

B2 and B3 show irregular shapes. Hence, we ponder: Is it possible to minimize these morphological variations arising from various factors mentioned before? If so, the subsequent segmentation task is simplified when dealing with nuclei of uniform morphology. This leads to higher segmentation accuracy across various tissue types, thereby improving the method's generalization performance.

In this paper, we propose a novel two-stage method as illustrated in Fig. 1. It deviates from the common approach of existing two-stage methods [1, 10, 16], which involves initial nucleus detection followed by precise segmentation. Inspired by the advancements in Artificial Intelligence Generated Content (AIGC) [6, 45], nuclei of various tissue types and imaging techniques are enhanced to an "ideal" morphology characterized by clear centers, solid filling and well-defined edges through generation in enhancement model. Subsequently, the nuclei with distinct features in texture and edges are input into the segmentation model to obtain precise contours. In contrast to existing one-stage [9, 15, 33] and two-stage [1, 10, 16] deep learning methods with task-specific network designs, a classic image-to-image translation generative adversarial network (GAN) [14] performs enhancement alongside a conventional network for segmentation. Since two-stage methods which are not designed to be end-to-end increase complexity to the entire framework [9], we develop a lightweight generator and discriminator based on ResNet and transformer [38], respectively, to improve both enhancement quality and computational efficiency.

The main contributions of this paper are as follows:

- This paper is the pioneering study on the general framework for nucleus segmentation in IHC images, with a focus on high generalization ability. We aim to accurately segment nuclei across various tissue types in both brightfield and fluorescence images by training on a small subset of types.
- We propose a novel two-stage method, GeNSeg-Net, which first enhances nuclei followed by segmentation. This process generates nuclei of various tissues with clearer texture and edges, effectively mitigating segmentation challenges.
- We design a lightweight generator and discriminator to improve enhancement quality by emphasizing the semantic relationship, texture, shape and size of generated nuclei further, while ensuring computational efficiency.
- We conduct experiments on a private systematic dataset of nuclei, covering diverse tissues with rich stain colors in both brightfield and fluorescence imaging, as well as public datasets DSB2018 [4] and BBBC006v1 [24] which include fluorescence images. Our method, GeNSeg-Net, exhibits SOTA performance in accuracy.

2 RELATED WORK

In this section, we divide current nucleus segmentation methods into traditional methods and deep learning-based methods.

Many traditional methods are based on watershed [27, 36, 39, 46]. For instance, Malpica et al. [27] introduced a morphological watershed-based method that utilizes both intensity and morphological information for nucleus segmentation. However, it tends to result in under- or over-segmentation when dealing with nucleus adhesion [36, 46]. Yang et al. [46] proposed a novel marker extraction method based on condition erosion to mitigate over-segmentation. Additionally, numerous other methods exist, such as threshold-based [5], contour-based [28], graph-based [3], region-based [40]

and others [26, 34, 43]. A common drawback of traditional methods is their reliance on manual feature extraction, which leads to good performance solely on specific datasets with rich features.

In recent years, deep learning-based methods have gained prominence. They can be categorized into one-stage [7, 9, 15, 22, 33] and two-stage [1, 10, 16, 23] methods. One-stage methods employ a single network and use post-processing to obtain precise contours, which can be further divided into classification-based and regression-based methods. Classification-based methods output classification probability maps. For example, DCAN [7] is a deep contour-aware network that predicts nuclei and boundaries simultaneously through two branches: a semantic segmentation branch and a boundary detection branch. BES-Net [29] and CIA-Net [49] establish connections between the two branches, further improving effectiveness. HARU-Net [8], the latest advancement, utilizes a hybrid attention-based residual U-blocks network to predict foreground regions and boundaries simultaneously. These methods derive final nucleus instances by subtracting boundaries from foreground regions, potentially resulting in certain loss of segmentation accuracy. Regression-based methods output regression maps. For instance, HoVer-Net [15] predicts the distances between nucleus pixels and their centroids in both the vertical and horizontal directions, followed by watershed post-processing. StarDist [33] predicts centroid probability maps and distances from each foreground pixel to its instance boundary along pre-defined directions. However, StarDist may lose information for large nuclei due to its reliance on only central pixel features in post-processing, which is addressed by CPP-Net [9]'s optimization. Two-stage methods typically involve a detection stage followed by a segmentation stage, i.e., first locating nucleus instances and then predicting precise masks. For example, BRP-Net [10] generates nucleus proposals based on instance boundaries and then refines the foreground masks. SAM [20] represents the emergence of large-scale segmentation models, excelling in natural image segmentation tasks, but its performance on special objects like nucleus is yet to be verified.

3 METHODOLOGY

The overall architecture of GeNSeg-Net is depicted in Fig. 2. It consists of two stages: an enhancement stage and a segmentation stage. In the first stage, the texture and edges of all nuclei are enhanced to obtain a uniform morphology with distinct features. The lightweight generator and discriminator in enhancement model improves both the quality of enhancement and computational efficiency. In the second stage, we perform class prediction and post-processing for segmentation.

3.1 Data Pre-processing

Data pre-processing is illustrated in the red area of Fig. 2. Considering texture and edge variations among nuclei, we initially enhance them through GAN [14, 17], achieving consistent "ideal" nuclei with clear centers, solid filling and well-defined edges, referred to as "nucleus enhancement". To obtain the ground-truth ideal nuclei, we calculate the pixel-to-boundary Euclidean distance for each nucleus in the annotated images. This categorizes each pixel into three classes: nucleus bodies, pending edges and background. More specifically, we adjust the staining intensity of each pixel within a

