DAN-NucNet: A Dual Attention Based Framework for Nuclei Segmentation in Cancer Histology Images Under Wild Clinical Conditions

Abstract

Nuclei segmentation plays an essential role in histology analysis. The nuclei segmentation in histology images is challenging in variable conditions (clinical wild), such as poor staining quality, stain variability, tissue variability, and conditions having higher morphological variability. Recently, some deep learning models have been proposed for nuclei segmentation. However, these models rarely solve the problems mentioned above simultaneously. Most of the information in Hematoxylin and Eosin (H&E) stained histology images is in its channel, and the remaining information is in the spatial domain. We observed that most problems could be solved by considering channel and spatial features simultaneously, e.g., the spatial and channel features provide the solution to the morphological variability and staining variability, respectively. Therefore, we propose a novel spatial-channel attention-based modified UNet architecture with ResNet blocks in encoder layers. The UNet baseline preserves coarse and fine features, thus proving the solution to the tissue variability. The proposed method significantly improves the segmentation performance compared to the state-of-the-art methods on three different benchmark datasets. We demonstrate that the proposed model is generalized for 20 cancer sites, more than any reported literature. The proposed model is less complex than most state-of-the-art models. The impact of the proposed model is that it will help improve further procedures such as nuclei instance segmentation, nuclei classification, and cancer grading.

Keywords: Nuclei segmentation, Histopathology segmentation, Attention UNet, Spatial Channel Attention
PACS: 0000, 1111
2000 MSC: 0000, 1111

Table 1: List of acronyms used in the document

Acronym	Definition
BCE	Binary Cross Entropy
CAD	Computer Assisted Diagnosis
CNN(s)	Convolutional Neural Network(s)
DCNN(s)	Deep Convolutional Neural Network(s)
FCN(s)	Fully Connected Networks(s)
H&E	Hematoxylin and Eosin
H&N	Head and Neck
ReLU	Rectified Linear Unit
WSI	Whole Slide Image

1. Introduction

Hematoxylin and Eosin (H&E) stained histology slides are a gold standard for micro-level cancer analysis. In medical practices, a human expert analyzes these slides. In case Computer-assisted diagnosis (CAD) systems are available, digital microscopic images are taken and analyzed. These digitized whole slide images (WSI) are usually of size 100,000 x 100,000. This process is tiresome and requires advanced technical knowledge. Although computer-based automated analysis is being used in the routine clinical process of cancer diagnosis, it is still subjective to the human interpretation for cases such as nuclei segmentation. Nuclei segmentation is a fundamental requirement for the analysis and diagnosis procedures. For example, leiomyosarcoma is identified with the help of blunt-ended nuclei. Abnormal epithelium nuclei distribution helps to identify squamous cell carcinoma (SCC). In prostate cancer, the enlargement of the nuclei helps to identify adenocarcinoma. Thus the importance of the precise segmentation of nuclei in the histology images can not be neglected.

The conventional nuclei segmentation in histology is a routine task for technicians. However, it is still a challenging task for CAD's in clinically wild conditions. This is due to the staining variability, poor staining quality, staining artifacts, variable magnification level, nuclear density, variation in tissue type, and image quality

(Xing and Yang, 2016). Early CAD systems were based on traditional image processing methods (Madabhushi and Lee, 2016). The most significant disadvantages of traditional methods are that they lack generalization and robustness. Recently, deep learning semantic segmentation methods have been developed for various cancers. These methods usually improve the generalization ability. The early class of such methods includes FCN's (Fully Connected Networks) and their variants (Sultana et al., 2020). Although these methods performed significantly better than the traditional techniques, these methods are unaware of global contexts, which are essential for tissue generalization and susceptibility to staining variation. Furthermore, they produce coarse output and have slow real-time inference performance. The encoder-decoder class of networks (SegNet, U-Net, DeConvNet, V-Net, and their variants) improves the tissue generalization capability, and they are less susceptible to staining variations and artifacts. However, they lose high-resolution representations through the encoding process (Shervin et al., 2021). Recently, a few methods using attention mechanisms have been developed for medical imaging (Shan and Yan, 2021; Lal et al., 2021; Jin et al., 2020; Guo et al., 2021; Schlemper et al., 2019). The attention mechanism for histology segmentation usually uses channel attention. The channel attention mechanism has improved performance concerning some of the wild conditions.

As discussed earlier, most of the information in H&E stained histology images is in its channel; the remaining information is in the spatial domain. Using this information accurately may further increase the nuclei segmentation performance, especially in wild clinical conditions. We believe that integrating multiple attention mechanisms provides further performance improvement based on the above observations. We propose a robust method for segmenting out nuclei from histology images. The proposed method is based on a modified UNet with dual attention mechanisms, a redesigned encoder, and a decoder. The spatial attention provides the solution to the morphological variability, while the channel attention provides the staining variability. The UNet baseline preserves coarse and fine features, thus proving the solution to the tissue variability. The resulting segmentation act as an essential stage for automatic computer-based cancer analysis. It provides prognostic value by acting as a pre-processing stage. The prognostic values include grading, classification, progression, and micro-environment analysis. It will also be a pre-processing step for an automated cancer grading system. The main contribution of this work is as follows,

1. A novel network that uses both channel and spatial features to overcome the drawbacks of the existing literature. The dual attention mechanism embedded into a UNet makes it less susceptible to harsh conditions caused due to morphological variability, tissue and staining variation, and staining quality.
2. We specifically demonstrate that the proposed network is widely generalized in terms of tissue variability. It is tested on 20 different tissue sites in three different largest available datasets, far more than any reported literature.
3. The network is also generalized for a standard range (20x and 40x) of magnification levels. The datasets used to test the network consist of both standard magnification levels.
4. The detailed studies of the method demonstrate the robustness of the method from a broader perspective.

