



Triple U-net: Hematoxylin-aware nuclei segmentation with progressive dense feature aggregation

Bingchao Zhao^{a,b,1}, Xin Chen^{b,d,1}, Zhi Li^{c,1}, Zhiwen Yu^{a,1}, Su Yao^c, Lixu Yan^c, Yuqian Wang^b, Zaiyi Liu^{b,*}, Changhong Liang^{b,*}, Chu Han^{b,*}

^a The School of Computer Science and Engineering, South China University of Technology, Guangzhou, Guangdong, 510006, China

^b Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong, 510080, China

^c Department of Pathology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong, 510080, China

^d Department of Radiology, Guangzhou First People's Hospital, the Second Affiliated Hospital of South China University of Technology, Guangzhou, Guangdong, 510180, China



ARTICLE INFO

Article history:

Received 9 February 2020

Revised 29 June 2020

Accepted 16 July 2020

Available online 18 July 2020

Keywords:

Digital pathology

Nuclei segmentation

Convolutional neural networks

ABSTRACT

Nuclei segmentation is a vital step for pathological cancer research. It is still an open problem due to some difficulties, such as color inconsistency introduced by non-uniform manual operations, blurry tumor nucleus boundaries and overlapping tumor cells. In this paper, we aim to leverage the unique optical characteristic of H&E staining images that hematoxylin always stains cell nuclei blue, and eosin always stains the extracellular matrix and cytoplasm pink. Therefore, we extract the Hematoxylin component from RGB images by Beer-Lambert's Law. According to the optical attribute, the extracted Hematoxylin component is robust to color inconsistency. With the Hematoxylin component, we propose a Hematoxylin-aware CNN model for nuclei segmentation without the necessity of color normalization. Our proposed network is formulated as a Triple U-net structure which includes an RGB branch, a Hematoxylin branch and a Segmentation branch. Then we propose a novel feature aggregation strategy to allow the network to fuse features progressively and to learn better feature representations from different branches. Extensive experiments are performed to qualitatively and quantitatively evaluate the effectiveness of our proposed method. In the meanwhile, it outperforms state-of-the-art methods on three different nuclei segmentation datasets.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Pathological sections are regarded as the gold standard for cancer diagnosis since they deliver massive tumor information. Digital pathology images have now been widely used for clinical predictions, like diagnosis (Zhou et al., 2019a), prognosis (Yu et al., 2016), metastasis (Liu et al., 2017) and etc. Nuclei segmentation plays a vital role among these clinical applications by analyzing the morphological signatures of nuclei patterns in the Whole Slide Image (WSI). With the segmented cells, we can further objectively and quantitatively evaluate the malignant tissue by measuring the distribution of tumor cells, epithelial cells, immune cells, etc. A successful segmentation result, especially with precise cellular boundaries, will greatly benefit the following prediction outcomes. How-

ever, nuclei segmentation is still a crucial task. First of all, the non-uniform operations during the collection of whole slide images may lead to image inconsistency. For example, manual hematoxylin and eosin staining or different slide scanners may cause color inconsistency and unclear nucleus boundaries. Second, since the growth rate of malignant cells is extremely high, the density of the malignant cells in the neoplastic tissue is usually much higher than that in the normal tissue. Squeezing among tumor and normal cells will bring a large amount of occluded and overlapping nuclei.

Traditional image segmentation techniques (Otsu, 1979; Vincent and Soille, 1991; Cheng et al., 2008) leverage low-level image features or a simple thresholding strategy to segment the nuclei. However, it is hard to select a suitable threshold for all pathological images. And these methods may also fail to predict a satisfactory segmentation result when the aforementioned difficulties appear. In recent years, convolutional neural network (CNN) approaches (He et al., 2017; Long et al., 2015) have achieved promising results in natural image segmentation tasks. However, the dif-

* Corresponding authors.

E-mail addresses: Zyliu@163.com (Z. Liu), cjr.lchh@vip.163.com (C. Liang), zq1992@gmail.com (C. Han).

¹ Contributed equally.

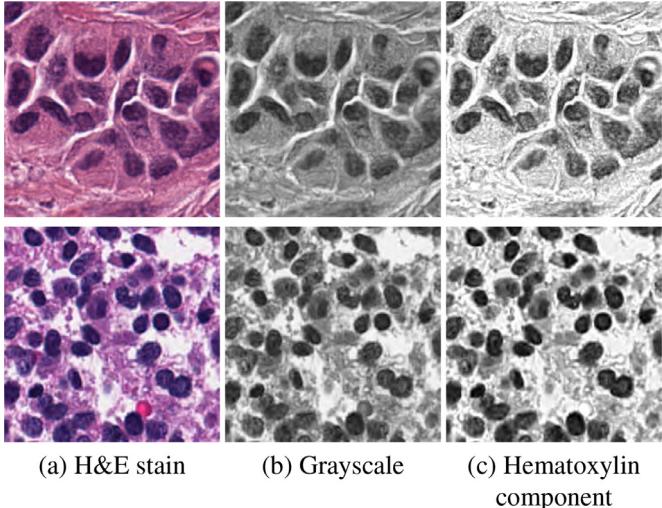


Fig. 1. Comparison of the Grayscale image (b) and the Hematoxylin component (c) extracted from H&E staining WSI patch (a). The contrast in (c) between nuclei and cytoplasm is much higher than (b), and is more robust even the color is inconsistent.

ference of the characteristics between natural images and medical images makes it hard to directly apply these methods to nuclei segmentation. U-net and its variants (Ronneberger et al., 2015; Zhou et al., 2018) have been proven their effectiveness in medical image segmentation. But these methods are still facing some challenges. The major problem is the lack of dataset with pixel-level labels. Because manually constructing pathological datasets requires high expertise and is labor-intensive. Moreover, a simple U-net architecture is not good enough to predict precise nucleus boundaries. With the recently proposed nuclei segmentation dataset (Kumar et al., 2017), more CNN models have been proposed and achieved outstanding performance. Some methods (Graham et al., 2019; Naylor et al., 2018) rely on distance maps to separate the touching or overlapping nuclei. Koohbanani et al. (2019) take spatial information into account and propose a proposal-free model. Zhou et al. (2019b) proposed a contour-aware CNN model, called CIA-Net, for predicting more precise nucleus boundaries. However, the segmentation performance still has room for improvement.

In this paper, we perform nuclei segmentation in a new perspective and propose a Hematoxylin-aware network. Our proposed method leverages the optical characteristics of H&E staining that hematoxylin always stains cell nuclei blue, and eosin always stains the extracellular matrix and cytoplasm pink, as shown in Fig. 1(a). According to Beer-Lambert's Law (Parson, 2007), we apply a color decomposition method (Ruifrok et al., 2001) to extract Hematoxylin (H) component from RGB color space, to obtain a much clearer nucleus view. It can be observed that the contrast between the nuclei and the cytoplasm & stroma in Fig. 1(c) is much higher than Fig. 1(b), even the color of two WSI patches are inconsistent. To make full use of the H component, we propose a Triple U-net architecture for nuclei segmentation. Our network consists of three U-net branches, an RGB branch, a hematoxylin (H) branch and a segmentation branch. The RGB branch serves for providing the raw feature representation of the original image. The H branch is introduced for more precise contour-aware supervision. The segmentation branch aggregates the feature from two other branches at every single convolution block and finally prediction segmentation results. Instead of directly concatenating all the features from different branches with different tasks, we propose a progressive dense feature aggregation module (PDFA). It allows the network to progressively merge and learn the features from different image domains. Extensive experiments are performed to qualitatively and

quantitatively evaluate the effectiveness of our proposed method, it outperforms state-of-the-art methods on three different nuclei segmentation datasets.

The main contributions of this paper are three-fold:

- We tailor a Hematoxylin-aware network with Triple U-net architecture for nuclei segmentation without the necessity of color normalization and achieve state-of-the-art nuclei segmentation performance.
- Hematoxylin component is extracted to provide a more precise contour awareness for the network supervision in order to predict segmentation results with more precise nucleus boundaries.
- The PDFA module is proposed to allow the network to aggregate features from different domains progressively.

2. Related works

2.1. CNN-based natural image segmentation

With the increasing computational power and the data explosion, convolutional neural networks have already dominated the field of natural image segmentation.