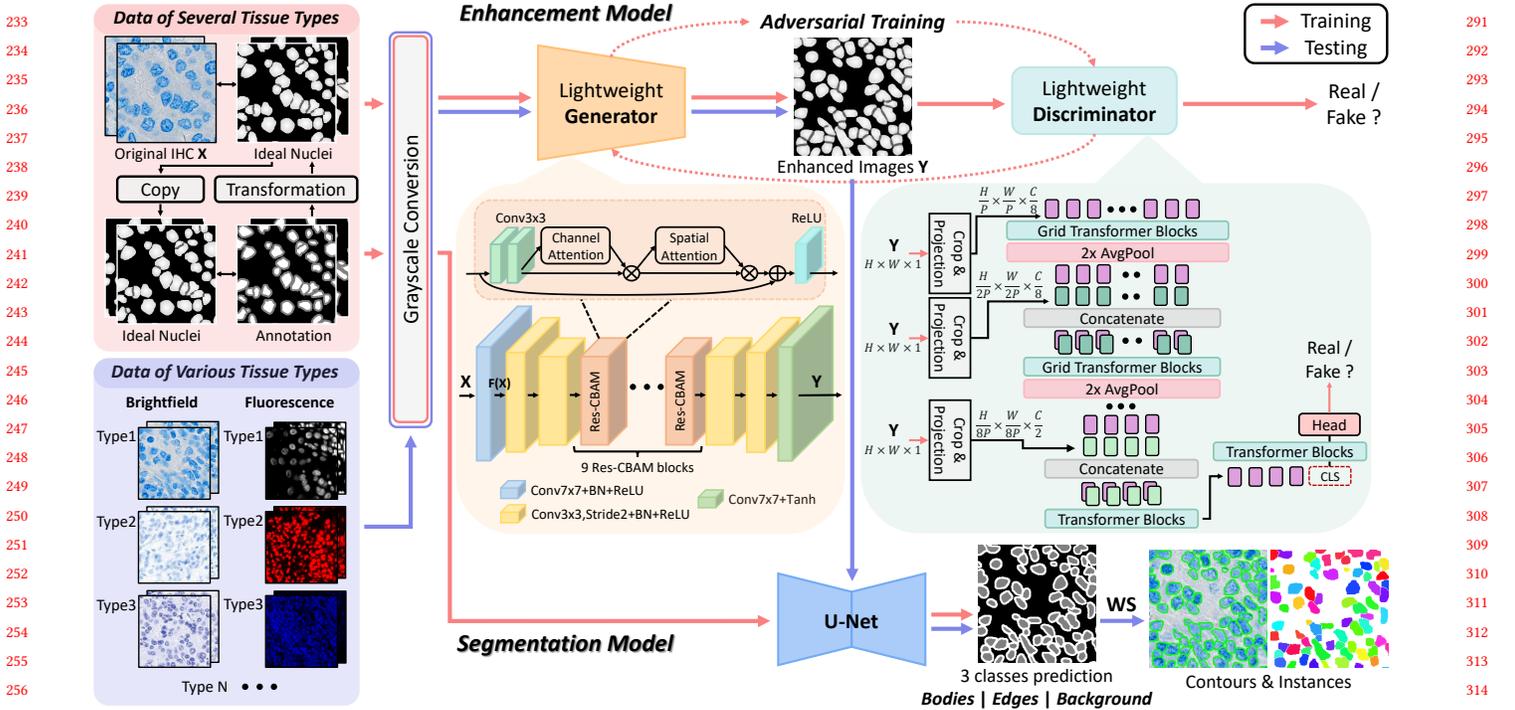


Figure 2: The overview of our proposed GeNSeg-Net. We take high-resolution images sized at 512×512 pixels as a typical example to illustrate our method. "WS" denotes watershed segmentation. For clarity, we display images with a size of 256×256 pixels and bold the precise contours generated by watershed.

nucleus based on its distance from the centroid, gradually decreasing towards the edge. Shared boundaries among nuclei receive a same staining intensity, termed pending edges. The transformed data serves as ground truth for the first stage and as input for the second stage.

3.2 Enhancement Model

3.2.1 Lightweight Generator. For the relatively straightforward generation task, we design a lightweight network as depicted in Fig. 2. Initially, the high-resolution image X is transformed into a feature map $F(X)$ with C channels using the convolution operation. Subsequently, spatial information of each nucleus is aggregated at multiple levels through downsampling. The aggregation effectively consolidates nucleus pixels while eliminating isolated background noise. Simultaneously, applying subsequent operations to the downsampled feature map improves computational efficiency. The feature map then passes through 9 custom-designed Res-CBAM blocks. In Res-CBAM blocks, channel attention and spatial attention are integrated after convolution to eliminate irrelevant information and noise [8, 44]. Finally, the feature map is gradually upsampled to the original size, yielding the enhanced image Y . The process mitigates the loss of fine details caused by direct upsampling at large ratios. In the experiment, we meticulously assess the impact of downsampling and Res-CBAM on model's performance.

3.2.2 Lightweight Discriminator. In modern GANs [31], the discriminator typically employs convolutional neural networks (CNN) as backbone. While it facilitates stable training for high-resolution images, the convolutional operation has a limited local receptive

field. In cases with extensive nucleus adhesion or large nuclei, insufficient network depth fails to capture long-range dependencies. However, increasing network depth can lead to feature and detail loss, complicating optimization and reducing computational efficiency, as demonstrated in previous studies [41, 47]. Therefore, we deviate from conventional CNN models while conducting a comparative analysis in our experiments.

In addition, we consider the following factors: (1) A nucleus covers multiple pixels, especially in cases of adhesive nuclei that span a significant area. Hence, rather than pixel-level discrimination, we opt for a coarser patch level. (2) Large patches in high-resolution images tend to lose low-level texture details, while small patches significantly increase computational demands and memory usage. (3) The adhesion and size of nuclei vary, making it challenging to adopt a fixed patch size for the entire image, which can affect model generalization.

Inspired by TransGAN [18], we introduce a transformer-based multi-scale lightweight discriminator, employing the transformer encoder [38] as the basic block. Our multi-scale discriminator handles local semantic relationships among nuclei, whether they are isolated or adhesive, regardless of their size. It also considers low-level texture, as well as global shape and size features by employing patches of varying sizes at different scales. In high-resolution feature maps, correlating two distant positions is unnecessary. Thus, we integrate grid self-attention [18] into the transformer block to improve computational efficiency.

In Fig. 2, the generator produces an enhanced image Y , which is then fed into the discriminator. At various stages, Y of size $H \times W \times 1$ is cropped into four sequences by selecting different patch size P ,

349 $2P$, $4P$ and $8P$. Considering the typical size of nuclei, the patch size
 350 can be set to 8. Subsequently, these sequences are linearly projected
 351 into dimensions $(\frac{H}{P} \times \frac{W}{P}) \times \frac{C}{8}$, $(\frac{H}{2P} \times \frac{W}{2P}) \times \frac{C}{8}$, $(\frac{H}{4P} \times \frac{W}{4P}) \times \frac{C}{4}$ and
 352 $(\frac{H}{8P} \times \frac{W}{8P}) \times \frac{C}{2}$. These sequences, combined with learnable positional
 353 encoding, serve as inputs to the transformer blocks. After each
 354 block, features are upsampled via average pooling and concatenated
 355 with the new sequence. A class token is introduced prior to the final
 356 transformer. It passes through the transformer block to determine
 357 whether Y is real or fake.

358 Grid self-attention [18] is adopted in high-resolution stages (reso-
 359 lution higher than 16×16). In contrast to conventional transformer
 360 blocks [13] that involve interactions between the individual to-
 361 ken and all others, we partition the entire feature map into grids
 362 based on a pre-defined window size. Attention operation is then
 363 performed within each grid, tailored to our requirements.
 364

365 3.3 Segmentation Model

366 In the second stage, class prediction is conducted firstly by U-Net
 367 [32], followed by watershed post-processing. To be specific, U-Net
 368 categorizes the image context into three classes: nucleus bodies,
 369 coarse edges (approximately 4 pixels wide) and background. The
 370 watershed algorithm then defines contours within the coarse edge
 371 regions. Our segmentation model follows a classic approach with
 372 a simple network design. It achieves both high accuracy and com-
 373 putational efficiency. Since the segmentation model operates inde-
 374 pendently of the enhancement model, GeNSeg-Net provides great
 375 flexibility, enabling the integration of more advanced segmentation
 376 methods. Despite such a classic method within our framework, the
 377 segmentation accuracy has already achieved SOTA performance.
 378

