77	93.05	86.78	83.72	67.82	71.95	53.55	**78.30**
	TPA-GAN	<u>90.53</u>	<u>95.18</u>	<u>88.67</u>	<u>92.02</u>	<u>72.73</u>	72.17	<u>67.50</u>	74.48
	ShaSpec	86.92	92.06	86.78	83.75	67.08	<u>73.19</u>	66.08	68.12
	MJL	85.54	91.06	85.41	87.60	60.05	63.12	61.24	57.88
	DCFMnet	86.09	94.00	80.67	90.43	68.70	70.16	55.10	<u>77.31</u>
	S^{2}GAIN	**92.47**	**96.12**	**91.10**	**93.55**	**74.17**	**77.76**	**70.00**	75.68
ADNI-2 training and ADNI-1 testing	JSRL	75.15	80.97	74.67	75.53	54.40	57.03	<u>59.37</u>	50.49
	TPA-GAN	81.68	90.22	82.86	<u>80.68</u>	<u>62.09</u>	<u>68.64</u>	56.87	66.18
	ShaSpec	75.83	82.02	76.61	73.91	60.16	67.27	50.62	<u>71.57</u>
	MJL	74.87	85.71	82.86	68.12	56.87	56.53	48.12	63.73
	DCFMnet	83.51	<u>90.37</u>	**85.14**	79.71	61.22	66.68	52.12	**73.59**
	S^{2}GAIN	**85.83**	**90.57**	<u>84.30</u>	**87.13**	**66.09**	**69.52**	**64.94**	67.01

Table 3 reports the key robustness mechanisms in S^{2}GAIN under missing-modality imputation. Three variants were considered: removing spatio-spectral joint alignment (w/o S^{2}JA), removing affinity-guided reliability-weighting mech-

Table 3: Results of S^{2}GAIN under different module combinations on AD vs. CN (ADNI-2 test set).

S^{2}JA	AGRW	R^2	ACC	AUC	SEN	SPE
✓			89.94	93.87	86.00	<u>93.09</u>
	✓		87.28	93.27	88.67	86.17
		✓	90.24	95.07	90.67	89.89
	✓	✓	<u>91.72</u>	<u>95.91</u>	**91.33**	92.02
✓	✓		90.53	94.43	87.33	88.83
✓	✓	✓	**92.47**	**96.12**	<u>91.10</u>	**93.55**

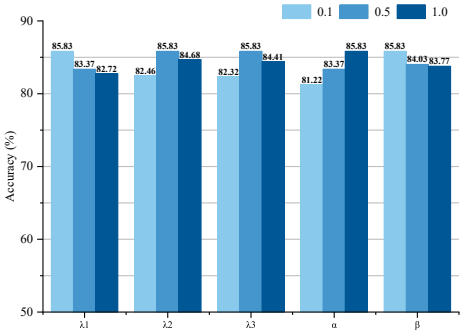

Fig. 2: Sensitivity analysis of S^{2}GAIN hyperparameters on the AD vs. CN task when tested on ADNI-1.

anism (w/o AGRW), and removing image-domain reconstruction regularization (w/o R^2). All three components provide consistent gains. Removing S^{2}JA causes overall degradation, indicating that spatio-spectral consistency is important for reducing cross-modal shift. Removing AGRW leads to larger drops in ACC and SPE, showing that AGRW and adaptive weighting suppress unreliable error propagation. Removing R^2 mainly reduces AUC and ACC, indicating that image-domain reconstruction regularization helps preserve spatial structural consistency. The best overall performance is achieved when all components are enabled, confirming their complementarity.

3.2 Parameter Sensitivity Analysis

S^{2}GAIN was evaluated on five loss weights ($\lambda_1, \lambda_2, \lambda_3, \alpha, \beta$) by varying one parameter while fixing the others. Each parameter was tested at 0.1, 0.5, and 1.0, and the resulting Acc changes are shown in Fig. 2. Increasing λ_1 consistently reduced performance (Acc: 85.83 $\rightarrow$ 82.72), suggesting that overweighting the alignment constraint biases optimization toward distribution consistency over diagnosis-relevant discrimination. In contrast, λ_2 and λ_3 showed mid-range optima, indicating that affinity-guided distillation and image-domain reconstruction should be moderately weighted, while excessive weights impose strong constraints and degrade performance. Within spatio-spectral joint alignment, α and β showed opposite trends: larger α improved performance, suggesting stronger spatial alignment benefits cross-modal structural correspondence, whereas larger β degraded performance, indicating overemphasized frequency-domain alignment may suppress diagnostically informative variations.

4 Conclusion

This work proposes S^{2}GAIN for early multimodal AD diagnosis with missing PET. It unifies multiple modules under a single training objective in a cascaded

dual-stream framework with lightweight multi-stage imputers. S^2JA reduces MRI-PET heterogeneity, AGRW suppresses unreliable error propagation, and image-domain R^2 preserves metabolic and structural consistency. Cross-cohort evaluations on ADNI-1/ADNI-2 and ablations show complementary gains. Limitations include reliance on affinity patterns learned under ground-truth PET supervision for reliability weighting and magnitude-spectrum-only frequency-domain alignment. Future work will extend S^2GAIN to more complex missing-modality settings and finer-grained frequency-domain modeling.

Acknowledgments. This work was partly supported by National Key Research and Development Program of China No. 2023YFC3304202, Ningbo Municipal Public Welfare Technology Research Project of China, No. 2023S023.

