MultiTransAD: Cross-Sequence Translation-Driven Anomaly Detection in Multi-Sequence Brain MRI

Qi Zhang[1], Yibo Hu[1], and Jianqi Sun[1,✉]

[1]School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China
milesun@sjtu.edu.cn

Abstract. Accurate anomaly detection in brain MRI is critical for early disease diagnosis, yet existing single-sequence reconstruction methods often fail to distinguish pathological anomalies from both normal anatomical variations and multi-sequence contrast discrepancies. We propose MultiTransAD, a novel framework that leverages inter-sequence contrast differences as primary biomarkers for unsupervised anomaly detection. Our approach introduces: (1) a disentangled architecture with anatomical edge constraints to decouple sequence-invariant anatomy from sequence features, (2) cross-sequence translation error analysis for direct anomaly quantification, and (3) dual-level anomaly detection combining pixel-level errors and patch-level feature dissimilarities. Evaluated on BraTS 2021, MultiTransAD achieves state-of-the-art performance with Dice scores of 0.6334 (14.2% improvement over reconstruction baselines) and AUROC of 0.9722, validating the effectiveness of multi-sequence contrast analysis in anomaly detection while establishing a extensible cross-sequence translation paradigm. The code is publicly available at: https://github.com/zhibaishouheilab/MT-AD

Keywords: Anomaly Detection · Multi-sequence MRI · Cross-sequence Translation · Self-supervised Learning.

1 Introduction

Early diagnosis of neurological disorders critically depends on precise anomaly detection in brain MRI. While supervised learning methods require labor-intensive lesion annotations that are prone to inter-observer variability, unsupervised anomaly detection (UAD) formulates the task as one-class classification (OCC) using only healthy data for training [2,9]. Reconstruction-based methods, such as autoencoders (AEs) and generative adversarial networks (GANs), identify anomalies in areas with high reconstruction errors. However, these methods suffer from three limitations: (1) elevated reconstruction errors in normal regions due to anatomical complexity (false positives) [13,5], (2) overfitting to pathology representation enabling perfect anomaly reconstruction (false negatives) [3], and

(3) underutilization of multi-sequence contrast information decreasing detection sensitivity.

Multi-sequence MRI analysis is standard in clinical practice, as pathological tissues exhibit distinct intensity profiles across sequences[4]. Recent advances in cyclic-translation frameworks[12,6] aim to mitigate anatomical variability by learning bidirectional mappings between sequences (e.g., T2 to FLAIR and vice versa). These methods train on paired healthy sequences to reconstruct a source image through a cyclic translation, hypothesizing that anomalies will exhibit high errors due to inconsistent translations. However, these methods may introduce new issues that incomplete disentanglement of anatomical and sequence features propagates pathologies during translation, which may lead to false positives. Furthermore, testing remains single-sequence inputs, lacking the direct utilization of inter-sequence contrasts.

To address these challenges, we propose MultiTransAD, a cross-sequence translation framework that directly utilizes inherent contrast discrepancies as primary anomaly biomarkers. Our contributions are threefold: (1) Redefinition of anomaly detection through cross-sequence translation error analysis, instead of secondary features from single-sequence reconstruction; (2) A disentangled architecture with anatomical edge constraints that explicitly separates sequence-invariant anatomy from sequence-specific features, preventing contamination by anomalies; (3) A dual-level evaluation combining pixel-level translation errors and patch-level feature dissimilarities, enhancing detection sensitivity greatly.

2 Methodology

2.1 Architecture Overview

MultiTransAD processes K MRI sequences $\{\mathbf{I}_{S_k}\}_{k=1}^{K}$ through three key stages: (1) **Disentangled translation** from source sequence S_s image $\mathbf{I}_{S_s}$ to target sequence S_t image $\mathbf{I}_{S_t}$ via content-style feature disentanglement and edge-constrained fusion; (2) **Feature dissimilarity map** extracting patch-level dissimilarities between source content $\mathbf{Z}^c_{S_s}$ and target content $\mathbf{Z}^c_{S_t}$; (3) **Anomaly detection** combining pixel-level translation errors $\mathbf{A}^{pixel}_{S_s \to S_t}$ with patch-level feature dissimilarities $\mathbf{A}^{feature}_{S_s \leftrightarrow S_t}$ to generate the final anomaly map $\mathbf{A}_{S_s \to S_t}$. The training and anomaly detection workflows are illustrated in Figure 1 and Figure 2, respectively.

