SAGED-Net: Structural Adaptive Gated Encoder-Decoder Network for Nuclei Segmentation in Histopathology Images

Arko Dasgupta[*,§], Arjeesh Palai[*,§], Alexander Voznesensky[†], Dmitrii Kaplun[‡,§,†], Ram Sarkar[*]

[*]*Dept. of Computer Science & Engineering, Jadavpur University*, Kolkata, India
[†]*Department of Automation and Control Processes, St. Petersburg Electrotechnical University "LETI"*, St. Petersburg, Russia
[‡]*School of CS and Tech./ School of AI, China University of Mining and Technology*, Xuzhou, China
[§]*Artificial Intelligence Center, Skolkovo Institute of Science and Technology*, Moscow, Russia

Email: arkodasgupta05@gmail.com, arjeeshpalai@gmail.com, dikaplun@etu.ru, asvoznesenskiy@etu.ru, ramjucse@gmail.com,

***Abstract*—Standard deep learning models for nuclei segmentation in histopathology often struggle with overlapping boundaries while remaining computationally heavy and sensitive to noise. A key limitation in many existing architectures, such as U-Net, is their reliance on static, hard-wired skip connections that indiscriminately merge low-level noise with semantic features. To overcome this, we propose SAGED-Net, a lightweight and efficient architecture that replaces rigid connectivity with a learnable Adaptive Gating mechanism to actively suppress background artifacts. Our framework also introduces a Tri-Domain loss function to enforce anatomical accuracy in cell shapes and boundaries. Evaluated on various histopathological benchmarks, SAGED-Net consistently outperformed state-of-the-art methods, achieving notable segmentation accuracy and precise instance delineation on TNBC (Dice score: 0.9103), CPM-15 (Dice score: 0.9286), CPM-17 (Dice score: 0.9057), and PanNuke (Dice score: 0.9058), validating its utility as a robust method for computational pathology. *Further details are available at:* https://github.com/shadowbeast0/SAGED-Net**

***Index Terms*—Nuclei Segmentation, Histopathology Images, Adaptive Gating, Tri-Domain Loss, U-Net Model**

I. Introduction

A crucial stage in computational pathology is the segmentation of microscopic nuclei, which helps with tasks like survival analysis and tumor grading. High accuracy is crucial because mistakes at this stage impact all measurements downstream. However, because tissue appearance varies greatly across datasets, accurate boundary detection is still challenging. Models often rely on superficial visual patterns instead of true nuclear structure. Nuclei in high-grade cancers are closely spaced, overlap, and have weak boundaries, making it challenging to distinguish touching nuclei and frequently leading to biologically implausible outcomes.

While U-Net–based models are widely used, their fixed skip connections often pass along background noise and irrelevant details, weakening meaningful features. In addition, common overlap-based losses can give high scores even when nuclear boundaries are poorly separated. To overcome these issues, we propose **SAGED-Net**, which selectively passes useful structural information while suppressing noise. We also use a training strategy that encourages consistency across shape, texture, and spatial cues, leading to sharper boundaries and better separation. The **main contributions** of this work are:

- **Development of SAGED-Net:** We propose an architecture that integrates **Half-Instance Normalization (HIN)** blocks within the encoder to reduce sensitivity to appearance-related textures and encourage the network to focus on nuclear structure rather than intensity variations.
- **Adaptive Gating Mechanism:** A dynamic feature routing scheme is introduced at the encoder–decoder interface. By balancing **Cross-Stage Context** and **Residual Refinement**, it suppresses background noise while preserving fine boundary information.
- **Tri-Domain Fidelity (TDF) Loss:** We design a composite objective that extends beyond spatial supervision by combining **Focal Frequency Loss [1]** for recovering high-frequency details with a differentiable **Morphological Loss** to enforce shape compactness and separation in dense nuclear clusters.

II. Related Work

Nuclei segmentation remains vital for computer-aided diagnosis in digital pathology. Region-proposal methods are attractive because they generate instance-level outputs directly. NuKit [2], based on Mask-RCNN [3], shows this promise through click-based refinement. However, its anchor reliance falters with extremely large or small nuclei. The final segmentation quality depends heavily on the initial bounding box. Moreover, by processing each cell in isolation, the model fails to use neighbor information to separate overlapping nuclei.

Attention mechanisms introduce a smart idea: teaching models where to look. Similarly, Roy et al. [4] embeds wavelet attention and HARU-Net [5] embeds Convolutional Block Attention Module (CBAM) into Residual U-blocks to

[§] Authors contributed equally to this work.

clarify nuclear features. However, they cannot adapt to scene complexity; empty fields get the same heavy processing as dense clusters. This spreads capacity thin, rather than focusing intensity on ambiguous boundaries. While attention sharpens refinement, mixing strategies emphasize context. To build wide receptive fields, CNM-ED [6] utilizes ConvMixers, whereas DeepLabV3+ employ pyramidal pooling. However, this global view has a key flaw: loose mixing often blends nuclear forms with background artifacts in dim regions, ultimately ruining crisp boundary delineation.

