Discriminative Directional Projection for Reverse Distillation in Breast Ultrasound Anomaly Detection

Shicheng Xu[1,2], Jincheng Li[1,2], Junhao Lin[3], Zuoyong Li[4], Hengrong Lan[1,2⋆], and Fei Gao[1,2,5*]

[1] School of Biomedical Engineering, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui, China
[2] Hybrid Imaging System Laboratory, Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou, Jiangsu, China
[3] College of Computer and Data Science, Fuzhou University, Fuzhou 350108, China
[4] Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, School of Computer and Big Data, Minjiang University, Fuzhou 350121, China
[5] School of Engineering Science, University of Science and Technology of China, Hefei, Anhui, China
xushicheng@mail.ustc.edu.cn, {hrlan,fgao}@ustc.edu.cn

Abstract. Breast ultrasound is widely used for screening, yet accurate annotations are costly, motivating anomaly detection methods trained only on normal data. Reverse Distillation (RD) is a popular teacher–student paradigm, but most RD approaches enforce global cosine consistency during training while relying on pixel-wise discrepancies at test time, leading to an objective mismatch. Moreover, directly optimizing cosine similarity is prone to poor local optima, which reduces anomaly sensitivity. To address these issues, we propose the Discriminative Directional Projection Loss (DDP Loss). First, we introduce a directional subspace projection mechanism, where a sampler generates directional bases along the channel dimension to project teacher and student features into a directional subspace. Pixel-wise consistency distillation is then enforced within this subspace to better align the training and testing objectives. Second, we design a Hard-Direction Mining strategy that alternates optimization to actively search for directions with the largest teacher–student discrepancy, thereby strengthening discriminative constraints. Third, we incorporate an orthogonality regularizer to promote directional diversity, mitigating direction collapse and improving subspace representational capacity. Experiments on BUS-UCLM and BUSI show that DDP Loss consistently improves AUC, F1, and FPR@TPR95 without extra architectures or pseudo anomalies. These results validate the effectiveness of improving the reverse distillation paradigm at the objective level and highlight its potential clinical value.

Keywords: Reverse Distillation · anomaly detection · breast ultrasound · directional subspace projection · hard-direction mining.

⋆ Corresponding authors: Hengrong Lan and Fei Gao.

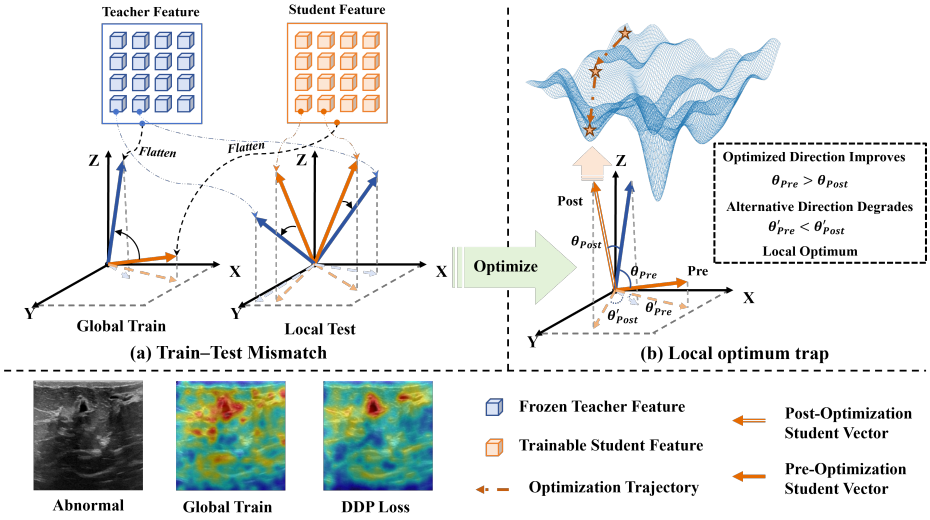


Fig. 1. Illustration of two limitations of reverse distillation (RD) and the improvements introduced by the Discriminative Directional Projection Loss (DDP Loss). (a) Training enforces global alignment on flattened features, whereas testing constructs anomaly maps via pixel-wise cosine discrepancy, leading to an objective mismatch. (b) Direct cosine optimization often follows the steepest update direction and can converge to poor local optima. DDP Loss performs pixel-wise alignment in a directional projected space and introduces discriminative constraints, reducing misclassification risk.