2. Related Work

Early traditional image processing techniques (Rogojanu et al., 2015; Landini et al., 2016; Fouad et al., 2017; Linder et al., 2012) use basic features of objects in histology images, including shape, size, texture, color, location, and local binary patterns (LBPs). The biggest drawbacks of these traditional techniques are that they are not generalized and have extremely low segmentation performance due to hand-crafted features. Supervised deep learning methods have recently been used to segment medical images due to their superior performance. The earliest deep learning methods were FCN's (Zhang et al., 2017; Zhao et al., 2017; Yuan et al., 2017; Rezaei et al., 2019b). These methods provide better generalization compared to the traditional methods. However, they show minimal performance for problems such as tissue-wise generalization, staining variability, etc. The development of the encoder-decoders class of models, specifically for medical image segmentation, outclassed their predecessor FCN style models. The most common is UNet (Ronneberger et al., 2015), specifically designed for biomedical image segmentation. (Yurttakal and Erbay, 2020) presents a method for segmenting the cells in combined P63 and H&E stained histology images for larynx cancer based on (Ronneberger et al., 2015). (Zhou et al., 2018b) introduces convolution layers in the skip connection path to further learn from the intermediate features set, thus further improving UNet capabilities. Other variants of the UNet (Dubost et al., 2017; Naylor et al., 2019; Li et al., 2018) were proposed to have a slight improvement in the performance. In general encoder-decoder models

can counter stain variability, poor staining, and morphological variation to an extent. Parallel to the encoder-decoder networks, another class of models for nuclear segmentation based on ResNet and its variants was proposed. (Lin et al., 2017) and (Zhou et al., 2018a) proposed a model which makes use of the residual structure (He et al., 2016a). (Simon et al., 2019) proposed a network is based on Preact-ResNet50 (He et al., 2016b), which outperformed the existing encoder-decoder and ResNet style networks.

Attention mechanisms have recently shown superior performance for segmentation tasks, particularly nuclear segmentation. (Lal et al., 2021) proposed a NucleiSegNet, which is a UNet-style architecture. It uses an attention mechanism developed by (Schlemper et al., 2019). The encoder is based on ResNet. The ResNet and attention gate combination produces highly improved results compared to the existing methods. The attention mechanism designed with a UNet baseline helps keep the benefits of the encoder-decoder and the attention benefits simultaneously.

3. Proposed Architecture

The literature review suggests that UNet-based deep learning architectures are highly effective for segmentation H&E stained histology images. In a UNet, the encoder extracts high semantic coarse features. In contrast, the decoder extracts low semantic fine features. These features are fused at every stage. This helps the network to reuse the spatial information, which is lost during the max-pooling operations. This produces improved performance compared to its predecessors. However, the network still has limitations. The encoder features are directly fused with decoder features. These features possess information that does not contribute to the performance of the network. It produces extra parameters, increasing the computation and memory requirements. Instead, only the required information should be fused. In the case of histology images, the required information is inside spatial and channel features. The histology images are H&E stained. Therefore a huge part of the information lies in the channels of the image. The second problem is that the contextual information is not enough. It leads to reduced performance due to the gap between the low and high-level contextual features. The third problem is the encoder block itself. The encoder block possesses simple convolution layers. These blocks introduce a gap between high-level and low-level features. The introduction of the intermediate layers at the skip connection path to tackle the first problem causes a minor vanishing gradient problem. This problem is solved by replacing these convolution blocks with ResNet blocks.

Using the earlier observations, we propose the following amendments. Inspired by (Khanh et al., 2020), we introduce spatial-channel attention blocks as intermediate layers between encoder and decoder. Furthermore, we replace the convolution blocks with a redefined ResNet Layer. The encoder is made up of a ResNet block and max-pooling layer. The attention block is made up of spatial attention and channel attention blocks. The decoder is made up of attention fusion and transposed convolution. Fig. 1 shows the network and its building blocks. Further details are presented in the following subsections.

3.1. Residual Block

The residual block extracts semantic rich features at each level (i). We have redesigned the original ResNet block proposed by (He et al., 2016a) to make it more robust. The input to the residual block at each level is passed through the 1x1 convolution layer, followed by the ReLU layer. This is followed by 3x3 convolution, ReLU, and a batch normalization layer, respectively. Another 1x1 convolution and ReLU are applied. The output is then fused with the input passed through 1x1 convolution. This fuses the shallow-level semantic features with the deep-level semantic features. This enables the residual blocks to learn from both shallow and deep-level semantic features. After fusion, a 3x3 convolution followed by ReLU and batch normalization is applied. The proposed residual block is shown in Fig. 1. Mathematically, the residual block is expressed as, Eq.(1).

$$\left.\begin{aligned} X_{i,0} &= H^F_{1,1}\left(in_i\right) \\ C_{i,0} &= f\left(H^F_{1,1}\left(in_i\right)\right) \\ C_{i,1} &= BN\left(f\left(H^F_{3,3}\left(C_{i,0}\right)\right)\right) \\ X_{i,1} &= f\left(H^F_{1,1}\left(C_{i,1}\right)\right) \\ out_i &= BN\left(f\left(H^F_{3,3}\left\{X_{i,0} \oplus X_{i.1}\right\}\right)\right) \end{aligned}\right\} \quad (1)$$

where in_i denotes input and Out_i denotes output of the residual block at any layer i. $C_{i,j}$ denotes intermediate features, $H^F_{k,k}$ denotes convolution, k denotes kernels, F denoted filters, f denotes ReLU activation, BN denotes batch normalization and $\oplus$ denotes concatenation.

3.2. Attention Mechanism

To solve the problems caused due to the direct skip connections, various attempts have been made (Zhou et al., 2018b; Zeng et al., 2019; Weng et al., 2019; Wan

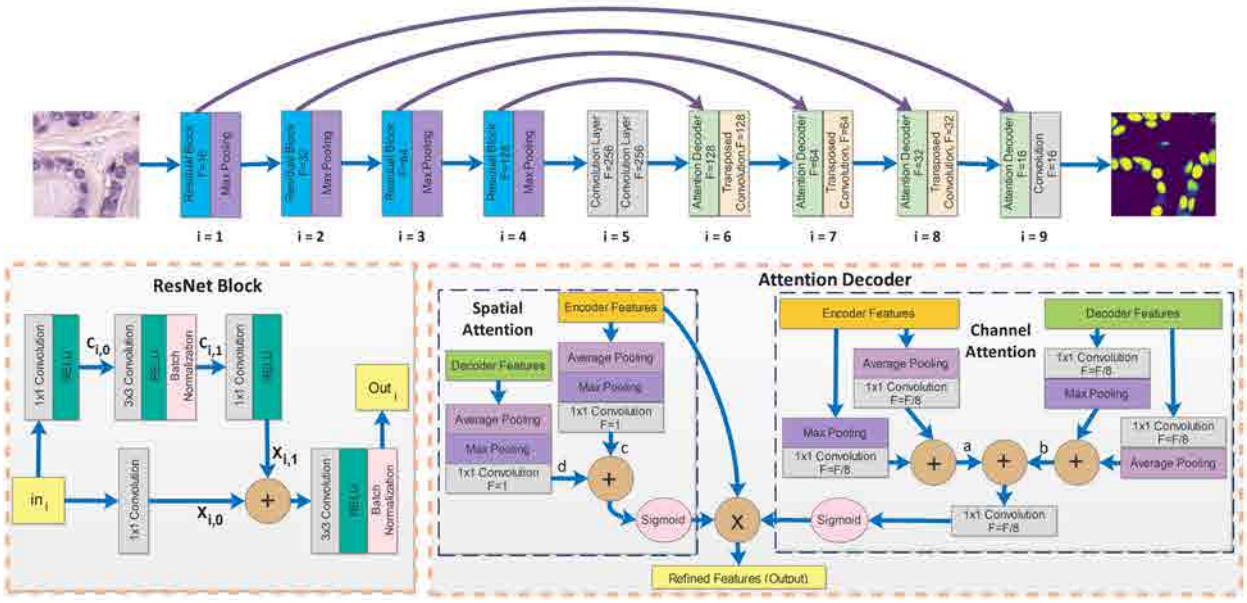

Figure 1: Overview of the proposed model architecture along with the detailed attention mechanism and ResNet Block.

