

Article

Efficient Staining-Invariant Nuclei Segmentation Approach Using Self-Supervised Deep Contrastive Network

Mohamed Abdel-Nasser ^{1,*} , Vivek Kumar Singh ² and Ehab Mahmoud Mohamed ^{1,3} ¹ Department of Electrical Engineering, Aswan University, Aswan 81542, Egypt² Computer Engineering and Mathematics Department, University Rovira i Virgili, 43007 Tarragona, Spain³ Department of Electrical Engineering, College of Engineering, Wadi Alddwasir, Prince Sattam Bin Abdulaziz University, Wadi Alddwasir 11991, Saudi Arabia

* Correspondence: m.nasser@ieee.org

Abstract: Existing nuclei segmentation methods face challenges with hematoxylin and eosin (H&E) whole slide imaging (WSI) due to the variations in staining methods and nuclei shapes and sizes. Most existing approaches require a stain normalization step that may cause losing source information and fail to handle the inter-scanner feature instability problem. To mitigate these issues, this article proposes an efficient staining-invariant nuclei segmentation method based on self-supervised contrastive learning and an effective weighted hybrid dilated convolution (WHDC) block. In particular, we propose a staining-invariant encoder (SIE) that includes convolution and transformers blocks. We also propose the WHDC block allowing the network to learn multi-scale nuclei-relevant features to handle the variation in the sizes and shapes of nuclei. The SIE network is trained on five unlabeled WSIs datasets using self-supervised contrastive learning and then used as a backbone for the downstream nuclei segmentation network. Our method outperforms existing approaches in challenging multiple WSI datasets without stain color normalization.



Citation: Abdel-Nasser, M.; Singh, V.K.; Mohamed, E.M. Efficient Staining-Invariant Nuclei Segmentation Approach Using Self-Supervised Deep Contrastive Network. *Diagnostics* **2022**, *12*, 3024. <https://doi.org/10.3390/diagnostics12123024>

Academic Editor: Jae-Ho Han

Received: 26 October 2022

Accepted: 29 November 2022

Published: 2 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: whole slide imaging; hematoxylin and eosin (H&E); stain color normalization; nuclei segmentation; deep learning

1. Introduction

The digital pathology revolution began using a whole slide imaging (WSI) scanner to digitize glass slides. Digital pathology has been used in various applications, including case diagnosis and management, education for all clinical and patient cases, and forensic pathology. However, pathologists devote significant efforts to manual WSI image analysis (i.e., visual assessment of WSIs), particularly for tasks such as nucleus cell segmentation and counting [1].

In the literature, many computer analysis methods have been developed to analyze histopathology images [2–5]. Kleczek et al. [6] combined statistical analysis, color thresholding, and binary morphology to segment histopathological images of skin tissues. Kleczek et al. [7] proposed an automated method for epidermis segmentation in histopathological images of human skin. They incorporated the domain-specific details of morphometric and biochemical characteristics of skin tissue regions in histopathology images.

In recent years, deep learning approaches have been used to analyze histopathology images for various diagnosis tasks [8,9], such as nuclei cell counting, cancer metastasis detection, and forensic pathology applications such as determining the cause of death after trauma and poisoning. The automatic segmentation of nuclei in WSI images has been studied extensively. In [10], a five-step segmentation approach for nuclei cells or nanoparticles was proposed. The five steps were (1) automatic gradient image formation, (2) automatic threshold selection, (3) manual calibration of the threshold selection method for each cell or nanoparticle image, (4) manual determination of the segmentation cases for each specific cell or nanoparticle image type, and (5) automatic quantification by iterative

morphological erosion. In [11], a selective-edge-enhancement-based nuclei segmentation method (SEENS) was proposed. In SEENS, a selective search algorithm was integrated with mathematical operators to segment cervical WSI images into small regions of interest while automatically evading duplicated segmentation and removing non-nuclei regions. An edge enhancement method based on the canny operator and mathematical morphology was used to extract edge information to enhance the nucleus edge.

In WSI image analysis, deep learning-based techniques, notably nucleus segmentation, are gaining popularity. In [12], various deep learning-based techniques were reviewed and assessed for breast tumor cell nuclei segmentation, including U-Net, Mask R-CNN, and GB U-Net. GB U-Net performed better in segmenting cell nuclei with an aggregated Jaccard index (AJI) score of 53%. Cui et al. [13] introduced an end-to-end deep learning network for nuclei segmentation that uses a nuclei boundary model to predict the inner nuclear instance, nuclear contour, and background in WSI images simultaneously. To improve and stabilize the inner nuclei and contour prediction, the authors used a weighted loss function based on the relative position of pixels inside the WSI image. They achieved an F1-score of 85.40% using the MoNuSeg dataset. Xie et al. [14] proposed the DIMAN method, a deep interval-marker-aware network, for nuclei segmentation. They integrated the convolutional neural networks with the marker-controlled watershed to delineate the foreground, marker, and interval of nuclei. DIMAN achieved an AJI score of 56.64% with the MoNuSeg dataset. Zhou et al. [15] introduced the U-Net++ model that combined UNets of various depths and restructured skip connections. They also used an architecture pruning approach to speed up inference while maintaining performance. On the MoNuSeg dataset, UNet++ had an F1-score of 88.17%.

Ilyas et al. [16] proposed a tissue-specific feature distillation network (TSFD-Net) trained with a combinational loss function to extract tissue-specific features from WSI images to produce better nuclei segmentation and classification. TSFD-Net was based on the fact that morphological features such as appearance, shape, and texture of nuclei in a tissue vary greatly depending upon the tissue type. With the PanNuke dataset, TSFD-Net obtained mean and binary panoptic quality of 50.4% and 63.77%, respectively. In an attempt to segment overlapped and clustered nuclei, Ref. [17] proposed the DenseRes-UNet model by integrating dense blocks in the last layers of the encoder block of U-Net, as well as distance map and binary threshold techniques to intensify the nuclei interior and contour information in WSI images. Rączkowski et al. [18] recommended an active (ARA) image classification method using Bayesian CNN that classifies colorectal cancer tissue. The authors designed a network that measures the uncertainty of the given test samples. This approach helped to identify the misclassified training samples and could improve the model performance. Hassan et al. [19] suggested a clustering-based stain selection technique. They trained a set of independent deep-learning models on several stain templates. The authors combined the segmentation masks of the individual models using an aggregation function based on the Choquet integral. Recently, self-supervised learning attained great success in analyzing histopathology images, where the trained models can extract rich features from the unlabeled data and later could be used to improve the downstream nuclei segmentation or classification performance [20–22].

Existing nuclei cell segmentation approaches require a stain color normalization step to reduce color variations in WSIs due to various stains used in laboratories and stain manufacturing processes across vendors. Selecting a proper staining normalization method is crucial for the nuclei segmentation methods. However, staining normalization methods have some limitations, such as (1) they cannot handle the problem of inter-scanner feature instability; (2) they modify the color of WSIs, which may yield a loss in the source domain information—they do not preserve the source intensity variation (notably, source color variation can help with WSI analysis, as it can reveal crucial differences in the tissue's underlying biochemical composition); and (3) they may produce inconsistent color normalization results when the number of stains increases (resulting WSIs deviate from the target staining template). In addition, most nuclei cell segmentation methods

still face challenges due to the variations in nuclei shapes and sizes and overlapping and clumped cell nuclei. Figure 1 shows WSI images obtained from different organs and multiple laboratories. As one can see, there is a big variation in the stain color, nuclei shapes and sizes, and the presence of overlapping and clumped nuclei. Such differences could greatly reduce the accuracy of automated nucleus cell segmentation systems.

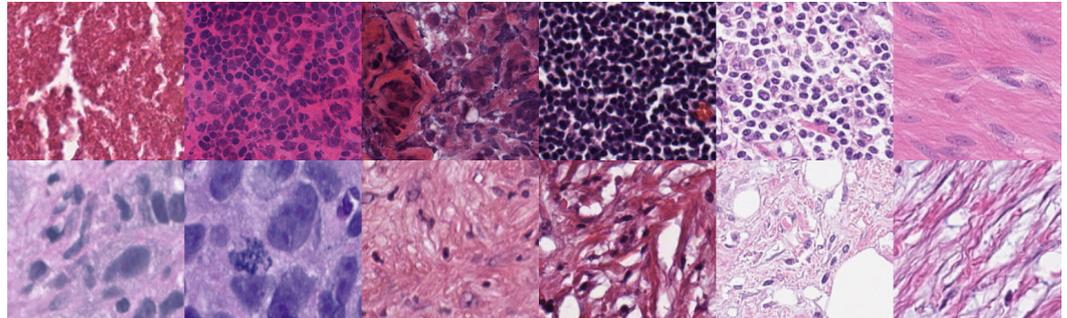


Figure 1. Examples of WSIs having various stains. The present nuclei have different shapes and sizes.