Multi-task frameworks tackle different goals by sharing features. Approaches like HI-Net [7], Lin et al. [8], and GSN-HVNET [9] use auxiliary tasks or advanced feature fusion to efficiently predict masks. For instance, GSN-HVNET utilizes HV-maps (horizontal and vertical maps), while Lin et al. focuses on boundary mining. However, these gradient-based or boundary-aware signals often fail when nuclei overlap and gradients flatten. Furthermore, combining features from different levels without adjustment confuses semantics, which ultimately hurts fine-grained details.

However, there are some limitations of existing methods: the reliance on static and hard-wired skip connections blindly mix low-level noise with semantic features. Supervision usually lacks frequency and morphological constraints. To this end, SAGED-Net fixes this by swapping rigid links for an adaptive gating mechanism to route feature flow dynamically, and uses disentangled feature encoding to split stain from structure across spatial, spectral, and morphological dimensions.

III. Proposed Method

A. Overview of SAGED-Net

We propose *SAGED-Net*, a modified U-Net architecture designed to enhance feature separability and improve the separation of overlapping nuclei. Fig. 1 illustrates the overall framework. To handle inconsistent color intensity across datasets, the encoder incorporates HIN blocks. The network integrates two primary contributions:

- **Cross-Stage Adaptive Gating Mechanism:** A filtering layer is added to skip connections. It uses learnable weights to down-scale background noise, ensuring that the decoder receives clearer feature maps focused on nuclear boundaries.
- **Tri-Domain Fidelity Loss:** A combined loss function is used that optimizes for spatial, spectral, and morphological consistency, helping the model learn smoother and more anatomically correct segmentation masks.

B. Encoder Backbone: Disentangled Feature Encoding

The encoding pathway of SAGED-Net is designed to address a common challenge in computational pathology: the inconsistency of segmentation performance due to texture and color variations. Standard CNNs can be sensitive to differences in color intensity. To resolve this, we utilize the HIN block, which separates appearance information from structural content.

Let $\mathbf{X}_l \in \mathbb{R}^{C\times H\times W}$ denote the input feature map at stage l. The block splits the tensor along the channel dimension:

$$\mathbf{X}_l^{\alpha} = \mathbf{X}_l[:C/2], \quad \mathbf{X}_l^{\beta} = \mathbf{X}_l[C/2:] \tag{1}$$

The first half, $\mathbf{X}_l^{\alpha}$, undergoes Instance Normalization (IN) to filter out global style statistics, while $\mathbf{X}_l^{\beta}$ retains the original contextual information. These are concatenated and refined via Squeeze-and-Excitation (SE) attention to form the output $\tilde{\mathbf{X}}_l$:

$$\tilde{\mathbf{X}}_l = \mathrm{SE}\left(\mathcal{C}(\mathrm{IN}(\mathbf{X}_l^{\alpha}), \mathbf{X}_l^{\beta})\right) \tag{2}$$

where $\mathcal{C}(\cdot)$ denotes channel-wise concatenation. This ensures the network prioritizes shape-related features over intensity-dependent color statistics.

*1) **Adaptive Gating Mechanism**:* To dynamically balance the global context with local feature refinement, we introduce a learnable gating mechanism situated at the **encoder-decoder interface**. It regulates the injection of feature maps into the decoding stage.

For each stage l, we define a learnable scalar $\gamma_l \in \mathbb{R}$, initialized to $+5.0$ (even stages) and -5.0 (odd stages). A gating coefficient $\alpha_l \in (0,1)$ is derived via sigmoid:

$$\alpha_l = \sigma(\gamma_l) \tag{3}$$

We define two auxiliary gated paths based on this coefficient:

- **Cross-Stage Context ($\mathbf{C}_l$):** Features from the preceding encoder stage $\tau(l)$ are projected via W_{cross} and weighted by the complement of the gate:

$$\mathbf{C}_l = (1-\alpha_l)\cdot\left(W_{cross} * \mathbf{X}_{\tau(l)}^{enc}\right) \tag{4}$$

- **Residual Refinement ($\mathbf{R}_l$):** Features from the current encoder stage $\mathbf{X}_l^{enc}$ are projected via W_{post} and weighted by the gate α_l:

$$\mathbf{R}_l = \alpha_l \cdot (W_{post} * \mathbf{X}_l^{enc}) \tag{5}$$

*2) **Decoder with Soft-Gated Fusion**:* The decoder at stage l aggregates the feature maps. We construct the decoder input $\mathbf{Y}_{in}^l$ by summing the upsampled features from the previous stage ($\mathbf{D}_{in}$), the gated cross-stage context ($\mathbf{C}_l$), and a dual-path connection from the current encoder (comprising both raw features $\mathbf{X}_l^{enc}$ and their projection via Ψ_{pre}):

$$\mathbf{Y}_{in}^l = \mathbf{D}_{in} + \mathbf{C}_l + \underbrace{\mathbf{X}_l^{enc} + \Psi_{pre}(\mathbf{X}_l^{enc})}_{\text{Encoder Dual-Path}} \tag{6}$$