Fully convolutional network (Long et al., 2015) was proposed for semantic segmentation which allows the network to take input of arbitrary size images. Zhao et al. (2017) proposed a PSP-Net which exploits the global context information by the proposed pyramid pooling module. DeepLab series (Chen et al., 2017a; Chen et al., 2017b; Chen et al., 2018) were proposed and achieved promising semantic segmentation results. Different from semantic segmentation, instance segmentation performs segmentation tasks at the instance level. He et al. (2017) proposed a well-known model Mask-RCNN which extends the region proposal based object detection task to instance segmentation. Liu et al. (2018) proposed an adaptive feature pooling method to propagate the features from the encoding phase to the proposal subnetworks. de Geus et al. (2018) merged the predictions of semantic and instance segmentation tasks using heuristics to achieve more effective panoptic segmentation. Yang et al. (2019) proposed a single-shot, bottom-up image parser to process the semantic and instance segmentation tasks jointly.

However, we cannot directly apply these approaches to medical image segmentation due to the domain discrepancy. Nevertheless, it is still worth to know these outstanding segmentation techniques.

2.2. CNN-based medical image segmentation

Medical image segmentation can be functionally separated according to the way that images are generated from, like CT (Christ et al., 2016; Roth et al., 2015), MRI (Havaei et al., 2017; Pereira et al., 2016), ultrasound (Milletari et al., 2017), whole slide image (Bejnordi et al., 2017) and etc. For section images like WSI, it can be further categorized into cell-level (Naik et al., 2008; Kumar et al., 2017) and tissue-level (Chen et al., 2016; Xu et al., 2016). The bottleneck of the learning-based medical image segmentation is the lack of training data. U-net (Ronneberger et al., 2015) has already shown great effectiveness on medical image segmentation with very few training images and thus many U-net hybrids were proposed upon it. Zhou et al. (2018) replaced the simple skip connections with dense skip connections to obtain more complete multi-level information. Li et al. (2018) proposed a hybrid densely connected U-net for CT volume segmentation. Man et al., 2019 associated reinforcement learning with geometry-aware U-net to activate the searching strategy for CT pancreas segmentation.

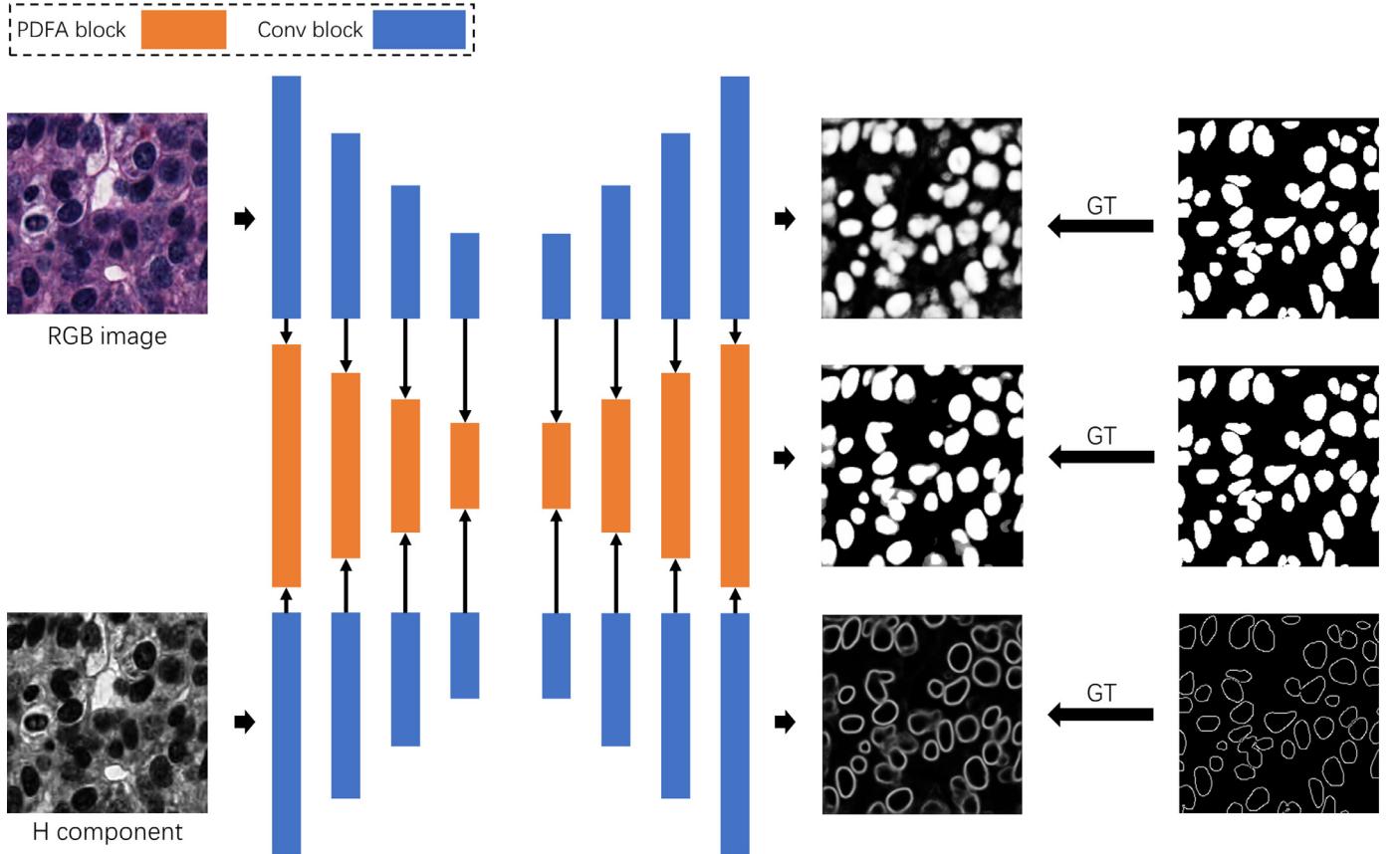


Fig. 2. Network Architecture. Our proposed network contains three branches. The RGB branch (upper) and the H branch (lower) are with the same structure with three downsamplings and aim to extract raw features and contour features respectively. The segmentation branch (middle) merges the features from two other branches and predicts the segmentation results. The block in orange is our proposed progressive dense feature aggregation module. Note that, all the skip connections in three branches are omitted for simplification.

Besides U-net architecture, many superior segmentation methods have also been proposed. Havaei et al. (2017) proposed multiple network architectures to fuse image features under different receptive fields to achieve high-quality segmentation of tumors in MR images. Chen et al. (2016) proposed an FCN-based model for gland segmentation by simultaneously predicting contours and instance probability maps to separate the touching glands. Zhang et al. (2017) published a dataset for gland instance segmentation and proposed a GAN-based segmentation network. Chan et al. (2019) proposed to use a class activation map for weakly supervised tissue segmentation.

2.3. Nuclei segmentation

Nuclei segmentation is a fundamental step for computer assistance diagnosis and tumor micro-environment analysis. Early approaches (Otsu, 1979) segment nuclei from histopathological images base on color thresholding and background subtraction. But it is hard to select a suitable threshold for all the scenarios. Marker controlled watershed segmentation (Vincent and Soille, 1991) based on initial seed points to growing region directly which often leads to the false positive segmentation. These traditional techniques usually have complex post-processing and are hard to process images where the challenging scenarios happen, such as overlapping and occlusion.

Recently, deep learning models have been widely used in nuclei segmentation. Kumar et al. (2017) introduced a challenging dataset for nuclei segmentation and enrich the data for the community. Naylor et al. (2018) formulated the segmentation prob-

lem as a regression task of distance maps to overcome the difficulty of touching nuclei. HoVer-Net (Graham et al., 2019) was proposed to perform nuclei segmentation and classification simultaneously. Chidester et al. (2019) proposed some enhancement strategies base on U-net structure including residual blocks and data augmentation. A novel stain-aware loss function (Graham and Raajpoot, 2018) is tailor-made for the stain variations in H&E images. Liu et al. (2019) designed a dual-branch panoptic model that incorporates instance segmentation with semantic segmentation. Koohbanani et al. (2019) leveraged spatial information and propose a proposal-free model. CIA-Net (Zhou et al., 2019b) utilized additional contour supervision to obtain segmentation results with more precise nucleus boundaries. However, the existing nuclei segmentation approaches still have room for improvement. In this paper, we take the representative color characteristic of an H&E stain image into consideration and propose a Hematoxylin-aware nuclei segmentation.