et al., 2020). These methods use features maps from the encoders without removing the redundant features. Recently, single-stage attention mechanisms have been proposed (Jin et al., 2020; Guo et al., 2021). These have significantly improved the segmentation performance of the UNet based networks. However, multiple attention mechanisms are required for histology images segmentation because of both spatial and channel features. Therefore we deploy spatial-channel attention inspired by (Khanh et al., 2020). However, we have added transposed convolution layer at the end of the decoder instead of naive upsample layer, after the feature fusion. This enables the network to learn the upsampling, which provides better performance for poorly stained slides. The channel and spatial attention blocks are shown in Fig. 1. These blocks are expressed mathematically as Eq.(2), and Eq.(3),

$$\begin{aligned} a &= H_{1,1}^{F/8}\left(AvgPool\left(F_e\right)\right) \oplus H_{1,1}^{F/8}\left(MaxPool\left(F_e\right)\right) \\ b &= H_{1,1}^{F/8}\left(AvgPool\left(F_d\right)\right) \oplus H_{1,1}^{F/8}\left(MaxPool\left(F_d\right)\right) \\ A_c &= \sigma\left(H_{1,1}^{F/8}\left(a \oplus b\right)\right) \end{aligned} \tag{2}$$

$$\begin{aligned} c &= H_{1,1}^{1}\left(MaxPool\left(AvgPool\left(F_e\right)\right)\right) \\ d &= H_{1,1}^{1}\left(MaxPool\left(AvgPool\left(F_d\right)\right)\right) \\ A_s &= \sigma\left(c \oplus d\right) \end{aligned} \tag{3}$$

where $H_{k,k}^{F}$ denotes convolution, k denotes kernels, F denoted filters, and $\oplus$ denotes concatenation. A_c, and A_s are the outputs of the channel attention and spatial attention blocks, F_d represent decoder features, F_e represents encoder features respectively.

Finally the refined features F_r are obtained by element wise multiplication ($\otimes$) of spatial attention map A_s and channel attention map A_c with the encoder features F_e. The feature fusion is mathematically represented as, Eq.(4).

$$F_r = F_e \otimes A_s \otimes A_c \tag{4}$$

3.3. Loss Function

Initially, we used binary cross-entropy loss function, which is widely used for segmentation tasks. However, when compared with the loss function reported by (Lal et al., 2021), for nuclei segmentation, we found that the loss function reported by (Lal et al., 2021) slightly performed better (see Section 6.3 Ablation Study for comparison). The loss function is represented as, Eq.(5).

$$L = \frac{Dice\ Loss * Jaccard\ Loss}{Dice\ Loss + Jaccard\ Loss} \tag{5}$$

4. Materials and Evaluation matrices

4.1. Dataset

The proposed architecture is evaluated on three different H&E stained digital histology datasets. The first dataset, PanNuke was recently presented by (Gamper et al., 2020). It is the most comprehensive histology dataset. The dataset is obtained from nineteen different tissues, mainly H&N, breast, liver, prostate, lungs,

etc. The dataset contains semi-automatically annotated 200,000 cancer and non-cancer nuclei. The dataset is available in three splits while maintaining the ratio of cancer and non-cancer cells. The second dataset we have used is Kumar (Kumar et al., 2017). The dataset contains seven different normal and cancerous tissues, i.e., breast, bladder, colon, liver, kidney, prostate, and stomach. The dataset contains a total of 30 images with around 22,000 nuclei. The third dataset we have used is CPM17 (Computational Precision Medicine - 2017) (Vu et al., 2019). It contains a total of 7570 nuclei. The images are taken at 40x and 20x magnification. Since the above-mentioned data is obtained from different sources, we apply pre-processing to convert it to a single standard format. The data is prepared under the following policy.

- The size of the image should be 256x256. If the image's size is greater than 256x256, it is resized by taking 256x256 patches from the image. For example, if the original data size is 1024x1024, then 16 patches of 256x256 are extracted from the image.
- The magnification should be 20x or 40x. If the magnification is different, the image is digitally zoomed in or zoomed out such that the magnification effect of 40x or 20x is produced. The 40x magnification is selected since the standard WSI is usually at 20x or 40x magnification (Mahbod et al., 2021)
- Data augmentation is applied to datasets for training. Datasets with more than 250 images are flipped vertically and horizontally. Additional data augmentation techniques, such as rotation and shifting, are applied to the datasets having less than 250 images.
- Blank images or images with no labels are removed.

It is to be noted that we only utilize other datasets, i.e., Kumar and CPM-17, to test our model. We do not train our model on these datasets. During our experiments, we trained our model on these smaller datasets but found that these datasets do not provide efficient results compared to when we trained the model on the PanNuke dataset. To test our model using these datasets, we utilize the test split provided by the datasets' authors.

4.2. Evaluation Criteria

We have used the F1 score, Jaccard index, precision, and recall as performance evaluation metrics. These matrices are commonly used as performance evaluation matrices in histology image segmentation (Tosta et al., 2019; Das et al., 2018; Rezaei et al., 2019a). The precision measures the purity of positive detection for the ground truth, while recall measures the completeness of positive predictions to the ground truth. F1 score is defined as the spatial overlap between the humanly labeled image and the computer-based segmented image. Jaccard index is used to measure the similarity between the humanly labeled image and the computer-based segmented image.

5. Experimentation and Results

5.1. Experimental Setup

The proposed architecture is developed using Keras, a python library for developing deep learning models. The implementation of the proposed model and its weights are accessible at this link. The model is trained using Nvidia GTX 1660 Ti. The learning rate is initiated as 0.003. During training, the learning rate is reduced by a factor of 0.5 when the improvement in the loss function remains below $1x10^{-6}$ for five consecutive epochs. We stop the training process when the validation loss does not improve for ten consecutive epochs.

