# Injecting a Low-Rank Background Prior: A Plugin for Medical Image Segmentation

Minxin Chen[1], Jing Li[1], Zian Zhao[1], Duo Wai-Chi Wong[1,2], and James Chung-Wai Cheung[1,2,3⋆]

[1] Department of Biomedical Engineering, Faculty of Engineering
The Hong Kong Polytechnic University, Hong Kong SAR, China
[2] Research Institute for Sports Science and Technology
The Hong Kong Polytechnic University, Hong Kong SAR, China
[3] Research Institute of Smart Ageing
The Hong Kong Polytechnic University, Hong Kong SAR, China
james.chungwai.cheung@polyu.edu.hk

**Abstract.** Accurate segmentation of pathological anomalies is often hindered by their high variability and camouflage against anatomically complex backgrounds. While discriminative models struggle with these irregularities, we propose a perspective shift inspired by the radiologist's cognitive process akin to "Search-by-Subtraction." We learn a compact subspace from bottleneck background features and use the resulting reconstruction residual energy as an empirical feature-space cue for potential lesions. To leverage this, we introduce the Low-Rank Background Prior (LRBP) Plugin, which isolates anomalies by measuring deviations from a reconstructed background (BG) subspace using a training-time Global Dynamic Memory Bank updated via hard-example mining. Furthermore, to handle varying lesion ambiguity without amplifying noise, we introduce a Discriminability-Guided Gate Supervision mechanism. Here, a lightweight gate predictor learns to adaptively gate residual injection from a ground-truth-derived separability target that estimates foreground–background discriminability during training. Extensive experiments on endoscopic (Kvasir-SEG) and ultrasound (BUSI) datasets demonstrate that our framework achieves strong performance, particularly in high-noise scenarios. Notably, our approach yields substantial reductions in Average Surface Distance, suggesting that modeling background regularity can facilitate precise boundary delineation.

**Keywords:** Medical Image Segmentation · Low-Rank Background Prior · Adaptive Gating · Subspace Learning.

## 1 Introduction

Accurate segmentation of anomalies (e.g., polyps, nodules, lesions) plays an important role in computer-aided diagnosis [14]. Compared with relatively structured normal anatomy, pathological findings often exhibit substantially higher

⋆ Corresponding author.

variability in size, shape, and appearance, and may be camouflaged or partially occluded, making them difficult to delineate reliably [25]. In endoscopic imaging, complex backgrounds and acquisition artifacts (e.g., specular highlights, bubbles, blur, and related occlusions) can further confound anomaly–background separation [2]. While Encoder–Decoder CNNs remain widely used backbones [18], with modern large-kernel designs like MedNeXt [19] improving efficiency, they can struggle to disentangle camouflaged noise without explicit priors. In parallel, Transformer variants (e.g., Swin-Unet [5, 15]) have emerged to model global dependencies [5, 15], offering improved context capture. However, regardless of the architectural paradigm, purely discriminative approaches may overlook the underlying regularity of the anatomical background, which is important for characterizing background patterns from which lesions may deviate.

To overcome these limitations, we look to human experts for inspiration. In clinical practice, radiologists do not merely "detect" lesions based on local features; they engage in a cognitive process akin to "Search-by-Subtraction" (i.e., detecting deviations from normal anatomy) [13, 17, 26]. Guided by the Gestalt principle of *Figure-Ground Segregation* [21], experts establish a mental model of normal anatomy (the background) and cognitively suppress repetitive patterns to isolate deviations (the lesion foreground). This insight suggests that robust segmentation may depend less on recognizing myriad lesion variations, and more on effectively filtering out known background regularities.

Bridging this cognitive mechanism with computation, we use a compact low-rank basis to model background regularities in the encoder's bottleneck feature space, rather than assuming raw images are intrinsically low-rank [4]. Unlike prior attempts that perform expensive test-time decomposition on temporal sequences [22], we argue that normal tissue representation should be learned globally across the entire training distribution. In such a global feature space, background features are encouraged to be well reconstructed by the learned basis, whereas lesions tend to yield larger projection residuals. Thus, this feature-space background prior highlights the elusive foreground (FG) by rejecting well-reconstructed background patterns.