2.2 Sequence-Agnostic Feature Disentanglement

The source image $\mathbf{I}_{S_s}$ is decomposed (via image patching ϕ_{patch}) into $N = \frac{HW}{p^2}$ patches (p is patch size), each projected to dimension d:

$$\mathbf{Z}^{\text{patch}}_{S_s} = \phi_{\text{patch}}(\mathbf{I}_{S_s}) \in \mathbb{R}^{N \times d} \tag{1}$$

A content encoder E_c with 12 ViT layers extracts sequence-invariant anatomical features:

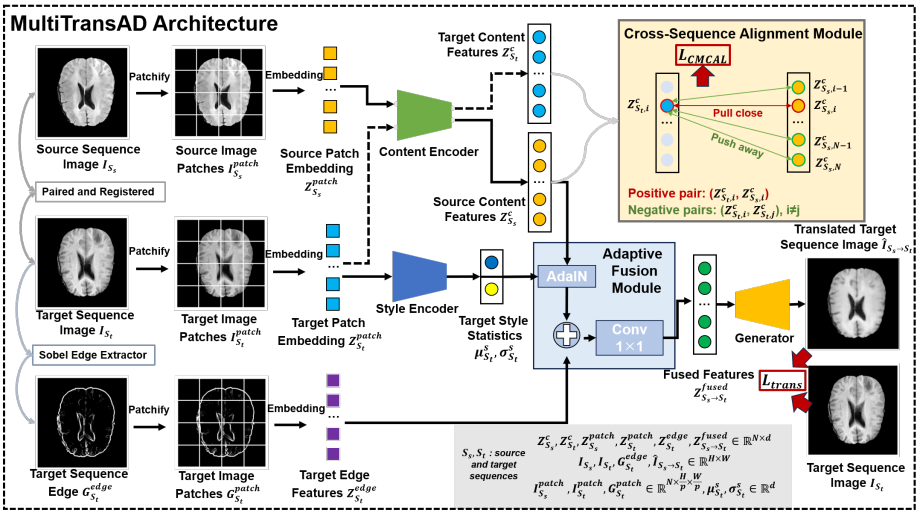

Fig. 1: Illustration of the MultiTransAD architecture. MultiTransAD leverages paired MR sequences during training and inference. Translated images integrate source content with target style.

$$\mathbf{Z}^c_{S_s} = E_c(\mathbf{Z}^{\text{patch}}_{S_s}) \in \mathbb{R}^{N \times d} \tag{2}$$

For style encoding of target sequence, style encoder E_s, which shares E_c's architecture, outputs sequence-specific statistics through parallel 1×1 convolutional heads ϕ_μ and ϕ_σ:

$$\mu^s_{S_t}, \sigma^s_{S_t} = \left[\text{AvgPool}(\phi_\mu(\mathbf{Z}^{\text{patch}}_{S_t})), \text{AvgPool}(\phi_\sigma(\mathbf{Z}^{\text{patch}}_{S_t}))\right] \in \mathbb{R}^d \tag{3}$$

2.3 Edge-Constrained Anatomical Preservation

The Anatomical Edge Constraint Module (AECM) uses dual-threshold gradient clipping to preserve structural boundaries and suppress noises:

$$\mathbf{Grad}_{S_t} = \text{Clip}(\text{Sobel}(\mathbf{I}_{S_t}), q_{50}(\nabla \mathbf{I}_{S_t}), q_{95}(\nabla \mathbf{I}_{S_t})) \tag{4}$$

where q_{50} and q_{95} represent the 50th/95th percentile of gradient magnitudes. Sobel indicates Sobel operator. Edge-aware features are encoded with the edge projection $\phi^{edge}_{\text{patch}}$:

$$\mathbf{Z}^{edge}_{S_t} = \phi^{edge}_{\text{patch}}(\mathbf{Grad}_{S_t}) \in \mathbb{R}^{N \times d} \tag{5}$$