5.2. Results

Tabel 2 and Tabel 3 shows tissue wise results obtained using PanNuke and Kumar dataset. The results of the proposed methods have been compared with NucleiSegNet (Lal et al., 2021). It is observed that in terms of F1 score and Jaccard index, the proposed model has outperformed NucleiSegNet for both of the datasets. The proposed network has shown better performance in all matrices for the adrenal gland, bile duct, bladder, breast, liver, prostate, testis, and pancreas. For stomach, however, the NucleiSegNet performs better in terms of precision only. On average, the proposed architecture has shown significantly improved performance in all performance evaluation matrices. The significance of the results is observed in a box plot shown in Fig. 2. The interquartile range (the colored boxes) shows a significant improvement in the average performance. The color dots depict the outliers. It is also observed that the proposed network significantly reduces the outliers. Furthermore, the line inside the box represents the median. Although the maximum F1 score and Jaccard index show little improvement, the median performance has been drastically improved.

The proposed model is also well generalized. The performance has been evaluated on a wide range of 20

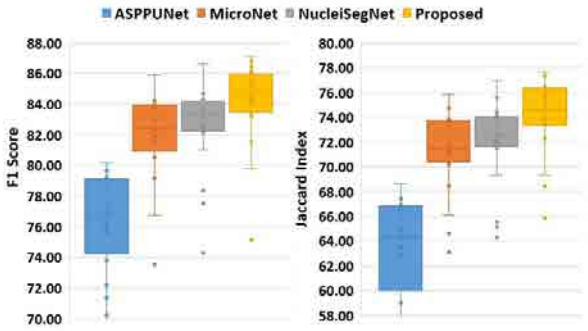

Figure 2: Box Plot represents a comparison of the proposed network with the best performing networks in the literature.

different tissues. The generalization performance of the network is clearly observed in Tabel 2 and Table 3. The visual performance of the network for various tissues within the PanNuke dataset is shown in Fig. 3. Fig. 3 shows various histology images obtained in variable conditions.

The proposed method has also been compared with the top-performing existing architectures using three different datasets (PanNuke, Kumar, and CPM dataset). These include UNet (Ronneberger et al., 2015), UNet++ (Ronneberger et al., 2015), DIST (Naylor et al., 2019), ASPPU Net (Wan et al., 2020), SCPP-Net (Chanchal et al., 2021), SCA-Net (Shan and Yan, 2021), and MicroNet (Rezaei et al., 2019b). Table 4 shows the detailed comparison of the proposed architecture with the existing ones. The results are reported by taking the average of all cancer types in the specific datasets.

6. Discussion

This section provides an in-depth discussion on the model performance in terms of complexity, robustness, and ablation studies performed upon the proposed model.

6.1. Complexity Analysis

We have compared the complexity of the proposed architecture with the existing literature. The proposed architecture has a total of 9.276 million parameters. The parameters are far much less than most of the existing architectures. Only DIST (Naylor et al., 2019), SCPP-Net (Chanchal et al., 2021), and ASPP UNet (Wan et al., 2020) have fewer parameters than the proposed network. The proposed method is more efficient and less complex at the same time. The parameters comparison is shown in Fig. 4. This comparison is necessary yet ignored in the previous studies. The size of the WSI is usually 100,00x100,00 pixels, meaning a more complex algorithm and ultimately more inference time.

6.2. Robustness Analysis

The networks perform well within the standard magnification range (20x and 40x) in terms of magnification. The example of variable magnification level is shown in Fig. 5. The network shows adequate performance in complex or variable situations. In complex situation such as Fig. 7, it is even extremely difficult for a professional pathologist to identify nuclei. Still, the proposed network performs in an acceptable range compared to the state-of-the-art. The network, however, produces false-positive edges for connective tissue in some cases. Examples of these false positives are highlighted in blue bounding boxes in Fig. 6. For both neoplastic and non-neoplastic nuclei, the network occasionally produces false-negative at the edge pixels of the nuclei. This could be observed in the green color in Fig. 3. These false positives are, however, produced by other approaches as well.

6.3. Ablation Study

We performed an ablation study in 4 stages. We evaluated our network without spatial attention mechanism, channel attention mechanism, and ResNet block, respectively, using the largest dataset, PanNuke. We also evaluated the model training with a cross-entropy loss function instead of the one proposed by (Lal et al., 2021). The results of the ablation study are shown in Tabel 5. These results are obtained by averaging the results of cancer types in the PanNuke dataset. The ablation study verifies our assumptions regarding the usage of the dual attention mechanism, the suggested ResNet usage, and the suggested loss function.

7. Conclusion

Computer-assisted semantic segmentation of histology images is of great importance. It improves and fastens the diagnosis process. Computer-assisted staging, grading, classification of cancer, and recurrence prediction depend upon the segmentation of cells. Existing methods lack generalization and are less accurate in harsh conditions. This research proposes spatial-channel attention-based UNet architecture, which uses spatial and channel features and the ResNet blocks' robustness to segment the target nuclei. The proposed architecture is tested comprehensively on the most extensive available dataset. The proposed architecture is

Table 2:
Segmentation quality metrics of the proposed architecture compared with NucleiSegNet for multi-organ PanNuke dataset.

	NucliSegNet				**Proposed**			
	F1	JI	Recall	Prec	F1	JI	Recall	Prec
Adrenal Gland	82.16	72.23	83.02	84.61	**83.51**	**73.94**	**84.52**	**85.01**
Bile-duct	82.47	72.04	86.72	80.58	**83.49**	**73.39**	**87.27**	**81.90**
Bladder	83.79	73.83	82.56	**87.02**	**86.44**	**77.34**	**86.61**	86.81
Breast	82.66	71.51	83.49	83.15	**83.23**	**72.27**	**83.88**	**83.62**
Cervix	84.53	75.59	**85.41**	85.06	**85.10**	**76.43**	85.19	**86.12**
Colon	77.51	65.53	77.77	80.18	**79.81**	**68.39**	**81.08**	**80.20**
Esophagus	84.33	74.40	83.99	**85.54**	**85.29**	**75.76**	**85.89**	85.53
Head & Neck	81.02	69.34	77.28	**89.20**	**86.79**	**75.46**	**86.45**	88.42
Kidney	83.41	73.22	86.21	**82.38**	**84.46**	**74.57**	**88.72**	81.44
Liver	84.73	74.42	85.45	85.05	**86.20**	**76.62**	**86.15**	**87.12**
Lung	78.39	65.16	75.18	83.00	**81.57**	**69.35**	**79.73**	**84.21**
Ovarian	84.69	74.15	**87.90**	82.26	**85.30**	**75.08**	87.85	**83.34**
Pancrease	83.33	72.29	87.93	80.12	**85.10**	**74.68**	**88.98**	**82.09**
Prostate	83.07	72.16	85.56	82.15	**84.29**	**73.83**	**87.04**	**82.67**
Skin	74.27	64.31	**84.68**	71.71	**75.11**	**65.87**	83.05	**73.87**
Stomach	86.68	76.99	**90.50**	83.73	**87.12**	**77.65**	89.48	**85.23**
Testis	83.70	73.19	87.30	81.74	**86.12**	**76.55**	**87.86**	**85.27**
Thyroid	83.77	73.96	**90.48**	79.69	**85.49**	**76.06**	90.36	**82.42**
Uterus	83.78	72.71	**87.52**	80.96	**84.23**	**73.38**	86.97	**82.27**
Average	82.54	71.95	84.68	82.53	**84.14**	**74.03**	**86.16**	**83.56**