Unlike most existing nuclei segmentation approaches that require a staining normalization algorithm, we propose an efficient staining-invariant nuclei segmentation method based on self-supervised contrastive learning and an effective weighted hybrid dilated convolution (WHDC) block. Specifically, we propose a staining-invariant encoder (SIE) that includes convolution (Conv) and transformers blocks [23], where Conv blocks help extract low-level nuclei features, and transformer blocks model their long-range relationships. We also propose a WHDC block to enable the network to learn multi-scale features to handle the variation in the size and shapes of nuclei. SIE is trained using the SimCLR contrastive learning framework [24] in a self-supervised manner that learns latent staining-invariant representations of WSIs without any labeled data. The staining invariant encoder is used as a backbone, followed by a supervised fine-tuning strategy for the nuclei segmentation task. The key contributions of this article are as follows:

- Proposing an efficient nuclei segmentation method for hematoxylin and eosin (H&E) WSI images using a deep staining-invariant self-supervised contrastive network. This method eliminates the need for a stain normalization step;
- Proposing an effective weighted hybrid dilated convolutional (WHDC) block that helps extract multi-scale nuclei-relevant representations;
- Achieving accurate nuclei segmentation on unseen single-organ and multi-organ datasets collected from different laboratories without employing stain color normalization or fine-tuning that demonstrate the proposed method's generalization capabilities.

It should be noted that the proposed method is not limited and could be employed for other applications such as nanoparticle segmentation.

2. Proposed Method

Figure 2 depicts the pipeline of the proposed method. A staining-invariant encoder is trained using self-supervised contrastive learning (Section 2.1). The encoder network includes the proposed WHDC block to handle the variation in the size and shapes of nuclei (Section 2.2). This encoder is a backbone for the downstream nuclei segmentation network trained using a supervised fine-tuning strategy (Section 2.3). Details are given below.

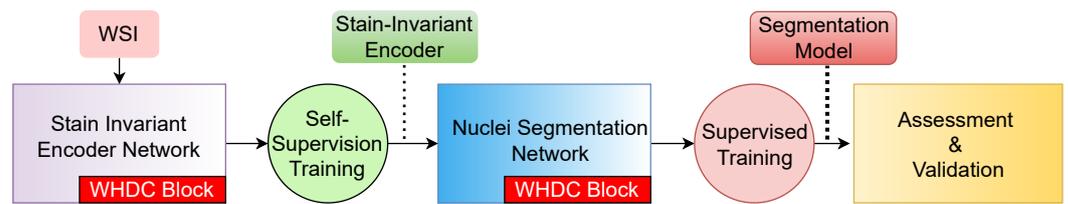


Figure 2. Illustration of the proposed method pipeline.

2.1. Staining-Invariant Encoder

Figure 3 presents the overview of the proposed staining-invariant encoder (SIE) network. The main components of SIE are convolution blocks, transformer blocks, and the proposed WHDC block. In particular, SIE is based on a convolutional-transformer neural network so-called CoAtNet [23]. SIE is trained using a self-supervised contrastive learning approach [24] that does not demand any labeled dataset prepared by pathologists. In other words, the training of SIE is completely based on the self-supervised learning technique, in which the model learns patterns by itself by extracting stain-invariant nuclei-relevant features.

As shown in Figure 3 (left), SIE extracts feature representations from pairs of augmented WSI images. We apply data augmentation techniques to construct pairs of WSI images, I_i , and I_j , to learn visual feature similarities between them. Specifically, we apply the following data augmentation techniques: flipping, rotation with 90 degrees, blurring, random brightness contrast with a probability of 0.2, and color jitter. SIE learns visual features through a contrastive loss function by increasing the agreement among different augmented views of the same WSI image patch example.

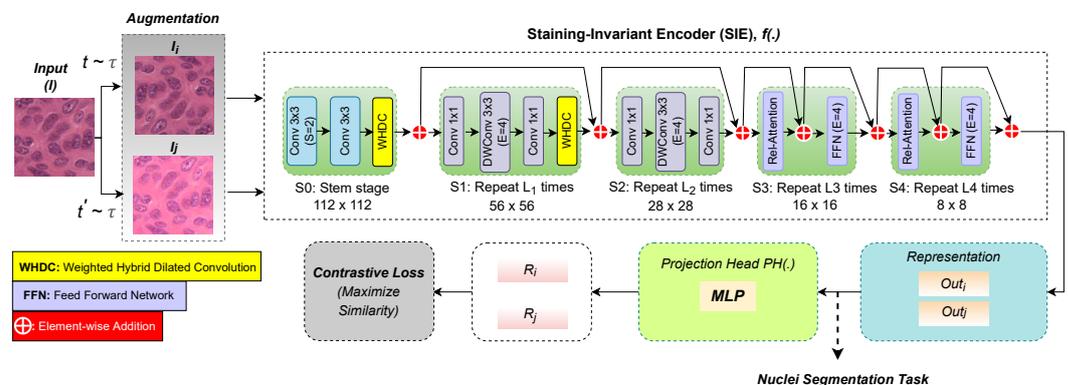


Figure 3. The architecture of proposed staining-invariant encoder (SIE) based on self-supervised contrastive learning.

The top branch shown in Figure 3 represents the backbone feature extractor (i.e., f), which is based on CoAtNet [23] that includes convolution (Conv) and transformer blocks [23]. SIE has five stages (S0, S1, S2, S3, and S4). The first three stages rely on convolution blocks, whereas the last two adopt transformer blocks. Stage S0 applies a down-sampling operation with a factor of 2 to reduce the patch image spatial size. The first stage has two standard convolution layers with a kernel size of 3×3 , allowing for extraction of nuclei-relevant features (e.g., shape, texture, and intensity) from WSI patches. As there is a wide variety of nuclei sizes, we use the WHDC block to encourage the model to learn multi-scale nuclei-relevant features. WHDC employs dilation rates of 3, 6, 9, and 18, where the small receptive fields capture the fine details of nuclei, and the larger receptive field provides contextual information (Figure 4). WHDC has a self-weighting mechanism that encourages the network to promote nuclei-relevant features.

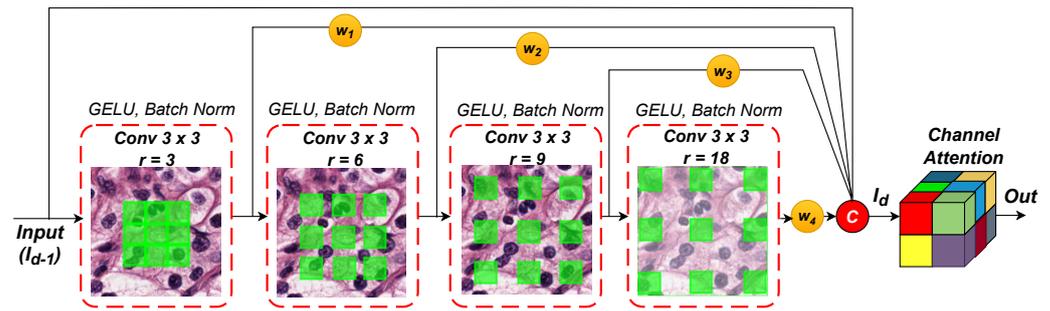


Figure 4. Illustration of the WHDC block.

The second stage, S1, contains an MBCConv block (i.e., inverted residual block) [23]. The MBCConv substitutes the conventional convolution with the depth-wise separable convolutions. WHDC is also added at the end of S1 with the same parameter setting used in S0. The architecture of S2 is identical to S1 but without WHDC. Stages S3 and S4 contain a transformer block that encourages the SIE network to establish long-range dependencies while avoiding overfitting using a 2D relation attention and feed-forward network (i.e., FFN) and a self-attention module. Stages S3 and S4 apply a max-pooling operation with a stride of 2 on the self-attention module’s constituents. The final size of the bottleneck of S4 is 8×8 (at the arrow connecting the top and bottom branches in Figure 3). The output of the top branch of the SIE network, f , can be formulated as follows:

$$Out_i = f(\tilde{I}_i) \tag{1}$$

$$Out_j = f(\tilde{I}_j) \tag{2}$$

where Out_i , and $Out_j \in \mathbb{R}^d$ stand for the output representations for the augmented image pairs I_i and I_j , respectively.