2.4 Adaptive Cross-sequence Fusion

The adaptive fusion module combines content, edge, and style through:

1. **Style Adaptation**: Applies Adaptive Instance Normalization (AdaIN) [7] to style and content features:

$$\mathbf{Z}^{\text{ada}}_{S_s \to S_t} = \sigma^s_{S_t} \cdot \frac{\mathbf{Z}^c_{S_s} - \mu(\mathbf{Z}^c_{S_s})}{\sigma(\mathbf{Z}^c_{S_s})} + \mu^s_{S_t} \tag{6}$$

2. **Edge Fusion**: Concatenates target edge features and fused features, followed by dimension reduction:

$$\mathbf{Z}^{fusion}_{S_s \to S_t} = \text{Conv}_{1\times1}([\mathbf{Z}^{\text{ada}}_{S_s \to S_t} \oplus \mathbf{Z}^{edge}_{S_t}]) \in \mathbb{R}^{N\times d} \tag{7}$$

A generator $\mathbf{G}$ with 8 ViT layers reconstructs the target image:

$$\hat{\mathbf{I}}_{S_s \to S_t} = \mathbf{G}(\mathbf{Z}^{\text{fusion}}_{S_s \to S_t}) \tag{8}$$

2.5 Objective Functions

Training minimizes translation fidelity and cross-sequence consistency:

1. **Translation Loss (L1 Loss)**: Enforces pixel accuracy:

$$\mathcal{L}_{\text{trans}} = \frac{1}{K(K-1)} \sum_{S_s \neq S_t} \|\mathbf{I}_{S_t} - \hat{\mathbf{I}}_{S_s \to S_t}\|_1 \tag{9}$$

2. **Cross-Sequence Contrastive Alignment Loss (CSCAL)**: Promotes anatomical consistency via patch-level content feature contrastive loss. Specifically, for a given patch index i, the positive sample pair is $(\mathbf{z}^c_{S_s,i}, \mathbf{z}^c_{S_t,i})$ and the negative sample pairs are $(\mathbf{z}^c_{S_s,i}, \mathbf{z}^c_{S_t,j})$ where $i \neq j$. The loss function is defined as:

$$\mathcal{L}_{\text{CSCAL}} = -\frac{1}{K(K-1)N} \sum_{S_s \neq S_t} \sum_{i=1}^{N} \log \frac{e^{s(\mathbf{z}^c_{S_s,i}, \mathbf{z}^c_{S_t,i})/\tau}}{\sum_{j=1}^{N} e^{s(\mathbf{z}^c_{S_s,i}, \mathbf{z}^c_{S_t,j})/\tau}} \tag{10}$$

where $s(\mathbf{a}, \mathbf{b}) = \mathbf{a}^\top \mathbf{b}/\|\mathbf{a}\|\|\mathbf{b}\|$ and $\tau = 0.1$.

The total loss combines these objectives with balancing factor $\alpha = 0.01$:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{trans}} + \alpha \mathcal{L}_{\text{CSCAL}} \tag{11}$$

2.6 Anomaly Quantification

The anomaly quantification workflow is illustrated in Figure 2. For each sequence pair (S_s, S_t), the anomaly score map $\mathbf{A}_{S_s \to S_t}$:

Pixel Error Map Quantifies pixel-level absolute translation errors:

$$\mathbf{A}^{pixel}_{S_s \to S_t}(x, y) = \|\mathbf{I}_{S_t}(x, y) - \hat{\mathbf{I}}_{S_s \to S_t}(x, y)\|_1 \tag{12}$$