1 Introduction

Breast ultrasound is widely used for screening and follow-up because it involves no ionizing radiation, is highly repeatable, and provides complementary information for dense breasts [2]. However, speckle noise, low contrast, and substantial morphological variability make reliable annotation costly [3], limiting the clinical deployability of strongly supervised detection methods in real-world settings [16]. Anomaly detection learns the distribution of normal data using only normal samples and, during testing, localizes suspicious regions by measuring the degree of deviation[19]. This setting better matches clinical reality, where abnormal cases are difficult to obtain and fine-grained annotations are costly to produce [21,10]. In recent years, reverse distillation paradigms have shown strong transferability and interpretability in medical imaging by freezing the teacher network and training a student to match multi-scale representations, such that teacher–student discrepancies naturally serve as evidence of anomalies[17].

In reverse distillation (RD)-based anomaly detection, recent studies mainly improve architecture design, multi-scale robustness, and optimization stability. RD4AD [5], Skip-TS [14], and SCRD4AD [12] improve pixel-level localization via multi-scale reconstruction, skip-connected multi-scale constraints, and scale-adaptive contrastive distillation with synthetic anomalies, respectively.

Meanwhile, ReContrast [8] and EDC [7] improve training stability and target-domain adaptation via stop-gradient and similarity-based end-to-end optimization, and UniNet [17] enhances robustness through domain-relevant feature selection and contrastive constraints. Despite these advances, existing reverse distillation methods in Fig. 1 still suffer from two limitations. First, training optimizes global similarity on flattened features. In contrast, testing relies on pixel-wise discrimination for anomaly detection, and this objective mismatch can cause detection errors. Second, directly optimizing cosine similarity tends to follow the steepest descent direction and is more likely to converge to poor local optima. To address these issues, we propose the Discriminative Directional Projection-Loss (DDP Loss), which aligns the objectives between training and testing and improves the optimization trajectory. Specifically, a learnable sampler generates channel-wise projection directions to map teacher and student features into a directional subspace, where pixel-wise cosine consistency is enforced to match the testing formulation. We further adopt an alternating optimization strategy: we first fix the directions to update the student, and then fix the student to update the sampler for hard-direction mining. This prevents the optimization from staying on easily aligned directions that can induce suboptimal solutions. Finally, an orthogonality regularizer promotes directional diversity, avoiding collapse and redundancy, enhancing subspace expressiveness, and strengthening discriminative anomaly representations.

2 Method

Reverse distillation freezes the teacher and trains a student on normal data to match multi-level features, using teacher–student discrepancies for anomaly detection. However, most methods optimize a global cosine loss on flattened features, causing a train–test mismatch and often converging to poor local optima. To address these issues, as shown in Fig. 2, we propose the DDP Loss. A learnable sampler generates channel-wise projection directions $\mathbf{D} \in \mathbb{R}^{K\times C}$ to project features $f \in \mathbb{R}^{C\times H\times W}$ into directional responses $\mathbf{P} \in \mathbb{R}^{K\times H\times W}$, where pixel-wise cosine consistency is enforced to align training and testing objectives. We optimize DDP Loss with a two-step alternating scheme. **Step 1:** Fix the projection directions and update the student to maximize pixel-wise consistency in the projected space. **Step 2:** Fix the student and update the sampler for hard-direction mining, with an orthogonality regularizer to encourage diverse directions and prevent collapse. The following sections detail directional projection, pixel-wise consistency distillation, and hard-direction mining.

2.1 Directional Subspace Projection

To better match pixelwise anomaly scoring at inference, we shift the distillation objective from flattened global cosine alignment to pixelwise local discrimination [22]. We propose Directional Subspace Projection, where a sampler encodes channel features into multiple directional responses to amplify subtle discrepancies