The fused representation $\mathbf{Y}_{in}^l$ is processed by the decoder block $\mathcal{F}_{dec}$. Finally, the gated residual refinement $\mathbf{R}_l$ is injected into the output to enforce semantic consistency:

$$\mathbf{D}_l = \mathcal{F}_{dec}(\mathbf{Y}_{in}^l) + \mathbf{R}_l \tag{7}$$

When $\alpha_l \to 1$, the residual path $\mathbf{R}_l$ dominates the output summation, prioritizing local feature preservation. Conversely, when $\alpha_l \to 0$, the network relies on the cross-stage context $\mathbf{C}_l$ integrated at the input level.

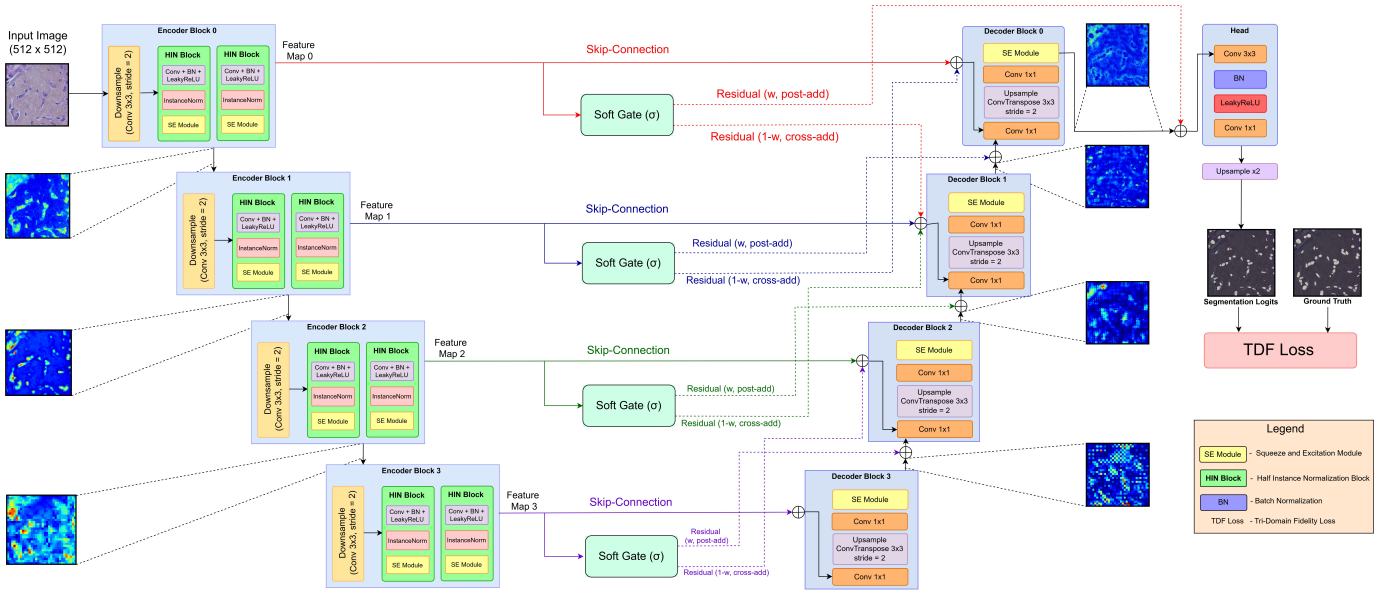

Fig. 1. The proposed SAGED-Net architecture. The framework features an encoder-decoder structure enhanced with HIN blocks and an Adaptive Gating mechanism that dynamically balances between two pathways: the **Residual Injection** (`post_add`), which preserves local spatial details via same-stage connections, and the **Cross-Stage Pathway** (`cross_add`), which injects global semantic context from deeper encoder layers into shallower decoder stages to resolve ambiguities.

C. *Tri-Domain Fidelity Loss*

Standard segmentation losses primarily optimize pixel-wise agreement in the spatial domain. However, nuclei boundaries contain high-frequency information that spatial losses fail to capture, and shape structures benefit from explicit topological constraints. To address this, we minimize a composite objective function that optimizes across spatial, spectral, and morphological domains:

$$\mathcal{L}_{\text{TDF}} = \lambda_{spatial} \cdot \mathcal{L}_{spatial} + \lambda_{spec} \cdot \mathcal{L}_{spec} + \lambda_{morph} \cdot \mathcal{L}_{morph} \quad (8)$$

where $\lambda_{spatial}$, λ_{spec}, and λ_{morph} are empirically determined weighting hyperparameters to balance the contributions of the primary segmentation objective and the regularization terms.