Feature Dissimilarity Map Evaluates cross-sequence content feature alignment. Given $\mathbf{Z}^c_{S_s}, \mathbf{Z}^c_{S_t} \in \mathbb{R}^{N\times d}$, compute cosine similarity $\mathbf{S}(m) = \frac{\mathbf{z}^{c\top}_{S_s,m} \mathbf{z}^c_{S_t,m}}{\|\mathbf{z}^c_{S_s,m}\|\|\mathbf{z}^c_{S_t,m}\|}$

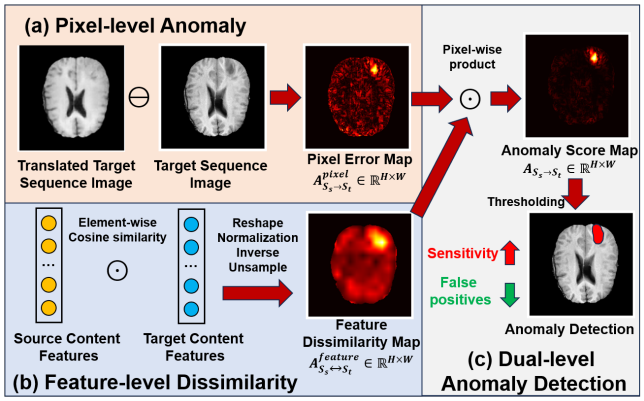

Fig. 2: Anomaly detection combines translation error maps with the feature dissimilarity map.

for each patch m. Reshape $\mathbf{S} \in \mathbb{R}^{N\times 1}$ into a 2D grid and upsample to $H\times W$, then normalize to $\mathbf{S}^{feature}_{S_s\leftrightarrow S_t} \in [0,1]^{H\times W}$. The feature dissimilarity map is: $\mathbf{A}^{feature}_{S_s\leftrightarrow S_t} = 1 - \mathbf{S}^{feature}_{S_s\leftrightarrow S_t}$.

Final Anomaly Map Combines pixel and feature anomalies via pixel-wise product:

$$\mathbf{A}_{S_s\to S_t} = \mathbf{A}^{pixel}_{S_s\to S_t} \odot \mathbf{A}^{feature}_{S_s\leftrightarrow S_t} \tag{13}$$

3 Experiments and Results

3.1 Datasets and Implementation

The framework was evaluated on the BraTS 2021 dataset [1], a public multi-sequence brain tumor MRI dataset with 1,251 subjects, each having T1, T2, T1ce, and FLAIR sequences. Data were split 4:1 into training and test sets. For training reconstruction-based methods, healthy T2 scans from 581 IXI dataset subjects [15] were used. Preprocessing included skull stripping, registration, and normalization to [0,1]. For each 3D image in BraTS 2021 and IXI, 2D slices 50-110 were extracted, center-cropped to 200×200, and resized to 256×256. To ensure a fair training and evaluation setup, we also extracted all normal slices (without tumor) from the BraTS 2021 training set, forming the BraTS 2021_{normal} subset. This subset mitigates domain shifts between training and test sets and excludes pathological data from the BraTS 2021 training set.

3.2 Model Configuration and Comparison Methods

Input images were resized to 256 × 256 and divided into 16 × 16 non-overlapping patches with 128-d embeddings, reduced to 64-d after fusion. Predictions were scaled to [0,1] via Sigmoid activation function. Training lasted 50 epochs with

data augmentation (random flipping, rotation, contrast/intensity/sharpness adjustments). An initial learning rate of $3e^{-3}$ was set, decreased by 0.1 every 10 epochs. The Adam optimizer was used with batch size 8. All experiments were conducted on an NVIDIA 4090 GPU with 24GB of memory. During each iteration, each sample generates 12 cross-sequence pairs through systematic source-target permutation ($C_4^3 \times 2 = 12$ for 4 sequences).

For comparison, we chose reconstruction-based and translation-based methods. Reconstruction-based UAD baselines included VAE [10], DAE [8], VAEGAN [11], f-AnoGAN [14], and AnoDDPM [16]. For cyclic translation, Cyclic-UNet [12] was used, employing dual U-Nets for sequence translation cycles. Anomalies were detected via cyclic error $\|x - \hat{x}\|$. All baselines report results for their optimal sequences and we think that's a fair comparison.