To realize this, we propose the Low-Rank Background Prior (LRBP) Plugin. While low-rank subspace priors have been explored for background modeling in video subtraction [3] and medical reconstruction (e.g., MRI) [12], with reconstruction used as an auxiliary regularizer [16], our method differs. Instead of using reconstruction only as a loss or requiring expensive online per-sequence decomposition [22], we implement an explicit residual-guided feature refinement via a globally learned subspace. We utilize a *Global Dynamic Memory Bank* with *Hard-Example Mining* to capture challenging textures without online SVD latency. To handle varying ambiguity without amplifying noise, we introduce *Discriminability-Guided Gate Supervision*, where a lightweight gate predictor learns to adaptively gate residual injection using a ground-truth-guided separability target.

In summary, our contributions are threefold: (1) We translate the clinical "Search-by-Subtraction" logic into a computational framework using background

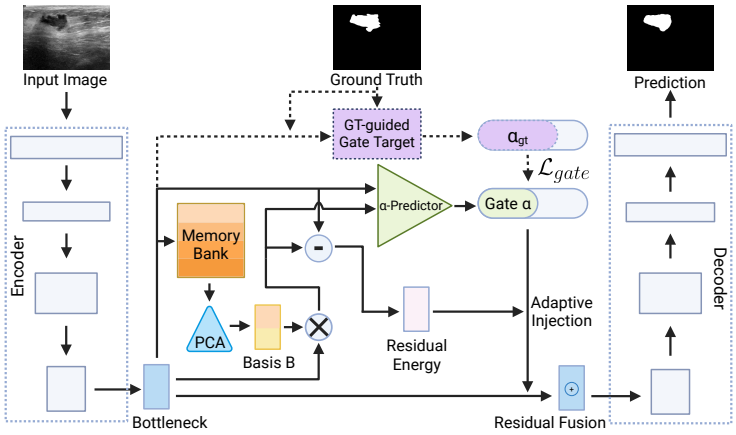



**Fig. 1. Proposed Framework.** LRBP models bottleneck background features with a memory-bank PCA basis and injects residual energy through a discriminability-guided gate. Dashed arrows denote training-only supervision; at inference, the bank and basis are frozen and no ground truth is used. Skip connections omitted for clarity.

regularity to isolate anomalies; (2) We propose an inference-time decomposition-free LRBP Plugin with hard-example mining for efficient, robust background modeling; and (3) We design a ground-truth-guided gate supervision mechanism to dynamically govern feature refinement during training. Extensive experiments on endoscopy (Kvasir-SEG [11]) and ultrasound (BUSI [1]) demonstrate consistent improvements over host backbones [18, 19, 5] and competitive performance against representative methods in high-noise scenarios.

## 2 Methodology

As shown in Fig. 1, we integrate LRBP Plugin into the encoder–decoder architecture: a global memory bank reconstructs the background subspace, residual energy highlights anomalies, and a supervision-guided gate controls residual injection. Specifically, the LRBP Plugin operates exclusively on the bottleneck features (i.e., the deepest encoder output), leaving skip connections intact. Therefore, the low-rank prior is imposed on learned feature representations rather than on raw image intensities.

### 2.1 Global Subspace Learning and Residual Extraction

We approximate bottleneck background features with a compact low-rank subspace and use poorly reconstructed regions as residual anomaly cues. To model this, we maintain a dataset-level Dynamic Memory Bank, $\mathcal{M} \in \mathbb{R}^{N \times C}$. In practice, each row of $\mathcal{M}$ stores a feature vector sampled from the bottleneck map at

ground-truth (GT) background locations (aligned via nearest-neighbor interpolation); we enqueue up to 256 vectors per iteration into a fixed-size bank. Principal Component Analysis (PCA) yields a rank-$K$ orthonormal basis $\mathbf{B} \in \mathbb{R}^{K\times C}$.

**Hard-Example Mining & Subspace Construction.** Instead of indiscriminate sampling, we employ an Image-Level Mining strategy to efficiently update the memory bank $\mathcal{M}$. We identify hard samples within each mini-batch by (i) low Dice, computed from sigmoid predictions binarized at 0.5 and (ii) for empty-mask images, large predicted foreground area ("hard-negative"); if no empty-mask images exist, we fill the quota with random background sampling. Background feature vectors are prioritized from these failure cases to capture challenging textures, supplemented by random samples to preserve global diversity. To maintain compactness, we enforce a *Redundancy-Based Replacement* policy: mined hard samples replace existing prototypes in $\mathcal{M}$ that are already well-reconstructed by the current basis. During training, $\mathbf{B}$ is a non-trainable buffer updated by PCA on $\mathcal{M}$ periodically (computed *gradient-free*); gradients only flow through the projection to $\mathbf{F}$. When the bank is full, we evict the most redundant prototypes (identified by the smallest reconstruction error $e(\mathbf{m}) = \|\bar{\mathbf{m}} - (\bar{\mathbf{m}}\mathbf{B}^\top)\mathbf{B}\|_2^2$, with $L_2$-normalized features $\bar{\mathbf{m}}$) rather than following a First-In-First-Out policy.