*1) **Spatial Domain Fidelity**:* The spatial component ($\mathcal{L}_{spatial}$) combines pixel-wise classification with region-based overlap optimization. We employ a summation of Cross-Entropy Loss ($\mathcal{L}_{\text{CE}}$) for probabilistic supervision and Dice Loss ($\mathcal{L}_{\text{Dice}}$) to mitigate class imbalance:

$$\mathcal{L}_{spatial} = \mathcal{L}_{\text{CE}}(\mathbf{Y}, \hat{\mathbf{Y}}) + \mathcal{L}_{\text{Dice}}(\mathbf{Y}, \hat{\mathbf{P}}) \quad (9)$$

where $\hat{\mathbf{Y}}$ represents the model logits, $\hat{\mathbf{P}}$ denotes the softmax-normalized probability maps, and $\mathbf{Y}$ is the ground truth. This combination ensures global segmentation consistency.

*2) **Spectral Domain Fidelity**:* To sharpen higher frequency details, we employ the **Focal Frequency Loss** [1] ($\mathcal{L}_{\text{Spec}}$). This loss operates in the frequency domain using the 2D real Fast Fourier Transform (2D Re-FFT). We minimize the distance between the spectrum of the predictions and the ground truth:

$$\mathcal{L}_{\text{Spec}} = \frac{1}{MN} \sum_{u,v} w(u,v) \cdot \left| \mathcal{F}(\hat{\mathbf{P}})(u,v) - \mathcal{F}(\mathbf{Y})(u,v) \right|^2 \quad (10)$$

where $\mathcal{F}(\cdot)$ denotes the 2D Re-FFT operation and $M \times N$ represents the spatial dimensions of the image. To prioritize hard-to-recover frequencies, a dynamic weight $w(u,v)$ scales the loss based on the current spectral error magnitude. It forces the network to refine edge details that are often smoothed out by standard spatial losses.

*3) **Morphological Domain Fidelity**:* To enforce structural consistency, we utilize a differentiable **Morphological Loss** ($\mathcal{L}_{\text{Morph}}$). We approximate the non-differentiable morphological erosion ($\ominus$) and dilation ($\oplus$) operations using min/max pooling with a fixed kernel size $\kappa = 5$, matching the approximate radius of nuclei in the dataset:

$$\text{Dilation:} \quad \hat{\mathbf{P}} \oplus \kappa \approx \text{MaxPool2d}(\hat{\mathbf{P}}, \kappa) \quad (11)$$

$$\text{Erosion:} \quad \hat{\mathbf{P}} \ominus \kappa \approx -\text{MaxPool2d}(-\hat{\mathbf{P}}, \kappa) \quad (12)$$

The loss calculates the Mean Squared Error (MSE) between the morphological representations of the predicted and ground truth masks. It is computed for each class k and averaged:

$$\mathcal{L}_{\text{Morph}} = \frac{1}{K} \sum_{k=1}^{K} \Big(\text{MSE}(\hat{\mathbf{P}}_k \ominus \kappa, \mathbf{Y}_k \ominus \kappa) + \text{MSE}(\hat{\mathbf{P}}_k \oplus \kappa, \mathbf{Y}_k \oplus \kappa) \Big) \quad (13)$$

This effectively penalizes under-segmentation in dense clusters by enforcing boundary separation and shape compactness.

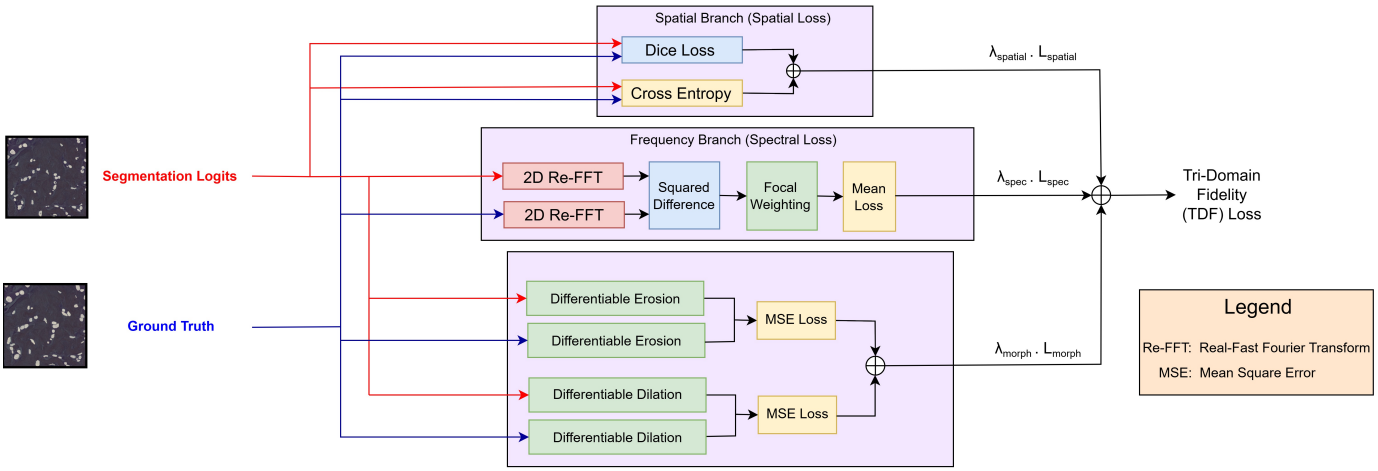

Fig. 2. The Tri-Domain Fidelity (TDF) Loss framework. The loss function optimizes segmentation performance across three distinct domains: (1) Spatial domain using Dice and Cross-Entropy, (2) Spectral domain using Focal Frequency Loss [1], and (3) Morphological domain using erosion and dilation consistency checks.