The projection head $PH(\cdot)$ projects the generated representations Out_i and Out_j into a 128-dimensional feature space using only a single-layer MLP (multi-layer perceptron). MLP computes the representation R_i for Out_i and a representation R_j for Out_j as follows:

$$R_i = PH(Out_i) = W^{(2)}\sigma(W^{(1)}PH_i) \tag{3}$$

$$R_j = PH(Out_j) = W^{(2)}\sigma(W^{(1)}PH_j) \tag{4}$$

where W stands for the weights of MLP, and σ is a non-linear rectified linear unit (ReLU) function.

SIE is built using a self-supervised contrastive learning approach [24], which does not require any labeled dataset. The contrastive loss can be formulated as

$$\mathcal{L}_{FINAL} = \frac{1}{2N} \sum_{k=1}^N [\ell(2k - 1, 2k) + \ell(2k, 2k - 1)] \tag{5}$$

where N stands for the mini-batch WSI patch images and contrastive prediction result to $2N$ data points computed through pairs of data-augmented patch samples. The ℓ can be computed as

$$\ell_{i,j} = -\log \frac{\exp(\text{CM}(R_i, R_j)/\tau)}{\sum_{k=1}^{2N} 1_{[k \neq i]} \exp(\text{CM}(R_i, R_k)/\tau)} \tag{6}$$

where τ stands for a temperature parameter set to 0.07 in our experiments; $1_{[k \neq i]} \in \{0, 1\}$ corresponds to an indicator function to measure if $k \neq i$; CM is the cosine similarity function that can be defined as

$$\text{CM}(R_i, R_j) = R_i^T R_j / \|R_i\| \|R_j\| \tag{7}$$

Contrastive loss decreases when projections from the same image are similar; otherwise, the error rate will increase.

2.2. Weighted Hybrid Dilated Convolution (WHDC) Block

Figure 4 presents the proposed WHDC block to widen the receptive fields of SIE with different dilation rates and encourage it to promote multi-scale nuclei features. WHDC consists of four cascaded convolutions layers, self-weighting, and channel attention mechanisms. It incorporates various spatial scales that provide fine pixel-level details and global contextual information without losing resolution size. In this study, we use dilation rates of 3, 6, 9, and 18. As shown in Figure 4, WHDC has four convolutional layers with a kernel size of 3×3 followed by batch normalization and a non-linear GELU activation function connected in a cascaded manner. The WHDC block allows for utilizing the extracted features from the previous layer with specific dilation rates and feeds them into the next convolutional layer. In WHDC, a convolutional layer of depth d and dilation rate with r can be expressed as follows:

$$\text{Dilated-Conv}_d^r : I_{d-1} \longrightarrow I_d^r \tag{8}$$

where the input of each dilated convolution has a size of $h, w,$ and c (i.e., height, width, and number of channels, respectively), and $I_{d-1} \in \mathbb{R}^{h' \times w' \times c'}$. WHDC produces feature maps of size $I_{d-1} \in \mathbb{R}^{h \times w \times c}$.

As shown in Figure 4, the proposed WHDC block has a self-attention mechanism to promote the nuclei-relevant features while ignoring other irrelevant features. It should be noted that the four weights of the self-attention mechanism, $w_1, w_2, w_3,$ and $w_4,$ are derived dynamically based on a *Sigmoid* activation function. The weighted features ($w_1 * I^{r=3}, w_2 * I^{r=6}, w_3 * I^{r=9},$ and $w_4 * I^{r=18}$) and original input I_{d-1} are concatenated and fed into a channel attention mechanism (CAM) [25] to advance channel interdependencies. CAM converts the concatenated nuclei feature maps output to a single vector through a global pooling layer named squeeze operation. Subsequently, CAM uses two fully connected (FC) layers with a channel reduction ratio of 16. For further details about the architecture of CAM and its FCs, the readers are recommended to see [25]. The weights of each channel are multiplied with the original input vector to boost nuclei-relevant features (i.e., excitation operation) automatically.

2.3. Nuclei Segmentation Network

Figure 5 presents the proposed nuclei segmentation network for WSI images. It includes an encoder and decoder network. SIE trained with contrastive learning (Section 2.1) is used as an encoder. The encoder’s bottleneck features (8×8 feature maps) are fed into the decoder network (i.e., the network at the bottom of Figure 5). The decoder consists of five layers. Each decoder layer utilizes a Conv-transpose layer with a kernel size of 4×4 and stride of 2. In this study, we adopt the attention mechanism [26] to initial four decoder layers that ignore irrelevant artifacts generated during feature reconstruction and concentrate only on nuclei-relevant features that lead to reducing the false positives. Except for the last layer, each decoder layer has batch normalization and ReLU activation functions. Skip connections between encoder and decoder networks are employed to narrow the semantics gaps in feature reconstruction. We use a threshold value of 0.5 to generate the final binary segmentation masks.

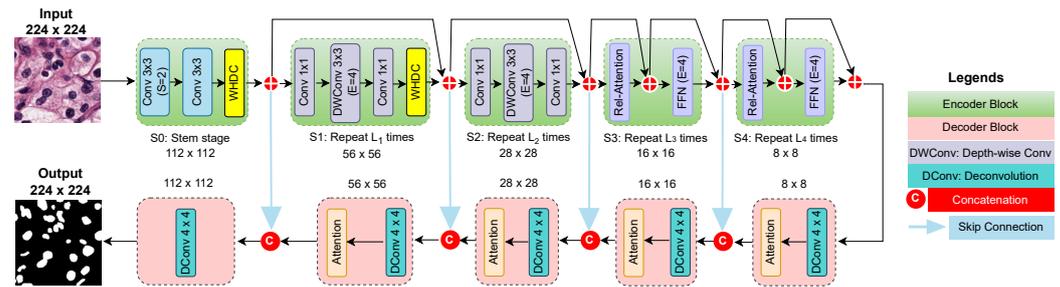


Figure 5. The framework of the proposed nuclei segmentation network.

We fuse two loss functions to train the segmentation network—binary cross-entropy (BCE) and Dice losses. This combined loss function could minimize the error and address the pixel imbalance issue between the nuclei and background areas. The BCE loss is defined as

$$\mathcal{L}^{\text{BCE}}(y, t) = -(y \cdot \log(t) + (1 - y) \cdot \log(1 - t)) \quad (9)$$

where y and t stand for the ground truth mask and mask generated by the proposed segmentation network. The Dice loss can be expressed as

$$\mathcal{L}^{\text{Dice}}(y, t) = 1 - \text{Dice}(y, t) = 1 - \frac{2|y| \cdot |t|}{|y|^2 + |t|^2} \quad (10)$$

The overall segmentation loss (OSL) is the weighted sum of \mathcal{L}^{BCE} and $\mathcal{L}^{\text{Dice}}$:

$$\mathcal{L}^{\text{OSL}}(y, t) = \gamma \mathcal{L}^{\text{BCE}}(y, t) + (1 - \gamma) \mathcal{L}^{\text{Dice}}(y, t) \quad (11)$$

where γ is a weighting factor experimentally set to 0.4.

3. Results and Discussion

3.1. Datasets

In this study, we leveraged a total of eight publicly available datasets, including Lizard [27], MHIST [28], BreCaHD [29], SPIE-BreastPathQ [30], Colorectal NCT-CRC-HE [31], MoNuSeg [32], CryoNuSeg [33], and BNS [34].

To build the staining-invariant encoder (i.e., SIE) based on contrastive learning, we used Lizard, MHIST, BreCaHD, SPIE-BreastPathQ, and Colorectal NCT-CRC. The Lizard dataset has a half-million labeled nuclei in H&E stained colon tissue with $20\times$ objective magnification. The entire set involves 291 images with an average resolution of 1016×917 . The MHIST has 3152 H&E-stained colorectal polyp images with 224×224 pixels. BreCaHAD contains 162 breast cancer WSIs with a resolution of 1360×1024 . SPIE-BreastPathQ has 96 H&E WSI scans acquired from 55 patients with residual invasive breast cancer. This dataset has a patch size of 512×512 and comprises training, validation, and test sets of 2394, 185, and 1119 images, respectively. Colorectal NCT-CRC-HE includes 100 thousand non-overlapping patches derived from 86 H&E stained human cancer 54 tissue slides of colorectal cancer and normal tissue. The size of the patches is 224×224 .