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

References

1. Better, M.A.: Alzheimer's disease facts and figures. Alzheimers Dement **19**(4), 1598–1695 (2023)
2. Gao, X., Shi, F., Shen, D., Liu, M.: Task-induced pyramid and attention gan for multimodal brain image imputation and classification in alzheimer's disease. IEEE journal of biomedical and health informatics **26**(1), 36–43 (2021)
3. Gauthier, S., Webster, C., Servaes, S., Morais, J., Rosa-Neto, P., et al.: Life after diagnosis: navigating treatment, care and support. Alzheimer's Disease International pp. 1–416 (2022)
4. Han, K., Hu, D., Zhao, F., Liu, T., Yang, F., Li, G.: Incomplete multi-modal disentanglement learning with application to alzheimer's disease diagnosis. IEEE Transactions on Medical Imaging (2025)
5. Holmes, C.J., Hoge, R., Collins, L., Woods, R., Toga, A.W., Evans, A.C.: Enhancement of mr images using registration for signal averaging. Journal of computer assisted tomography **22**(2), 324–333 (1998)
6. Jack Jr, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L. Whitwell, J., Ward, C., et al.: The alzheimer's disease neuroimaging initiative (adni): Mri methods. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine **27**(4), 685–691 (2008)
7. James, B.D., Leurgans, S.E., Hebert, L.E., Scherr, P.A., Yaffe, K., Bennett, D.A.: Contribution of alzheimer disease to mortality in the united states. Neurology **82**(12), 1045–1050 (2014)
8. Li, Y., Dong, Y., Wu, Y., Yan, H., Gao, L.: Alzheimer's disease recognition based on adaptive graph normalization flow for incomplete multimodal data fusion. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 64–73. Springer (2025)
9. Lian, C., Liu, M., Pan, Y., Shen, D.: Attention-guided hybrid network for dementia diagnosis with structural mr images. IEEE transactions on cybernetics **52**(4), 1992–2003 (2020)

10. Liu, Y., Pan, Y., Yang, W., Ning, Z., Yue, L., Liu, M., Shen, D.: Joint neuroimage synthesis and representation learning for conversion prediction of subjective cognitive decline. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 583–592. Springer (2020)
11. Liu, Y., Yue, L., Xiao, S., Yang, W., Shen, D., Liu, M.: Assessing clinical progression from subjective cognitive decline to mild cognitive impairment with incomplete multi-modal neuroimages. Medical image analysis **75**, 102266 (2022)
12. Pan, Y., Liu, M., Lian, C., Xia, Y., Shen, D.: Spatially-constrained fisher representation for brain disease identification with incomplete multi-modal neuroimages. IEEE transactions on medical imaging **39**(9), 2965–2975 (2020)
13. Pan, Y., Liu, M., Lian, C., Zhou, T., Xia, Y., Shen, D.: Synthesizing missing pet from mri with cycle-consistent generative adversarial networks for alzheimer's disease diagnosis. In: International conference on medical image computing and computer-assisted intervention. pp. 455–463. Springer (2018)
14. Peng, L., Cai, S., Wu, Z., Shang, H., Zhu, X., Li, X.: Mmgpl: Multimodal medical data analysis with graph prompt learning. Medical Image Analysis **97**, 103225 (2024)
15. Qiu, Z., Yang, P., Xiao, C., Wang, S., Xiao, X., Qin, J., Liu, C.M., Wang, T., Lei, B.: 3d multimodal fusion network with disease-induced joint learning for early alzheimer's disease diagnosis. IEEE Transactions on Medical Imaging **43**(9), 3161–3175 (2024)
16. Sled, J.G., Zijdenbos, A.P., Evans, A.C.: A nonparametric method for automatic correction of intensity nonuniformity in mri data. IEEE transactions on medical imaging **17**(1), 87–97 (2002)
17. Sperling, R.A., Aisen, P.S., Beckett, L.A., Bennett, D.A., Craft, S., Fagan, A.M., Iwatsubo, T., Jack Jr, C.R., Kaye, J., Montine, T.J., et al.: Toward defining the preclinical stages of alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. Alzheimer's & dementia **7**(3), 280–292 (2011)
18. Tarawneh, R., Holtzman, D.M.: The clinical problem of symptomatic alzheimer disease and mild cognitive impairment. Cold Spring Harbor perspectives in medicine **2**(5), a006148 (2012)
19. Ting, H., Liu, M.: Multimodal transformer of incomplete mri data for brain tumor segmentation. IEEE journal of biomedical and health informatics **28**(1), 89–99 (2023)
20. Wang, C., Piao, S., Huang, Z., Gao, Q., Zhang, J., Li, Y., Shan, H., Initiative, A.D.N., et al.: Joint learning framework of cross-modal synthesis and diagnosis for alzheimer's disease by mining underlying shared modality information. Medical Image Analysis **91**, 103032 (2024)
21. Wang, H., Chen, Y., Ma, C., Avery, J., Hull, L., Carneiro, G.: Multi-modal learning with missing modality via shared-specific feature modelling. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 15878–15887 (2023)
22. Xiong, W., Wang, T., Chen, X., Zhang, Y., Zhang, W., Feng, Q., Huang, M., Initiative, A.D.N., et al.: Disentanglement and codebook learning-induced feature match network to diagnose neurodegenerative diseases on incomplete multimodal data. Pattern Recognition **165**, 111597 (2025)
23. Xu, H., Wang, J., Feng, Q., Zhang, Y., Ning, Z.: Domain-specific information preservation for alzheimer's disease diagnosis with incomplete multi-modality neuroimages. Medical Image Analysis **101**, 103448 (2025)

24. Zhou, T., Liu, M., Thung, K.H., Shen, D.: Latent representation learning for alzheimer's disease diagnosis with incomplete multi-modality neuroimaging and genetic data. IEEE transactions on medical imaging **38**(10), 2411–2422 (2019)