3. Proposed method

The network architecture of our proposed Triple U-net is illustrated in Fig. 2. Our network contains three U-net branches that serve different purposes. The *RGB branch* serves for extracting the raw features of the segmentation task. The *H branch* focuses on the Hematoxylin-aware feature extraction of the nucleus edge detection task. Since the hematoxylin component has strong robustness to the color inconsistency, we do not apply any color normalization to the original image to avoid any information loss during the normalization process. The *segmentation branch* fuses the RGB

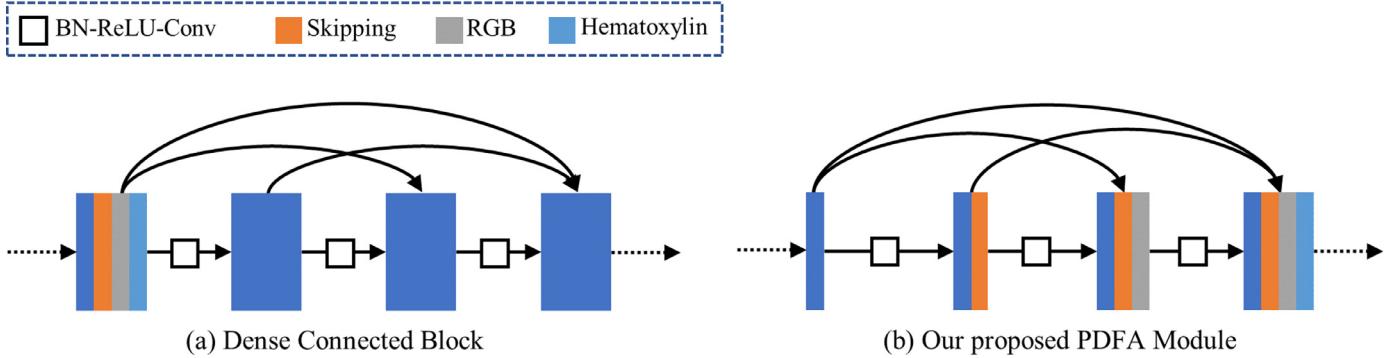


Fig. 3. Densely connected block (Huang et al., 2017) (a) and our proposed progressive dense feature aggregation module (b). PDFA not only inherits the benefits from densely connected block, but also provides a softer feature aggregation strategy when new features from different domains come.

raw features and the Hematoxylin-aware contour features and then predict the final segmentation results. Note that, all the skip connection of three U-net branches in Fig. 2 are omitted for the simplification of the figure.

Next, we will introduce the Hematoxylin component extraction in Section 3.1. A novel Progressive Dense Feature Aggregation Module (PDFA) is proposed in Section 3.2. The loss functions of our network are demonstrated in Section 3.3. Section 3.4 shows the implementation and training details.

3.1. Hematoxylin component extraction

Hematoxylin and eosin (H&E) stains have been used for at least a century for cancer diagnosis because it can essentially differentiate cytoplasmic, nuclei, and extracellular matrix features. According to the principle of H&E stain, hematoxylin stains cell nuclei blue, and eosin stains the extracellular matrix and cytoplasm pink. In this paper, we aim to use the above unique characteristic of H&E stain as Hematoxylin-aware guidance for the segmentation network. To achieve this, we apply a color decomposition technique (Ruirok et al., 2001) to decompose the Hematoxylin Component from the original RGB image. This approach is commonly utilized as a color normalization preprocessing in traditional methods due to its robustness of color inconsistency in the H&E stained WSI.

We assume that the gray level in each RGB channel is linear with light transmission rate $T = I/I_0$, where I_0 is the incident light and I is the transmitted light. So each specific stain will be characterized by a specific absorption factor c for the light in each of the three RGB channels. Then we can model the relationship between the amount of stain and its absorption using Beer-Lambert's Law (Parson, 2007). The energy I of each RGB channel can be written as follows:

$$I = I_0 \exp(-Ac) \quad (1)$$

where I_0 is the intensity of the light entering the specimen, I is the intensity of the light passed through the specimen. A is the amount of stain and c is the absorption factor.

Due to the non-linearity of the relative intensity in each of the channels, the intensity values I of the image cannot be directly used for the stain color decomposition. Thus, we calculate the specimen optical density o , which is linear with the concentration of absorbing stain, as follows:

$$o = -\log \frac{I}{I_0} \quad (2)$$

Hence Eq. (1), we have got:

$$o = Ac \quad (3)$$

So each specific stain can be characterized by a specific optical density vector. Here, the hematoxylin resulted in o values [0.18,

0.20, 0.08] for R, G and B channels respectively. With this color representation model, we are able to map RGB color space to any stain specific color space. Based on this model, we extracted the Hematoxylin component by applying the color deconvolution method proposed by Ruirok et al. (2001).

3.2. Progressive dense feature aggregation module

In our proposed Triple U-net, the RGB branch and the H branch serve for feature extraction in the segmentation task and nuclei contour detection task respectively. The RGB branch takes the original image as input and extracts raw image features. The H branch is designed as a Hematoxylin-aware supervision branch. And additional segmentation branch is introduced to fuse the raw features and the Hematoxylin-aware contour features together, from the encoding phase to the decoding phase, to predict the final segmentation results.

To fuse and learn a better feature representation from different branches with different goals, we propose a softer feature fusion strategy with densely connections, called Progressive Dense Feature Aggregation (PDFA). The intention of PDFA is simple and intuitive. When the new features from different image domains or tasks come, we want the network to fuse the previous features with the latter ones progressively, instead of directly concatenating all the features together.

Fig. 3(b) demonstrates our proposed PDFA module. Compared with the normal densely connected block (Huang et al., 2017) in Fig. 3(a), our proposed PDFA module not only inherits the advantages of densely connected block, e.g., strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters, but also provides a more reasonable feature fusion strategy. Note that, the number of layers in one PDFA module is determined by the number of new features to be fused. During the decoding phase in our network, the PDFA modules contain four layers as shown in Fig. 3(b). But in the encoding phase, since there are no skipping features at the beginning of the network, the PDFA module only contains three layers.

Ablation studies in Section 4.4 have been conducted to evaluate the effectiveness of our proposed PDFA compared with normal densely connected block. In addition, we also discuss how the order of feature aggregation affects the network in ablation studies.

3.3. Loss function

The total objective function is the weighted summation of the losses from three branches:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{RGB} + \lambda_2 \mathcal{L}_H + \lambda_3 \mathcal{L}_{seg_ST} + \lambda_4 \mathcal{L}_{seg_SD} \quad (4)$$

where λ_i is the hyper-parameter of the corresponding loss. The details of each loss is described as follows.

3.3.1. RGB branch

Since the original image in RGB color space contains the raw information, which can provide the richest semantic information for the segmentation task. Therefore, the direct output of this branch is supervised by the segmentation ground truth using binary cross entropy loss, to obtain the most precise pixel-level features when performing segmentation. Hence, the loss \mathcal{L}_{RGB} is defined as follows:

$$\mathcal{L}_{RGB} = -\frac{1}{N} \sum_{i=1}^N (y_i \cdot \log x_i + (1 - y_i) \cdot \log(1 - x_i)) \quad (5)$$

where x_i denotes the predicted probability of the i th pixel, y_i is the corresponding ground truth. N is the number of pixels.

3.3.2. H branch

The input Hematoxylin component in the H branch directly reflects the concentration of hematoxylin in the H&E stained whole slide image, highlights the contour of the nucleus, and is robust to color inconsistency. Therefore, we leverage the contour information from Hematoxylin component and assign the nuclei contour detection task in the H branch to guide the model to better learn the contour information of the nucleus. Hence the network can predict segmentation results with more precise nucleus boundaries. Here, we utilize Soft Dice Loss for this branch so that the model can focus on learning the information of the nuclear contour. The loss function is as follows:

$$\mathcal{L}_H = 1 - \frac{2 \sum_{i=1}^N x_i y_i}{\sum_{i=1}^N x_i^2 + \sum_{i=1}^N y_i^2} \quad (6)$$

where x_i is the predicted probability of the i th pixel and y_i is the ground truth.