TABLE I
SUMMARY OF DATASETS AND EXPERIMENTAL PROTOCOLS

Dataset	Scale	Protocol	Epochs	Split Ratio
TNBC [10]	50 Samples	5-Fold CV	200 (per fold)	80 : 20 (CV : Test)
CPM-15[1]	15 Samples	3-Fold CV	200 (per fold)	80 : 20 (CV : Test)
CPM-17[1]	32 Samples	5-Fold CV	200 (per fold)	50 : 50 (CV : Test)
PanNuke [11]	∼7,901 Patches	Train-Test split	30	85 : 7.5 : 7.5

IV. EXPERIMENTS AND RESULTS

A. *Datasets and Implementation*

We tested SAGED-Net for its generalization capability on four diverse datasets prepared on different tissues under H&E staining conditions. The images were resized to 512×512 pixels and normalized. The train augmentations used were horizontal and vertical flips, rotations, and elastic transformation.

B. *Quantitative Analysis*

Implementation Environment: All models were implemented using Python 3.12 and PyTorch 2.8.0 framework with CUDA 12.6 integration. Training was accelerated using a single NVIDIA Tesla P100-PCIE GPU (16GB VRAM) within a Kaggle computational environment. Data augmentation and preprocessing were handled via Albumentations 2.0.8 library. To optimize data throughput, the number of worker threads was dynamically adjusted based on the available CPU cores.

Performance Comparison: We compared SAGED-Net against state-of-the-art methods across four datasets. Performance was evaluated using Dice Similarity Coefficient (DSC), Aggregated Jaccard Index (AJI), Precision, and Recall. As detailed in Table II, it demonstrates superior segmentation accuracy and instance separation capabilities.

Performance on Dense Nuclei (TNBC [10]): It is characterized by high cellularity and overlapping nuclei. SAGED-Net achieved a DSC of **0.9103** and an AJI of **0.8431**. Notably, this represents a significant improvement over the baseline HI-Net [7], which utilizes similar HIN blocks but lacks our adaptive gating. This performance gap confirms that our *Adaptive Gating* mechanism is the critical factor for extracting dense nuclear clusters that standard models fail to separate.

Multi-Organ Generalization (CPM-15 & CPM-17): The model demonstrated exceptional generalization across multi-organ tissues. On CPM-15, it achieved a remarkable AJI of **0.8711**, significantly outperforming HI-Net [7] and the closest competitor Lin et al. [2]. On CPM-17, it surpasses the heavier HI-Net [7] with a DSC of **0.9057**, while maintaining a lower parameter count. This consistency proves that SAGED-Net adapts robustly to diverse organ textures whereas standard baselines struggle with domain shifts.

Large-Scale Robustness (PanNuke [11]): The robustness of the proposed framework is evident on the large-scale PanNuke [11] benchmark. While HI-Net [7] achieves a competitive DSC of 0.8867, SAGED-Net establishes a new state-of-the-art with a DSC of **0.9058** and AJI of **0.8339**. By surpassing both the baseline HI-Net [7] and recent transformers like Youssef et al. [21], we validate that the *TDF-Loss* enforces structural consistency across 19 different tissue types, preventing overfitting as observed in other normalization-based methods.

Sensitivity-Precision Trade-off: Component metrics reveal that SAGED-Net consistently **prioritizes Recall over Precision**. This behavior stems from the topological constraints of the *TDF-Loss*, which penalize structural gaps and encourage the recovery of faint nuclei. Clinically, this high-sensitivity profile is advantageous, as missing malignant nuclei (false negatives) poses a significantly higher diagnostic risk than correctable false positives.