**Projection & Residual Extraction.** Given feature maps $\mathbf{F}$, we $L_2$-normalize each spatial vector to obtain $\bar{\mathbf{f}}_i$ and reconstruct the background via subspace projection: $\mathbf{f}_i^{bg} = (\bar{\mathbf{f}}_i\mathbf{B}^\top)\mathbf{B}$. The anomaly signal is defined as the deviation from this subspace, quantified as the Residual Energy map $E$:

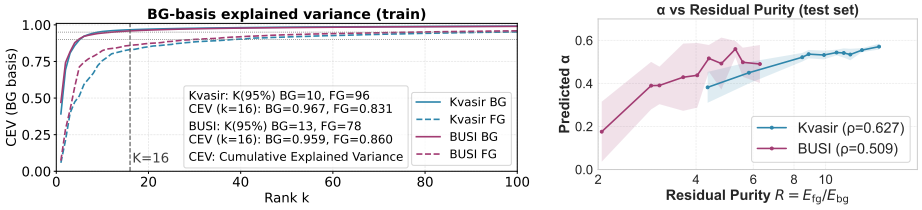
$$E_i = \|\bar{\mathbf{f}}_i - \mathbf{f}_i^{bg}\|_2^2 . \tag{1}$$

Intuitively, high $E_i$ indicates structures violating statistical regularity (potential lesions). During inference, $\mathcal{M}$ and $\mathbf{B}$ are frozen, with no GT masks, test-time PCA/SVD, or memory updates.

**Empirical support for the feature-space prior.** Fig. 2 (a) provides empirical support that the learned bottleneck BG features are compact enough for a PCA basis to be useful: a PCA basis learned from training BG features captures BG energy with a small rank (e.g., 95% energy at $K\approx$10–13), while FG requires a substantially higher rank under the same BG basis. Furthermore, pixel-wise residual energy $E$ effectively discriminates between FG and BG ($K$=16), achieving AUCs of 0.971 and 0.890 on Kvasir and BUSI, respectively. The median $E$ exhibits substantial shifts from BG to FG: 0.007 $\rightarrow$ 0.093 (Kvasir) and 0.016 $\rightarrow$ 0.065 (BUSI).

### 2.2 Discriminability-Guided Gate Supervision

Blindly injecting residuals risks amplifying noise in camouflaged regions where the low-rank feature-space prior may not be fully reliable. To address this, we propose a Discriminability-Guided Gate Supervision mechanism acting as a signal-to-noise gate. A GT-guided separability target evaluates foreground-background separability during training to supervise the injection gate: high separability promotes the prior ($\alpha \rightarrow 1$), while low separability suppresses it ($\alpha \rightarrow 0$), forcing reliance on backbone features to avoid noise contamination.




**(a)** Feature-space prior analysis. **(b)** Gating behavior analysis.

**Fig. 2. Mechanism analysis.** (a) A BG-learned PCA basis explains most BG bottleneck energy with a small rank. (b) Predicted gate $\alpha$ increases with residual purity $R$ (Spearman's $\rho$ is computed over positive test cases).

**GT-Guided Gate Target (Training Only).** Using GT, we construct the gate target by assessing sample difficulty via the cosine similarity between average foreground ($\mathbf{f}_{fg}$) and background ($\mathbf{f}_{bg}$) features. Here, $\mathbf{f}_{fg}$ and $\mathbf{f}_{bg}$ are obtained by masked average pooling on the same bottleneck feature map using GT (downsampled via nearest-neighbor interpolation). High similarity implies strong camouflage. This GT-guided target generates a conservative gate value $\alpha_{gt}$ to prevent artifact amplification:

$$\alpha_{gt} = \tanh(S_{scale} \cdot (1 - \text{CosineSim}(\mathbf{f}_{fg}, \mathbf{f}_{bg}))) \,, \tag{2}$$

where $S_{scale}$ is a scaling factor. Thus, $\alpha_{gt} \to 1$ only when the lesion is distinct. If an image has no foreground pixels, we set $\alpha_{gt}$=0.