3.3.3. Segmentation branch

Due to the stain color inconsistency and the blurred nucleus boundaries, mislabelled nuclei and inaccurate boundaries are inevitable during manual annotations for pixel-level nuclei segmentation. As a result, these inaccurate labels may dominate the gradients and hence prevent the network from learning informative samples. Therefore, in the segmentation branch, we apply a Smooth Truncated Loss proposed by Zhou et al. (2019b) to suppress the interference from the outliers. The loss is defined as follows:

$$\mathcal{L}_{seg_ST} = \begin{cases} -\log(\gamma) + \frac{1}{2}(1 - \frac{x_i^2}{\gamma^2}) & , x_i < \gamma \\ -\log(x_i) & , x_i \geq \gamma \end{cases} \quad (7)$$

where x_i denotes the predicted probability of the i th pixel, $x_i = x$ if $y_i = 1$, and $x_i = 1 - x$ otherwise. $\gamma = 0.3$ is the truncated point.

Besides, we also introduce a Soft Dice Loss for the segmentation branch as follows:

$$\mathcal{L}_{seg_SD} = 1 - \frac{2 \sum_{i=1}^N x_i y_i}{\sum_{i=1}^N x_i^2 + \sum_{i=1}^N y_i^2} \quad (8)$$

3.4. Implementation and training details

We implement our network using Pytorch 1.2.0 on a workstation equipped with an NVIDIA GeForce RTX 2080 Ti. During training, we use several data augmentations including, elastic transformation, random cropping, mirror, rotation and flipping. We crop each original image (1000×1000 pixels) into 16 patches (256×256 pixels) and then feed them into the network. We use AdamW optimizer with the batch size of 2. The learning rate is

initialized as 10^{-4} with a learning rate decay of 0.96. The weight decay is set as 10^{-4} . The hyper-parameters of losses λ_i were set as $\{0.3, 0.3, 1, 0.3\}$. We did not use transfer learning and independent branch training during the training process.

4. Experiments

We have conducted extensive experiments to quantitatively and qualitatively evaluate the effectiveness of our proposed Triple U-net on nuclei segmentation. Section 4.1 introduces the details of the dataset and evaluation metrics. Section 4.2 and Section 4.3 show the quantitative comparisons with the existing methods. We have also conducted ablation studies in Section 4.4 to evaluate the effectiveness of the Triple U-net structure, the superiority of the proposed Progressive Dense Feature Aggregation module and the impact of different losses. Section 4.5 demonstrates the qualitative results. We show the limitation of our model and have some discussions in Section 4.6.

4.1. Dataset and evaluation metrics

We evaluate our model on three nuclei segmentation dataset, (1) Multi-Organ Nuclei Segmentation (MoNuSeg) dataset (Kumar et al., 2017), (2) Colorectal Nuclear Segmentation and Phenotypes (CoNSeP) dataset (Graham et al., 2019) and (3) CPM-17 (Vu et al., 2019).

MoNuSeg contains 7 organs, 30 images under the resolution of 1000×1000 with the annotations of more than 21,000 nuclear boundaries. We divide all images into a training set and a test set. The training set contains 16 image pieces from 4 organs, and the test set contains 14 image pieces from 7 organs. We use the same image split with the existing methods (Zhou et al., 2019b; Graham et al., 2019; Kumar et al., 2017).

CoNSeP involves 41 H&E stained images, each of which is 1000×1000 pixels, scanned at 40X magnification. The images were extracted from 16 WSIs from different colorectal adenocarcinoma (CRA) patients but involved different tissues and cell types. We conduct the training and the test set using the same image split with (Graham et al., 2019), where the training set contains 27 images, the test set contains 14 images.

CPM-17 contains 40 pathological images with pixel-level annotations, of which 32 were in the training set and 8 were in the test set. Each image is 500×500 pixels, scanned at 40X magnification. This dataset contains more than one tumor type, including non-small cell lung cancer (NSCLC), head and neck squamous cell carcinoma (HNSCC), glioblastoma multiforme (GBM), and lower grade glioma (LGG) tumors.

In the experiments, we use three common metrics to evaluate the nuclei segmentation results of each model, including Aggregated Jaccard Index (AJI) (Kumar et al., 2017), Dice score and Panoptic Quality (PQ) (Kirillov et al., 2020).

AJI is defined as follows:

$$AJI = \frac{\sum_{i=1}^n G_i \cap P_j}{\sum_{i=1}^n G_i \cup P_j + \sum_{k \in N} P_k} \quad (9)$$

where $j = \text{argmax}_k \frac{G_i \cap P_k}{G_i \cup P_k}$, $G = \{G_1, G_2, \dots, G_n\}$ and $P = \{P_1, P_2, \dots, P_m\}$ denote the ground truth and the prediction results respectively. N is the set of indices of prediction results without any corresponding ground truth.

DICE score is defined as follows:

$$DICE = \frac{2 \times |G \cap P|}{|G| + |P|} \quad (10)$$

P and G denote the prediction results and the ground truth respectively.

Table 1
Quantitative comparison with existing methods.

Method	MoNuSeg			CoNSeP			CPM-17		
	AJI	DICE	PQ	AJI	DICE	PQ	AJI	DICE	PQ
(1) CNN3 (Kumar et al., 2017)	0.508	0.762	-	-	-	-	-	-	-
(2) Mask R-CNN (He et al., 2017)	0.546	0.760	0.509	0.474	0.740	0.460	0.684	0.850	0.674
(3) DIST (Naylor et al., 2018)	0.560	0.786	0.443	0.502	0.804	0.398	0.616	0.826	0.504
(4) DCAN (Chen et al., 2016)	0.525	0.793	0.492	0.289	0.733	0.256	0.561	0.828	0.545
(5) Deep Panoptic (Liu et al., 2019)	0.585	0.794	-	-	-	-	-	-	-
(6) HoVer-Net (Graham et al., 2019)	0.618	0.826	0.597	0.571	0.853	0.547	0.705	0.869	0.697
(7) CIA-Net (Zhou et al., 2019b)	0.620	0.818	0.577	-	-	-	-	-	-
Ours	0.621	0.837	0.601	0.579	0.843	0.562	0.711	0.888	0.685

Table 2
Cross validation.

Experiment	MoNuSeg			CoNSeP			CPM-17		
	AJI	DICE	PQ	AJI	DICE	PQ	AJI	DICE	PQ
HoVer-Net	0.595	0.822	0.541	0.578	0.815	0.518	0.610	0.825	0.589
	0.591	0.803	0.525	0.569	0.850	0.545	0.642	0.862	0.641
	0.583	0.822	0.524	0.535	0.824	0.536	0.623	0.866	0.621
	0.544	0.786	0.509	0.583	0.839	0.570	0.585	0.841	0.579
	0.539	0.810	0.520	0.566	0.830	0.537	0.599	0.868	0.611
Ours	Mean \pm SD 0.570 \pm 0.024	0.809 \pm 0.013	0.524 \pm 0.010	0.566 \pm 0.017	0.832 \pm 0.012	0.541 \pm 0.017	0.612 \pm 0.020	0.852 \pm 0.017	0.608 \pm 0.022
	(1) fold1	0.615	0.829	0.577	0.593	0.813	0.524	0.661	0.851
	(2) fold2	0.617	0.815	0.556	0.563	0.848	0.550	0.688	0.872
	(3) fold3	0.587	0.817	0.549	0.547	0.829	0.530	0.647	0.849
	(4) fold4	0.547	0.799	0.512	0.581	0.841	0.573	0.652	0.852
Ours	(5) fold5	0.553	0.825	0.525	0.569	0.828	0.547	0.641	0.875
	Mean \pm SD 0.584 \pm 0.030	0.817 \pm 0.010	0.544 \pm 0.023	0.571 \pm 0.016	0.832 \pm 0.012	0.545 \pm 0.017	0.658 \pm 0.017	0.860 \pm 0.011	0.631 \pm 0.013

Since the AJI score may over-penalize the overlapping region. To avoid this problem, Panoptic Quality (PQ) (Graham et al., 2019) is introduced to evaluate the performance of nuclei segmentation in instance-level, which is defined as follows:

$$PQ = \frac{|TP|}{|TP| + \frac{1}{2}|FP| + \frac{1}{2}|FN|} \times \frac{\sum_{(x,y) \in TP} IoU(x, y)}{|TP|} \quad (11)$$

x and y denote the ground truth and the prediction segment, respectively, in instance-level. IoU denotes the intersection over the union. When the $IoU > 0.5$ of each (x, y) pair, the result can be regarded to be unique (Kirillov et al., 2020). TP , FP , and FN denote matched pairs of prediction, unmatched prediction, and unmatched ground truth prediction respectively.

