



Meta multi-task nuclei segmentation with fewer training samples

Chu Han^{a,b,c,1}, Huasheng Yao^{a,c,1}, Bingchao Zhao^{a,c,1}, Zhenhui Li^{a,b,c}, Zhenwei Shi^{a,b,c}, Lei Wu^{a,b,c}, Xin Chen^f, Jinrong Qu^e, Ke Zhao^{a,b,c}, Rushi Lan^{d,*}, Changhong Liang^{a,c,*}, Xipeng Pan^{a,b,c,d,**}, Zaiyi Liu^{a,c,*}

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

^b Guangdong Cardiovascular Institute, Guangzhou, China

^c Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China

^d School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin 541004, China

^e Department of Radiology, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou 450008, China

^f 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 6 January 2021

Revised 5 May 2022

Accepted 13 May 2022

Available online 18 May 2022

Keywords:

Nuclei segmentation

Meta learning

Multi-task learning

Convolutional neural networks

ABSTRACT

Cells/nuclei deliver massive information of microenvironment. An automatic nuclei segmentation approach can reduce pathologists' workload and allow precise of the microenvironment for biological and clinical researches. Existing deep learning models have achieved outstanding performance under the supervision of a large amount of labeled data. However, when data from the unseen domain comes, we still have to prepare a certain degree of manual annotations for training for each domain. Unfortunately, obtaining histopathological annotations is extremely difficult. It is high expertise-dependent and time-consuming. In this paper, we attempt to build a generalized nuclei segmentation model with less data dependency and more generalizability. To this end, we propose a meta multi-task learning (Meta-MTL) model for nuclei segmentation which requires fewer training samples. A model-agnostic meta-learning is applied as the outer optimization algorithm for the segmentation model. We introduce a contour-aware multi-task learning model as the inner model. A feature fusion and interaction block (FFIB) is proposed to allow feature communication across both tasks. Extensive experiments prove that our proposed Meta-MTL model can improve the model generalization and obtain a comparable performance with state-of-the-art models with fewer training samples. Our model can also perform fast adaptation on the unseen domain with only a few manual annotations. Code is available at <https://github.com/ChuHan89/Meta-MTL4NucleiSegmentation>

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Computational pathology has been one of the hottest concepts in biological and clinical research in recent years. As the basic unit of the life system, cells/nuclei microenvironment delivers massive information. Many research works attempt to explore both explicit and implicit patterns of nuclei for disease diagnosis (Yamashita et al., 2020), prognostic prediction (Skrede et al.,

2020; Lu et al., 2020), tissue phenotyping (Javed et al., 2020) and etc. However, it is impractical to segment millions of nuclei manually due to the huge resolution of whole slide images. Therefore, it is valuable to construct a fully automatic model for nuclei segmentation.

Convolutional neural network has demonstrated the superiority in medical image segmentation (Moeskops et al., 2016; Zhou et al., 2018; Chan et al., 2019). Meanwhile, Various CNN-based models (Graham et al., 2019; Liu et al., 2019) have been developed for nuclei segmentation and achieved outstanding performance under full supervision. However, most of the existing models are data-dependent. They have to optimize the neural networks for each individual dataset. When data from the different domains come, we still have to prepare a new dataset for the new training. Unfortunately, manual annotations are expensive, especially

* Corresponding author at: Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China.

** Co-corresponding authors.

E-mail addresses: rslan2016@guet.edu.cn (R. Lan), liangchanghong@gdph.org.cn (C. Liang), pxp201@guet.edu.cn (X. Pan), Zyliu@163.com (Z. Liu).

¹ Contributed equally.

for histopathology images. It requires high expertise and is extremely tedious and time-consuming. Moreover, the neural network is often trained on top of a massive amount of training data with manual labels. Otherwise, it may drop into the overfitting problem. It is impractical to collect a large amount of annotated data for every single task. Therefore, it will be convenient and effective if fewer training samples can be used for model training while also greatly relieving the tremendous annotation workload of pathologists.