To construct and evaluate the nuclei segmentation model, we used the MoNuSeg, CryoNuSeg, and BNS datasets. MoNuSeg is a multi-center multi-organ dataset containing 30 WSIs with a resolution of 1000×1000 . It has a total of 21 thousand manually annotated nuclei. MoNuSeg involves WSI of seven organs—breast, kidney, colon, stomach, prostate, liver, and bladder. A total of 23 WSIs are used for training and 7 WSIs (i.e., one WSI per organ) for a fair assessment.

To train the segmentation model, we resized the original WSIs to the size of 1024×1024 . Then, we applied non-overlapping cropping with patch size 512×512 . To increase the number of training samples, for each non-overlapping patch, we applied random crops, generating 200 samples of patch size 256×256 . In total, we generated 18,400 ($23 \times 4 \times 200$) training samples. BNS has 33 WSIs with a resolution of 512×512 for breast cancer

(2754 labeled nuclei). CryoNuSeg has WSIs of 10 human organs—adrenal gland, larynx, lymph node, mediastinum, pancreas, pleura, skin, testis, thymus, and thyroid gland. It has 30 WSIs with a resolution of 512×512 . It is worth noting that BNS and CryoNuSeg datasets are independently used for evaluating the proposed model, meaning they are not used for training or fine-tuning the segmentation model.

3.2. Implementation Details

The training process is two-fold—training the staining-invariant encoder network (i.e., SIE) based on contrastive learning (Section 2.1) and training the nuclei segmentation network (Section 2.3). We applied data augmentation techniques, including flipping, 90-degree rotation, blurring, random brightness contrast with a probability of 0.2, and color jitter. The input image size was 224×224 . The SIE network is trained using an SGD optimizer with an initial learning rate of 0.001. The number of epochs is 50, with a mini-batch size of 2. For training the nuclei segmentation network, the number of epochs is set to 100 with a mini-batch size of 4. ADAM optimizer is used with $\beta_1 = 0.5$ and $\beta_2 = 0.999$ and a learning rate of 0.0002. In our experiments, all hyperparameters are manually tuned. We used the same hyperparameter settings for the proposed model and state-of-the-art models presented in this study. The proposed method is implemented on PyTorch 1.7.0, CUDA 11.2 on Intel Core-i9 with 32 GB RAM, and GeForce RTX 3090 GPU with 24 GB memory.

3.3. Evaluation Metrics

To assess the performance of segmentation methods, we used the dice coefficient (Dice), aggregated Jaccard index (AJI) [32], precision, and recall. These metrics can be expressed as follows:

$$Dice = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}, \tag{12}$$

$$Precision = \frac{TP}{(TP + FP)}, \tag{13}$$

$$Recall = \frac{TP}{(TP + FN)}, \tag{14}$$

$$AJI = \frac{\sum_{i=1}^L |GT_i \cap N\psi_j^*(i)|}{\sum_{i=1}^K |GT_i \cup N\psi_j^*(i)| + \sum_{K \in LIP} N\psi_k} \tag{15}$$

In these expressions, TP, FP, FN, and TN rates stand for true positive, false positive, false negative, and true negative, respectively; GT_i , $N\psi_k$, and $N\psi_j^*(i)$ stand for the i th ground-truth mask of nuclei pixels, the predicted nuclei segmentation mask, and the connected component from the predicted mask that maximizes the Jaccard index, respectively; LIP stands for the list of indices of pixels that do not belong to any element in GT.

3.4. Ablation Study

Here, we conducted a thorough ablation study to demonstrate the efficacy of the proposed model’s specific components, where we investigated the effect of various configurations of the proposed segmentation model and the nuclei segmentation results of different loss functions.

3.4.1. Analysis of Various Configurations

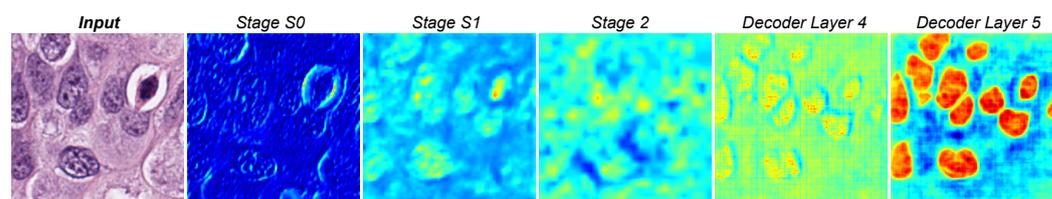
Table 1 presents the ablation study of different configurations of the proposed nuclei segmentation network—baseline (i.e., BL), baseline+WHDC, SIE without WHDC, SIE without contrastive learning (CL) approach (i.e., proposed w/o CL), and the proposed SIE network (i.e., network with all components). In this table, we present the mean and the standard deviation (SD) of all evaluation metrics across all the test samples.

Table 1. Ablation study on MoNuSeg. The best results are highlighted in bold.

| Model | Dice (%) \uparrow | AJI (%) \uparrow | Precision (%) \uparrow | Recall (%) \uparrow |
|-----------------|------------------------------------|------------------------------------|-----------------------------------|-----------------------------------|
| BL | 83.32 \pm 11.80 | 66.69 \pm 13.17 | 82.11 \pm 15.12 | 84.58 \pm 18.03 |
| BL + WHDC | 84.46 \pm 10.52 | 68.56 \pm 12.73 | 83.97 \pm 13.93 | 84.95 \pm 16.87 |
| Proposed w/o CL | 85.82 \pm 9.98 | 69.11 \pm 11.41 | 85.19 \pm 11.84 | 86.46 \pm 12.76 |
| Proposed | 88.64 \pm 6.20 | 73.14 \pm 8.19 | 88.2 \pm 7.38 | 89.1 \pm 8.09 |

We assessed the performance of the BL network that entirely relies on the encoder and decoder network without adopting the WHDC block or CL. BL obtained Dice and AJI scores of 83.32% and 66.69%, respectively. We added the WHDC block with the BL network, which boosted the Dice and AJI scores by 1.1% to 2%. This configuration allows for the extraction of spatial nuclei-relevant features comprising shape, texture, and intensity while avoiding irrelevant ones. In turn, feature reconstruction is an important step in creating segmentation maps. We leverage the spatial attention mechanism in the decoder to prevent losing the semantic correlations. BL+WHDC with spatial attention mechanism in the encoder (proposed w/o CL) improved the segmentation performance by 2% to 3% in Dice, AJI, precision, and recall scores when compared to BL. The proposed model incorporated the pre-trained SIE model trained with contrastive learning and WHDC. Pre-trained SIE provided staining-invariant nuclei-relevant features, while the proposed WHDC block helped generate multi-scale nuclei features. This led to a significant gain of 5% to 6% in all evaluation metrics. We also observed the proposed model generated fewer false positive pixels leading to a lower SD of 5% in Dice and IoU scores against the BL model.

Figure 6 presents heatmaps from the encoding and decoding layers of the proposed segmentation model. As one can see, stage S0 extracts the spatial nuclei features such as shape, texture, and intensity. Stages S1 and S2 emphasize nuclei features with finer details, due to the proposed WHDC block, which encourages the network to learn multi-scale nuclei-relevant feature representation. Due to low spatial resolution (16×16 and 8×8), we do not show the feature maps of stages S3, S4, and the early decoder layer output. In addition, the decoder layers 4 and 5 also show that the segmentation model correctly identified the nuclei region with sharp boundaries (highlighted in red) while ignoring the background.