3.3 Post-Processing and Evaluation

Post-processing on anomaly score maps involved median filtering (kernel size 6) and thresholding of anomaly maps, discarding components $<$55 pixels. These are common steps to mitigate noises. In terms of evaluation, we calculated set-level metrics for the test set, including the Dice coefficient, AUROC (Area Under the Receiver Operating Characteristic Curve), and AUPRC (Area Under the Precision - Recall Curve). The threshold maximizing Dice was selected from 100 candidates spanning the anomaly score range.

3.4 Results and Discussion

Quantitative Performance Analysis As shown in Table 1, MultiTransAD using T1ce→FLAIR achieves state-of-the-art performance with Dice scores of 0.5080/0.6334 and AUROC of 0.9517/0.9722 when trained on BraTS 2021 and BraTS 2021_{normal}, respectively. This represents 7.3% (0.5080 vs 0.4734) and 14.2% (0.6334 vs 0.5548) improvement in Dice over the best reconstruction-based methods (AnoDDPM and DAE). The framework's clinical relevance is further validated by its superior AUPRC (0.5558). Besides, T1→FLAIR achieved competitive performance (the second highest Dice and AUROC), confirming architectural robustness.

Cyclic UNet (FLAIR-T2-FLAIR) achieved the highest AUPRC of 0.5559 when trained on BraTS 2021_{normal}. However, its Dice (0.4966) is obviously lower than VAEGAN, DAE, AnoDDPM, and ours. Furthermore, its Dice scores decreased from 0.4537/0.4966 to 0.3203/0.3646 when adopting T2-FLAIR-T2, revealing critical instability. The failure may stem from exclusive reliance on pixel-level cycle consistency failing to capture cross-sequence feature relationships. While MultiTransAD directly utilizes the cross-sequence contrast in both pixel and feature level, reducing the false negatives in anomaly regions.

Ablation Study Table 2 quantifies the contributions of each component: (1) Cross-sequence contrastive alignment ($\mathcal{L}_{\text{CSCAL}}$) contributes 8.4% ($T1ce \rightarrow$

Table 1: Model comparison results on BraTS 2021 test set.

	Method	**IXI/BraTS**(train)			**BraTS 2021$_{\text{normal}}$**(train)		
		Dice↑	AUROC↑	AUPRC↑	Dice↑	AUROC↑	AUPRC↑
Recon.	VAE[10]	0.3602	0.9147	0.2561	0.4176	0.9266	0.3057
	DAE[8]	0.3890	0.8896	0.3064	0.5548	0.9538	0.5531
	VAEGAN[11]	0.4647	0.9339	0.3595	0.5248	0.9464	0.4616
	f-AnoGAN[14]	0.4262	0.9190	0.3510	0.4509	0.9263	0.3571
	AnoDDPM[16]	0.4734	0.9321	**0.4308**	0.5512	0.9491	0.5268
Trans.	Cyclic[12](FL-T2-FL)	0.4537	0.9287	0.3566	0.4966	0.9449	**0.5559**
	Cyclic[12](T2-FL-T2)	0.3203	0.8710	0.3167	0.3646	0.9166	0.3252
	Ours(T1→FL)	<u>0.5075</u>	<u>0.9509</u>	0.3929	<u>0.6027</u>	<u>0.9660</u>	0.5169
	Ours(T1ce→FL)	**0.5080**	**0.9517**	<u>0.4022</u>	**0.6334**	**0.9722**	<u>0.5558</u>

Recon. means Reconstruction-based methods and **Trans.** means Translation-based. Methods are trained on two datasets. **IXI/BraTS**: Reconstruction-based methods are trained on IXI T2 and Translation-based methods are trained on BraTS 2021 training set. **BraTS 2021$_{\text{normal}}$**: All methods are trained on BraTS 2021$_{\text{normal}}$. **Bold** values indicate the best performance and <u>underlined</u> indicates the second best. **FL** indicates FLAIR.