**Gate Predictor.** Without GT at inference, a lightweight Gate Predictor estimates $\alpha$ using two global cues: Global Average Pooled features $\mathbf{f}_{gap}$ (derived by spatially averaging $\mathbf{F}$) and Global Reconstruction Similarity, defined as $s_{prior} = \text{CosineSim}(\mathbf{f}_{gap}, (\mathbf{f}_{gap}\mathbf{B}^\top)\mathbf{B})$. The gate is predicted via an MLP:

$$\alpha_{pred} = \text{Sigmoid}(\text{MLP}([\mathbf{f}_{gap}, s_{prior}])) \,, \tag{3}$$

where $s_{prior}$ implies how well the global context fits the normal subspace.

**Feature Injection.** Directly injecting the raw residual vector may introduce a mismatch with the encoder feature distribution and distort learned semantics. We therefore inject only the scalar residual energy as a projection-error field, preserving efficiency and architecture-agnostic compatibility. The residual energy $E$ is normalized to $\hat{E} \in [0, 1]$ via instance-level min-max scaling: $\hat{E} = (E - E_{\min})/(E_{\max} - E_{\min})$, and injected under the predicted gate:

$$\mathbf{F}_{out} = \mathbf{F} \odot (1 + \alpha_{pred}\hat{E}) \,. \tag{4}$$

### 2.3 Loss Functions

The objective combines segmentation loss with subspace constraints: $\mathcal{L}_{total} = \mathcal{L}_{seg} + 0.1(\mathcal{L}_{bg} + \mathcal{L}_{gate})$, where $\mathcal{L}_{seg} = \mathcal{L}_{bce} + \mathcal{L}_{dice}$. The weight 0.1 is chosen empirically to balance the gradient magnitudes.

**Background Reconstruction ($\mathcal{L}_{bg}$):** This term constrains the basis $\mathbf{B}$ to capture only normal tissue by minimizing residual energy strictly in background regions ($M_{gt} = 0$):

$$\mathcal{L}_{bg} = \frac{\sum_i (1 - M_{gt}^{(i)}) \cdot E_i}{\sum_i (1 - M_{gt}^{(i)})} \,. \tag{5}$$

**Gate Supervision ($\mathcal{L}_{gate}$):** The Gate Predictor learns the GT-guided gate target via Smooth L1 loss:

$$\mathcal{L}_{gate} = \text{SmoothL1}(\alpha_{pred}, \alpha_{gt}) \,. \tag{6}$$

## 3 Experiments and Results

*Datasets.* Primary: **Kvasir-SEG** [11] (1,000; 8:1:1 split) and **BUSI** [1] (780; 7:2:1 split). Auxiliary: **ISIC-2018** [6, 24] (2594; 8:1:1 split) and **TN3K** [8] (3493; official test set + 9:1 train/val) for generalization/ablation. We establish a fixed partition for each dataset and apply these respective splits consistently across all methods and ablations. All datasets are public and used following their licenses.

*Experimental setup and implementation details.* All methods are re-implemented with the same fixed seeds. **Hosts:** ResNet-34 U-Net [18, 9] (ImageNet initialization [7]) is the primary host; we additionally report MedNeXt-Small [19] and Swin-Unet [5, 15](ImageNet initialization), both with/without LRBP, on Kvasir-SEG and BUSI for architectural compatibility. Training protocols are unified (100 epochs, batch size 8) at $256^2$ resolution (Swin-Unet at $224^2$). Optimization: our frameworks use AdamW (learning rate: $10^{-3}$ for CNNs, $10^{-4}$ for Swin-Unet; weight decay: $10^{-4}$; cosine decay), whereas baselines adhere to their original optimization schemes. LRBP Settings: The memory bank utilizes $N$=10000 and $K$=16, with PCA updates every 10 iterations. We set $K$=16 as a conservative empirical rank, supported by the 95% BG explained variance at $K \approx$10–13 in Fig. 2(a) and the TN3K rank sweep in Table 1(b). Dice$<$ 0.5 is used only for training-time hard-positive mining, not inference. We empirically set the injection factor $S_{scale}$ based on backbone capacity: we use a conservative $S_{scale}$=1 for modern architectures or high-noise data (BUSI), and a stronger $S_{scale} \in [3, 5]$ for the ResNet-34 U-Net to maximize prior guidance. The plugin adds negligible overhead (+0.30% params, +0.01% GFLOPs). Metrics are Dice and Average Surface Distance (ASD) (via MedPy on binarized predictions; no post-processing).