379 3.4 Training and Inference

380 Our training data consists of several tissue types, while the testing
 381 data includes nuclei from both the same tissue types as in the
 382 training and different tissue types. To focus the networks' attention
 383 on nucleus morphology and minimize interference from imaging
 384 techniques and other factors, both the training and testing data are
 385 uniformly converted into grayscale images, as shown in the gray
 386 area of Fig. 2. The enhancement model and segmentation model
 387 are trained individually using their respective paired data for full
 388 supervision.
 389

390 During the training of our enhancement model, the loss function
 391 for the image-to-image translation GAN is defined as:

$$392 L_{GAN}(G, D) = \mathbb{E}_{y \sim p_{data}(y)} \{\log D(y)\} \\ 393 + \mathbb{E}_{x \sim p_{data}(x)} \{\log(1 - D(G(x)))\} \quad (1)$$

394 where G and D refer to the generator and discriminator, while
 395 x and y represent the input and ground truth, respectively. Pre-
 396 vious research suggests that introducing noise to the input can
 397 prevent deterministic outputs and address issues related to fitting
 398 restricted distributions [42]. However, nucleus images inherently
 399 contain background noise due to various factors. The intensity and
 400 distribution of pixel values are crucial for the model to learn local-
 401 ization and texture, with subtle variations playing a significant role
 402 during learning. Therefore, we refrain from introducing additional
 403 information besides the original input. Nevertheless, incorporat-
 404 ing a traditional loss, such as L1, has proven beneficial for model
 405
 406

407 training [30]. While the discriminator's task remains unchanged,
 408 the generator now has a dual objective: to generate images with
 409 realism in the semantic relationship, texture, shape and size, and to
 410 approach the ground truth in an L1 sense. The overall loss function
 411 of enhancement model is:

$$412 L_T(G) = \mathbb{E}_{x,y} \{\|y - G(x)\|_1\} \quad (2)$$

$$413 L_{enh}(G, D) = L_T(G) + \lambda L_{GAN}(G, D) \quad (3)$$

414 In the fully supervised U-Net training, a combined Cross Entropy
 415 and Dice loss function is utilized for predicting three classes. The
 416 segmentation model's loss function is formulated as:

$$417 L_{seg} = \alpha \cdot L_{CE} + \beta \cdot L_{Dice} \quad (4)$$

418 The weight of each term is controlled by α and β .
 419

420 During inference, test images of diverse tissue types initially
 421 undergo enhancement by the trained generator. Subsequently, the
 422 trained U-Net model predicts the three classes. Precise contours
 423 are then obtained by watershed post-processing.
 424

425 4 EXPERIMENTAL SETUP

426 4.1 Dataset

427 *DSB2018 and BBBC006v1.* Data Science Bowl 2018 (DSB2018) [4]
 428 is a competition focused on nucleus detection and segmentation. It
 429 comprises 670 images with manual annotations, sized from 256×256
 430 to 520×696 pixels. Fluorescence imaging is utilized, encompassing
 431 DAPI and Hoechst 33342 stains. We follow the dataset division
 432 protocol outlined in [9], allocating 380 images for training, 67 for
 433 validation and 50 for testing. BBBC006v1 [24] consists of 768 fluo-
 434 rescence images containing Hoechst 33342-stained U2OS cells, each
 435 with the size 696×520 pixels. The dataset is randomly divided into
 436 462 training images, 153 validation images and 153 testing images.
 437

438 *Private Systematic Dataset.* To address the lack of a systematic
 439 public IHC dataset, we build a nucleus dataset encompassing di-
 440 verse tissue types and quantities. It consists of IHC images acquired
 441 through two primary imaging techniques: brightfield and fluores-
 442 cence. In brightfield imaging, hematoxylin stains nuclei, yielding
 443 vibrant images under transmitted light. Fluorescence imaging em-
 444 ploys DAPI to label nuclei. The nuclei emit fluorescence signal when
 445 excited at specific wavelengths, which is then captured using a fluo-
 446 rescence microscope. To emulate real-world scenarios, we refrain
 447 from strictly controlling staining and imaging conditions, allowing
 448 for natural variations introduced by experts and environments. In
 449 clinical and industrial settings, pseudo-colors are commonly ap-
 450 plied to fluorescence images for better visualization. Hence, we
 451 randomly assign pseudo-colors (e.g. red, blue and orange) while
 452 ensuring consistent color assignment within the same tissue. All
 453 stained whole slide images are segmented into patches of 512×512
 454 pixels. Five pathologists with over three years of clinical experience
 455 provided pixel-wise annotations, as illustrated in Fig. 3. In the train-
 456 ing dataset, images are annotated into three classes: nucleus bodies,
 457 edges and background. To mitigate annotation discrepancies among
 458 experts, edges are required to be coarsely annotated with a 4-pixel
 459 width, covering the entire actual edges. Two pixels are within the
 460 nuclei, while the other two pixels lie outside. In the testing dataset,
 461 each instance is precisely annotated.
 462
 463
 464

The dataset composition and division are summarized in supplementary materials. Training and validation data originates from four tissue types: lung, liver, colon and stomach cancers, each with 60 images, alongside 20 testing images per tissue. Additionally, the testing data includes four different tissues: cervix, osteosarcoma, tuberculoma and lymph cancers, each comprising 20 images. They serve to evaluate the generalization ability of segmentation methods. During training, both brightfield and fluorescence data are combined, resulting in a total of 480 images. Among these, 384 images are randomly selected for training, with the remaining 96 reserved for validation.

It's noteworthy that our private dataset, compared to DSB2018 [9] and BBBC006v1 [24], contains a larger number of nuclei with more diverse morphology and more complex environment, as depicted in Fig. 4 and 5. As a result, the segmentation task is significantly more challenging.

4.2 Evaluation Metrics

We employ six evaluation metrics to fully assess method performance: (1) Dice: It evaluates the separation of nuclei from the background. (2) Aggregated Jaccard Index (AJI) [21] and AJI+: AJI is based on instance-wise IoU between predictions and ground truth [10, 15, 48]. AJI+ improves AJI by ensuring maximal unique pairing to obtain overall intersection, thus addressing AJI's over-penalisation. (3) Panoptic Quality (PQ) [19]: It has been widely employed in panoptic segmentation tasks [19]. PQ was first introduced into nucleus segmentation by [15] as the most comprehensive and persuasive metric. It evaluates both detection quality (DQ) and instance-wise segmentation quality (SQ) comprehensively.

4.3 Implementation Details

All methods are implemented using PyTorch 1.12.1 on a system equipped with a single NVIDIA GeForce RTX 3090 GPU. During training, we maintain the original image size and apply basic data augmentation techniques for all methods, including flip, rotation and brightness adjustment. For GeNSeg-Net, the enhancement model adopts Adam optimization with a learning rate of 10^{-4} . A linear decay strategy is employed in the latter half of epochs. Batch size is set to 2 and training proceeds for 400 epochs. Considering both nucleus size and computational efficiency, the discriminator's window size is configured to 8. The loss weight coefficient λ is empirically set to 0.01. The enhancement model's backbone is Pix2pix [17] without any condition. In the segmentation model, the U-Net architecture is standard as proposed in [32], with a fixed learning rate of 10^{-2} . α and β are respectively set to 1 and 3. For other methods, we adhere to the settings in their respective papers.