**Figure 6.** Heatmaps from encoding and decoding layers of the proposed segmentation model.

3.4.2. Analysis of the Loss Function

Table 2 presents the effect of different loss functions (i.e., \mathcal{L}^{BCE} , $\mathcal{L}^{\text{Dice}}$, and $\mathcal{L}^{\text{BCE}} + \mathcal{L}^{\text{Dice}}$) on the efficiency of the proposed segmentation network evaluated with the MoNuSeg dataset. We used two loss functions consisting of \mathcal{L}^{BCE} and $\mathcal{L}^{\text{Dice}}$ losses. We initialized our training by only using the \mathcal{L}^{BCE} loss that provides a Dice score of 84.76% and an AJI score of 69.2%. The $\mathcal{L}^{\text{Dice}}$ loss was employed to focus more on dense pixel prediction by generating fewer false positives; $\mathcal{L}^{\text{Dice}}$ achieves the 83.71% and 68.97% Dice and AJI scores, respectively. Both \mathcal{L}^{BCE} and $\mathcal{L}^{\text{Dice}}$ performed well, and thus we combined them to achieve better results with lower false-positive rates. We set the weighting factor γ to 0.4 (Equation (11)). The combined loss functions reduce SD in the range of 3 – 6% for all the evaluated metrics against \mathcal{L}^{BCE} . The ablation study confirmed that each loss function reasonably contributed to the final nuclei segmentation (4% improvement in Dice and AJI scores).

Table 2. Ablation study of the loss functions on MoNuSeg. The best results are highlighted in bold.

| Loss Function | Dice (%) \uparrow | AJI (%) \uparrow | Precision (%) \uparrow | Recall (%) \uparrow |
|--|------------------------------------|------------------------------------|-----------------------------------|-----------------------------------|
| \mathcal{L}^{BCE} | 84.76 \pm 11.46 | 69.2 \pm 13.67 | 83.47 \pm 12.94 | 86.03 \pm 11.88 |
| $\mathcal{L}^{\text{Dice}}$ | 83.71 \pm 11.81 | 68.97 \pm 14.93 | 82.01 \pm 13.02 | 85.48 \pm 12.20 |
| $\mathcal{L}^{\text{BCE}} + \mathcal{L}^{\text{Dice}}$ | 88.64 \pm 6.20 | 73.14 \pm 8.19 | 88.2 \pm 7.38 | 89.1 \pm 8.09 |

3.5. Comparison with Existing Methods

Table 3 compares the proposed method with 12 state-of-the-art networks on the MoNuSeg dataset. We trained 5 networks of these 12 networks from scratch, meaning the 5 networks were completely trained without utilizing any pre-trained ImageNet weights. The five networks trained from scratch using the same hyperparameters used for training the proposed model are U-Net, fully convolutional network (FCN), DeepLabv3+, Attention U-Net [26], and U-Net++ [35] with the same hyperparameter settings. The findings of the other seven methods are taken from recently published nuclei segmentation studies. As tabulated in Table 3, the proposed model outperformed state-of-the-art methods by a significant margin and achieved a SD ranging from 6 to 9% in all evaluation metrics, which is much lower than the other compared methods. It achieves Dice, AJI, precision, and recall scores of 88.64%, 73.14%, 88.2%, and 89.1%, respectively. The U-Net, DeepLabv3+, and FCN obtained Dice scores of 77.94%, 76.59%, and 76.36% respectively, which are 10% lower than the proposed method. Both Attention U-Net [26] and U-Net++ [35] obtain an average Dice of 79.5%. The proposed model attains 2% and 1% improvements in the Dice and AJI scores, respectively, which are higher than the second-best method cGAN [36]. The cGAN-based approach generated synthetic nuclei images and combined them with original training data to segment the nuclei areas. The RIC-UNet [37], DIST [38], MedT [39], Chanchal et al. [40], and BiO-Net [41] achieve Dice scores lower than 83%. Although MSAL-Net [42] used a multi-scale attention learning network with dense dilated convolution, it provides a Dice score of 83.9%, which is 4.5% lower than our method.

Table 3. Comparing the proposed method with 12 existing methods on the MoNuSeg dataset. The – symbol represents the missing value that was not reported in the study. The best results are highlighted in bold.

| Model | Dice (%) \uparrow | AJI (%) \uparrow | Precision (%) \uparrow | Recall (%) \uparrow |
|----------------------|------------------------------------|------------------------------------|-----------------------------------|-----------------------------------|
| U-Net | 77.94 \pm 17.71 | 64.01 \pm 17.23 | 76.69 \pm 20.65 | 79.24 \pm 23.96 |
| Attention U-Net | 79.52 \pm 19.44 | 65.22 \pm 18.82 | 78.36 \pm 19.89 | 80.73 \pm 24.80 |
| DeepLabv3+ | 76.59 \pm 25.75 | 60.86 \pm 22.43 | 74.28 \pm 24.81 | 79.05 \pm 29.47 |
| FCN | 76.36 \pm 9.60 | 65.01 \pm 10.19 | 73.86 \pm 12.34 | 79.05 \pm 14.59 |
| U-Net++ | 79.57 \pm 18.80 | 63.92 \pm 18.01 | 78.02 \pm 19.81 | 81.19 \pm 24.85 |
| RIC-UNet [37] | 82.78 | 56.35 | – | – |
| DIST [38] | 78.63 | 55.98 | – | – |
| Chanchal et al. [40] | 80.65 | 67.95 | – | – |
| cGANs [36] | 86.60 | 72.10 | – | – |
| MedT [39] | 79.55 | 66.17 | – | – |
| MSAL-Net [42] | 83.9 | 70.6 | 82.1 | 85.3 |
| BiO-Net [41] | 82.5 | 70.4 | – | – |
| Proposed | 88.64 \pm 6.20 | 73.14 \pm 8.19 | 88.2 \pm 7.38 | 89.1 \pm 8.09 |

Figure 7 presents statistics of AJI score of the proposed method, U-Net, Attention U-Net [26], DeepLabv3+, FCN, and U-Net++ [35]. Our model achieves the highest mean and median scores and lowest standard deviation among other compared methods. It only has three outliers, whereas other methods have many outliers with large standard deviations.

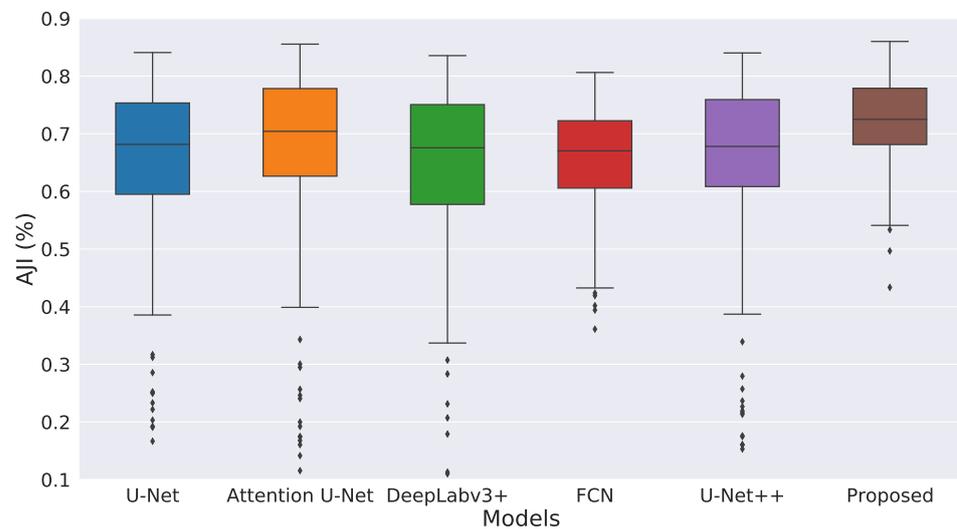


Figure 7. Boxplot of AJI scores of the proposed nuclei segmentation method on MoNuSeg dataset.

Figure 8 presents the nuclei segmentation results of the proposed method with WSI images collected from different laboratories for kidney, bladder, stomach, and prostate organs. We provided the color maps to easily interpret the segmentation results compared to the ground truth. With the MoNuSeg dataset, the proposed method could precisely capture the different nuclei sizes and segment the nuclei boundaries (orange color) with very few false positives (green color).