There have been a few attempts to develop neural network models with fewer training samples. Transfer learning (Pan and Yang, 2009) was proposed to address the lack of labeled data problem by transferring the data distribution from the source domain to the target domain. This kind of model fine-tuning strategy also relies on a certain degree of training data of the target domain. Recently, multi-task learning (Chen et al., 2016; Zhao et al., 2020) has been widely used for medical image segmentation by setting up auxiliary tasks from the same domain to improve the target task. This strategy can effectively increase model generalization. Meta-learning (Naik and Mammone, 1992; Hospedales et al., 2020) aims to optimize the model itself instead of optimizing the final goal of the model. The learning to learn spirit is much closer to how humans learn the world.

In this paper, we aim to build a generalized model for nuclei segmentation with fewer training samples. Consequently, we absorb the spirits of meta-learning as well as multi-task learning and associate them together as a *Meta Multi-Task Learning* (Meta-MTL) scheme. To improve the model generalizability, we apply a model-agnostic meta-learning approach (Finn et al., 2017) as the outer optimization algorithm for the complete model. Then we construct a contour-aware multi-task learning model as the inner model for nuclei segmentation. There are two tasks in the inner model, a segmentation task and a contour detection task. These two parallel tasks can complement each other and let the neural network focus on more generalized features instead of task-specific ones. We construct an intermediate component with several feature fusion and interaction blocks (FFIB) to support feature interactions and feature sharing across two tasks. A feature attention strategy is introduced in the decoding phase to further amplify contour information in the segmentation task.

Extensive experiments are conducted to evaluate the effectiveness of our proposed model in three aspects, data dependency, domain adaptation and domain generalization. (1) For each dataset, we progressively deduct the amount of training samples randomly with a 20% descending proportion. (2) Given a well-trained model on a specific dataset, we transfer this model to an unseen dataset with fewer training samples. (3) Given a model train on a specific dataset, we directly apply this model to an unseen dataset. Experimental results suggest that our model achieves comparable performances with state-of-the-art nuclei segmentation models with fewer training data. And our model demonstrates a better potential for domain adaptation and domain generalization. With such properties, our proposed model could be an initial segmentation model for the pathologist-in-the-loop labeling strategy with only limited start-up pixel-level labels, to save the annotation efforts. The contributions of this paper are three-fold:

- We propose a novel meta multi-task learning (Meta-MTL) scheme for nuclei segmentation with better domain adaptation and domain generalization abilities as well as less data dependency.
- In the contour-aware multi-task learning model, we propose a feature fusion and interaction block (FFIB) for features communication between two parallel tasks.
- Our proposed model achieves comparable performances with state-of-the-art models with fewer training samples.

2. Related works

2.1. Nuclei segmentation

Nuclei segmentation is a fundamental step in histopathology image analysis. Kumar et al. (2017) published a nuclei segmentation data set of histopathology images originating from multiple tissue sites, and proposed a simple CNN-based method to segment nuclei, background and cell boundaries, respectively. In addition, in order to improve the objectivity of the model evaluation, a new segmentation performance evaluation metric (Aggregated Jaccard Index, AJI) was proposed. This metric considered both pixel-level and instance-level errors.

Deep neural networks are now the most popular techniques in the research field of medical image segmentation. Chen et al. (2016) proposed a contour-aware network to separate the attached objects with both appearance and contour information. However, its performance highly depends on the quality of nuclei boundaries, which can be unclear occasionally. Oda et al. (2018) employed contour-specific features to learn complementary information in nuclei branch. Zhou et al. (2019) proposed a contour-aware informative aggregation network (CIA-Net) to address the subjective annotating and mislabeling issue to promote the generalization capability for robustly segmenting unseen organ nuclei. Spatial and texture information was aggregated and smooth truncated loss was presented to reduce the perturbation from outliers. Cheng and Qu (2020) proposed a box-based method for nucleus instance segmentation by combining detection and segmentation process. The fusion module based on a feature pyramid network was employed to detect the nuclei position by combining the shallow and deep information. Then, the feature maps were cropped according to the bounding boxes and fed into the U-net network as a guide to separate clustered nuclei. A stain-aware loss was proposed by Graham and Rajpoot (2018) for nuclei segmentation. Hover-Net (Graham et al., 2019) performed nuclei segmentation and classification at the same time. This work utilized horizontal and vertical distance maps to segment attached nuclei, which achieved state-of-the-art performance. Zhao et al. (2020) developed a Triple U-net model to segment nuclei without color normalization, and designed a new feature aggregation strategy to gradually fuse features and learn better feature representation from different domains. Graham et al. (2021) proposed a pathologist-in-the-loop pipeline for nuclei segmentation and classification. Over 490,000 colonic nuclei were fast generated by using this pipeline.