Table 3:
Segmentation quality metrics of the proposed architecture compared with NucleiSegNet for multi-organ Kumar dataset.

	NucliSegNet				**Proposed**			
	F1	JI	Recall	Prec	F1	JI	Recall	Prec
Bladder	81.27	69.52	88.02	76.64	**83.22**	**72.20**	**88.41**	**79.44**
Breast	79.48	66.59	87.93	73.32	**80.03**	**67.37**	**87.95**	**74.28**
Colon	73.78	59.70	**74.50**	76.01	**75.83**	**62.37**	77.12	**77.16**
Kidney	81.35	68.95	**88.13**	75.86	**81.55**	**69.24**	87.92	**76.40**
Liver	78.02	64.73	79.27	79.71	**78.83**	**65.80**	**79.77**	**80.39**
Prostate	77.97	65.20	79.73	78.50	**79.57**	**67.19**	**82.14**	**78.93**
Stomach	88.33	79.46	**91.59**	**85.56**	**88.76**	**80.12**	90.15	87.64
Average	80.03	67.74	84.17	77.94	**81.11**	**69.18**	**84.78**	**79.18**

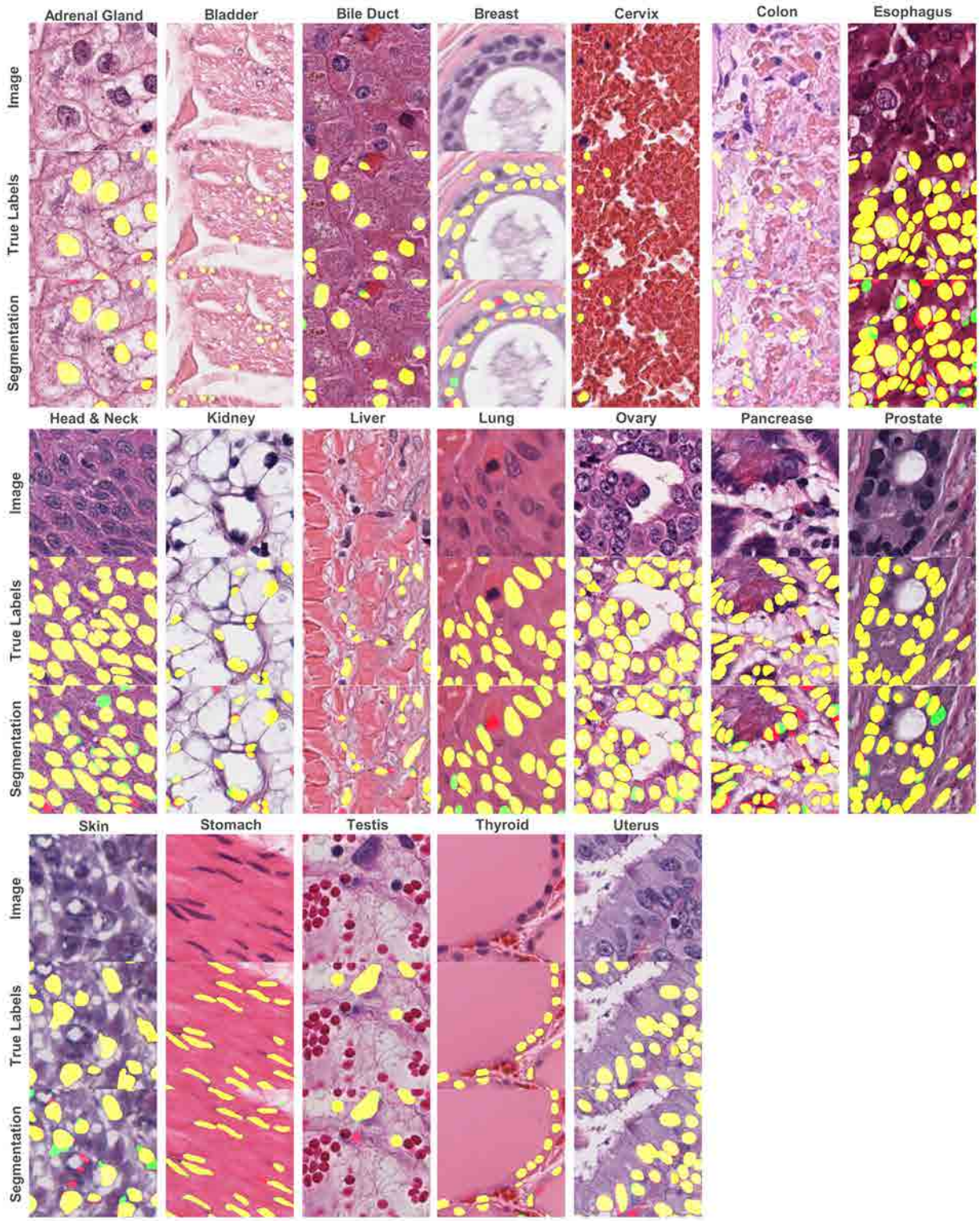

Figure 3: Example of visual results from PanNuke dataset (Gamper et al., 2020). in variable conditions. The first row of each tissue shows the original image; the second row shows the true labels overlayed in yellow, while the third row shows the segmentation performed by the proposed method. In the third row, the yellow color shows true positive, the green color shows false negative, and the red color shows false positive.