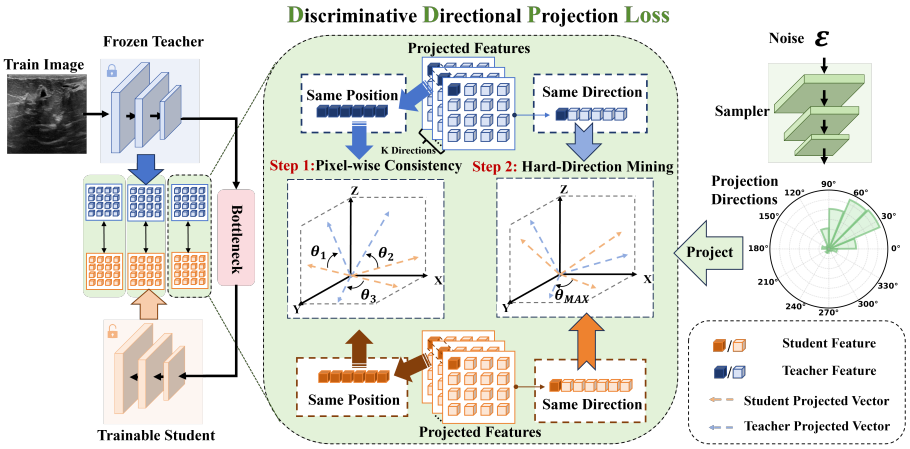


Fig. 2. The two-step alternating optimization strategy of the DDP Loss. Step 1: Fix the projection directions and update the student to maximize pixel-wise consistency in the projected space. Step 2: Fix the student and update the sampler to perform hard-direction mining. This strategy enables more stable distillation alignment in the directional subspace and mitigates poor local optima caused by overly easy-to-align directions.

across directions and reduce noise dominated alignment [11,4]. Let the teacher and student features at the l-th distillation layer be $\mathbf{F}_l^t, \mathbf{F}_l^s \in \mathbb{R}^{C_l \times H_l \times W_l}$, where, C_l is the number of channels and H_l, W_l denote the spatial resolution. The sampler outputs a set of directional bases $\mathbf{D}_l = [\mathbf{d}_{l,1}; \ldots; \mathbf{d}_{l,K}] \in \mathbb{R}^{K \times C_l}$, where each direction vector $\mathbf{d}_{l,k} \in \mathbb{R}^{C_l}$ is ℓ_2-normalized to eliminate scale effects. Given D_l, we define the directional projection operator $\Pi(\cdot; \mathbf{D}_l)$, which projects the channel vector at each spatial location onto K directions:

$$\mathbf{P}_l = \Pi(\mathbf{F}_l; \mathbf{D}_l) \in \mathbb{R}^{K \times H_l \times W_l}, \quad \mathbf{P}_l[k,h,w] = \sum_{c=1}^{C_l} \mathbf{D}_l[k,c]\mathbf{F}_l[c,h,w], \tag{1}$$

where $\mathbf{F}_l$ can be either $\mathbf{F}_l^t$ or $\mathbf{F}_l^s$. Intuitively, this applies the same set of directions to compute K weighted inner products of the channel features at each pixel for both teacher and student, thereby preserving spatial structure while compressing the original C_l-dimensional representation into a directional subspace $\mathbf{P}_l^t$, $\mathbf{P}_l^s$ that is more suitable for local alignment and subsequent anomaly map construction.

2.2 Pixel-wise Consistency Distillation

RD aligns flattened features via a global similarity objective [6,20]. In contrast, anomaly detection typically relies on pixel-wise anomaly maps at testing [5], which leads to a mismatch between training signals and test-time discrimination.

Using the projected representations $\mathbf{P}_l^t$, $\mathbf{P}_l^s$, we fix the teacher parameters θ_t and sampler parameters θ_D, and update only the student parameters θ_s to maximize consistency within the projected subspace. For a normal sample $x \sim \mathcal{X}_{\text{normal}}$, we compute $\mathbf{F}_l^t(x), \mathbf{F}_l^s(x)$ and obtain $\mathbf{P}_l(x) = \mathit{\Pi}\left(\mathbf{F}_l(x); \mathbf{D}_l(\theta_D)\right) \in \mathbb{R}^{K \times H_l \times W_l}$. We then compute the pixel-wise cosine similarity

$$\mathbf{S}_l(x)[h,w] = \cos\left(\mathbf{P}_l^s(x)[:,h,w], \mathbf{P}_l^t(x)[:,h,w]\right), \mathbf{E}_l(x) = 1 - \mathbf{S}_l(x). \tag{2}$$

The student-stage objective is defined as

$$\min_{\theta_s} \mathcal{L}_{\text{PCD}}(\theta_s; \theta_D) = \frac{1}{|\mathcal{L}|} \sum_{l \in \mathcal{L}} \mathbb{E}_x \left[\mathcal{A}_\rho\left(\mathbf{E}_l(x)\right)\right], \tag{3}$$

where $\mathcal{A}_\rho(\cdot)$ is a robust spatial aggregation operator. In our implementation, $\mathcal{A}_\rho$ selects the bottom-k pixel errors according to a ratio ρ to suppress noise and locally unstable points. Importantly, gradients in this stage are back-propagated only to θ_s, while $\mathbf{D}_l(\theta_D)$ is applied with stop-gradient.