4.2. Quantitative comparisons with existing methods

We compare our model with several semantic segmentation baseline models and existing nuclei segmentation models, including: (1) CNN3 (Kumar et al., 2017), (2) Mask R-CNN (He et al., 2017), (3) DIST (Naylor et al., 2018), (4) DCAN (Chen et al., 2016), (5) Deep Panoptic (Liu et al., 2019), (6) HoVer-Net (Graham et al., 2019) and (7) CIA-Net (Zhou et al., 2019b). Since CNN has demonstrated its superiority on segmentation task, we only compare our proposed network with existing CNN-based models. Note that, the results of Mask R-CNN is implemented by the author of HoVer-Net (Graham et al., 2019). And the DICE score of (7) is calculated by the released model from the author of CIA-Net (Zhou et al., 2019b). The quantitative results of other models are directly from their respective papers.

Table 1 demonstrates the quantitative results. CNN3 (1) is the baseline model of nuclei segmentation. Mask R-CNN (2) performs better than (1) in the AJI score thanks to the region proposal strategy. DIST model (3) leverages distance maps for nuclei segmentation and achieves promising results. However, their results are still much lower than other methods because neither distance maps nor region proposal can guarantee precise nuclei boundaries. The

contour-aware model in DCAN (4) allows it to predict more accurate boundaries. However, the low network capacity of simple FCN backbone leads to non-satisfactory AJI scores. Deep Panoptic (5) introduces an additional semantic branch with an instance branch for both global and local feature consideration. It can predict more precise nucleus locations to avoid the penalty of false positive in AJI measurement and thus achieves a higher AJI score than (1) to (3). Since the morphology of the nucleus is an important feature for clinical researches. HoVer-Net (6) proposed to use horizontal and vertical distance maps simultaneously and formulate the segmentation task and the classification task in the same network. CIA-Net (7) introduces contour supervision to obtain the segmentation results with more precise nucleus boundaries and proposes a smooth truncated loss to suppress non-overlapping regions with ground truth. The overall segmentation performances of our approach in three different metrics and datasets are superior compared with the state-of-the-art models (6) and (7). Our proposed Hematoxylin-aware network not only achieves state-of-the-art segmentation performance but also avoids any color normalization pre-processing step.

4.3. Cross validation

We also conduct a 5-fold cross validation to statistically evaluate the stability of our proposed network compared with the state-of-the-art algorithm. Since our quantitative results shown in Table 1 are apparently higher than other methods. In this experiment, we only compare with HoVer-Net (Graham et al., 2019). Note that, we trained HoVer-Net using their released code and completely followed the training strategy shown in the paper. Statistical results are shown in Table 2. We compare our proposed model with HoVer-Net in three different datasets, MoNuSeg, CoNSeP and CPM-17. The upper part of the table demonstrates the results from HoVer-Net, and the lower part shows ours. Our model demonstrates its superiority of nuclei segmentation in all datasets under all metrics compared with the state-of-the-art model.

Table 3

Effectiveness of Hematoxylin-aware model. H, R, G and B indicate the Hematoxylin component, R channel, G channel and B channel respectively.

Method	AJI	DICE	PQ
(1) SingleU with RGB	0.552	0.805	0.481
(2) SingleU with Grayscale	0.538	0.791	0.463
(3) SingleU with H	0.547	0.798	0.477
(4) SingleU with RGBH	0.561	0.808	0.493
(5) Multi-task with RGBH	0.574	0.812	0.536
(6) TripleU with RGB & R	0.582	0.829	0.557
(7) TripleU with RGB & G	0.584	0.834	0.555
(8) TripleU with RGB & B	0.583	0.818	0.559
(9) TripleU with RGB & RGBH	0.595	0.815	0.563
(10†) TripleU with RGB & Grayscale	0.584	0.828	0.566
(10) TripleU with RGB & Grayscale	0.581	0.825	0.569
(11†) TripleU with RGB & RGB	0.614	0.836	0.586
(11) TripleU with RGB & RGB	0.612	0.836	0.587
(12†) TripleU with RGB & H	0.608	0.834	0.574
(12) TripleU with RGB & H (Ours)	0.621	0.837	0.601

4.4. Ablation studies

We conduct ablation studies to evaluate the effectiveness of the Hematoxylin-aware model, the superiority of the proposed progressive dense feature aggregation module and the impact of different loss configurations.

4.4.1. Effectiveness of hematoxylin-aware model

In this experiment, we aim to evaluate the effectiveness of our proposed network in three aspects, Triple U-net structure, Hematoxylin-aware model and color normalization. So we compare our complete model with different network structures with different input configurations. The network settings are shown in Table 3 are illustrated as follows: (1)–(4) A single U-net structure with the input RGB, Grayscale, Hematoxylin component, RGBH images, respectively. (5) A multi-task framework with two tasks, contour detection task and nuclei segmentation task. (6)–(11) Triple U-net structures and replacing the input image of H branch with R channel, G channel, B channel, RGBH, grayscale image and RGB image respectively. (12) Our complete model. (10†)–(12†) The models with (†) adopted a color normalization method (Vahadane et al., 2016) to the input RGB image in RGB branch. Note that, all the U-net structures in this experiment down- and up-samples only three times. The loss configurations in all Triple U-net variants are the same as our final model. And the loss configurations in the single U-nets are the same with the segmentation branch with smoothed truncated loss and soft dice loss.

Table 3 shows the statistical results. As can be seen, the models (1) to (4) with only one single U-net are overall poorer than the ones (6) to (12) with Triple U-net structure due to the limited network capability and the lack of contour supervision. Among models (1) to (4), (2) with the input grayscale image performs worst because of color information loss. It is interesting that even the Hematoxylin component also loses color information, (3) still shows comparable results with (1) since the light absorbing model has already decomposed the Hematoxylin stained colors. It also evidences that the extracted Hematoxylin component is robust for H&E stained nuclei segmentation task. Model (4) with RGBH input image demonstrates the best results among four single U-net structure baselines. It proves that the hematoxylin component can provide additional guidance for segmentation. We also compare our model with a multi-task learning baseline (5) with two tasks, a contour detection task and a segmentation task. Although it outperforms single U-net models (1) to (4) thanks to the multi-task training. It still lacks a feature aggregation strategy comparing with models (6) to (12). (6) to (12) in Table 3 shows the results with

Table 4

Effectiveness of PDFA.

Method	AJI	DICE	PQ
(1) Ours with conv block	0.581	0.781	0.536
(2) Ours with dense block	0.590	0.795	0.555
(3) Ours with PDFA (H-Skip-RGB)	0.625	0.836	0.584
(4) Ours with PDFA (H-RGB-Skip)	0.624	0.831	0.586
(5) Ours with PDFA (RGB-Skip-H)	0.625	0.834	0.587
(6) Ours with PDFA (RGB-H-Skip)	0.626	0.832	0.585
(7) Ours with PDFA (Skip-H-RGB)	0.627	0.833	0.585
(8) Ours with PDFA (Skip-RGB-H)	0.628	0.838	0.588

Triple U-net with the contour supervision using R channel, G channel, B channel, RGBH, RGB, grayscale and Hematoxylin component respectively. As can be seen, with an additional stand alone contour supervision branch, the performance boosted compared with (1) to (4) models with only one U-net. However, the segmentation results of these models are still worse than (12). Because the H branch is originally designed for detecting the contours of the nuclei according to the color specialty of H&E staining. When replacing the Hematoxylin component with other images without such color specialty, the performance of the H branch decreases.