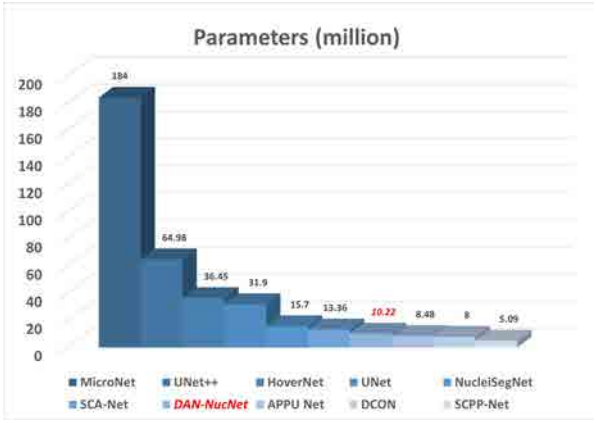

Figure 4: Complexity comparison of the proposed architecture with the existing literature in terms of the model parameters.

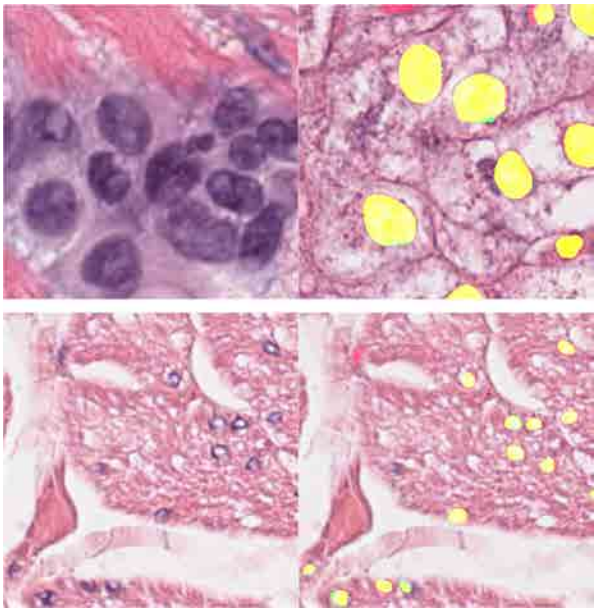

Figure 5: Example of different magnification level images (to the left) and the segmented image (to the right).

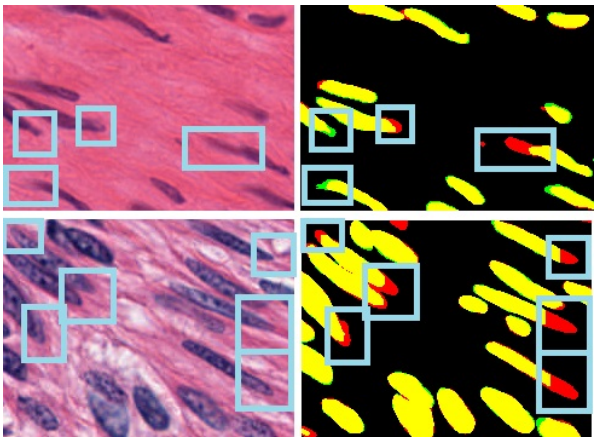

Figure 6: The model produces higher false positive edges in the case of connecting tissues, the examples of which are highlighted within the blue boxes.

widely generalized in terms of tissue location and a wide range of variations in histology caused by stain artifacts and variations in the staining process. Apart from the maximum achieved F1 and JI, the reduction in outliers has greatly improved the average segmentation performance while having minimal complexity in terms of parameters. Due to the lack of a dataset, the model is tested for nuclei of epithelial tissue, connective/soft tissue, and dead tissue. The future work will include the extension to other tissue types.

Acknowledgments

We are thankful to Dr.Muhmmad Taimur at Khyber Medical University, Pakistan for his support and help in oncology.

Funding Sources

This research work was supported by the National Natural Science Foundation of China under Grant No. 62171377.

Declaration of Competing Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

Chanchal, A.K., Kumar, A., Lal, S., Kini, J., 2021. Efficient and robust deep learning architecture for segmentation of kidney and breast histopathology images. Computers & Electrical Engineering 92, 107177. doi:10.1016/j.compeleceng.2021.107177.

Table 4:
The table compares results obtained from DAN-NucNet with the relevant existing methods in terms of standard evaluation metrics. The results are reported by taking the average of all cancer types in the specific dataset.

	PanNuke				**Kumar**				**CPM**			
Previous Work	F1	JI	Prec	Recall	F1	JI	Prec	Recall	F1	JI	Prec	Recsll
UNet (Ronneberger et al., 2015)	77.56	65.22	79.00	79.03	75.51	62.29	78.31	74.90	71.52	56.17	75.18	70.46
UNet++ (Zhou et al., 2018b)	78.13	66.03	81.64	77.10	76.07	62.29	80.29	74.20	72.83	56.71	76.99	69.07
DIST (Naylor et al., 2019)	71.79	55.47	68.43	76.48	68.40	53.27	65.15	74.76	70.24	55.91	80.57	66.05
MicroNet (Rezaei et al., 2019b)	81.32	70.17	82.71	81.16	78.30	65.34	82.18	76.74	73.26	58.36	85.85	72.40
ASPP UNet (Wan et al., 2020)	75.97	63.15	79.28	77.02	73.66	59.40	76.97	77.29	72.64	59.42	85.82	64.89
Enhanced SC-UNet (Khanh et al., 2020)	82.30	71.53	84.17	82.12	77.57	64.40	82.54	75.27	76.52	64.17	87.44	71.78
SCPP-Net (Chanchal et al., 2021)	77.36	64.70	77.73	78.81	76.35	62.57	78.76	75.65	71.56	59.78	80.43	69.01
SCA-Net (Shan and Yan, 2021)	82.31	71.69	**87.19**	78.85	79.14	67.92	84.65	78.12	76.29	77.96	**87.61**	71.91
NucleiSegNet (Lal et al., 2021)	82.54	71.95	84.68	82.53	80.03	67.74	84.17	77.94	**77.13**	78.18	87.57	72.44
Proposed	**84.14**	**74.03**	86.16	**83.56**	**81.11**	**69.18**	**84.78**	**79.18**	77.10	**78.22**	87.46	**72.59**