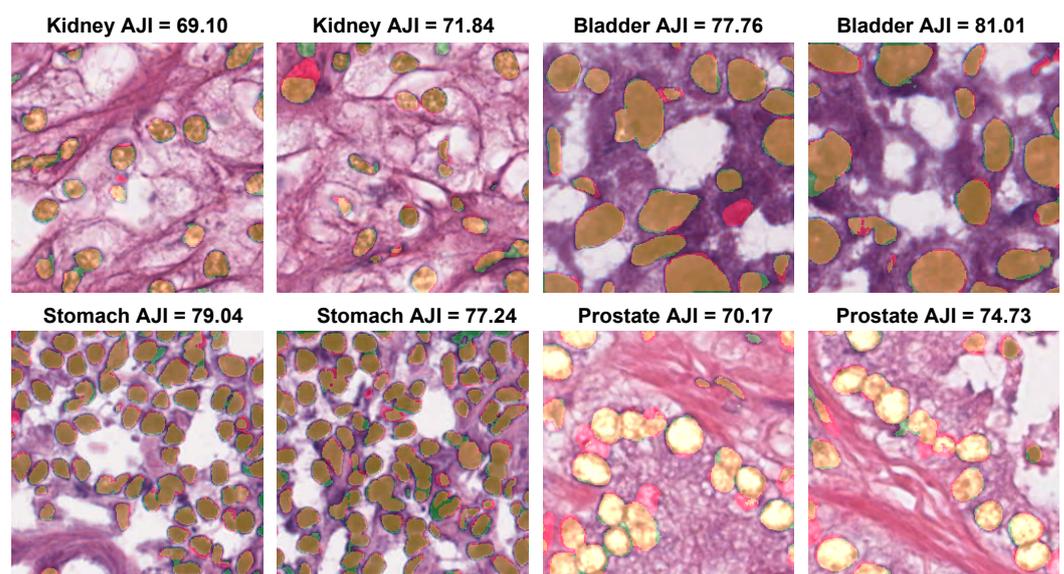


Figure 8. Nuclei segmentation by the proposed method in WSIs of four organs. The color maps are displayed as follows: true positives (TP: orange), false positives (FP: green), false negatives (FN: red), and true negatives (TN: background).

Figure 9 shows qualitative segmentation results of the proposed method compared to U-Net, Attention U-Net [26], DeepLabv3+, FCN, and U-Net++ [35]. The AJI scores demonstrated the quantitative improvement produced by the proposed segmentation method (71.84%) compared to other approaches. The proposed model produces excellent segmentation results with fewer false positives of small nuclei, whereas other methods do not completely segment many nuclei. In addition, the Wilcoxon signed-rank test demonstrated that the results of the proposed model and second-best U-Net++ on the MoNuSeg dataset were statistically significant (p -value < 0.001).

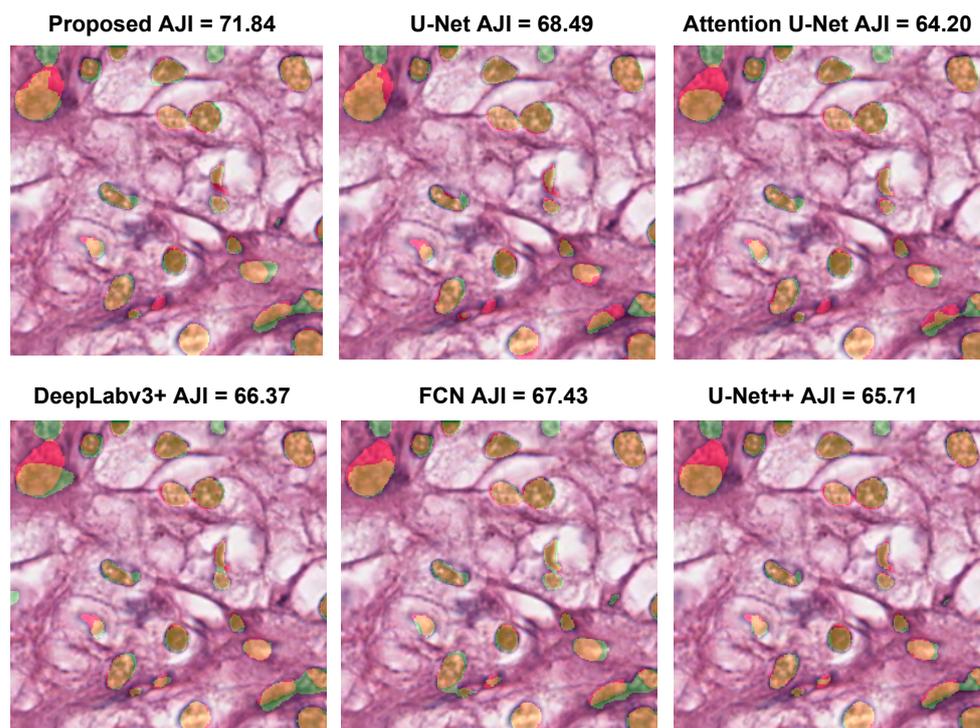


Figure 9. Comparison of the proposed model with five existing methods on the MoNuSeg dataset. The color maps are displayed as follows: true positives (TP: orange), false positives (FP: green), false negatives (FN: red), and true negatives (TN: background).

3.6. Evaluating the Proposed Method on Other Datasets

Herein, we evaluate the effectiveness of the proposed model trained on the MoNuSeg dataset using the CryoNuSeg and BNS datasets without retraining or fine-tuning (the complete dataset is used as the test set). We independently trained the U-Net, Attention U-Net [26], DeepLabv3+, FCN, and U-Net++ [35] on the CryoNuSeg and BNS datasets from scratch. As tabulated in Table 4, the proposed method outperformed other segmentation approaches with Dice, AJI, precision, and recall scores of 86.53%, 64.7%, 85.48%, and 87.62%, respectively. In the case of the proposed method, the estimation errors (i.e., SD) of the Dice and AJI scores are 1% lower than for U-Net. Hassan et al. [19] achieved the second-best results, which was 1% less than our method. As tabulated in Table 5, the proposed method outperformed the other methods. DeepLabv3+ obtained poor segmentation results with limited samples. The U-Net++ achieved the second-best results with an 83.39% Dice score and a 62.72% AJI. As one can see, the estimation errors of the proposed method in terms of the Dice and AJI scores are 1% lower than for U-Net++. Although CryoNuSeg and BNS datasets were entirely unseen by the proposed segmentation model, they achieved the best results, thanks to the robust multi-scale nuclei-relevant staining-invariant feature representations learned by the model.

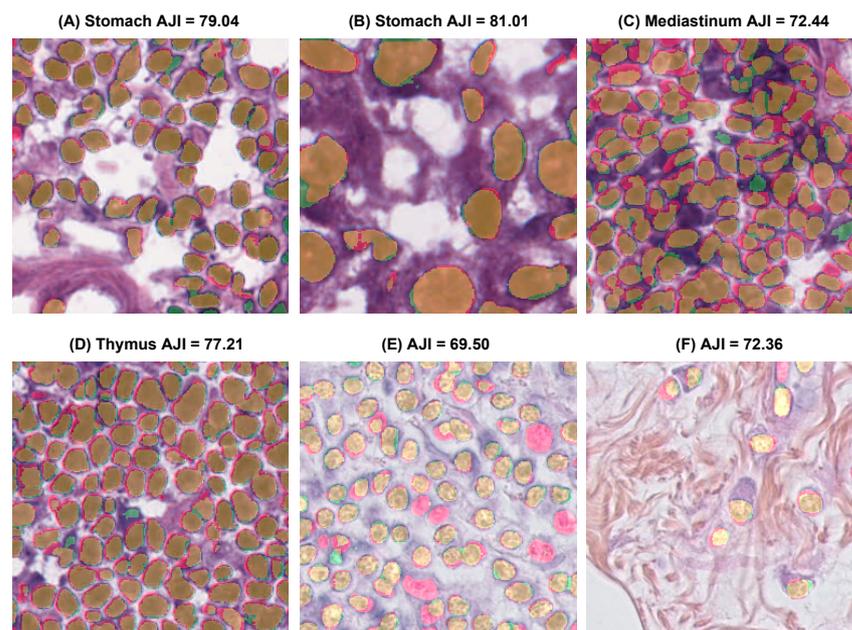
Table 4. Comparing the proposed method with existing approaches on the CryoNuSeg dataset. The best results are highlighted in bold.