At test time, we upsample each discrepancy map $\mathbf{E}_l(x)$ to the input resolution and average over $l \in \mathcal{L}$ to obtain the aggregated response $\mathbf{M}(x)$. We then apply Gaussian smoothing with kernel size 4 and define the anomaly score as the maximum response:

$$\mathbf{M}(x) = \frac{1}{|\mathcal{L}|} \sum_{l \in \mathcal{L}} \text{Up}\left(\mathbf{E}_l(x)\right), \quad s(x) = \max_{h,w} \mathcal{G}_4(\mathbf{M}(x))\,[h,w]. \tag{4}$$

where $\text{Up}(\cdot)$ denotes upsampling to the input resolution and $\mathcal{G}_4(\cdot)$ denotes Gaussian smoothing with kernel size 4.

2.3 Hard-Direction Mining

Although pixel-wise consistency in the directional subspace aligns the training and testing objectives, the distillation quality still depends heavily on the response subspace covered by D_l. If D_l mostly lies in easy-to-align directions, the student can quickly reduce the loss early in training, leading to insufficient discriminative pressure and convergence to suboptimal solutions. To address this, we introduce a second stage where we fix θ_t and θ_s and update only the sampler parameters θ_D, so that D_l actively mines the most inconsistent hard directions between teacher and student. For $x \sim \mathcal{X}_{\text{normal}}$ and $l \in \mathcal{L}$, we obtain $\mathbf{P}_l^t(x), \mathbf{P}_l^s(x) \in \mathbb{R}^{K \times H_l \times W_l}$,, and flatten each directional response spatially $\mathbf{p}_{l,k}(x) = \text{Flatten}(\mathbf{P}_l(x)[k,:,:]) \in \mathbb{R}^{H_l W_l}$. We define the sampler-stage hardness objective as the mean directional cosine distance:

$$\mathcal{L}_{\text{HDM}}(\theta_D) = \frac{1}{|\mathcal{L}|} \sum_{l \in \mathcal{L}} \mathbb{E}_x \left[\frac{1}{K} \sum_{k=1}^{K} \left(1 - \cos\left(\mathbf{p}_{l,k}^s(x), \mathbf{p}_{l,k}^t(x)\right)\right)\right]. \tag{5}$$

To prevent direction collapse (i.e., multiple directions degenerating into highly correlated redundant bases), we add an orthogonality regularizer [9]. After ℓ_2-normalizing the rows of D_l to obtain $\hat{\mathbf{D}}_l$, we impose a soft inter-direction decorrelation regularizer on the normalized projection directions:

$$\mathcal{R}_{\text{orth}}(\theta_D) = \frac{1}{|\mathcal{L}|} \sum_{l \in \mathcal{L}} \left\| \hat{\mathbf{D}}_l(\theta_D) \hat{\mathbf{D}}_l(\theta_D)^\top - \mathbf{I} \right\|_F^2 . \tag{6}$$

The sampler objective is then $\max_{\theta_D} \mathcal{L}_{\text{sam}}(\theta_D) = \mathcal{L}_{\text{HDM}}(\theta_D) - \lambda_{\text{orth}} \mathcal{R}_{\text{orth}}(\theta_D)$, where λ_{orth} controls the regularization strength. In practice, we detach the teacher and student features and update the sampler by back-propagating $-\mathcal{L}_{\text{sam}}$.