Although existing models have achieved outstanding performance under the supervision of dense annotations. How to reduce training samples while maintaining a reasonable performance is still an open problem.

2.2. Meta learning

Deep neural networks have achieved great success in many image understanding tasks with dense annotations, such as image recognition, image super-resolution and machine translation (He et al., 2016; Lan et al., 2020; Wan et al., 2020). However, humans are able to learn things with a very limited amount of samples. Therefore, researchers are now seeking a smarter way to mimic how humans learn things in the real world scenario. Meta-learning was proposed to learn more generalized features rather than being trapped in a specific task or dataset itself. It is well known for its learning to learn spirit, which can improve the model generalization and adaptation abilities. Meta learning can be roughly categorized into two types, metric-based and optimization-based meta learning. Metric-based meta learning aims to measure the distance among samples in different distance spaces, such as siamese

network (Koch et al., 2015), matching network (Vinyals et al., 2016), prototypical network (Snell et al., 2017) and relation network (Sung et al., 2018). Optimization-based meta learning focuses on how to optimize the model by separating the task into a support set and a query set instead of training the network with the entire dataset. MAML (Finn et al., 2017) and Reptile (Nichol and Schulman, 2018) are two representative optimization-based meta learning algorithms.

Currently, more researchers applied meta learning to image segmentation tasks. Liu et al. (2020) proposed a shape-aware meta-learning method to address the domain shift issue and perform domain generalization for different MR images. In order to reduce the dependence on labeled data and obtain good model generalization, Cao et al. (2019) proposed a generalized meta-learning framework (Meta-Seg) for few-shot semantic segmentation, which consisted of a base-learner and a meta-learner. The meta-learning deep visual words was presented for fast video object segmentation (Behl et al., 2018). A continuous and online meta-learning method was proposed to complete multiple video object segmentation tasks (Xiao et al., 2019). The meta-learning model could capture their common knowledge and acquire the ability to quickly adapt to the new video sequence segmentation model.

In this paper, we aim to leverage the learning to learn spirit for nuclei segmentation task to increase the model generalizability and finally reduce the data dependency of the model.

2.3. Multi-task learning

Single-task learning typically learns things from a specific dataset which may lack model adaptability and generalizability when facing data from the unseen domain. Multi-task learning (MTL) is generally used to increase the model generalizability by solving multiple tasks with similar goals at the same time in one single model. It can maximize the value of the training data, alleviate the overfitting of the model with good generalization ability, and achieve fast learning using auxiliary information.

In recent years, the combination of multi-task learning and deep learning has been widely used in computer vision, natural language processing and medical image processing (Crawshaw, 2020; Zhang and Yang, 2017; Ruder, 2017). An end-to-end deep learning algorithm (Chamanzar and Nie, 2020) was developed to perform cell detection and segmentation by combining different task-oriented point label coding methods with multi-task training. Hou et al. (2019) proposed a sparse convolutional autoencoder (CAE) to detect and encode nuclei in histopathology tissue images. Sun et al. (2019) constructed a multi-scale region proposal network using mask R-CNN for nuclei detection and instance segmentation. Multi-scale features and anchor mechanisms were conducive to the detection of tiny objects. Besides, various auxiliary tasks have been introduced in MTL for nuclei segmentation, such as contour detection task (Zhao et al., 2020; Zhou et al., 2019), distance map prediction task (Naylor et al., 2018), nuclei classification task (Graham et al., 2019) and etc.

In this paper, we combine multi-task learning with meta learning to further improve the model generalization and adaptation abilities while reducing the data dependency.

3. Meta-MTL: meta multi-task nuclei segmentation

The huge resolutions and the expertise dependency of histopathology images make pixel-wise manual annotations extremely difficult. We want to fully discover the potential of the nuclei segmentation model, in order to minimize the dependency of the labeled data and improve model generalization, thereby finally alleviate the workloads of pathologists. To

this end, we associate two very important techniques in meta-learning (MAML and multi-task learning) and propose a *Meta-MTL* model.