5 EXPERIMENTS AND ANALYSIS

5.1 Comparisons with the SOTA Methods

To evaluate both the segmentation accuracy and speed of our proposed GeNSeg-Net, we conduct experiments on our private systematic dataset, as well as public datasets DSB2018 [4] and BBBC006v1 [24]. GeNSeg-Net is compared against a large-scale segmentation model and SOTA nucleus segmentation methods, i.e., SAM [20], NucleiSegNet [22], HoVer-Net [15], StarDist [33] and CPP-Net [9]. To ensure fair comparison, we follow the same data split protocol

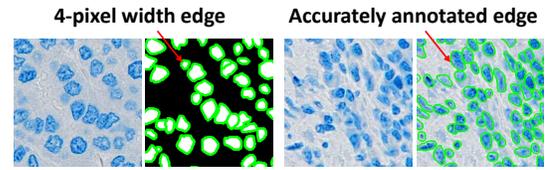


Figure 3: Annotation examples of brightfield training data (left) and testing data (right).

as outlined in Section 4.1 and apply consistent data augmentation techniques for all methods. During training, for methods other than GeNSeg-Net, we adopt the midpoint position of the 4-pixel width annotation as ground truth since they require precise edge annotation.

5.1.1 Comparisons on the Private Systematic Dataset. In Table 1, the results of brightfield same tissue testing are shown. Notably, NucleiSegNet [22], HoVer-Net [15], StarDist [33] and CPP-Net [9] exhibit robust performance. SAM [20] exhibits advantages in natural images but tends to under-segment due to the significant adhesion of nuclei and minimal target variations in our task. Moreover, SAM fails to identify nuclei of low signal and blurred edges, leading to increased false negatives. NucleiSegNet's well-designed architecture facilitates efficient object localization and high-level semantic map extraction while consciously mitigating false positives and negatives. HoVer-Net performs well in identifying foreground regions. It effectively addresses adhesive and blurred nuclei through distance-based predictions. Although StarDist is proficient in foreground region identification and adhesive nuclei differentiation, its instance metrics are compromised due to the boundary influence. StarDist's reliance solely on central pixel features results in incomplete contours, particularly obvious for large nuclei. Additionally, its inclination towards representing nucleus shapes with relatively regular ellipses leads to inaccuracies when depicting irregularly shaped nuclei. CPP-Net improves upon StarDist by more accurately identifying foreground nuclei, distinguishing adhesive nuclei and representing diverse shapes. GeNSeg-Net outperforms these methods across all metrics, with improvements of approximately 2.6% in AJI, 3.1% in AJI+, 0.1% in Dice, 3.1% in DQ, 2.7% in SQ and 5.7% in PQ. The quantitative comparison underscores its superior accuracy.

In brightfield different tissue testing, notable variations in nucleus morphology not presented during training pose segmentation challenges. Testing metrics are presented in Table 1. NucleiSegNet [22] and HoVer-Net [15] exhibit poor segmentation performance. In contrast, StarDist [33] and CPP-Net [9] show superior generalization ability, albeit with a performance decrease ranging from 4% to 13% across metrics. The decline is mainly due to their difficulty in detecting nuclei with different texture, shapes, sizes and signal strength. CPP-Net's reliance on similar nucleus shapes as a design prior further constrains its performance on datasets with diverse shapes. GeNSeg-Net maintains robustness and demonstrates superior results across all metrics compared to current SOTA methods.

In fluorescence tissue testing, most methods surpass their brightfield counterparts as shown in Table 2. This superiority arises from differences in staining and imaging, which result in fluorescence images with reduced background noise, heightened signal intensity and clearer nucleus edges compared to brightfield images. As a result, while maintaining consistent trends with brightfield testing,

Table 1: Comparisons on private brightfield data.

Methods	Metrics					
	AJI↑	AJI+↑	Dice↑	DQ↑	SQ↑	PQ↑
Brightfield same tissue testing						
SAM [20]	0.561	0.592	0.854	0.748	0.795	0.595
NucleiSegNet [22]	0.700	0.730	0.877	0.886	0.823	0.729
HoVer-Net [15]	0.654	0.690	0.840	0.852	0.830	0.707
StarDist [33]	0.655	0.684	0.850	0.857	0.800	0.686
CPP-Net [9]	0.700	0.722	0.873	0.875	0.826	0.723
GeNSeg-Net (ours)	0.726	0.761	0.878	0.917	0.857	0.786
Brightfield different tissue testing						
SAM [20]	0.470	0.494	0.713	0.590	0.736	0.434
NucleiSegNet [22]	0.423	0.454	0.749	0.479	0.746	0.357
HoVer-Net [15]	0.422	0.457	0.575	0.589	0.753	0.444
StarDist [33]	0.591	0.624	0.758	0.811	0.737	0.598
CPP-Net [9]	0.587	0.620	0.749	0.801	0.750	0.601
GeNSeg-Net (ours)	0.701	0.729	0.783	0.895	0.843	0.754

Table 2: Comparisons on private fluorescence data.

Methods	Metrics					
	AJI↑	AJI+↑	Dice↑	DQ↑	SQ↑	PQ↑
Fluorescence same tissue testing						
SAM [20]	0.363	0.423	0.772	0.442	0.843	0.373
NucleiSegNet [22]	0.500	0.541	0.892	0.632	0.808	0.511
HoVer-Net [15]	0.615	0.682	0.842	0.833	0.853	0.711
StarDist [33]	0.677	0.722	0.869	0.902	0.819	0.739
CPP-Net [9]	0.709	0.753	0.875	0.944	0.833	0.786
GeNSeg-Net (ours)	0.737	0.779	0.886	0.936	0.863	0.808
Fluorescence different tissue testing						
SAM [20]	0.328	0.448	0.826	0.553	0.800	0.442
NucleiSegNet [22]	0.526	0.560	0.882	0.628	0.764	0.48
HoVer-Net [15]	0.636	0.690	0.832	0.817	0.839	0.685
StarDist [33]	0.674	0.713	0.863	0.869	0.799	0.694
CPP-Net [9]	0.695	0.736	0.875	0.889	0.816	0.725
GeNSeg-Net (ours)	0.721	0.766	0.891	0.900	0.848	0.763

metrics in fluorescence different tissue testing exhibit no significant decrease compared to those in fluorescence same tissue testing. Overall, GeNSeg-Net maintains an advantage over other methods.

5.1.2 Comparisons on the DSB2018 and BBBC006v1 Dataset. In contrast to our private dataset, images in DSB2018 [4] and BBBC006v1 [24] feature nuclei with lower density and more consistent morphology. This leads to the decent performance of HoVer-Net [15], StarDist [33] and CPP-Net [9], as presented in Table 3. Compared to the testing results of private fluorescence data, nearly all metrics show improvement. However, upon closer examination, we notice the presence of adhesive nuclei that these methods fail to fully distinguish. Our method effectively reduces false positives and negatives, achieving higher segmentation accuracy.

Table 3: Comparisons on DSB2018 and BBBC006v1.