Thanks to the Hematoxylin-aware model, color normalization is no longer necessary in our network. To prove such strength, we also compare the models with (†) and without color normalization. We can observe that the models with color normalization (10†) and (11†) show slightly better performance than the models (10) and (11) without it. It is because the input images for the H branch do not contain the staining color specialty in the Hematoxylin component. And that is the reason why additional color normalization may help. In contrast, when we feed the Hematoxylin component into the H branch in the model (12†), color normalization for the RGB image becomes a burden. Because Hematoxylin component itself has already been independent of the color changes. Further color normalization for the RGB image may lead to information losses which may unexpectedly harm the segmentation performance. Our final model without color normalization (12) demonstrates the best performance and proves the robustness of color changes.

4.4.2. Effectiveness of progressive dense feature aggregation

In the second ablation study, we evaluate the performance of our proposed PDFA module. We compare our complete model with the following two variants: (1) We replace the PDFA module with the normal convolutional block. (2) We replace the PDFA module with the Densely Connected block (Huang et al., 2017). In the meanwhile, it is also worth discussing how the feature coming and merging orders affect the performance of the network. So we also try all the combinations of the orders of three different features in the decoding phase, including the features from skip connections, the RGB branch and the H branch, indexing from (3) to (8) in Table 4.

To avoid the subjective selection of the final model, we conduct this experiment with a completely random image split for the training and the test sets, which are isolated from the datasets have been used in other experiments. It can be seen from Table 4, our model with the dense block (2) achieves better results than the convolutional block (1) indisputably. Our proposed PDFA module proves its superiority compared with the normal dense block (2), no matter in which feature aggregation orders (3) to (8). Results evidence that aggregating features, especially those features from different domains or branches, in a progressive manner allow the network to learn features better. Models from (3) to (8) are constructed to discuss how the feature aggregation orders affect the network performance. As can be seen, even three measurement scores vibrate in different orders, the statistical results are in

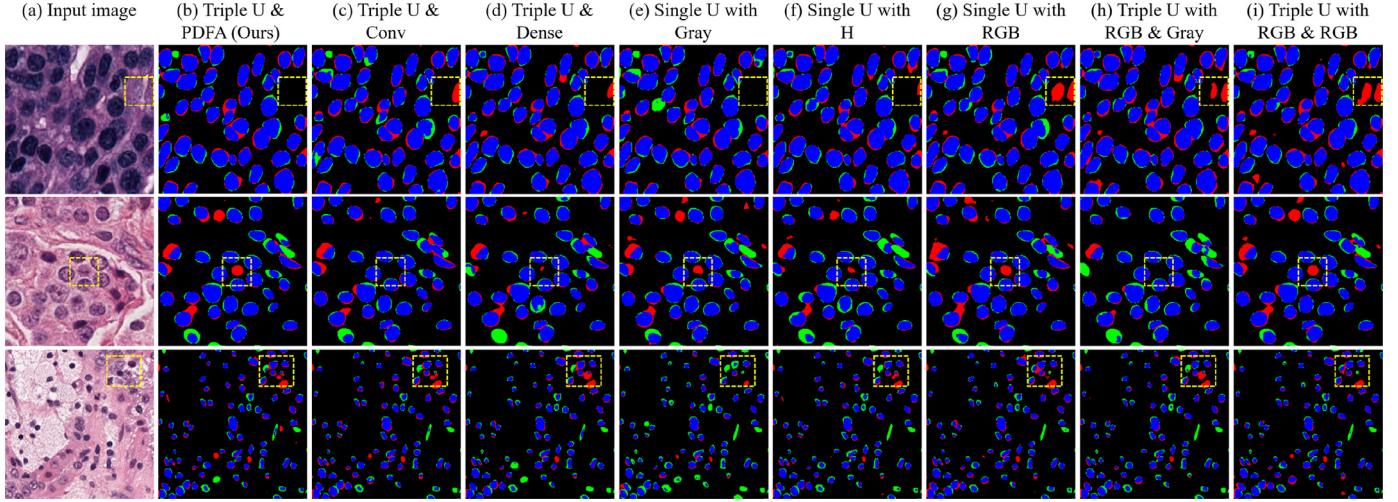


Fig. 4. Qualitative results of ablation studies. We show the difference maps between each result and the ground truth image for clearer visualization. Blue areas in the difference maps indicate the intersection of the result and the ground truth. Red and green areas indicate the false positive and the false negative segmentation respectively. Yellow boxes highlight the difference of all the results. H indicates the Hematoxylin component. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Different loss configurations.

Method	AJI	DICE	PQ
(1) Smooth Truncated (ST) Loss only	0.621	0.832	0.577
(2) Soft Dice Loss only	0.618	0.836	0.570
(3) ST & SD (Ours)	0.628	0.838	0.588
(4) HoVer-Net with ST Loss	0.592	0.815	0.547

a relatively high standard which proves that the feature aggregation order actually does not affect the network performance much. We choose the model with the best performance in the following order (Skip-RGB-H) and all our final results shown in this paper are generated using this feature aggregation order.

Fig. 4 also shows the qualitative comparison of two ablation studies by visualizing the difference maps between each result and the ground truth. Here we selectively demonstrate the results from the representative models for better layout and visualization. As we can see from the difference maps, our complete model gives the most accurate results than other variants compared with the ground truth. Besides, our results are with the most accurate nucleus boundaries and with less false positive and false negative segmentation.

4.4.3. Different loss configurations

In this experiment, we aim to discuss the impact of different losses. Since the goals of the RGB branch and the H branch are clear and the losses in these two branches are simple, we only discuss the loss configurations in the segmentation branch. (1) Smooth truncated loss only. (2) Soft dice loss only. (3) Smooth truncated loss and soft dice loss (ours). (4) HoVer-Net with smooth truncated loss.

The results shown in Table 5 are quite intuitive. Since the smooth truncated loss is originally designed to suppress outliers, which is the non-overlapping regions with ground truth. (1) and (3) with smooth truncated loss shows better AJI and PQ performance than (2). The soft dice loss is introduced for pixel-wise segmentation precision. (2) and (3) with soft dice loss shows a better DICE score than (1). Our final model (3) with both losses achieves the best performance among these three models. We have also applied smooth truncated loss in HoVer-Net (4). The performance of HoVer-Net with ST loss drastically decreased. It is because HoVer-

Net is designed as a multi-task network and all the tasks share the same encoder. Adding too more constraints to one specific task may cause imbalance of network training and make the network overemphasis this specific task. But this problem will not happen in our proposed network because ST loss was introduced in the main segmentation branch.

4.5. Instance segmentation results

Fig. 5 shows the instance segmentation results. We can observe that the nuclei are well segmented, and the segmentation results are similar to the ground truth in Fig. 5(b) according to the difference maps in Fig. 5(d). Although there are still some false positive and false negative segmentation on the nucleus boundaries, the overall shapes of our predicted results are still highly consistent with the ground truth and even more accurate than some of the manual labels.

4.6. Discussions

During the experiments, we found some imperfect ground truth labels. One is the inaccurate nucleus boundaries, which are highlighted in yellow boxes in Fig. 6(b). Another is some miss-labeled nuclei in blue boxes. However, these imperfections are inevitable because pathologists will face many difficulties when manual labeling. Huge resolution WSI images make it hard to label all the nuclei. The ambiguous tumor nuclei with blurred boundaries and low color contrast lead to inaccurate manual labels, especially on the nucleus boundaries. Fortunately, these imperfections can be the noises of data which can somehow avoid model overfitting. In some cases, our proposed model can detect mislabeled nuclei and show more accurate boundaries. It can be a complement of manual annotations. Fig. 6(c) shows our results of Fig. 6(a). We can observe that our proposed model segmented some miss-labeled nuclei and predicted more accurate boundaries than the ground truth in the yellow and blue boxes. However, our proposed model has an apparent drawback that we cannot well separate the attached nuclei, which is also a common limitation of existing approaches.