3 Experiment

3.1 Datasets and Implementation Details

We evaluate on two anomaly detection splits constructed from public breast ultrasound datasets, BUSI [1] and BUS-UCLM [15]. **BUSI** was collected in 2018 at Baheya Hospital for Early Detection and Treatment of Women's Cancer, Cairo, Egypt, from female patients aged 25 to 75, and was released as an anonymized dataset with institutional ethical approval. It contains 780 images, including 133 normal, 437 benign, and 210 malignant cases. **BUS-UCLM** was collected at Hospital General Universitario de Ciudad Real during 2022-2023 using a Siemens ACUSON S2000 system, with approval under PID2021-127567NB-I00. It contains 683 anonymized images from 38 subjects, including 419 normal, 174 benign, and 90 malignant cases. In our anomaly detection setting, normal breast ultrasound images are treated as the normal class, while all lesion-containing images, including benign and malignant cases, are treated as anomalies. We report AUC, F1-score (F1), and Accuracy (ACC), and implement all experiments in Python 3.8.20 on Ubuntu 20.04.5 LTS with 3 RTX 4090 GPUs, using batch size 32 for 2,000 epochs. We set the number of projection directions to K=4,096 and attach a lightweight sampler to each multi-scale feature level with channels 256,

Table 1. Quantitative comparison on the BUS-UCLM and BUSI datasets, with the **best** and second-best results highlighted.

Method	BUS-UCLM			BUSI		
	AUC (%) ↑	F1 (%) ↑	ACC (%) ↑	AUC (%) ↑	F1 (%) ↑	ACC (%) ↑
RD4AD [5]	83.52	83.93	77.39	81.97	98.70	97.48
UniNet [17]	82.32	83.44	73.40	79.96	98.55	97.19
Skip-TS [14]	85.42	78.76	74.47	**84.88**	68.02	53.33
SCRD4AD [12]	64.82	83.28	72.87	62.38	97.88	95.85
EDC [7]	79.06	76.33	**83.30**	79.87	96.89	96.48
ReContrast [8]	83.60	83.96	76.33	83.09	98.70	97.48
Ours	**87.50**	**88.32**	82.98	84.41	**98.85**	**97.78**

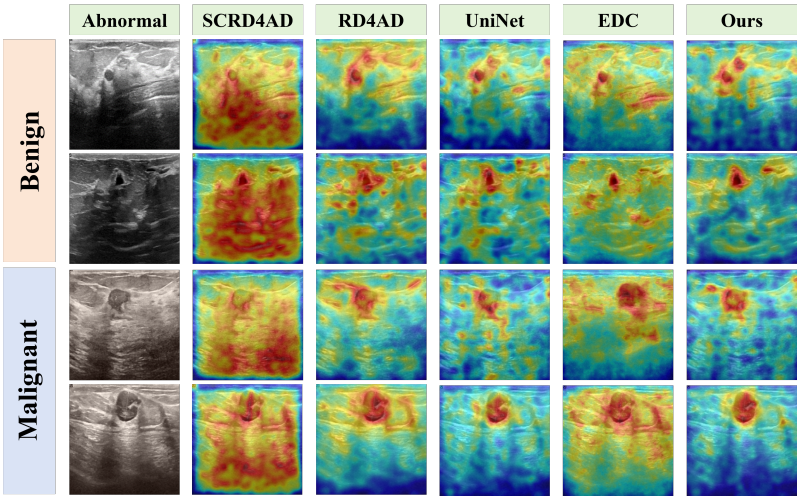


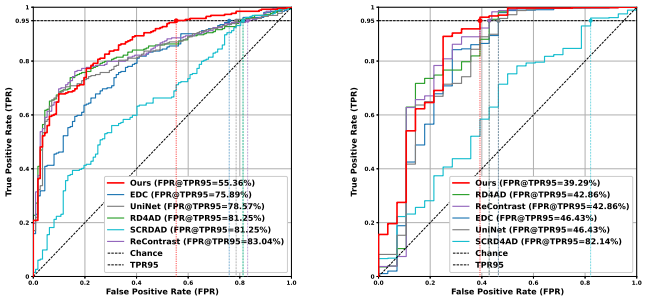
Fig. 3. Visualization comparisons of breast ultrasound anomaly detection methods on benign and malignant nodules.

512, and 1,024. The sampler is a 3-layer MLP consisting of Linear, BN, and LeakyReLU layers, with a final Tanh layer to generate channel-wise projection directions from Gaussian noise. We set $\lambda_{\text{orth}} = 0.1$ and adopt an alternating optimization schedule with 10 student updates followed by 1 sampler update. The spatial aggregation operator $\mathcal{A}_\rho(\cdot)$ uses ρ=70%.