Methods	Metrics					
	AJI↑	AJI+↑	Dice↑	DQ↑	SQ↑	PQ↑
DSB2018						
SAM [20]	0.674	0.698	0.941	0.798	0.859	0.693
NucleiSegNet [22]	0.662	0.689	0.934	0.787	0.848	0.678
HoVer-Net [15]	0.777	0.788	0.905	0.868	0.877	0.767
StarDist [33]	0.795	0.806	0.915	0.902	0.850	0.770
CPP-Net [9]	0.832	0.843	0.936	0.934	0.873	0.818
GeNSeg-Net (ours)	0.841	0.862	0.930	0.940	0.882	0.829
BBBC006v1						
SAM [20]	0.701	0.727	0.935	0.845	0.861	0.727
NucleiSegNet [22]	0.638	0.688	0.943	0.793	0.868	0.690
HoVer-Net [15]	0.915	0.921	0.976	0.961	0.958	0.920
StarDist [33]	0.914	0.918	0.970	0.966	0.950	0.917
CPP-Net [9]	0.961	0.963	0.993	0.990	0.978	0.969
GeNSeg-Net (ours)	0.968	0.969	0.991	0.988	0.981	0.969

Table 4: Cross-dataset evaluation.

Tasks	Methods	Metrics					
		AJI↑	AJI+↑	Dice↑	DQ↑	SQ↑	PQ↑
Private ↓ DSB2018	HoVer-Net[15]	0.727	0.735	0.843	0.848	0.826	0.711
	StarDist[33]	0.764	0.773	0.895	0.917	0.814	0.747
	CPP-Net[9]	0.823	0.831	0.925	0.943	0.850	0.802
	GeNSeg-Net	0.837	0.845	0.935	0.929	0.868	0.808
Private ↓ BBBC 006v1	HoVer-Net[15]	0.775	0.778	0.92	0.901	0.850	0.766
	StarDist[33]	0.763	0.765	0.914	0.879	0.838	0.752
	CPP-Net[9]	0.790	0.792	0.929	0.910	0.855	0.778
	GeNSeg-Net	0.801	0.803	0.933	0.899	0.868	0.781
DSB2018 ↓ BBBC 006v1	HoVer-Net[15]	0.754	0.756	0.908	0.868	0.837	0.727
	StarDist[33]	0.762	0.763	0.908	0.876	0.839	0.735
	CPP-Net[9]	0.765	0.766	0.915	0.883	0.846	0.748
	GeNSeg-Net	0.784	0.789	0.923	0.889	0.863	0.767
BBBC 006v1 ↓ DSB2018	HoVer-Net[15]	0.570	0.588	0.812	0.660	0.784	0.531
	StarDist[33]	0.537	0.558	0.792	0.644	0.762	0.504
	CPP-Net[9]	0.426	0.434	0.666	0.469	0.729	0.378
	GeNSeg-Net	0.665	0.673	0.821	0.755	0.836	0.631

5.1.3 Cross-dataset Evaluation. To further assess GeNSeg-Net’s generalization performance, we design cross-dataset evaluation experiments. Specifically, we evaluate GeNSeg-Net with HoVer-Net [15], StarDist [33] and CPP-Net [9] across four tasks: our private dataset (Private)→DSB2018 [4], Private→BBBC006v1 [24], DSB2018→BBBC006v1 and BBBC006v1→DSB2018. Here, datasets listed before the arrow represent the training data, while those listed after denote the testing data. We present results in Table 4. The metrics reveal that models trained on our private dataset and DSB2018, which encompass diverse tissue types, exhibit superior generalization ability. Conversely, the limited diversity of tissue types in BBBC006v1 impedes the model’s generalization ability. Across all four tasks, GeNSeg-Net consistently outperforms other

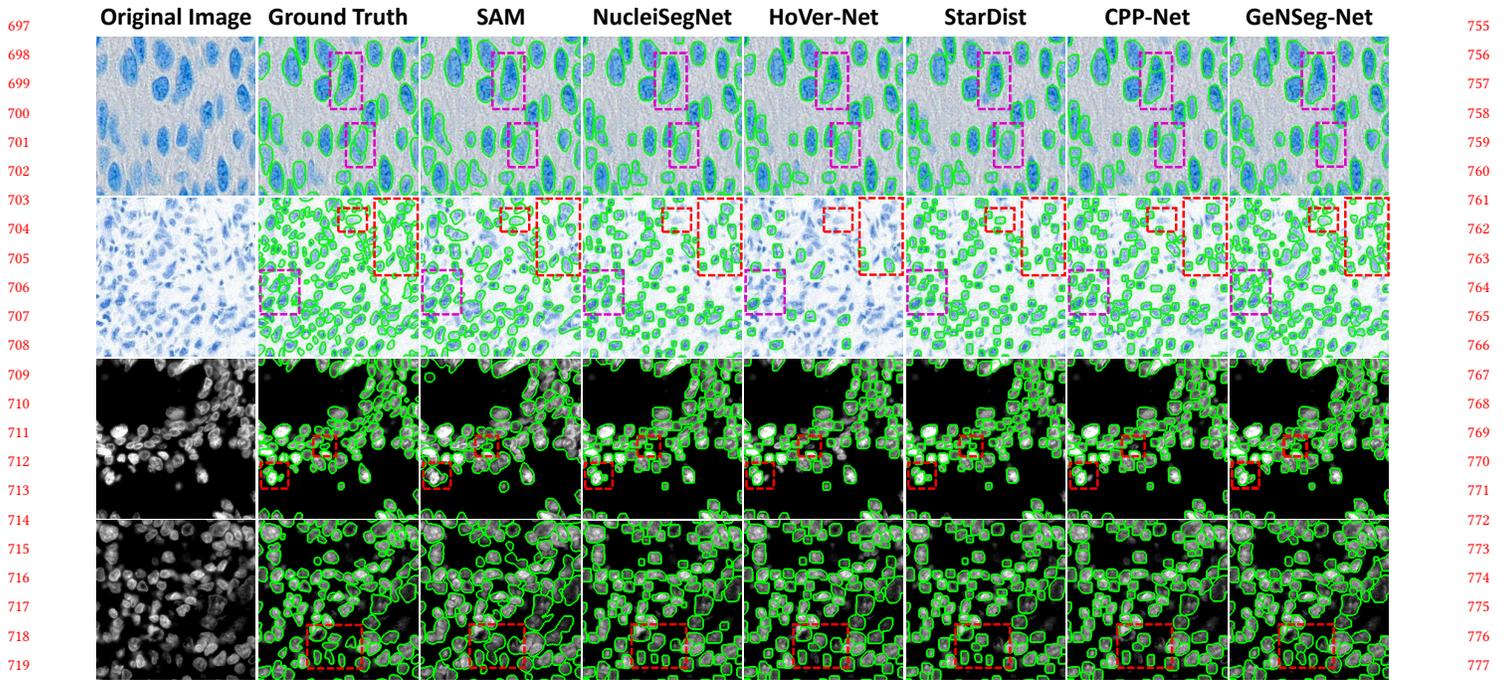


Figure 4: Qualitative comparisons with SOTA methods on our private dataset. The four lines from top to bottom present images from brightfield same tissue testing, brightfield different tissue testing, fluorescence same tissue testing and fluorescence different tissue testing, respectively. All images are of size 265×265 pixels.