To measure the experts' agreement of the segmentation results, we also invited five experienced pathologists for an inter-rater evaluation. We randomly selected 20 images from each dataset, a total of 60 images were selected. Each group contains two images,

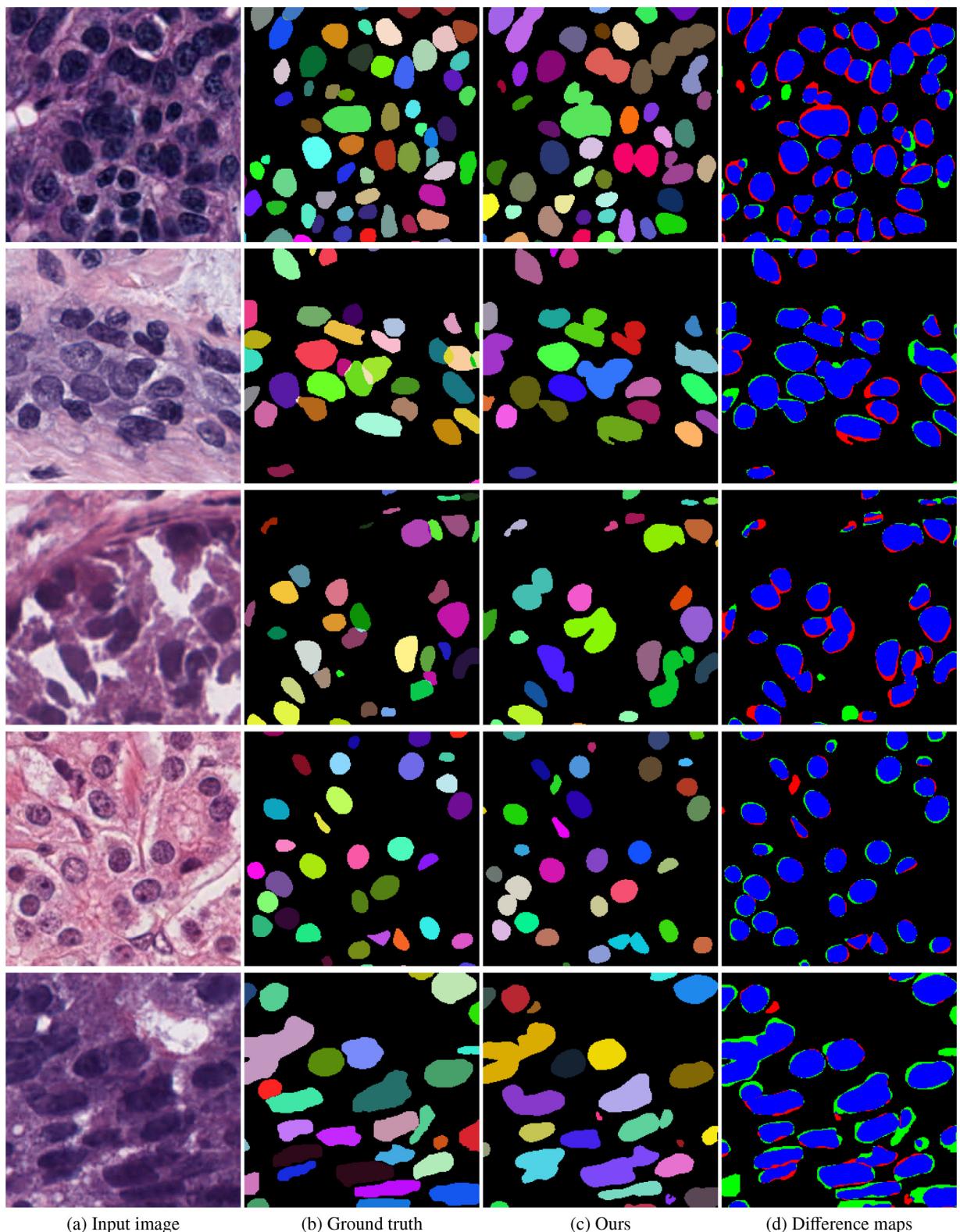


Fig. 5. The result of instance segmentation. (a) is the input image. (b) is the ground truth. (c) is the result of our model. (d) shows the difference maps between the ground truth and our results. Blue areas indicate the intersection of our result and the ground truth. Red and green areas indicate the false positive and the false negative segmentation respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

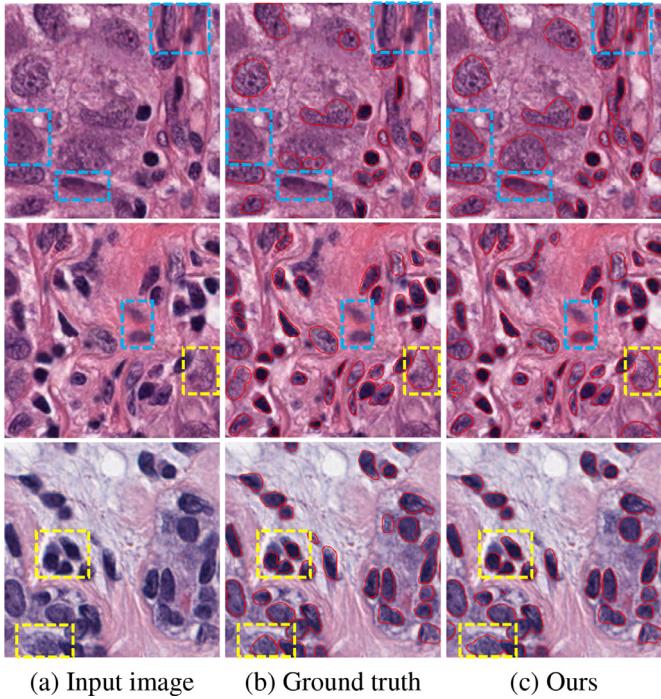


Fig. 6. Examples of inaccurate labels in the MoNuSeg dataset (Kumar et al., 2017). The overlooked nucleus labels in ground truth (b) are marked in blue boxes. The inaccurate nucleus boundaries in ground truth (b) are marked in yellow boxes. The correctness of the observation of these inaccurate labels is confirmed by the experienced pathologists. Our results in (c) are with more accurate boundaries than the ground truth in some cases. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

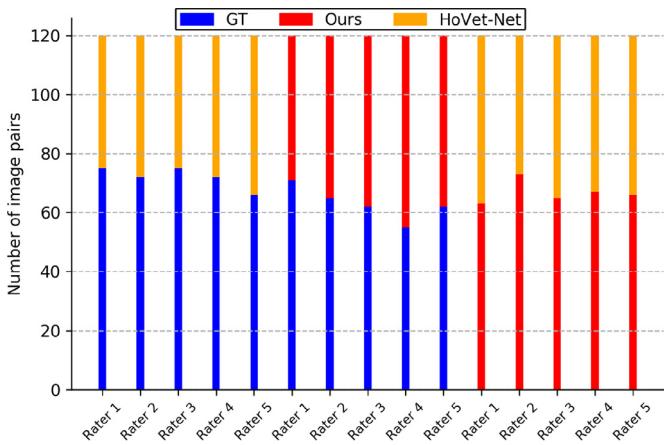


Fig. 7. HoVer-Net vs. GT, Ours vs. GT and HoVer-Net vs. Ours in the randomly selected 120 nuclei segmentation pairs per rater.

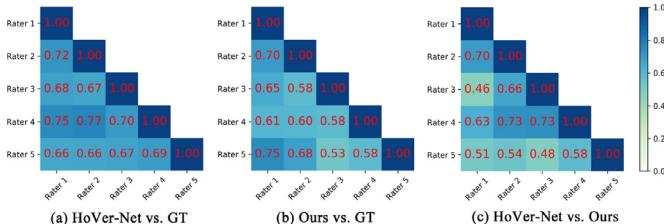


Fig. 8. Confusion matrix of inter-rater reliability. Each block indicates one-versus-one similarity calculated by Cohen's Kappa coefficient.

HoVer-Net vs. GT, Ours vs. GT and HoVer-Net vs. Ours. For each segmentation pair, pathologists were asked to choose a better segmentation result in two aspects, the completeness of segmentation in instance-level and the precision of nuclei boundaries. So each rater has to make 120 choices for each comparison. Fig. 7 demonstrates the overall voting scores of five raters for all the comparisons. And Fig 8(a)-(c) shows the inter-rater reliability of HoVer-Net vs. GT, Ours vs. GT and HoVer-Net vs. Ours, measured by Cohen's Kappa scores.