$FLAIR$) and 7.6% ($T1 \rightarrow FLAIR$) Dice improvements (row 4 vs 2), validating its critical role in suppressing normal anatomical variations; (2) The AECM enhances detection precision with 8.9%/4.4% AUPRC gains (row 2 vs 1) for respective pathways through edge-constrained boundary preservation; (3) Feature dissimilarity integration provides 16.5%/14.3% Dice improvements and 19.9%/16.8% AUPRC improvements (row 4 vs 3) by capturing anatomical anomalies undetectable through pixel-level analysis alone. These findings demonstrate two fundamental advantages: (i) The CSCAL-AECM synergy ensures robust anatomical consistency across sequences; (ii) Dual-level (pixel and feature) anomaly mapping provides complementary detection sensitivity, that pixel errors quantify intensity deviations while feature dissimilarities identify abnormal patterns.

Table 2: Ablation study (train and test both on BraTS 2021).

Module			**T1ce→FLAIR**			**T1→FLAIR**		
$\mathcal{L}_{\text{CSCAL}}$	AECM	$\mathbf{A}^{feature}_{S_s \leftrightarrow S_t}$	Dice↑	AUROC↑	AUPRC↑	Dice↑	AUROC↑	AUPRC↑
×	×	✓	0.4409	0.9213	0.3242	0.4508	0.9219	0.3369
×	✓	✓	0.4685	0.9474	0.3529	0.4714	0.9462	0.3518
✓	✓	×	0.4357	0.9350	0.3355	0.4439	0.9371	0.3364
✓	✓	✓	**0.5080**	**0.9517**	**0.4022**	**0.5075**	**0.9509**	**0.3929**

Bold values indicate the best performance. $\mathcal{L}_{\text{CSCAL}}$: Cross-sequence contrastive alignment loss. AECM: Anatomical Edge Constraint Module. $\mathbf{A}^{feature}_{S_s \leftrightarrow S_t}$: Feature dissimilarity map.

Qualitative Evaluation Figure 3 illustrates qualitative results through three representative cases. For small lesions (row 1), DAE, f-AnoGAN, AnoDDPM and Cyclic UNet generate false negatives by perfectly reconstructing anomalies. VAE's poor reconstruction introduces false positives (yellow arrows). In contrast, MultiTransAD successfully translates the lesions with normal representations and highlight the anomaly. For medium lesions (row 2), most methods including our pixel-level method produce false positives due to imperfect reconstruction/translation in normal regions, leading to decreased specificity. Our feature dissimilarity maps (orange outlines in the final column) effectively highlight the anomaly regions and reduce the false positives. For large lesions (row 3), MultiTransAD preserves normal tissue structure while corrupting anomaly regions during translation. The qualitative results further demonstrate the effectiveness of the proposed model and the dual-level anomaly detection paradigm.