C. *Visualizations*

To interpret SAGED-Net's decision-making, we visualized gradient-weighted activation maps using HiResCAM (Fig. 3). These maps reveal three key behaviors:

[1]https://www.kaggle.com/datasets/ashutoshnaik67/hovernetdataset

TABLE II

PERFORMANCE COMPARISON ON TNBC, CPM-15, CPM-17, AND PANNUKE USING DSC, AJI, PRECISION, AND RECALL. VALUES FOR PROPOSED METHOD ARE REPORTED AS MEAN ± STANDARD DEVIATION FOR TNBC, CPM-15 AND CPM-17 AND ONLY MEAN FOR PANNUKE

TNBC					CPM-15				
Model	**DSC**	**AJI**	**Precision**	**Recall**	**Model**	**DSC**	**AJI**	**Precision**	**Recall**
Roy et al. [4]	0.8165	0.6918	-	-	Lin et al. [8]	0.7513	0.5454	-	-
Kang et al. [12]	0.8290	0.6110	0.8260	0.8330	Mahbod et al. [13]	0.7570	0.5460	-	-
Mahmood et al. [14]	0.8441	0.7332	0.8352	0.8306	Lin et al. [2]	0.8070	0.6800	-	-
Guan et al. [15]	0.8500	0.6920	-	-	U-Net	0.7427	0.5940	0.6834	0.8293
Fırat et al. [6]	0.9075	0.8381	0.9035	0.9115	U-Net++	0.7542	0.6073	0.6993	0.8290
HI-Net [7]	0.8242	0.7468	0.8359	0.8521	HI-Net [7]	0.8480	0.7406	0.8962	0.8246
SAGED-Net (Ours)	**0.9103 ± 0.0050**	**0.8431 ± 0.0076**	**0.8325 ± 0.0174**	**0.8522 ± 0.0268**	**SAGED-Net (Ours)**	**0.9286 ± 0.0147**	**0.8711 ± 0.0244**	**0.8659 ± 0.0224**	**0.8950 ± 0.0310**
CPM-17					**PanNuke**				
Model	**DSC**	**AJI**	**Precision**	**Recall**	**Model**	**DSC**	**AJI**	**Precision**	**Recall**
Lin et al. [2]	0.8390	0.7250	-	-	Liu et al. [16]	0.8325	0.7425	0.8192	0.8675
Li et al. [17]	0.8660	0.7010	-	-	Fan et al. [18]	0.8346	0.7909	-	-
Chen et al. [19]	0.8740	0.7220	-	-	Tang et al. [20]	0.8350	0.7185	-	-
Chen et al. [5]	0.8950	0.7210	-	-	Youssef et al. [21]	0.8760 ± 0.0046	0.8261 ± 0.0044	-	0.8891 ± 0.0030
Zhao et al. [9]	0.8990	0.7910	-	-	DeepLabV3+	0.8353 ± 0.0068	0.7795 ± 0.0079	-	0.8653 ± 0.0031
HI-Net [7]	0.8894	0.7953	0.8692	0.8954	HI-Net [7]	0.8867	0.7773	0.8293	0.8490
SAGED-Net (Ours)	**0.9057 ± 0.0030**	**0.8328 ± 0.0048**	**0.8317 ± 0.0128**	**0.8720 ± 0.0135**	**SAGED-Net (Ours)**	**0.9058**	**0.8339**	**0.9053**	**0.9062**

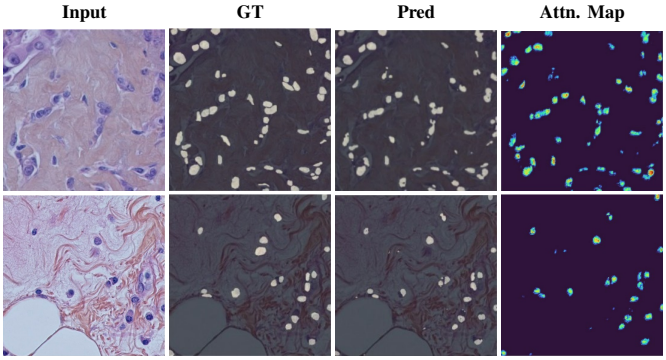

Fig. 3. Qualitative analysis of SAGED-Net. Rows exhibit representative samples from the TNBC [10] dataset featuring dense nuclear clusters. The HiResCAM attention maps (right) confirm that the model learns to prioritize high-frequency nuclear chromatin features while aggressively suppressing the stromal background (deep blue regions), demonstrating that segmentation is driven by relevant biological semantics.

Semantic Specificity: Activation hotspots align precisely with nuclear regions while the stromal background remains suppressed (deep blue), confirming that the Adaptive Gating mechanism effectively filters irrelevant noise.

Texture-driven Learning: The model focuses on high-frequency heterochromatin aggregates rather than uniform shapes. It leverages biological textures to distinguish nuclei.

Instance Separation: In dense clusters (second row), distinct activation peaks are maintained for individual nuclei, validating that the *TDF-Loss* successfully enforces boundary integrity between touching instances.

Further visualizations are available at: https://github.com/shadowbeast0/SAGED-NET/tree/main/assets/visualizations

D. Ablation Study

To show the contributions of our proposed loss formulation, we conducted a component-wise analysis of *TDF Loss* across all four datasets (Table III). We examined the behavior of Spectral Loss ($\mathcal{L}_{spec}$), Morphological Loss ($\mathcal{L}_{morph}$), and their interaction with the standard Spatial Loss ($\mathcal{L}_{spatial}$).