5. Conclusion

In this paper, we present a Hematoxylin-aware Triple U-net for nuclei segmentation from WSI images, which is a new paradigm and perspective to effectively leverage the unique characteristic of H&E staining for network supervision. Modeling the WSI as a light absorbing phenomenon and extracting the Hematoxylin component by Beer-Lambert's law, our proposed model is much more robust to color inconsistency, hence the color normalization is no longer necessary. In addition, we propose a more reasonable feature aggregation module, called Progressive Dense Feature Aggregation, which allows the network to progressively merge and learn features from different domains or different tasks. Extensive experiments and ablation studies on three datasets demonstrated the superiority of our proposed network. Quantitative experiments show that our proposed method has achieved state-of-the-art nuclei segmentation performance.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Bingchao Zhao: Software, Methodology, Formal analysis, Validation, Writing - original draft. **Xin Chen:** Investigation, Formal analysis. **Zhi Li:** Investigation, Resources. **Zhiwen Yu:** Investigation, Resources. **Su Yao:** Data curation, Investigation. **Lixu Yan:** Data curation, Investigation. **Yuqian Wang:** Investigation. **Zaiyi Liu:** Supervision, Resources, Project administration, Funding acquisition. **Changhong Liang:** Supervision, Resources, Project administration, Funding acquisition. **Chu Han:** Supervision, Writing - original draft, Writing - review & editing, Methodology, Conceptualization, Visualization, Validation.

Acknowledgements

This work was supported by the National Key R&D Program of China (No.2017YFC1309100), the National Science Fund for Distinguished Young Scholars (No.81925023), the National Natural Science Foundation of China (No. 81771912, 81601469 and 81702322), the National Science Foundation for Young Scientists of China (No. 81701662) and Science and Technology Planning Project of Guangdong Province (No. 2017B020227012), Guangzhou Science and Technology Project of Health (No. 2019A011002).

References

- Bejnordi, B.E., Veta, M., Van Diest, P.J., Van Ginneken, B., Karssemeijer, N., Litjens, G., Van Der Laak, J.A., Hermsen, M., Manson, Q.F., Balkenhol, M., et al., 2017. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA* 318 (22), 2199–2210.
- Chan, L., Hosseini, M.S., Rowsell, C., Plataniotis, K.N., Damaskinos, S., 2019. HistosegNet: semantic segmentation of histological tissue type in whole slide images. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 10662–10671.

- Chen, H., Qi, X., Yu, L., Heng, P.-A., 2016. DCAN: Deep contour-aware networks for accurate gland segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2487–2496.
- Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2017. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans Pattern Anal Mach Intell* 40 (4), 834–848.
- Chen, L.-C., Papandreou, G., Schroff, F., Adam, H., 2017b. Rethinking atrous convolution for semantic image segmentation. arXiv:1706.05587.
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 801–818.
- Cheng, J., Rajapakse, J.C., et al., 2008. Segmentation of clustered nuclei with shape markers and marking function. *IEEE Trans. Biomed. Eng.* 56 (3), 741–748.
- Chidester, B., Ton, T.-V., Tran, M.-T., Ma, J., Do, M.N., 2019. Enhanced rotation-equivariant U-net for nuclear segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 0–0.
- Christ, P.F., Elshaer, M.E.A., Ettlinger, F., Tatavarty, S., Bickel, M., Bilic, P., Rempfler, M., Armbruster, M., Hofmann, F., D'Anastasi, M., et al., 2016. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3d conditional random fields. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 415–423.
- de Geus, D., Meletis, P., Dubbelman, G., 2018. Panoptic segmentation with a joint semantic and instance segmentation network. arXiv:1809.02110.
- Graham, S., Rajpoot, N.M., 2018. SAMS-NET: Stain-aware multi-scale network for instance-based nuclei segmentation in histology images. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 590–594.
- Graham, S., Vu, Q.D., Raza, S.E.A., Azam, A., Tsang, Y.W., Kwak, J.T., Rajpoot, N., 2019. HoVer-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med. Image Anal.* 58, 101563.
- Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M., Larochelle, H., 2017. Brain tumor segmentation with deep neural networks. *Med. Image Anal.* 35, 18–31.
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969.
- Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700–4708.
- Kirillov, A., He, K., Girshick, R., Rother, C., Dollár, P., 2020. Panoptic segmentation.
- Koohbanani, N.A., Jahanifar, M., Gooya, A., Rajpoot, N., 2019. Nuclear instance segmentation using a proposal-free spatially aware deep learning framework. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 622–630.
- 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 Trans. Med. Imaging* 36 (7), 1550–1560.
- 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 Trans. Med. Imaging* 37 (12), 2663–2674.
- Liu, D., Zhang, D., Song, Y., Zhang, C., Zhang, F., O'Donnell, L., Cai, W., 2019. Nuclei segmentation via a deep panoptic model with semantic feature fusion. In: Proceedings of the 28th International Joint Conference on Artificial Intelligence. AAAI Press, pp. 861–868.
- Liu, S., Qi, L., Qin, H., Shi, J., Jia, J., 2018. Path aggregation network for instance segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8759–8768.
- Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G. E., Kohlberger, T., Boyko, A., Venugopalan, S., Timofeev, A., Nelson, P. Q., Corrado, G. S., et al., 2017. Detecting cancer metastases on gigapixel pathology images. arXiv:1703.02442.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440.
- Man, Y., Huang, Y., Li, J.F.X., Wu, F., 2019. Deep Q learning driven CT pancreas segmentation with geometry-aware u-net. *IEEE Trans. Med. Imaging* 38 (8), 1971–1980.
- Milletari, F., Ahmadi, S.-A., Kroll, C., Plate, A., Rozanski, V., Maiostre, J., Levin, J., Dietrich, O., Ertl-Wagner, B., Böttzel, K., et al., 2017. Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound. *Comput. Vis. Image Underst.* 164, 92–102.
- Naik, S., Doyle, S., Agner, S., Madabhushi, A., Feldman, M., Tomaszewski, J., 2008. Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology. In: 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, pp. 284–287.
- Naylor, P., Laé, M., Reyal, F., Walter, T., 2018. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Trans. Med. Imaging* 38 (2), 448–459.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern* 9 (1), 62–66.
- Parson, W.W., 2007. Modern Optical Spectroscopy, Vol. 2. Springer.
- Pereira, S., Pinto, A., Alves, V., Silva, C.A., 2016. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans. Med. Imaging* 35 (5), 1240–1251.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.
- Roth, H.R., Lu, L., Farag, A., Shin, H.-C., Liu, J., Turkbey, E.B., Summers, R.M., 2015. DeepOrgan: Multi-level deep convolutional networks for automated pancreas segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 556–564.
- Ruifrok, A.C., Johnston, D.A., et al., 2001. Quantification of histochemical staining by color deconvolution. *Anal. Quant. Cytol. Histol.* 23 (4), 291–299.
- Vahadane, A., Peng, T., Sethi, A., Albarqouni, S., Wang, L., Bautz, M., Steiger, K., Schlitter, A.M., Esposito, I., Navab, N., 2016. Structure-preserving color normalization and sparse stain separation for histological images. *IEEE Trans. Med. Imaging* 35 (8), 1962–1971.
- Vincent, L., Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* (6) 583–598.
- 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. *Front. Bioeng. Biotechnol.* 7.
- Xu, J., Luo, X., Wang, G., Gilmore, H., Madabhushi, A., 2016. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. *Neurocomputing* 191, 214–223.
- Yang, T.-J., Collins, M. D., Zhu, Y., Hwang, J.-J., Liu, T., Zhang, X., Sze, V., Papandreou, G., Chen, L.-C., 2019. DeeperLab: Single-shot image parser. arXiv:1902.05093.
- Yu, K.-H., Zhang, C., Berry, G.J., Altman, R.B., Ré, C., Rubin, D.L., Snyder, M., 2016. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat. Commun.* 7, 12474.
- Zhang, Y., Yang, L., Chen, J., Fredericksen, M., Hughes, D.P., Chen, D.Z., 2017. Deep adversarial networks for biomedical image segmentation utilizing unannotated images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 408–416.
- 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, pp. 2881–2890.
- Zhou, Y., Graham, S., Alemi Koohbanani, N., Shaban, M., Heng, P.-A., Rajpoot, N., 2019. CGC-Net: Cell graph convolutional network for grading of colorectal cancer histology images. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, 0–0.
- Zhou, Y., Onder, O.F., Dou, Q., Tsougenis, E., Chen, H., Heng, P.-A., 2019. CIA-Net: Robust nuclei instance segmentation with contour-aware information aggregation. In: International Conference on Information Processing in Medical Imaging. Springer, pp. 682–693.
- Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J., 2018. 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, pp. 3–11.