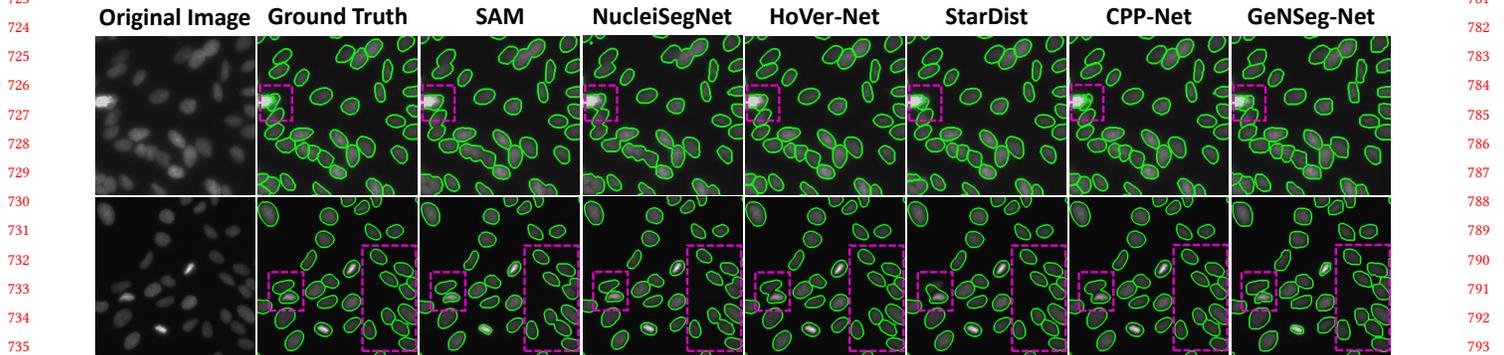


Figure 5: Qualitative comparisons with SOTA methods on the DSB2018 [4] and BBBC006v1 [24] dataset. The two lines from top to bottom present images from DSB2018 and BBBC006v1, respectively. All images are of size 265×265 pixels.

methods. For instance, in the DSB2018→BBBC006v1 task, GeNSeg-Net exhibits improvements of 1.9%, 2.3%, 0.8% and 1.9% in AJI, AJI+, Dice and PQ, respectively. It's noteworthy that GeNSeg-Net significantly improves all metrics in the BBBC006v1→DSB2018 task, though the scores remain below those in the Private→DSB2018 task. This highlights the importance of increasing the diversity of tissue types in the training data, even if the types in testing are not present during training.

5.1.4 Inference Time. We evaluate the inference speed of all methods on our private dataset using the same computer as outlined in Section 4.1 and 4.3. Applying same post-processing method to 512×512 images, the average inference time per image for SAM [20], NucleiSegNet [22], HoVer-Net [15], StarDist [33], CPP-Net [9] and GeNSeg-Net is 1.938, 0.548, 1.863, 0.203, 0.353 and 0.242 seconds, respectively, covering both prediction and post-processing

time. Obviously, StarDist exhibits the fastest inference speed. Despite being a two-stage method, GeNSeg-Net's lightweight design enables faster processing compared to most existing methods.

5.2 Qualitative Comparisons

In this section, we qualitatively compare all methods on our private dataset, DSB2018 [4] and BBBC006v1 [24]. Results for the private dataset are shown in Fig. 4, while those for DSB2018 and BBBC006v1 are in Fig. 5, in line with our previous analysis. GeNSeg-Net, incorporating the enhancement model, excels in identifying weakly stained nuclei, as indicated by the red boxes in Fig. 4. It also effectively distinguishes adhesive nuclei and reduces over-segmentation, as shown by the purple boxes in Fig. 4 and 5. Furthermore, thanks to our segmentation model, GeNSeg-Net produces more natural contours, better capturing the true shapes of nuclei.

Table 5: Ablation study of each component in GeNSeg-Net.

Enh model		Seg	Metrics					
Gen	Dis	model	AJI↑	AJI+↑	Dice↑	DQ↑	SQ↑	PQ↑
-	-	✓	0.580	0.650	0.820	0.762	0.844	0.643
CPP-Net [9]			0.673	0.708	0.843	0.877	0.806	0.707
basic	basic	✓	0.682	0.715	0.844	0.853	0.835	0.712
basic	ours	✓	0.685	0.732	0.855	0.872	0.852	0.743
ours	ours	✓	0.721	0.759	0.860	0.912	0.853	0.778

Table 6: Ablation study on the generator.

Gen		Dis	Metrics					
Backbone	DS	US	AJI↑	AJI+↑	Dice↑	DQ↑	SQ↑	PQ↑
basic	-	ours	0.685	0.732	0.855	0.872	0.852	0.743
ResNet9	✓	ours	0.711	0.755	0.852	0.908	0.853	0.775
Res-CBAM	✓	ours	0.721	0.759	0.860	0.912	0.853	0.778

5.3 Ablation Study

We conduct ablation studies on all brightfield and fluorescence testing data in our private dataset. Initially, the efficacy of the enhancement model is demonstrated. We then assess the discriminator's effectiveness and determine the generator's structure. Qualitative comparisons depicted in Fig. 6 show enhanced images in the first row, predicted class maps in the second row and segmentation results in the final row, with the top-left image representing the ground truth. Evaluation metrics from the ablation study of each component in GeNSeg-Net and the generator are presented in Table 5 and 6, where "gen", "dis", "enh model", "seg model", "ds" and "us" represent the generator, discriminator, enhancement model, segmentation model, downsampling and upsampling, respectively.

5.3.1 Ablation Study of Each Component in GeNSeg-Net. When solely relying on the segmentation model, it effectively identifies foreground nucleus regions, as indicated by the relatively high Dice. However, due to the network's general design and lack of prior information, it struggles to distinguish adhesive nuclei, resulting in noticeable under-segmentation, as shown in the first column of Fig. 6. To address this issue, we introduce an enhancement model aimed at enhancing morphological features, i.e., ensuring clear centers, solid filling and well-defined edges to simplify subsequent segmentation. Initially, the enhancement model employs ResNet as the generator's backbone, comprising 9 blocks, denoted as "basic" in Table 5 and 6. We adopt a CNN-based discriminator proposed in [17], which discriminates at the patch level, as the "basic" discriminator. Integrating the enhancement model significantly improves segmentation performance, with AJI, AJI+, Dice, DQ, SQ and PQ reaching 0.682, 0.715, 0.844, 0.853, 0.835 and 0.712, respectively, slightly surpassing CPP-Net [9]. In our straightforward generation task, increasing the number of blocks adds complexity to training, while reducing blocks leads to insufficient model learning, reflected in declining evaluation metrics. We further attempt to employ a standard U-Net [32] as the generator's backbone, but it results in enhanced images with increased artifacts and noise. It performs worse in distinguishing the two foreground classes, i.e., nucleus