3.2 Quantitative and Qualitative Analysis

Quantitative Analysis: As shown in Table 1, we compare ours with six representative methods on BUS-UCLM and BUSI. On BUS-UCLM, our method achieves the best AUC and F1 of 87.50% and 88.32%, while remaining competitive in ACC. On BUSI, we obtain the highest F1 and ACC of 98.85% and 97.78%, and we also maintain a competitive AUC, indicating strong overall detection performance. Although Skip-TS attains a slightly higher AUC on BUSI, its much lower F1 and ACC imply less reliable threshold-based decisions. In contrast, our method is more balanced across metrics. To isolate the effect of the objective, we only replace the loss in the RD4AD backbone and keep the rest of the architecture and the test-time pipeline unchanged. Aside from an extremely lightweight sampler used to implement the proposed loss, the improvements mainly stem from the optimized training objective.

Qualitative analysis: To qualitatively evaluate anomaly detection on ultrasound images, we compare anomaly heatmaps for benign and malignant samples in Fig. 3. Speckle noise, low contrast, and blurred boundaries often cause background over-activation and spurious highlights. Compared with previous methods, our DDP Loss produces cleaner heatmaps with reduced background activations, enabling more reliable anomaly discrimination.



(a) BUS-UCLM (b) BUSI

Fig. 4. ROC curves comparisons of different methods on BUS-UCLM and BUSI. At TPR 0.95, the value at the intersection between each ROC curve and the black horizontal line indicates the corresponding FPR, denoted as FPR@TPR95.

Table 2. We evaluate the impact of different loss components on performance on BUS-UCLM.

$\mathcal{L}_{\text{PCD}}$	$\mathcal{L}_{\text{HDM}}$	$\mathcal{R}_{\text{orth}}$	AUC (%)	F1 (%)	ACC (%)
			83.52	83.93	77.39
✓			84.03	84.48	78.99
✓	✓		84.81	86.43	81.38
✓	✓	✓	87.50	88.32	82.98

3.3 Ablation Studies

Component Ablation: To verify the contribution of each DDP Loss component, we perform a stepwise ablation on BUS-UCLM and report the results in Table 2. Starting from the RD4AD baseline, adding $\mathcal{L}_{\text{PCD}}$ improves performance consistently, indicating that enforcing pixel-wise consistency in the directional projection space aligns the training and testing objectives. Introducing $\mathcal{L}_{\text{HDM}}$ yields further gains, suggesting that hard-direction mining provides more discriminative alignment signals. With the orthogonality regularizer $\mathcal{R}_{\text{orth}}$, the model achieves the best overall performance and reaches an AUC of 87.50%, showing that promoting directional diversity effectively avoids direction collapse and strengthens anomaly detection.

False-positive analysis: To meet the high-sensitivity requirement in clinical screening, we report FPR@TPR95, the false positive rate when the true positive rate is fixed at 0.95, which quantifies false alarms under minimized missed detections [18,13]. On the ROC curve, it is read at TPR equals 0.95, where a lower value indicates fewer unnecessary follow-up examinations. As shown in Fig. 4, DDP Loss achieves the lowest FPR@TPR95 on BUS-UCLM and BUSI, reaching 55.36% and 39.29%, demonstrating stronger false-positive suppression and better clinical practicality.

4 Conclusion

We propose the DDP Loss to address two core limitations of reverse distillation for breast ultrasound anomaly detection, namely the Train-Test Mismatch and the local optimum trap. DDP Loss projects teacher and student features into a directional subspace and enforces pixel-wise consistency to better align the training and testing objectives, while hard-direction mining and orthogonality regularization encourage discriminative and diverse directions for more robust anomaly cues. Extensive experiments and ablations demonstrate consistent improvements over representative baselines across key metrics.

Acknowledgments

This work was supported in whole, or in part, by the Suzhou Innovation and Entrepreneurship Leading Talent Program (ZXL2025307), the National Natural Science Foundation of China (62401323, 62471207), the Natural Science Foundation of Fujian Province (2024J02029), and the Central Government Guided Local Science and Technology Development Fund Project (2024L3014).