### 3.1 Quantitative Comparison

**Performance on Hard Benchmarks.** As summarized in Table 1, our framework compares favorably against representative baselines. More importantly, it yields distinct improvements over the respective host backbones, particularly in boundary precision (ASD). For ResNet-34 U-Net, while the Dice scores remain comparable, the ASD is substantially reduced across both datasets (e.g.,

from 12.03 to **8.34** on Kvasir-SEG and 18.45 to **11.97** on BUSI). This suggests that our method effectively refines ambiguous boundaries where standard CNNs struggle, without compromising overall semantic overlap. For MedNeXt-Small, we observe comprehensive gains in both metrics: on Kvasir-SEG, Dice improves by **1.88%** with ASD dropping to **5.97**; similarly, on BUSI, ASD decreases by over **23%** (12.55 to **9.64**) alongside steady Dice improvements. For Swin-Unet, we observe clear gains on BUSI, and a modest ASD improvement on Kvasir-SEG with comparable Dice. These trends are consistent in the averaged results.

**Robustness to random seeds.** On BUSI, we repeated training of the primary ResNet-34 U-Net host with 5 fixed random seeds as a sensitivity analysis; ResNet-34 U-Net with LRBP Plugin consistently reduces ASD (baseline $17.81 \pm 6.41 \rightarrow 14.67 \pm 5.54$) while maintaining comparable Dice.

**Generalizability Check.** For the high-contrast ISIC task (Table 1 (a)), our method (with ResNet-34 U-Net as host) maintains parity with the baseline (e.g., 89.97% vs. 89.84% Dice), which confirms the plugin's selectivity: it activates to suppress complex background noise but preserves feature fidelity when the host's discriminative features are sufficiently effective.

**Ablation Study.** We conduct a controlled analysis on TN3K, the largest auxiliary dataset, using a fixed ResNet-34 U-Net, data split, and training protocol while varying only LRBP components (Table 1 (b)). Kvasir-SEG and BUSI remain the primary cross-backbone benchmarks. (1) Necessity of Adaptive Gating: Enforcing a Static Injection strategy (fixing $\alpha \equiv 0.5$) results in a performance drop (Dice: 85.44% $\rightarrow$ 83.33%). This confirms that a constant injection factor is insufficient; the model must dynamically attenuate the residual signal for noisy samples while amplifying it for camouflaged lesions. (2) Necessity of GT-Guided Gate Supervision: Removing the gate supervision loss ($\lambda_{gate} = 0$) degrades the performance to 82.89%. This indicates the Gate Predictor cannot learn an effective gating policy solely through implicit gradients from the segmentation loss. Explicit supervision from the GT-guided separability target is crucial for guiding the network to distinguish between helpful residuals and harmful noise. We further compare against widely-used SE block [10] and simple heuristics (rank $K$ sweeps, heuristic gate), which yield limited gains (Table 1 (b)).

**Gating behavior.** Since residual injection may amplify artifacts when background reconstruction is imperfect, our Gate Predictor learns to modulate the injection strength. On the held-out test set, the predicted gate $\alpha$ positively correlates with residual purity $R = E_{fg}/E_{bg}$ (Fig. 2 (b)), indicating more aggressive injection when residuals are foreground-dominant and suppression when background leakage is high. This behavior is consistent with Table 1 (b).

### 3.2 Qualitative Visualization

The "Search-by-Subtraction" effect is illustrated in Fig. 3. In the endoscopy case (Row 1), the baseline fails to separate the polyp from nearby mucosal folds, leading to over-segmentation. In contrast, the Residual Map ($E$) highlights the polyp as a high-energy deviation from the reconstructed background, allowing our model ($\alpha$=0.52) to trace the boundary more accurately. In the ultrasound

**Table 1. Quantitative results.** Kvasir-SEG and BUSI on three host backbones (baseline vs. +LRBP). Bold indicates improvement over the corresponding host. (a) ISIC sanity check. (b) Controlled TN3K component ablation using the same ResNet-34 U-Net host and training protocol.