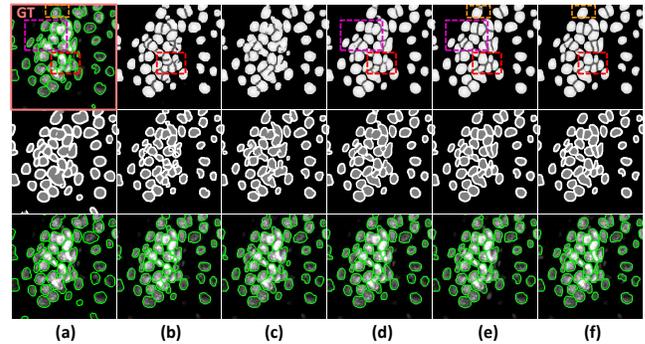


Figure 6: Qualitative comparisons in the ablation study. The six columns from left to right represent -/- (a), basic/basic (b), U-Net/basic (c), basic/ours (d), ResNet9 DS&US/ours (e) and Res-CBAM DS&US/ours (f). Left and right of "/" indicate the structure of generator and discriminator respectively.

bodies and pending edges, leading to increased adhesion, as illustrated in the third column of Fig. 6. When both the generator and discriminator follow a "basic" structure, as seen in the red box in the second column of Fig. 6, the model struggles to differentiate the two foreground classes, resulting in erroneous enhancement. Abrupt texture and noise is obvious. Our designed discriminator focuses more on the semantic relationships among classes, enabling better differentiation. It also improves texture and shape in the enhanced image. With our discriminator, AJI, AJI+, Dice and PQ are increased by 0.3%, 1.7%, 1.1% and 3.1%. Compared to the basic generator, our generator improves all metrics to optimal levels, achieving scores of 0.721, 0.759, 0.860, 0.912, 0.853 and 0.778, respectively.

5.3.2 Ablation Study on the Generator in Enhancement Model. We demonstrate each component's role in the generator of the enhancement model. Integrating downsampling into the basic generator effectively aggregates nucleus pixel features, leading to a clearer separation between classes, as evidenced by the purple boxes in the fourth and fifth columns of Fig. 6. The improved differentiation directly reduces false positives and false negatives. Notably, Table 6 shows that SQ remains unchanged while DQ exhibits a 3.6% improvement, which validates our findings. Additionally, the introduction of the Res-CBAM module further suppresses irrelevant information and noise. Improvement in non-nucleus positions can be observed from the orange boxes in the fifth and sixth columns of Fig. 6. All evaluation metrics for this structure reach optimal levels.

6 CONCLUSION

In this study, we propose a general framework for nucleus segmentation in IHC images. It comprises two stages: initial enhancement followed by segmentation. A classic generative architecture and segmentation network demonstrate the effectiveness of our framework. In the enhancement model, we design a lightweight generator and discriminator to improve both enhancement robustness and computational efficiency. Comparisons against existing methods on both the private and public datasets highlight our method's SOTA accuracy. Cross-dataset evaluation further validates its superior generalization ability. Additionally, our method exhibits highly competitive processing speed. Further discussions on methodology and future work are presented in supplementary materials.