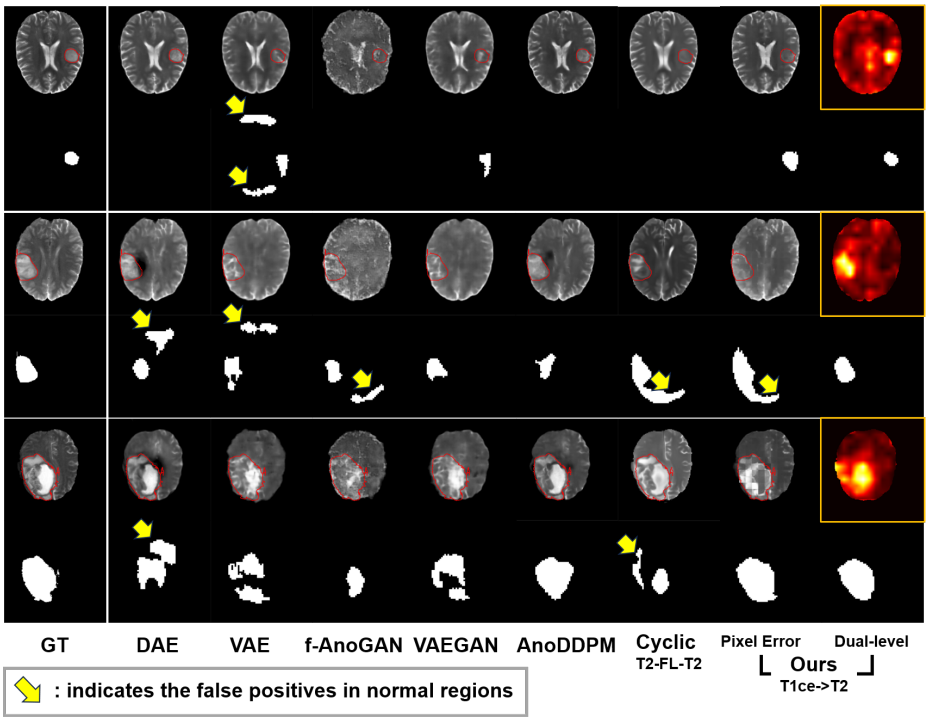

Fig. 3: **Column 1**: Reference T2 images with radiological annotations (red contours). **Columns 2-7**: Comparative results from baselines, including reconstructed/translated T2 images and predicted anomaly segmentation. **Columns 8-9**: MultiTransAD's detection process: (8) Pixel-level translation errors highlighting intensity discrepancies; (9) Dual-level anomaly segmentation combining feature dissimilarities (orange outlines) and pixel errors.

4 Conclusion and Future Work

We presented MultiTransAD, a cross-sequence translation-based framework for brain MRI anomaly detection. It utilizes inter-sequence contrast discrepancies as critical biomarkers. The framework's innovations, including disentangled content-style encoding, edge-constrained anatomical preservation, and dual-level anomaly evaluation, address the shortcomings of traditional reconstruction and cyclic-translation approaches. On the BraTS 2021 dataset, MultiTransAD outperformed state-of-the-art reconstruction methods, which demonstrates the promise of the cross-sequence translation-based UAD paradigm in medical lesion analysis. In the future, we plan to extend this paradigm to a wider range of pathologies (e.g., multiple sclerosis, stroke) and image modalities (CT/MR or CT/PET).

Acknowledgment. This work was supported in part by National Natural Science Foundation of China Grant 62471297, in part by the Ministry of Science and Technology of the People's Republic of China under Grant 2023YFC2411401, in part sponsored by the Fundamental Research Funds for the Central Universities (No.YG2025ZD03).

Disclosure of Interests. The authors have no competing interests to declare relevant to this article's conten.

References

1. Baid, U., Ghodasara, S., Mohan, S., Bilello, M., Calabrese, E., Colak, E., Farahani, K., Kalpathy-Cramer, J., Kitamura, F.C., Pati, S., Prevedello, L.M., Rudie, J.D., Sako, C., Shinohara, R.T., Bergquist, T., Chai, R., Eddy, J., Elliott, J., Reade, W., Schaffter, T., Yu, T., Zheng, J., Moawad, A.W., Coelho, L.O., McDonnell, O., Miller, E., Moron, F.E., Oswood, M.C., Shih, R.Y., Siakallis, L., Bronstein, Y., Mason, J.R., Miller, A.F., Choudhary, G., Agarwal, A., Besada, C.H., Derakhshan, J.J., Diogo, M.C., Do-Dai, D.D., Farage, L., Go, J.L., Hadi, M., Hill, V.B., Iv, M., Joyner, D., Lincoln, C., Lotan, E., Miyakoshi, A., Sanchez-Montano, M., Nath, J., Nguyen, X.V., Nicolas-Jilwan, M., Jimenez, J.O., Ozturk, K., Petrovic, B.D., Shah, C., Shah, L.M., Sharma, M., Simsek, O., Singh, A.K., Soman, S., Statsevych, V., Weinberg, B.D., Young, R.J., Ikuta, I., Agarwal, A.K., Cambron, S.C., Silbergleit, R., Dusoi, A., Postma, A.A., Letourneau-Guillon, L., Perez-Carrillo, G.J.G., Saha, A., Soni, N., Zaharchuk, G., Zohrabian, V.M., Chen, Y., Cekic, M.M., Rahman, A., Small, J.E., Sethi, V., Davatzikos, C., Mongan, J., Hess, C., Cha, S., Villanueva-Meyer, J., Freymann, J.B., Kirby, J.S., Wiestler, B., Crivellaro, P., Colen, R.R., Kotrotsou, A., Marcus, D., Milchenko, M., Nazeri, A., Fathallah-Shaykh, H., Wiest, R., Jakab, A., Weber, M.A., Mahajan, A., Menze, B., Flanders, A.E., Bakas, S.: The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification (2021)
2. Baur, C., Denner, S., Wiestler, B., Navab, N., Albarqouni, S.: Autoencoders for unsupervised anomaly segmentation in brain mr images: a comparative study. Medical Image Analysis **69**, 101952 (2021)
3. Cai, Y., Chen, H., Cheng, K.T.: Rethinking autoencoders for medical anomaly detection from a theoretical perspective. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 544–554. Springer (2024)