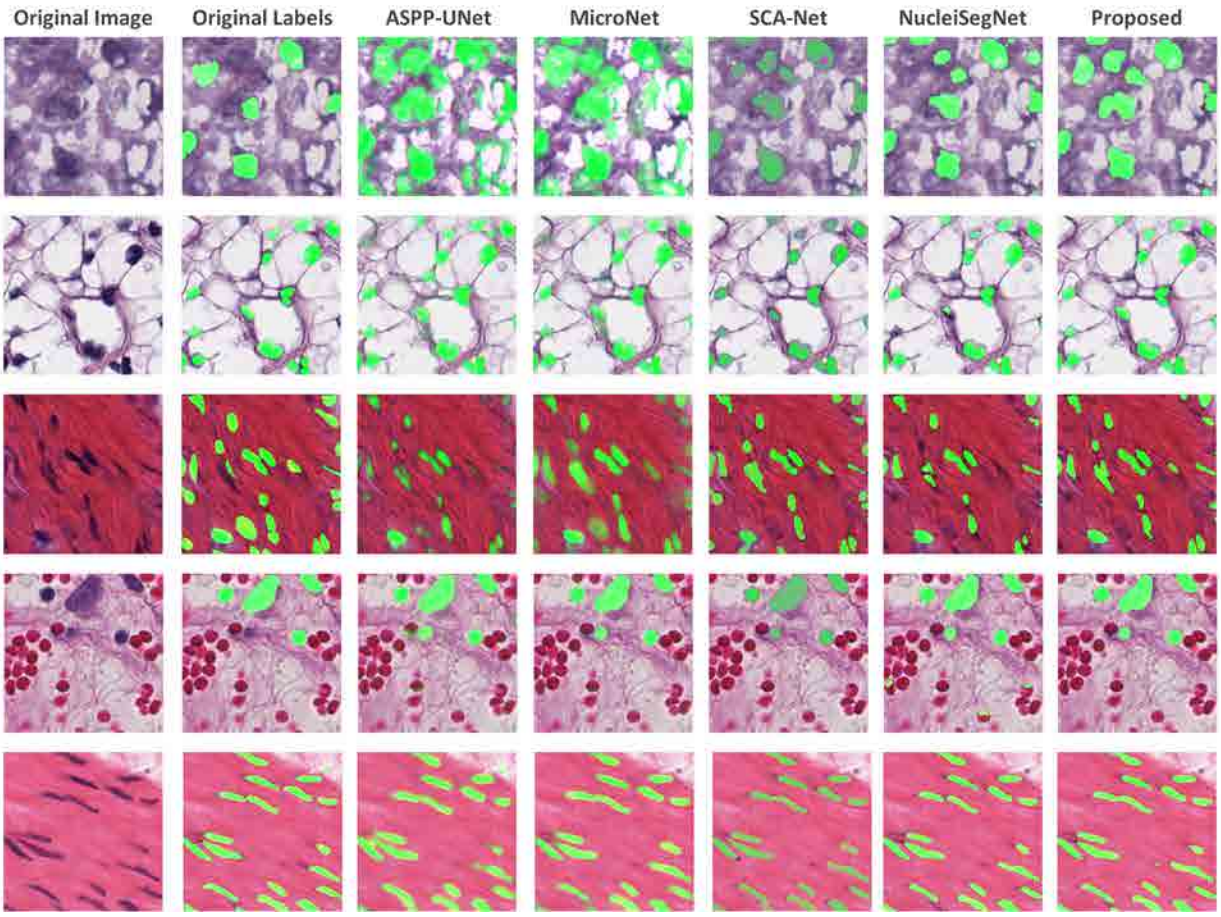

Figure 7: Each row in the image show example of a histology image in variable conditions along with the original labels, segmented image using ASPP UNet, MicroNet, SCA-Net, NucleiSegNet, and proposed network, respectively.

Table 5:
The table summarizes the ablation studies performed on the proposed network.

Ablation studies	F1	JI	Prec	Recall
Model performance without spatial attention	81.45	70.48	85.25	80.31
Model performance without channel attention	81.64	70.76	81.41	82.46
Model performance without ResNet blocks	85.31	72.10	86.13	81.55
Model performance with BCE as loss function	84.00	73.04	85.92	82.84
Proposed	84.14	74.03	86.16	83.36

Das, D.K., Bose, S., Maiti, A.K., Mitra, B., Mukherjee, G., Dutta, P.K., 2018. Automatic identification of clinically relevant regions from oral tissue histological images for oral squamous cell carcinoma diagnosis. Tissue and Cell 53, 111–119. doi:10.1016/j.tice.2018.06.004.

Dubost, F., Bortsova, G., Adams, H., Ikram, A., Niessen, W.J., Vernooij, M., Bruijne, M.D., 2017. GP-Unet: Lesion detection from weak labels with a 3D regression network, in: Medical Image Computing and Computer Assisted Intervention - MICCAI 2017. Springer International Publishing, pp. 214–221. doi:10.1007/978-3-319-66179-7_25.

Fouad, S., Randell, D., Galton, A., Mehanna, H., Landini, G., 2017. Unsupervised morphological segmentation of tissue compartments in histopathological images. PloS one 12, e0188717. doi:10.1371/journal.pone.0188717.

Gamper, J., Koohbanani, N.A., Graham, S., Jahanifar, M., Khurram, S.A., Azam, A., Hewitt, K., Rajpoot, N., 2020. PanNuke dataset extension, onsights and baselines. arXiv preprint arXiv:2003.10778 doi:10.48550/arXiv.2003.10778.

Guo, C., Szemenyei, M., Yi, Y., Wang, W., Chen, B., Fan, C., 2021. SA-UNet: Spatial attention U-Net for retinal vessel segmentation, in: 2020 25th International Conference on Pattern Recognition (ICPR), IEEE. doi:10.1109/icpr48806.2021.9413346.

He, K., Zhang, X., Ren, S., Sun, J., 2016a. Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. doi:10.1109/CVPR.2016.90.

He, K., Zhang, X., Ren, S., Sun, J., 2016b. Identity mappings in deep residual networks, in: European conference on computer vision, Springer. pp. 630–645. doi:10.1007/978-3-319-46493-0_38.

Jin, Q., Meng, Z., Sun, C., Cui, H., Su, R., 2020. RA-UNet: A hybrid deep attention-aware network to extract liver and tumor in CT scans. Frontiers in Bioengineering and Biotechnology 8. doi:10.3389/fbioe.2020.605132.

Khanh, T.L.B., Dao, D.P., Ho, N.H., Yang, H.J., Baek, E.T., Lee, G., Kim, S.H., Yoo, S.B., 2020. Enhancing U-Net with spatial-channel attention gate for abnormal tissue segmentation in medical imaging. Applied Sciences 10, 5729. doi:10.3390/app10175729.

Kumar, N., Verma, R., Sharma, S., Bhargava, S., Vahadane, A., Sethi, A., 2017. A dataset and a technique for generalized nuclear segmentation for computational pathology. IEEE Transactions on Medical Imaging 36, 1550–1560. doi:10.1109/tmi.2017.2677499.

Lal, S., Das, D., Alabhya, K., Kanfade, A., Kumar, A., Kini, J., 2021. NucleiSegNet: Robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images. Computers in Biology and Medicine 128, 104075. doi:10.1016/j.compbiomed.2020.104075.