| Method | Kvasir-SEG | | BUSI | | Average | |
|---|---|---|---|---|---|---|
| | Dice ↑ | ASD ↓ | Dice ↑ | ASD ↓ | Dice ↑ | ASD ↓ |
| *Host backbones (baseline vs. +LRBP)* | | | | | | |
| ResNet-34 U-Net [18, 9] | 88.94 | 12.03 | 80.47 | 18.45 | 84.71 | 15.24 |
| +LRBP (Ours) | **89.00** | **8.34** | **81.03** | **11.97** | **85.02** | **10.16** |
| MedNeXt-Small [19] | 87.95 | 7.61 | 79.95 | 12.55 | 83.95 | 10.08 |
| +LRBP (Ours) | **89.83** | **5.97** | **81.17** | **9.64** | **85.50** | **7.81** |
| Swin-Unet [5, 15] | 92.05 | 4.44 | 79.90 | 13.92 | 85.98 | 9.18 |
| +LRBP (Ours) | 91.98 | **4.01** | **81.43** | **7.68** | **86.71** | **5.85** |
| *Other representative methods* | | | | | | |
| CMUNeXt [23] | 85.96 | 7.07 | 80.59 | 13.74 | 83.28 | 10.41 |
| EGE-UNet [20] | 78.12 | 11.61 | 72.83 | 20.71 | 75.48 | 16.16 |

**(a) Sanity check on ISIC.**

| Method | ISIC | |
|---|---|---|
| | Dice ↑ | ASD ↓ |
| ResNet-34 U-Net | 89.84 | 5.38 |
| **Ours** | **89.97** | **5.15** |

**(b) Ablation on TN3K.**

| Config. | Dice ↑ | ASD ↓ |
|---|---|---|
| Base (ResNet-34 U-Net) | 84.17 | 8.96 |
| **Ours (Full)** | **85.44** | **6.35** |
| Static Injection ($\alpha \equiv 0.5$) | 83.33 | 8.49 |
| w/o GT-Guided Gate Supervision | 82.89 | 7.42 |
| Heuristic Gate ($\alpha = \text{clip}(1 - s_{\text{prior}})$) | 84.18 | 8.21 |
| Rank $K$=4 | 83.55 | 7.51 |
| Rank $K$=64 | 84.38 | 7.03 |
| ResNet-34 Unet-SE [18, 9, 10] | 84.12 | 6.99 |

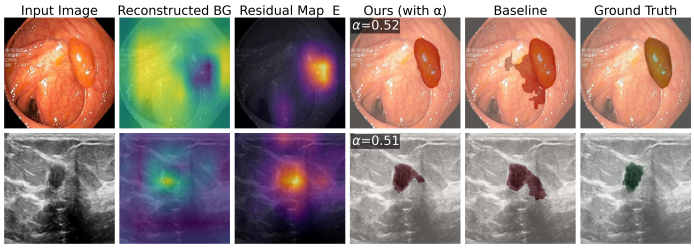



**Fig. 3. Qualitative results with ResNet-34 U-Net.** LRBP suppresses mucosal-fold over-segmentation in Kvasir-SEG and shadow-induced false positives in BUSI, yielding cleaner masks.

case (Row 2), the baseline is distracted by the acoustic shadow beneath the tumor and erroneously includes it as foreground. Our background reconstruction

absorbs the regular tissue pattern, leaving the tumor prominent in residual space; the injected prior ($\alpha$=0.51) then suppresses this shadow-induced noise, consistent with the ASD gains above.

## 4 Conclusion

In this work, we proposed a bio-inspired "Search-by-Subtraction" framework for medical anomaly segmentation. By using a compact background basis in bottleneck feature space, LRBP complements discriminative recognition with a residual-based "rejecting the known" cue. Experiments and gating analysis support its effectiveness in separating anomalies from complex backgrounds, while adaptive gate supervision improves robustness and preserves feature fidelity on high-contrast data. Future work will examine directional residual cues, annotation-noise robustness, broader plug-and-play comparisons, and statistical significance on larger datasets.