In this section, we introduce our proposed Meta-MTL model for nuclei segmentation, demonstrated in Fig. 1. Our proposed model contains two nested parts, (1) an outer algorithm of the training strategy, model-agnostic meta-learning (Finn et al., 2017), for fast model adaptation and minimal data dependency and (2) an inner model of the multi-task backbone with two parallel tasks. We propose a feature fusion and interaction block (FFIB) to connect both tasks. An attention mechanism is applied to FFIB in the decoding phase to amplify the contour information. We demonstrate the details of the Meta-MTL model as follows.

3.1. Problem setting

We aim to build a nuclei segmentation model that only require a small amount of manual annotations, and can be easily extended to the unseen domain. The same with the existing meta-learning approaches, we split the dataset into meta train set \mathcal{T}_{train} for model training and meta test set \mathcal{T}_{test} for testing. As shown in Fig. 1, $\mathcal{T}_{train} = \{\mathcal{T}_i\}_{i=1}^n$ and $\mathcal{T}_{test} = \{\mathcal{T}_i\}_{i=n+1}^m$ are composed of several sub-tasks \mathcal{T} which contains a set of support images S and a set of query images Q , define as $\mathcal{T} = \{S, Q\}$. Since the inner model is a dual tasks model, there are two corresponding ground truth masks for one image, $S = \{(\mathcal{I}_i, \mathcal{M}_i^{seg}, \mathcal{M}_i^{edge})\}_{i=1}^{N_{support}}$ and $Q = \{(\mathcal{I}_i, \mathcal{M}_i^{seg}, \mathcal{M}_i^{edge})\}_{i=1}^{N_{query}}$ where \mathcal{M}^{seg} is the segmentation map, \mathcal{M}^{edge} is the edge map. We set $N_{support} = 2$ and $N_{query} = 1$ in our model.

3.2. Outer algorithm: model-agnostic meta-learning

Considering the inherent difficulties of collecting medical image labels, we aim to improve the model generalization with minimal data dependency. To this end, we introduce a model-agnostic meta-learning approach (Finn et al., 2017) as the outer pre-train algorithm of the network.

Different from conventional machine learning schemes that minimize the loss function \mathcal{L} between ground truth labels and the predicted values, MAML tends to learn the internal characteristics of the tasks distribution $p(\mathcal{T})$. Therefore, it encourages machine learning models to focus on internal features instead of any task-specific features.

Let $f(\mathcal{T}, x; \theta)$ denote the model function f with parameters θ given the training sample $x \in S$ from the sub-task \mathcal{T} , we can compute an intermediate parameter set θ' :

$$\theta' = \theta - \alpha \nabla_{\theta} \mathcal{L}_{\mathcal{T}}(f(\mathcal{T}, x; \theta)) \quad (1)$$

where α can be regarded as the learning rate of this iteration. Then the model is optimized and updated by the intermediate model performance $f(\mathcal{T}, y; \theta')$ with respect to θ using the test sample $y \in Q$. The MAML objective is defined as follows:

$$\min_{\theta} \sum_{\mathcal{T} \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}}(f(\mathcal{T}, y; \theta')) \quad (2)$$

Here, Eq. (2) is equivalent to Eq. (3)

$$\min_{\theta} \sum_{\mathcal{T} \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}}(f(\mathcal{T}, y; \theta - \alpha \nabla_{\theta} \mathcal{L}_{\mathcal{T}}(f(\mathcal{T}, x; \theta)))) \quad (3)$$

After a meta-optimization of the query set, the parameters θ^* will be updated using stochastic gradient descent as follows:

$$\theta^* \leftarrow \theta - \beta \nabla_{\theta} \sum_{\mathcal{T} \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}}(f(\mathcal{T}, y; \theta')) \quad (4)$$

where β is the meta step size.