Disclosure of Interests

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

References

1. Al-Dhabyani, W., Gomaa, M., Khaled, H., Fahmy, A.: Dataset of breast ultrasound images. Data in Brief **28**, 104863 (2020)
2. Berg, W.A., Rafferty, E.A., Friedewald, S.M., Hruska, C.B., Rahbar, H.: Screening algorithms in dense breasts: Ajr expert panel narrative review. American Journal of Roentgenology **216**(2), 275–294 (2021)
3. Chen, G., Li, L., Zhang, J., Dai, Y.: Rethinking the unpretentious u-net for medical ultrasound image segmentation. Pattern Recognition **142**, 109728 (2023)
4. Dehaene, D., Frigo, O., Combrexelle, S., Eline, P.: Iterative energy-based projection on a normal data manifold for anomaly localization. In: International Conference on Learning Representations (2020)
5. Deng, H., Li, X.: Anomaly detection via reverse distillation from one-class embedding. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 9737–9746 (2022)
6. Fang, Q., Su, Q., Lv, W., Xu, W., Yu, J.: Boosting fine-grained visual anomaly detection with coarse-knowledge-aware adversarial learning. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 16532–16540 (2025)
7. Guo, J., Lu, S., Jia, L., Zhang, W., Li, H.: Encoder-decoder contrast for unsupervised anomaly detection in medical images. IEEE Transactions on Medical Imaging **43**(3), 1102–1112 (2023)

8. Guo, J., Lu, S., Jia, L., Zhang, W., Li, H.: Recontrast: Domain-specific anomaly detection via contrastive reconstruction. Advances in Neural Information Processing Systems **36**, 10721–10740 (2023)
9. He, J., Du, J., Ma, W.: Preventing dimensional collapse in self-supervised learning via orthogonality regularization. Advances in Neural Information Processing Systems **37**, 95579–95606 (2024)
10. Kim, H., Wang, Y., Ahn, M., Choi, H., Zhou, Y., Hong, C.: Harnessing ehrs for diffusion-based anomaly detection on chest x-rays. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 235–245. Springer (2025)
11. Kim, K.H., Shim, S., Lim, Y., Jeon, J., Choi, J., Kim, B., Yoon, A.S.: Rapp: Novelty detection with reconstruction along projection pathway. In: International Conference on Learning Representations (2019)
12. Li, C., Shi, Y., Hu, J., Zhu, X.X., Mou, L.: Scale-aware contrastive reverse distillation for unsupervised medical anomaly detection. In: International Conference on Learning Representations (2025)
13. Liang, S., Li, Y., Srikant, R.: Enhancing the reliability of out-of-distribution image detection in neural networks. In: International Conference on Learning Representations (2018)
14. Liu, M., Jiao, Y., Lu, J., Chen, H.: Anomaly detection for medical images using teacher–student model with skip connections and multiscale anomaly consistency. IEEE Transactions on Instrumentation and Measurement **73**, 1–15 (2024)
15. Vallez, N., Bueno, G., Deniz, O., Rienda, M.A., Pastor, C.: Bus-uclm: Breast ultrasound lesion segmentation dataset. Scientific Data **12**(1), 242 (2025)
16. Wang, J., Qiao, L., Zhou, S., Zhou, J., Wang, J., Li, J., Ying, S., Chang, C., Shi, J.: Weakly supervised lesion detection and diagnosis for breast cancers with partially annotated ultrasound images. IEEE Transactions on Medical Imaging **43**(7), 2509–2521 (2024)
17. Wei, S., Jiang, J., Xu, X.: Uninet: A contrastive learning-guided unified framework with feature selection for anomaly detection. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 9994–10003 (2025)
18. Xia, Y., Zhang, Y., Liu, F., Shen, W., Yuille, A.L.: Synthesize then compare: Detecting failures and anomalies for semantic segmentation. In: European Conference on Computer Vision. pp. 145–161. Springer (2020)
19. Xu, S., Li, W., Li, Z., Zhao, T., Zhang, B.: Facing differences of similarity: Intra-and inter-correlation unsupervised learning for chest x-ray anomaly detection. IEEE Transactions on Medical Imaging **44**(2), 801–814 (2024)
20. Zhang, J., Suganuma, M., Okatani, T.: Contextual affinity distillation for image anomaly detection. In: Proceedings of the IEEE Winter Conference on Applications of Computer Vision. pp. 149–158 (2024)
21. Zhang, Q., Hu, Y., Sun, J.: Multitransad: Cross-sequence translation-driven anomaly detection in multi-sequence brain mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 389–398. Springer (2025)
22. Zhang, X., Li, S., Li, X., Huang, P., Shan, J., Chen, T.: Destseg: Segmentation guided denoising student-teacher for anomaly detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3914–3923 (2023)