| Model | Dice (%) \uparrow | AJI (%) \uparrow | Precision (%) \uparrow | Recall (%) \uparrow |
|--------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| U-Net | 81.01 \pm 9.15 | 59.62 \pm 11.13 | 78.33 \pm 8.42 | 83.89 \pm 6.97 |
| Attention U-Net | 81.87 \pm 8.70 | 60.58 \pm 10.48 | 79.81 \pm 7.77 | 84.06 \pm 7.63 |
| DeepLabv3+ | 84.20 \pm 10.87 | 62.07 \pm 12.74 | 81.96 \pm 7.63 | 86.57 \pm 5.74 |
| FCN | 84.94 \pm 10.20 | 62.34 \pm 12.19 | 82.77 \pm 7.19 | 87.23 \pm 5.57 |
| U-Net++ | 83.41 \pm 11.05 | 61.51 \pm 12.66 | 80.59 \pm 6.55 | 86.44 \pm 5.98 |
| Hassan et al. [19] | 85.55 | 64.51 | 85.02 | 86.16 |
| Proposed | 86.53 \pm 8.67 | 64.7 \pm 10.22 | 85.48 \pm 5.13 | 87.62 \pm 6.72 |

Table 5. Comparing the proposed method with existing approaches on the BNS dataset. The best results are highlighted in bold.

| Model | Dice (%) \uparrow | AJI (%) \uparrow | Precision (%) \uparrow | Recall (%) \uparrow |
|-----------------|------------------------------------|-----------------------------------|------------------------------------|------------------------------------|
| U-Net | 82.40 \pm 10.52 | 60.86 \pm 11.15 | 81.78 \pm 8.65 | 83.02 \pm 7.85 |
| Attention U-Net | 83.39 \pm 9.87 | 61.20 \pm 10.69 | 82.94 \pm 8.03 | 83.71 \pm 7.50 |
| DeepLabv3+ | 78.44 \pm 14.58 | 55.29 \pm 17.82 | 77.53 \pm 11.48 | 79.38 \pm 10.83 |
| FCN | 81.38 \pm 11.53 | 59.63 \pm 12.94 | 80.50 \pm 9.11 | 82.28 \pm 8.46 |
| U-Net++ | 83.89 \pm 7.47 | 62.72 \pm 9.07 | 85.24 \pm 8.33 | 87.92 \pm 7.39 |
| Proposed | 88.82 \pm 6.69 | 65.2 \pm 7.81 | 87.64 \pm 7.36 | 90.04 \pm 5.72 |

Figure 10 shows the segmentation results of the proposed method on the MoNuSeg, CryoNuSeg, and BNS datasets. These WSI were collected from different organs in laboratories employing various stain colors. However, the proposed method could accurately segment nuclei. These findings proved the generalization abilities of the proposed method, and it could segment nuclei in WSI images without employing stain color normalization or fine-tuning the model.

**Figure 10.** The segmentation results of the proposed method on (A,B) MoNuSeg, (C,D) CryoNuSeg, and (E,F) BNS. The color maps are displayed as follows: true positives (TP: orange), false positives (FP: green), false negatives (FN: red), and true negatives (TN: background).

3.7. Discussion and Limitations

Although the existing methods achieved acceptable results, they required a stain color normalization, which can differ from one dataset to another and may yield a loss in the source domain information (other limitations stated in Section 1). In addition, existing methods could not distinctly delineate the nuclei boundary, leading to several false positives. Adopting a self-supervising contrastive learning approach and the proposed WHDC block to build a stain-invariant encoder encouraged the segmentation model to concentrate on segmenting and separating the nuclei boundaries. It is evident that the proposed model showed great potential and provided an efficient solution to segment nuclei in WSIs of different stains and multiple organs and surpasses the existing deep learning approaches by a significant margin.

One of the limitations of the proposed method is that it produces poor segmentation results in the case of overlapped and clumped nuclei.

4. Conclusions and Future Work

This paper proposed an effective staining-invariant nuclei segmentation method based on a self-supervised contrastive learning approach. In particular, we introduced a staining-invariant method that does not use color normalization before processing H&E WSI. The proposed staining invariant encoder (SIE) leveraged the convolution, WHDC, and transformer blocks in a self-supervised training setting that facilitates learning better nuclei feature representation. A trained SIE model was used as the backbone for the downstream nuclei segmentation task. We used eight WSI datasets, including five datasets for training the self-supervised SIE network, and the remaining three were applied to assess the effectiveness of the proposed nuclei segmentation model. The proposed method achieved state-of-the-art AJI scores of 73.14%, 64.7%, and 65.20% with MoNuSeg, CryoNuSeg, and BNS datasets, respectively. Our analysis showed that the proposed method achieved accurate nuclei segmentation on a completely unseen independent dataset, due to the robust multi-scale nuclei-relevant staining-invariant feature representations learned by the SIE model. It also demonstrated the generalization capabilities of the proposed method on multiple datasets, and it could segment nuclei in WSI images without employing stain color normalization or fine-tuning the model. It is worth noting that the proposed staining-invariant method is not limited and can be applied to other applications, such as nanoparticle segmentation, which will be the focus of future work.