**Acknowledgments.** Figure 1 was created in BioRender. Polyu, B. (2026) `https://BioRender.com/6lneukm`.

**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. Ali, S., Dmitrieva, M., Ghatwary, N., Bano, S., Polat, G., Temizel, A., Krenzer, A., Hekalo, A., Guo, Y.B., Matuszewski, B., et al.: Deep learning for detection and segmentation of artefact and disease instances in gastrointestinal endoscopy. Medical image analysis **70**, 102002 (2021)
3. Bouwmans, T., Javed, S., Sultana, M., Jung, S.K.: Deep neural network concepts for background subtraction: A systematic review and comparative evaluation. Neural Networks **117**, 8–66 (2019)
4. Candès, E.J., Li, X., Ma, Y., Wright, J.: Robust principal component analysis? Journal of the ACM (JACM) **58**(3), 1–37 (2011)
5. Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., Wang, M.: Swin-unet: Unet-like pure transformer for medical image segmentation. In: European conference on computer vision. pp. 205–218. Springer (2022)
6. Codella, N.C., Rotemberg, V., Tschandl, P., Celebi, M.E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., et al.: Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). arXiv preprint arXiv:1902.03368 (2019), `https://arxiv.org/abs/1902.03368`
7. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 248–255 (2009). https://doi.org/10.1109/CVPR.2009.5206848

8. Gong, H., Chen, G., Wang, R., Xie, X., Mao, M., Yu, Y., Chen, F., Li, G.: Multi-task learning for thyroid nodule segmentation with thyroid region prior. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). pp. 257–261. IEEE (2021). https://doi.org/10.1109/ISBI48211.2021.9434087
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90
10. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 7132–7141. IEEE (2018)
11. Jha, D., Smedsrud, P.H., Riegler, M.A., Halvorsen, P., De Lange, T., Johansen, D., Johansen, H.D.: Kvasir-seg: A segmented polyp dataset. In: MultiMedia Modeling. pp. 451–462. Springer (2020)
12. Ke, Z., Huang, W., Cui, Z.X., Cheng, J., Jia, S., Wang, H., Liu, X., Zheng, H., Ying, L., Zhu, Y., et al.: Learned low-rank priors in dynamic mr imaging. IEEE Transactions on Medical Imaging **40**(12), 3698–3710 (2021)
13. Kundel, H.L., Nodine, C.F.: Interpreting chest radiographs without visual search. Radiology **116**(3), 527–532 (1975)
14. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Medical Image Analysis **42**, 60–88 (2017)
15. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 10012–10022 (2021)
16. Myronenko, A.: 3d mri brain tumor segmentation using autoencoder regularization. Lecture Notes in Computer Science pp. 311–320 (2019). https://doi.org/10.1007/978-3-030-11726-9_28
17. Nodine, C.F., Kundel, H.L.: The cognitive side of visual search in radiology, pp. 573–582. Elsevier, Amsterdam (1987)
18. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
19. Roy, S., Koehler, G., Ulrich, C., Baumgartner, M., Petersen, J., Isensee, F., Jaeger, P.F., Maier-Hein, K.H.: Mednext: transformer-driven scaling of convnets for medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 405–415. Springer (2023)
20. Ruan, J., Xie, M., Gao, J., Liu, T., Fu, Y.: Ege-unet: An efficient group enhanced unet for skin lesion segmentation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. Lecture Notes in Computer Science, vol. 14223, pp. 473–482. Springer, Cham (October 2023)
21. Rubin, E.: Synsoplevede figurer: studier i psykologisk analyse. Gyldendal, Nordisk forlag (1915)
22. Solomon, O., Cohen, R., Zhang, Y., Yang, Y., He, Q., Luo, J., van Sloun, R.J.G., Eldar, Y.C.: Deep unfolded robust pca with application to clutter suppression in ultrasound. IEEE Transactions on Medical Imaging **39**(4), 1051–1063 (2020). https://doi.org/10.1109/TMI.2019.2941271
23. Tang, F., Ding, J., Quan, Q., Wang, L., Ning, C., Zhou, S.K.: Cmunext: An efficient medical image segmentation network based on large kernel and skip fusion. In: IEEE International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE, Athens, Greece (May 2024). https://doi.org/10.1109/ISBI56570.2024.10635609

24. Tschandl, P., Rosendahl, C., Kittler, H.: The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific Data **5**, 180161 (2018). https://doi.org/10.1038/sdata.2018.161
25. Wang, H., Hu, T., Zhang, Y., Zhang, H., Qi, Y., Wang, L., Ma, J., Du, M.: Unveiling camouflaged and partially occluded colorectal polyps: Introducing cpsnet for accurate colon polyp segmentation. Computers in Biology and Medicine **171**, 108186 (2024)
26. Wolfe, J.M., Võ, M.L.H., Evans, K.K., Greene, M.R.: Visual search in scenes involves selective and nonselective pathways. Trends in Cognitive Sciences **15**(2), 77–84 (2011). https://doi.org/10.1016/j.tics.2010.12.001