- $\mathcal{L}_{spec}$**:** Utilizing only the spectral loss component yielded competitive performance.
- $\mathcal{L}_{morph}$**:** Focusing on morphological constraints for improved boundary delineation.
- $\mathcal{L}_{spec} + \mathcal{L}_{morph}$**:** Integrating both losses provided a stable improvement across all metrics.
- $\mathcal{L}_{spatial} + \mathcal{L}_{spec} + \mathcal{L}_{morph}$**:** The complete framework, combining spatial losses with our spectral and morphological constraints, achieved the highest performance. This shows that preserving topological structure alongside spatial accuracy is critical for optimal segmentation.

Global Topology vs. Local Precision: While $\mathcal{L}_{spec}$ effectively captures global structural connectivity, it lacks fine-grained definition. The addition of $\mathcal{L}_{morph}$ compensates for this by explicitly sharpening boundaries, directly driving the significant IoU improvements observed across all datasets.

The Spatial Anchor Effect: Topological constraints act as a powerful regularizer but require a firm foundation. $\mathcal{L}_{spatial}$ serves as this critical "anchor," grounding the structural losses to pixel intensity. This synergy is what unlocked the massive performance leap on CPM-15 (DSC: $90.74 \rightarrow 92.86$).

Robustness in Instance Separation: The full SAGED-Net configuration maximizes the stricter IoU metric, distinguishing it from partial baselines. This proves that our framework does not merely classify pixels but actively preserves the structural integrity of individual nuclei, effectively preventing the merging of adjacent instances in dense clusters.

E. Computational Efficiency

As summarized in Table IV, SAGED-Net is designed for deployment in resource-constrained clinical environments. With only **3.79 million** trainable parameters, our model achieves a **45% reduction** in complexity compared to the baseline HI-Net [7]. Furthermore, SAGED-Net is significantly lighter than standard baselines, being **88% smaller** than U-Net and orders of magnitude more efficient than transformer-based heavyweights like TransUNet and SAM (ViT-B).

TABLE III
ABLATION STUDY OF LOSS FUNCTION COMPONENTS ACROSS TNBC, CPM-15, CPM-17, AND PANNUKE. VALUES FOR PROPOSED METHOD ARE REPORTED AS MEAN ± STANDARD DEVIATION FOR TNBC, CPM-15 AND CPM-17 AND ONLY MEAN FOR PANNUKE

Loss Function Components	Datasets							
	TNBC		CPM-15		CPM-17		PanNuke	
	DSC	AJI	DSC	AJI	DSC	AJI	DSC	AJI
$\mathcal{L}_{spec}$	0.8714 ± 0.0043	0.7860 ± 0.0060	0.8692 ± 0.0066	0.7808 ± 0.0091	0.8880 ± 0.0025	0.8051 ± 0.0038	0.8910	0.8114
$\mathcal{L}_{morph}$	0.8976 ± 0.0058	0.8239 ± 0.0085	0.9068 ± 0.0081	0.8359 ± 0.0127	0.9035 ± 0.0024	0.8292 ± 0.0038	0.9003	0.8309
$\mathcal{L}_{spec} + \mathcal{L}_{morph}$	0.9067 ± 0.0051	0.8375 ± 0.0077	0.9074 ± 0.0066	0.8368 ± 0.0103	0.8987 ± 0.0050	0.8218 ± 0.0078	0.9058	0.8338
$\mathcal{L}_{spatial} + \mathcal{L}_{spec} + \mathcal{L}_{morph}$	**0.9103 ± 0.0050**	**0.8431 ± 0.0076**	**0.9286 ± 0.0147**	**0.8711 ± 0.0244**	**0.9057 ± 0.0030**	**0.8328 ± 0.0048**	**0.9058**	**0.8339**

TABLE IV
COMPARISON OF MODEL COMPLEXITY (TRAINABLE PARAMETERS)

Model	Parameters (M)
HI-Net [7]	∼6.9
U-Net	∼31.03
Attention U-Net	∼34.87
U-Net++	∼36.63
DeepLabV3+ (ResNet50)	∼40.35
SAM (ViT-B)	∼91.00
TransUNet	∼105.28
SAGED-Net (Ours)	**3.79**

V. CONCLUSION

In this work, SAGED-Net is presented, a parameter-efficient framework that integrates adaptive residual pathways and tri-domain loss to achieve precise nuclear boundary delineation with significantly lower computational cost than high-capacity baselines.

Despite robust generalization, challenges persist in ambiguous regions. Faint, low-contrast nuclei may be under-segmented due to background similarity, while the model's focus on separating dense clusters can occasionally lead to minor over-segmentation. These edge cases highlight the inherent trade-off between detection sensitivity and topological precision. Future work will aim to validate SAGED-Net's versatility by extending its application to broader medical segmentation domains beyond histopathology, such as radiological imaging (MRI and CT). Furthermore, motivated by its compact footprint of 3.79M parameters, we will explore deployment on resource-constrained edge platforms to assess its suitability for portable and low-power diagnostic applications.