Author Contributions: Conceptualization, M.A.-N. and V.K.S.; methodology, M.A.-N. and V.K.S.; software, M.A.-N. and V.K.S.; validation, M.A.-N. and V.K.S.; formal analysis, M.A.-N. and V.K.S.; investigation, M.A.-N. and V.K.S.; resources, M.A.-N. and V.K.S.; data curation, M.A.-N. and V.K.S.; writing—original draft preparation, M.A.-N. and V.K.S., writing—review and editing, M.A.-N., V.K.S., and E.M.M.; visualization, M.A.-N. and V.K.S.; supervision, M.A.-N., V.K.S., and E.M.M.; project administration, M.A.-N. and E.M.M.; funding acquisition, M.A.-N. and E.M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The authors confirm that all datasets used in this study are publicly available and cited in the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hanna, M.G.; Parwani, A.; Sirintrapun, S.J. Whole slide imaging: technology and applications. *Adv. Anat. Pathol.* **2020**, *27*, 251–259. [[CrossRef](#)] [[PubMed](#)]
2. Roullier, V.; Lézoray, O.; Ta, V.T.; Elmoataz, A. Multi-resolution graph-based analysis of histopathological whole slide images: Application to mitotic cell extraction and visualization. *Comput. Med. Imaging Graph.* **2011**, *35*, 603–615. [[CrossRef](#)] [[PubMed](#)]
3. Doyle, S.; Madabhushi, A.; Feldman, M.; Tomaszewski, J. A boosting cascade for automated detection of prostate cancer from digitized histology. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Copenhagen, Denmark, 1–6 October 2006; Springer: Berlin/Heidelberg, Germany, 2006; pp. 504–511.
4. Janowski, P.; Strzelecki, M.; Brzezinska-Blaszczyk, E.; Zalewska, A. Computer analysis of normal and basal cell carcinoma mast cells. *Med. Sci. Monit.* **2001**, *7*, 260–265. [[PubMed](#)]
5. Lu, C.; Ma, Z.; Mandal, M. Automated segmentation of the epidermis area in skin whole slide histopathological images. *IET Image Process.* **2015**, *9*, 735–742. [[CrossRef](#)]
6. Kleczek, P.; Jaworek-Korjakowska, J.; Gorgon, M. A novel method for tissue segmentation in high-resolution H&E-stained histopathological whole-slide images. *Comput. Med. Imaging Graph.* **2020**, *79*, 101686.
7. Kleczek, P.; Dyduch, G.; Jaworek-Korjakowska, J.; Tadeusiewicz, R. Automated epidermis segmentation in histopathological images of human skin stained with hematoxylin and eosin. In Proceedings of the Medical Imaging 2017: Digital Pathology, Orlando, FL, USA, 11–16 February 2017; Volume 10140, pp. 140–158.
8. Wu, Y.; Cheng, M.; Huang, S.; Pei, Z.; Zuo, Y.; Liu, J.; Yang, K.; Zhu, Q.; Zhang, J.; Hong, H.; et al. Recent Advances of Deep Learning for Computational Histopathology: Principles and Applications. *Cancers* **2022**, *14*, 1199. [[CrossRef](#)]
9. Bándi, P.; van de Loo, R.; Intezar, M.; Geijs, D.; Ciompi, F.; van Ginneken, B.; van der Laak, J.; Litjens, G. Comparison of different methods for tissue segmentation in histopathological whole-slide images. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, Australia, 18–21 April 2017; pp. 591–595.
10. Wang, Z. A new approach for segmentation and quantification of cells or nanoparticles. *IEEE Trans. Ind. Inform.* **2016**, *12*, 962–971. [[CrossRef](#)]
11. Zhao, M.; Wang, H.; Han, Y.; Wang, X.; Dai, H.N.; Sun, X.; Zhang, J.; Pedersen, M. Seens: Nuclei segmentation in pap smear images with selective edge enhancement. *Future Gener. Comput. Syst.* **2021**, *114*, 185–194. [[CrossRef](#)]
12. Lagree, A.; Mohebpour, M.; Meti, N.; Saednia, K.; Lu, F.I.; Slodkowska, E.; Gandhi, S.; Rakovitch, E.; Shenfield, A.; Sadeghi-Naini, A.; et al. A review and comparison of breast tumor cell nuclei segmentation performances using deep convolutional neural networks. *Sci. Rep.* **2021**, *11*, 8025. [[CrossRef](#)]
13. Cui, Y.; Zhang, G.; Liu, Z.; Xiong, Z.; Hu, J. A deep learning algorithm for one-step contour aware nuclei segmentation of histopathology images. *Med. Biol. Eng. Comput.* **2019**, *57*, 2027–2043. [[CrossRef](#)]
14. Xie, L.; Qi, J.; Pan, L.; Wali, S. Integrating deep convolutional neural networks with marker-controlled watershed for overlapping nuclei segmentation in histopathology images. *Neurocomputing* **2020**, *376*, 166–179. [[CrossRef](#)]
15. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Trans. Med. Imaging* **2020**, *39*, 1856–1867. [[CrossRef](#)] [[PubMed](#)]
16. Ilyas, T.; Mannan, Z.I.; Khan, A.; Azam, S.; Kim, H.; De Boer, F. TSFD-Net: Tissue specific feature distillation network for nuclei segmentation and classification. *Neural Netw.* **2022**, *115*, 1–15. [[CrossRef](#)] [[PubMed](#)]
17. Kiran, I.; Raza, B.; Ijaz, A.; Khan, M.A. DenseRes-Unet: Segmentation of overlapped/clustered nuclei from multi organ histopathology images. *Comput. Biol. Med.* **2022**, *143*, 105267. [[CrossRef](#)] [[PubMed](#)]
18. Rączkowski, Ł.; Możejko, M.; Zambonelli, J.; Szczurek, E. ARA: Accurate, reliable and active histopathological image classification framework with Bayesian deep learning. *Sci. Rep.* **2019**, *9*, 14347. [[CrossRef](#)] [[PubMed](#)]
19. Hassan, L.; Abdel-Nasser, M.; Saleh, A.; A Omer, O.; Puig, D. Efficient Stain-Aware Nuclei Segmentation Deep Learning Framework for Multi-Center Histopathological Images. *Electronics* **2021**, *10*, 954. [[CrossRef](#)]
20. Ciga, O.; Xu, T.; Martel, A.L. Self supervised contrastive learning for digital histopathology. *Mach. Learn. Appl.* **2022**, *7*, 100198. [[CrossRef](#)]
21. Ye, H.L.; Wang, D.H. Stain-Adaptive Self-Supervised Learning for Histopathology Image Analysis. *arXiv* **2022**, arXiv:2208.04017.
22. Boserup, N.; Selvan, R. Efficient Self-Supervision using Patch-based Contrastive Learning for Histopathology Image Segmentation. *arXiv* **2022**, arXiv:2208.10779.
23. Dai, Z.; Liu, H.; Le, Q.; Tan, M. Coatnet: Marrying convolution and attention for all data sizes. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 3965–3977.
24. Chen, T.; Kornblith, S.; Norouzi, M.; Hinton, G. A simple framework for contrastive learning of visual representations. In Proceedings of the International Conference on Machine Learning, PMLR, Online, 26–28 August 2020; pp. 1597–1607.
25. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
26. Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N.Y.; Kainz, B.; et al. Attention U-Net: Learning where to look for the pancreas. *arXiv* **2018**, arXiv:1804.03999.
27. Graham, S.; Jahanifar, M.; Azam, A.; Nimir, M.; Tsang, Y.W.; Dodd, K.; Hero, E.; Sahota, H.; Tank, A.; Benes, K.; et al. Lizard: A Large-Scale Dataset for Colonic Nuclear Instance Segmentation and Classification. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 684–693.

28. Wei, J.; Suriawinata, A.; Ren, B.; Liu, X.; Lisovsky, M.; Vaickus, L.; Brown, C.; Baker, M.; Tomita, N.; Torresani, L.; et al. A Petri Dish for Histopathology Image Analysis. In Proceedings of the International Conference on Artificial Intelligence in Medicine, Virtual Event, 15–18 June 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 11–24.
29. Aksac, A.; Demetrick, D.J.; Ozyer, T.; Alhadj, R. BreCaHAD: A dataset for breast cancer histopathological annotation and diagnosis. *BMC Res. Notes* **2019**, *12*, 82. [[CrossRef](#)] [[PubMed](#)]
30. Petrick, N.; Akbar, S.; Cha, K.H.; Nofech-Mozes, S.; Sahiner, B.; Gavrielides, M.A.; Kalpathy-Cramer, J.; Drukker, K.; Martel, A.L.; et al. SPIE-AAPM-NCI BreastPathQ Challenge: An image analysis challenge for quantitative tumor cellularity assessment in breast cancer histology images following neoadjuvant treatment. *J. Med. Imaging* **2021**, *8*, 034501. [[CrossRef](#)] [[PubMed](#)]
31. Kather, J.N.; Krisam, J.; Charoentong, P.; Luedde, T.; Herpel, E.; Weis, C.A.; Gaiser, T.; Marx, A.; Valous, N.A.; Ferber, D.; et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS Med.* **2019**, *16*, e1002730. [[CrossRef](#)]
32. Kumar, N.; Verma, R.; Anand, D.; Zhou, Y.; Onder, O.F.; Tsougenis, E.; Chen, H.; Heng, P.A.; Li, J.; Hu, Z.; et al. A multi-organ nucleus segmentation challenge. *IEEE Trans. Med. Imaging* **2019**, *39*, 1380–1391. [[CrossRef](#)] [[PubMed](#)]
33. Mahbod, A.; Schaefer, G.; Bancher, B.; Löw, C.; Dorffner, G.; Ecker, R.; Ellinger, I. CryoNuSeg: A dataset for nuclei instance segmentation of cryosectioned H&E-stained histological images. *Comput. Biol. Med.* **2021**, *132*, 104349.
34. Naylor, P.; Laé, M.; Reyat, F.; Walter, T. Nuclei segmentation in histopathology images using deep neural networks. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, Australia 18–21 April 2017; pp. 933–936.
35. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 3–11.
36. Mahmood, F.; Borders, D.; Chen, R.J.; McKay, G.N.; Salimian, K.J.; Baras, A.; Durr, N.J. Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE Trans. Med. Imaging* **2019**, *39*, 3257–3267. [[CrossRef](#)]
37. Zeng, Z.; Xie, W.; Zhang, Y.; Lu, Y. RIC-Unet: An Improved Neural Network Based on Unet for Nuclei Segmentation in Histology Images. *IEEE Access* **2019**, *7*, 21420–21428. [[CrossRef](#)]
38. Naylor, P.; Laé, M.; Reyat, F.; Walter, T. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Trans. Med. Imaging* **2018**, *38*, 448–459. [[CrossRef](#)]
39. Valanarasu, J.M.J.; Oza, P.; Hacıhaliloglu, I.; Patel, V.M. Medical transformer: Gated axial-attention for medical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Virtual Event, 27 September–1 October 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 36–46.
40. Chanchal, A.K.; Lal, S.; Kini, J. Deep structured residual encoder-decoder network with a novel loss function for nuclei segmentation of kidney and breast histopathology images. *Multimed. Tools Appl.* **2022**, *81*, 9201–9224. [[CrossRef](#)]
41. Xiang, T.; Zhang, C.; Liu, D.; Song, Y.; Huang, H.; Cai, W. BiO-Net: Learning recurrent bi-directional connections for encoder-decoder architecture. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 74–84.
42. Ali, H.; Cui, L.; Feng, J. MSAL-Net: Improve accurate segmentation of nuclei in histopathology images by multiscale attention learning network. *BMC Med. Inform. Decis. Mak.* **2022**, *22*, 22. [[CrossRef](#)] [[PubMed](#)]