4. Ellingson, B.M., Bendszus, M., Boxerman, J., Barboriak, D., Erickson, B.J., Smits, M., Nelson, S.J., Gerstner, E., Alexander, B., Goldmacher, G., et al.: Consensus recommendations for a standardized brain tumor imaging protocol in clinical trials. Neuro-oncology **17**(9), 1188–1198 (2015)
5. Gu, Y., Lin, Y., Cheng, K.T., Chen, H.: Revisiting deep ensemble uncertainty for enhanced medical anomaly detection. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 520–530. Springer (2024)
6. Huang, H., Shen, Z., Wang, J., Wang, X., Lu, J., Lin, H., Ge, J., Zuo, C., Wang, Q.: Metaad: Metabolism-aware anomaly detection for parkinson's disease in 3d 18 f-fdg pet. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 291–301. Springer (2024)
7. Huang, X., Belongie, S.: Arbitrary style transfer in real-time with adaptive instance normalization. In: Proceedings of the IEEE international conference on computer vision. pp. 1501–1510 (2017)
8. Kascenas, A., Pugeault, N., O'Neil, A.Q.: Denoising autoencoders for unsupervised anomaly detection in brain mri. In: International Conference on Medical Imaging with Deep Learning. pp. 653–664. PMLR (2022)
9. Kascenas, A., Sanchez, P., Schrempf, P., Wang, C., Clackett, W., Mikhael, S.S., Voisey, J.P., Goatman, K., Weir, A., Pugeault, N., et al.: The role of noise in denoising models for anomaly detection in medical images. Medical Image Analysis **90**, 102963 (2023)
10. Kingma, D.P.: Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114 (2013)
11. Larsen, A.B.L., Sønderby, S.K., Larochelle, H., Winther, O.: Autoencoding beyond pixels using a learned similarity metric. In: International conference on machine learning. pp. 1558–1566. PMLR (2016)
12. Liang, Z., Anthony, H., Wagner, F., Kamnitsas, K.: Modality cycles with masked conditional diffusion for unsupervised anomaly segmentation in mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 168–181. Springer (2023)
13. Liang, Z., Guo, X., Noble, J.A., Kamnitsas, K.: Itermask 2: Iterative unsupervised anomaly segmentation via spatial and frequency masking for brain lesions in mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 339–348. Springer (2024)
14. Schlegl, T., Seeböck, P., Waldstein, S.M., Langs, G., Schmidt-Erfurth, U.: f-anogan: Fast unsupervised anomaly detection with generative adversarial networks. Medical image analysis **54**, 30–44 (2019)
15. Team, I.D.: Information extraction from images (ixi) dataset. http://brain-development.org/ixi-dataset/ (2006), accessed: 2023-08-20, Contains 600 healthy subjects' MRI scans (T1, T2, PD, MRA, DTI) from 3 London hospitals
16. Wyatt, J., Leach, A., Schmon, S.M., Willcocks, C.G.: Anoddpm: Anomaly detection with denoising diffusion probabilistic models using simplex noise. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. pp. 650–656 (June 2022)