REFERENCES

- [1] Navid Alemi Koohbanani, Mostafa Jahanifar, Ali Gooya, and Nasir Rajpoot. 2019. Nuclear instance segmentation using a proposal-free spatially aware deep learning framework. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 622–630.
- [2] T Bonacho, Francisco Rodrigues, and Joana Liberal. 2020. Immunohistochemistry for diagnosis and prognosis of breast cancer: A review. *Biotechnic & Histochemistry* 95, 2 (2020), 71–91.
- [3] Yuri Y Boykov and M-P Jolly. 2001. Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. In *Proceedings of the IEEE International Conference on Computer Vision*, Vol. 1. IEEE, 105–112.
- [4] Juan C Caicedo, Allen Goodman, Kyle W Karhohs, Beth A Cimmini, Jeanelle Ackerman, Marzieh Haghighi, CherKeng Heng, Tim Becker, Minh Doan, Claire McQuin, et al. 2019. Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nature Methods* 16, 12 (2019), 1247–1253.
- [5] Juan C Caicedo, Jonathan Roth, Allen Goodman, Tim Becker, Kyle W Karhohs, Matthieu Broisin, Csaba Molnar, Claire McQuin, Shantanu Singh, Fabian J Theis, et al. 2019. Evaluation of deep learning strategies for nucleus segmentation in fluorescence images. *Cytometry Part A* 95, 9 (2019), 952–965.
- [6] Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S Yu, and Lichao Sun. 2023. A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. *ArXiv Preprint ArXiv:2303.04226* (2023).
- [7] Hao Chen, Xiaojuan Qi, Lequan Yu, and Pheng-Ann Heng. 2016. DCAN: Deep contour-aware networks for accurate gland segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2487–2496.
- [8] Junzhou Chen, Qian Huang, Yulin Chen, Linyi Qian, and Chengyuan Yu. 2023. Enhancing nucleus segmentation with haru-net: a hybrid attention based residual u-blocks network. *ArXiv Preprint ArXiv:2308.03382* (2023).
- [9] Shengcong Chen, Changxing Ding, Minfeng Liu, Jun Cheng, and Dacheng Tao. 2023. CPP-net: Context-aware polygon proposal network for nucleus segmentation. *IEEE Transactions on Image Processing* 32 (2023), 980–994.
- [10] Shengcong Chen, Changxing Ding, and Dacheng Tao. 2020. Boundary-assisted region proposal networks for nucleus segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 279–288.
- [11] Kin-Hoe Chow, Rachel E Factor, and Katharine S Ullman. 2012. The nuclear envelope environment and its cancer reviews. *Nature Reviews Cancer* 12, 3 (2012), 196–209.
- [12] Leandro Luongo De Matos, Damila Cristina Truffelli, Maria Graciela Luongo De Matos, and Maria Aparecida da Silva Pinhal. 2010. Immunohistochemistry as an important tool in biomarkers detection and clinical practice. *Biomarker Insights* 5 (2010), BMI–S2185.
- [13] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiuhua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *ArXiv Preprint ArXiv:2010.11929* (2020).
- [14] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, Vol. 27.
- [15] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot. 2019. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical Image Analysis* 58 (2019), 101563.
- [16] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. 2017. Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*. 2961–2969.
- [17] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1125–1134.
- [18] Yifan Jiang, Shiyu Chang, and Zhangyang Wang. 2021. Transgan: Two pure transformers can make one strong gan, and that can scale up. In *Advances in Neural Information Processing Systems*, Vol. 34. 14745–14758.
- [19] Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, and Piotr Dollár. 2019. Panoptic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9404–9413.
- [20] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. 2023. Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 4015–4026.
- [21] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. 2017. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Transactions on Medical Imaging* 36, 7 (2017), 1550–1560.
- [22] Shyam Lal, Devikalyan Das, Kumar Alabhya, Anirudh Kanfode, Aman Kumar, and Jyoti Kini. 2021. NucleiSegNet: Robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images. *Computers in Biology and Medicine* 128 (2021), 104075.
- [23] Dongnan Liu, Donghao Zhang, Yang Song, Heng Huang, and Weidong Cai. 2021. Panoptic feature fusion net: a novel instance segmentation paradigm for biomedical and biological images. *IEEE Transactions on Image Processing* 30 (2021), 2045–2059.
- [24] Vebjorn Ljosa, Katherine L Sokolnicki, and Anne E Carpenter. 2012. Annotated high-throughput microscopy image sets for validation. *Nature Methods* 9, 7 (2012), 637–637.
- [25] Cheng Lu, David Romo-Bucheli, Xiangxue Wang, Andrew Janowczyk, Shridar Ganesan, Hannah Gilmore, David Rimm, and Anant Madabhushi. 2018. Nuclear shape and orientation features from H&E images predict survival in early-stage estrogen receptor-positive breast cancers. *Laboratory Investigation* 98, 11 (2018), 1438–1448.
- [26] Zhi Lu, Gustavo Carneiro, and Andrew P Bradley. 2015. An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. *IEEE Transactions on Image Processing* 24, 4 (2015), 1261–1272.
- [27] Norberto Malpica, Carlos Ortiz De Solórzano, Juan José Vaquero, Andrés Santos, Isabel Vallcorba, José Miguel García-Sagredo, and Francisco Del Pozo. 1997. Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry: The Journal of the International Society for Analytical Cytology* 28, 4 (1997), 289–297.
- [28] Csaba Molnar, Ian H Jermyn, Zoltan Kato, Vesa Rahkama, Päivi Östling, Piia Mikkonen, Vilja Pietiäinen, and Peter Horvath. 2016. Accurate morphology preserving segmentation of overlapping cells based on active contours. *Scientific Reports* 6, 1 (2016), 32412.
- [29] Hirohisa Oda, Holger R Roth, Kosuke Chiba, Jure Sokolić, Takayuki Kitasaka, Masahiro Oda, Akinari Hinoki, Hiroo Uchida, Julia A Schnabel, and Kensaku Mori. 2018. BESNet: Boundary-enhanced segmentation of cells in histopathological images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 228–236.
- [30] Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A Efros. 2016. Context encoders: Feature learning by inpainting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2536–2544.
- [31] Alec Radford, Luke Metz, and Soumith Chintala. 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *ArXiv Preprint ArXiv:1511.06434* (2015).
- [32] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 234–241.
- [33] Uwe Schmidt, Martin Weigert, Coleman Broaddus, and Gene Myers. 2018. Cell detection with star-convex polygons. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 265–273.
- [34] David Stutz, Alexander Hermans, and Bastian Leibe. 2018. Superpixels: An evaluation of the state-of-the-art. *Computer Vision and Image Understanding* 166 (2018), 1–27.
- [35] Wei Chang Colin Tan, Sanjna Nilesh Nerurkar, Hai Yun Cai, Harry Ho Man Ng, Duoduo Wu, Yu Ting Felicia Wee, Jeffrey Chun Tatt Lim, Joe Yeong, and Tony Kiat Hon Lim. 2020. Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. *Cancer Communications* 40, 4 (2020), 135–153.
- [36] Afaf Tareef, Yang Song, Heng Huang, Dagan Feng, Mei Chen, Yue Wang, and Weidong Cai. 2018. Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. *IEEE Transactions on Medical Imaging* 37, 9 (2018), 2044–2059.
- [37] Janis M Taube, Guray Akturk, Michael Angelo, Elizabeth L Engle, Sacha Gnjatich, Shirley Greenbaum, Noah F Greenwald, Cyrus V Hedvat, Travis J Hollmann, Jonathan Juco, et al. 2020. The society for immunotherapy in cancer statement on best practices for multiplex immunohistochemistry (IHC) and immunofluorescence (IF) staining and validation. *Journal for Immunotherapy of Cancer* 8, 1 (2020).
- [38] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*, Vol. 30.
- [39] Tomas Vicar, Jan Balvan, Josef Jaros, Florian Jug, Radim Kolar, Michal Masarik, and Jaromir Gumulec. 2019. Cell segmentation methods for label-free contrast microscopy: review and comprehensive comparison. *BMC Bioinformatics* 20 (2019), 1–25.
- [40] Carolina Wählby, I-M Sintorn, Fredrik Eerlandsson, Gunilla Borgfors, and Ewert Bengtsson. 2004. Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. *Journal of Microscopy* 215, 1 (2004), 67–76.
- [41] Xiaolong Wang, Ross Girshick, Abhinav Gupta, and Kaiming He. 2018. Non-local neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7794–7803.
- [42] Xiaolong Wang and Abhinav Gupta. 2016. Generative image modeling using style and structure adversarial networks. In *European Conference on Computer Vision*. Springer, 318–335.
- [43] Mark Winter, Walter Mankowski, Eric Wait, Edgar Cardenas De La Hoz, Angeline Aguinaldo, and Andrew R Cohen. 2018. Separating touching cells using pixel

929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044

1045	replicated elliptical shape models. <i>IEEE Transactions on Medical Imaging</i> 38, 4 (2018), 883–893.	
1046	[44] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. 2018. Cbam: Convolutional block attention module. In <i>Proceedings of the European Conference on Computer Vision</i> . 3–19.	
1047	[45] Jiayang Wu, Wensheng Gan, Zefeng Chen, Shicheng Wan, and Hong Lin. 2023. Ai-generated content (aigc): A survey. <i>ArXiv Preprint ArXiv:2304.06632</i> (2023).	
1048	[46] Xiaodong Yang, Houqiang Li, and Xiaobo Zhou. 2006. Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy. <i>IEEE Transactions on Circuits and Systems I: Regular Papers</i> 53, 11 (2006), 2405–2414.	
1049		
1050		
1051		
1052		
1053		
1054		
1055		
1056		
1057		
1058		
1059		
1060		
1061		
1062		
1063		
1064		
1065		
1066		
1067		
1068		
1069		
1070		
1071		
1072		
1073		
1074		
1075		
1076		
1077		
1078		
1079		
1080		
1081		
1082		
1083		
1084		
1085		
1086		
1087		
1088		
1089		
1090		
1091		
1092		
1093		
1094		
1095		
1096		
1097		
1098		
1099		
1100		
1101		
1102		
	[47] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. 2019. Self-attention generative adversarial networks. In <i>International Conference on Machine Learning</i> . PMLR, 7354–7363.	1103
	[48] Bingchao Zhao, Xin Chen, Zhi Li, Zhiwen Yu, Su Yao, Lixu Yan, Yuqian Wang, Zaiyi Liu, Changhong Liang, and Chu Han. 2020. Triple U-net: Hematoxylin-aware nuclei segmentation with progressive dense feature aggregation. <i>Medical Image Analysis</i> 65 (2020), 101786.	1104
	[49] Yanning Zhou, Omer Fahri Onder, Qi Dou, Efstratios Tsougenis, Hao Chen, and Pheng-Ann Heng. 2019. Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation. In <i>Information Processing in Medical Imaging</i> . Springer, 682–693.	1105
		1106
		1107
		1108
		1109
		1110
		1111
		1112
		1113
		1114
		1115
		1116
		1117
		1118
		1119
		1120
		1121
		1122
		1123
		1124
		1125
		1126
		1127
		1128
		1129
		1130
		1131
		1132
		1133
		1134
		1135
		1136
		1137
		1138
		1139
		1140
		1141
		1142
		1143
		1144
		1145
		1146
		1147
		1148
		1149
		1150
		1151
		1152
		1153
		1154
		1155
		1156
		1157
		1158
		1159
		1160