VI. ACKNOWLEDGEMENTS

This work was supported by the Russian Science Foundation (Project № 25-71-30008).

REFERENCES

[1] L. Jiang, B. Dai, W. Wu, and C. C. Loy, "Focal frequency loss for image reconstruction and synthesis," 2021.

[2] C.-N. Lin, C. H. Chung, and A. C. Tan, "Nukit: A deep learning platform for fast nucleus segmentation of histopathological images," *Journal of Bioinformatics and Computational Biology*, vol. 21, no. 01, p. 2350002, 2023, pMID: 36958934.

[3] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," 2018.

[4] A. Roy, P. Pramanik, D. Kaplun, S. Antonov, and R. Sarkar, "Awgunet: Attention-aided wavelet guided u-net for nuclei segmentation in histopathology images," 2024.

[5] J. Chen, Q. Huang, Y. Chen, L. Qian, and C. Yu, "Enhancing nucleus segmentation with haru-net: A hybrid attention based residual u-blocks network," 2023.

[6] H. Fırat, H. Üzen, D. Hanbay, and A. Sengur, "Convnext mixer-based encoder decoder method for nuclei segmentation in histopathology images," *International Journal of Imaging Systems and Technology*, vol. 34, 09 2024.

[7] L. Chen, X. Lu, J. Zhang, X. Chu, and C. Chen, "Hinet: Half instance normalization network for image restoration," 2022.

[8] Y. Lin, Z. Wang, D. Zhang, K.-T. Cheng, and H. Chen, "Bonus: Boundary mining for nuclei segmentation with partial point labels," 2024.

[9] T. Zhao, C. Fu, Y. Tian, W. Song, and C.-W. Sham, "Gsn-hvnet: A lightweight, multi-task deep learning framework for nuclei segmentation and classification," *Bioengineering*, vol. 10, no. 3, 2023.

[10] R. Raju, A. M. Paul, V. Asokachandran, and etc., "The triple-negative breast cancer database: an omics platform for reference, integration and analysis of triple-negative breast cancer data," *Breast Cancer Research*, vol. 16, no. 6, Dec. 2014.

[11] J. Gamper, N. A. Koohbanani, K. Benet, A. Khuram, and N. Rajpoot, "Pannuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification," in *European Congress on Digital Pathology*. Springer, 2019, pp. 11–19.

[12] Q. Kang, Q. Lao, and T. Fevens, "Nuclei segmentation in histopathological images using two-stage learning," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, and A. Khan, Eds. Cham: Springer International Publishing, 2019, pp. 703–711.

[13] A. Mahbod, G. Dorffner, I. Ellinger, R. Woitek, and S. Hatamikia, "Improving generalization capability of deep learning-based nuclei instance segmentation by non-deterministic train time and deterministic test time stain normalization," *Computational and Structural Biotechnology Journal*, vol. 23, pp. 669–678, Dec 2024.

[14] T. Mahmood, M. Owais, K. J. Noh, and etc., "Accurate segmentation of nuclear regions with multi-organ histopathology images using artificial intelligence for cancer diagnosis in personalized medicine," *Journal of Personalized Medicine*, vol. 11, no. 6, 2021.

[15] B. Guan, G. Chu, Z. Wang, J. Li, and B. Yi, "Instance-level semantic segmentation of nuclei based on multimodal structure encoding," *BMC Bioinformatics*, vol. 26, no. 1, p. 42, Feb 2025.

[16] Q. Liu, Z. Zhao, Y. Wu, S. Wu, Y. He, H. Wang, and S. Wang, "Multi-module unet++ for colon cancer histopathological image segmentation," *Scientific Reports*, vol. 15, no. 1, p. 28895, Aug 2025.

[17] J. Li, X. Li, M. Li, and P. Yan, "A dual-path instance segmentation network based on nuclei contour in histology image," *Discover Artificial Intelligence*, vol. 3, no. 1, p. 35, Oct 2023.

[18] D. Fan, J. Li, C. Cai, L. Lin, R. Chen, Y. Chen, and L. Wei, "Ipmmg: Information propagation with multi-granularity morphology-guided for nuclear segmentation and classification," *Expert Systems with Applications*, vol. 298, p. 129676, 2026.

[19] J. Chen, R. Wang, W. Dong, H. He, and S. Wang, "Histonext: dual-mechanism feature pyramid network for cell nuclear segmentation and classification," *BMC Medical Imaging*, vol. 25, no. 1, p. 9, Jan 2025.

[20] Z. Tang, Z. Yang, X. Cai, Z. Li, L. Wei, P. Fan, and X. Yao, "Cellkan: Cellular multi-attention kolmogorov-arnold networks for nuclei segmentation in histopathology images," *Displays*, vol. 91, p. 103246, 2026.

[21] A. Youssef, A. Youssif, and W. El Behaidy, "CancerSeg-XA: Medical histopathology segmentation system based on xception backbone and attention mechanisms," *Statistics, Optimization & Information Computing*, Dec 2025.