Landini, G., Randell, D., Fouad, S., Galton, A., 2016. Automatic thresholding from the gradients of region boundaries. Journal of Microscopy 265, 185–195. doi:10.1111/jmi.12474.

Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.W., Heng, P.A., 2018. H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Transactions on Medical Imaging 37, 2663–2674. doi:10.1109/tmi.2018.2845918.

Lin, G., Milan, A., Shen, C., Reid, I., 2017. RefineNet: Multi-path refinement networks for high-resolution semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Linder, N., Konsti, J., Turkki, R., Rahtu, E., Lundin, M., Nordling, S., Haglund, C., Ahonen, T., Pietikäinen, M., Lundin, J., 2012. Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. Diagnostic Pathology 7, 22. doi:10.1186/1746-1596-7-22.

Madabhushi, A., Lee, G., 2016. Image analysis and machine learning in digital pathology: Challenges and opportunities. Medical Image Analysis 33, 170–175. doi:10.1016/j.media.2016.06.037.

Mahbod, A., Schaefer, G., Bancher, B., Löw, C., Dorffner, G., Ecker, R., Ellinger, I., 2021. CryoNuSeg: A dataset for nuclei instance segmentation of cryosectioned H&E-stained histological images. arXiv preprint arXiv:2101.00442 doi:10.1016/j.compbiomed.2021.104349.

Naylor, P., Lae, M., Reyal, F., Walter, T., 2019. Segmentation of nuclei in histopathology images by deep regression of the distance map. IEEE Transactions on Medical Imaging 38, 448–459. doi:10.1109/tmi.2018.2865709.

Rezaei, S., Emami, A., Zarrabi, H., Rafiei, S., Najarian, K., Karimi, N., Samavi, S., Soroushmehr, S.M.R., 2019a. Gland segmentation in histopathology images using deep networks and handcrafted features, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. doi:10.1109/embc.2019.8856776.

Rezaei, S., Emami, A., Zarrabi, H., Rafiei, S., Najarian, K., Karimi, N., Samavi, S., Soroushmehr, S.M.R., 2019b. Micro-Net: A unified model for segmentation of various objects in microscopy images. Medical image analysis 52, 160–173. doi:10.1016/j.media.2018.12.003.

Rogojanu, R., Thalhammer, T., Thiem, U., Heindl, A., Mesteri, I., Seewald, A., Jäger, W., Smochina, C., Ellinger, I., Bises, G., 2015. Quantitative image analysis of epithelial and stromal area in histological sections of colorectal cancer: An emerging diagnostic tool. BioMed Research International 2015, 1–9. doi:10.1155/2015/569071.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Lecture Notes in Computer Science. Springer International Publishing, pp. 234–241. doi:10.1007/978-3-319-24574-4_28.

Schlemper, J., Oktay, O., Schaap, M., Heinrich, M., Kainz, B., Glocker, B., Rueckert, D., 2019. Attention gated networks: Learning to leverage salient regions in medical images. Medical Image Analysis 53, 197–207. doi:10.1016/j.media.2019.01.012.

Shan, T., Yan, J., 2021. SCA-Net: A spatial and channel attention network for medical image segmentation. IEEE Access 9, 160926–160937. doi:10.1109/access.2021.3132293.

Shervin, M., Yuri, Y, B., Fatih, P., Antonio, J, P., Nasser, K., Demetri, T., 2021. Image segmentation using deep learning: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence , 1–1doi:10.1109/tpami.2021.3059968.

Simon, G., Quoc Dang, V., Shan E Ahmed, R., Ayesha, A., Yee Wah, T., Jin Tae, K., Nasir, R., 2019. Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Medical Image Analysis 58, 101563. doi:10.1016/j.media.2019.101563.

Sultana, F., Sufian, A., Dutta, P., 2020. Evolution of image seg-

mentation using deep convolutional neural network: A survey. Knowledge-Based Systems 201-202, 106062. doi:10.1016/j.knosys.2020.106062.

Tosta, T.A.A., de Faria, P.R., Neves, L.A., do Nascimento, M.Z., 2019. Computational normalization of H&E-stained histological images: Progress, challenges and future potential. Artificial Intelligence in Medicine 95, 118–132. doi:10.1016/j.artmed.2018.10.004.

Vu, Q.D., Graham, S., Kurc, T., To, M.N.N., Shaban, M., Qaiser, T., Koohbanani, N.A., Khurram, S.A., Kalpathy-Cramer, J., Zhao, T., et al., 2019. Methods for segmentation and classification of digital microscopy tissue images. Frontiers in bioengineering and biotechnology , 53.doi:10.3389/fbioe.2019.00053.

Wan, T., Zhao, L., Feng, H., Li, D., Tong, C., Qin, Z., 2020. Robust nuclei segmentation in histopathology using ASPPU-Net and boundary refinement. Neurocomputing 408, 144–156. doi:10.1016/j.neucom.2019.08.103.

Weng, Y., Zhou, T., Li, Y., Qiu, X., 2019. NAS-UNet: Neural architecture search for medical image segmentation. IEEE Access 7, 44247–44257. doi:10.1109/access.2019.2908991.

Xing, F., Yang, L., 2016. Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: A comprehensive review. IEEE Reviews in Biomedical Engineering 9, 234–263. doi:10.1109/rbme.2016.2515127.

Yuan, Y., Chao, M., Lo, Y.C., 2017. Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. IEEE Transactions on Medical Imaging 36, 1876–1886. doi:10.1109/tmi.2017.2695227.

Yurttakal, A.H., Erbay, H., 2020. Segmentation of Larynx Histopathology Images via Convolutional Neural Networks, in: Advances in Intelligent Systems and Computing. Springer International Publishing, pp. 949–954. doi:10.1007/978-3-030-51156-2_110.

Zeng, Z., Xie, W., Zhang, Y., Lu, Y., 2019. RIC-UNet: An improved neural network based on unet for nuclei segmentation in histology images. IEEE Access 7, 21420–21428. doi:10.1109/access.2019.2896920.

Zhang, L., Sonka, M., Lu, L., Summers, R.M., Yao, J., 2017. Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei, in: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), IEEE. doi:10.1109/isbi.2017.7950548.

Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2017. Pyramid scene parsing network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.48550/arXiv.1612.01105.

Zhou, Y., Dou, Q., Chen, H., Qin, J., Heng, P.A., 2018a. Sfcn-opi: Detection and fine-grained classification of nuclei using sibling fcn with objectness prior interaction, in: Proceedings of the AAAI Conference on Artificial Intelligence. doi:10.1609/aaai.v32i1.11900.

Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J., 2018b. 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 International Publishing, pp. 3–11. doi:10.1007/978-3-030-00889-5_1.