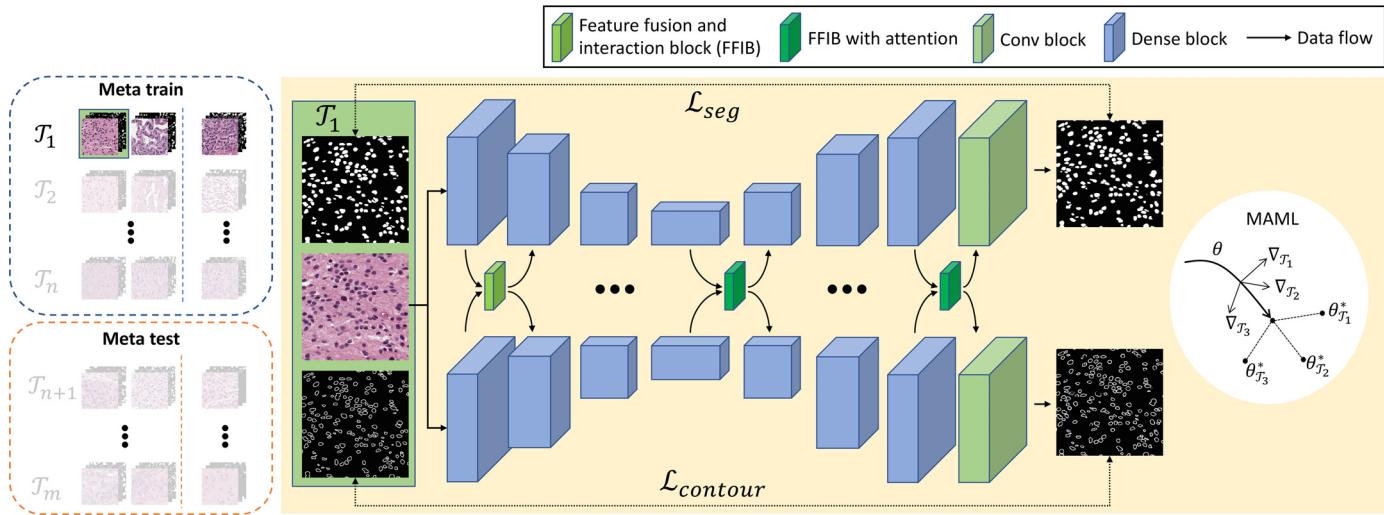


Fig. 1. System overview of meta multi-task learning model. This model contains two tasks with two U-net structures, a segmentation task and a contour detection task. These two tasks influence each other by passing the features into feature fusion and intersection blocks (FFIB). A model-agnostic meta-learning is applied to a multi-task network to improve model generalization. For each sub-task T in meta train and meta test, we have two support images and one query image. Skip connections are hidden for simplification. There are three FFIBs in the encoding phase and four FFIBs with attention in the decoding phase.

Here, the loss function in the MAML objective is the total loss of the multi-task learning network which is defined in Section 3.3.3.

3.3. Inner model: contour-aware multi-task network

Multi-task learning (Caruana, 1997) is also a kind of meta-learning technique. Introducing additional auxiliary tasks can make the network learn more general features rather than task-specific features. It improves the model diversity and can obtain a better transferable model. In nuclei segmentation, we are able to generate contour labels of the nuclei from the segmentation masks without additional cost for human annotations. So in the inner network, we set up two completely parallel tasks, a segmentation task and a contour detection task, which both share a similar goal to detect nuclei. The difference is that one focuses on segmenting the complete nucleus and the other attempts to extract the contour of the nucleus. They can complement each other while further improving the model generalization and refining the contour of the segmentation results.

3.3.1. Network structure

Fig. 1 demonstrates the network structure of our proposed model. We use U-net (Ronneberger et al., 2015) structure equipped with densely-connect blocks (Huang et al., 2017) as the backbone of the segmentation task as well as the contour detection task. These two tasks do not share the same encoder because we want each task can focus on both internal features and task-specific features simultaneously.

3.3.2. Feature fusion and interaction block

To build the communication between two tasks, we propose a feature fusion and interaction block (FFIB). As demonstrated in Fig. 2(a), features from each dense-block at the same level in two U-nets are aggregated in FFIB. The aggregated features will be passed back to the next level of both U-nets after a convolutional layer. Such a feature interaction scheme allows two tasks to complement each other, resulting in performance improvement on both tasks.

In the decoding phase, we introduce a contour attention module to further amplify the contour information in the segmenta-

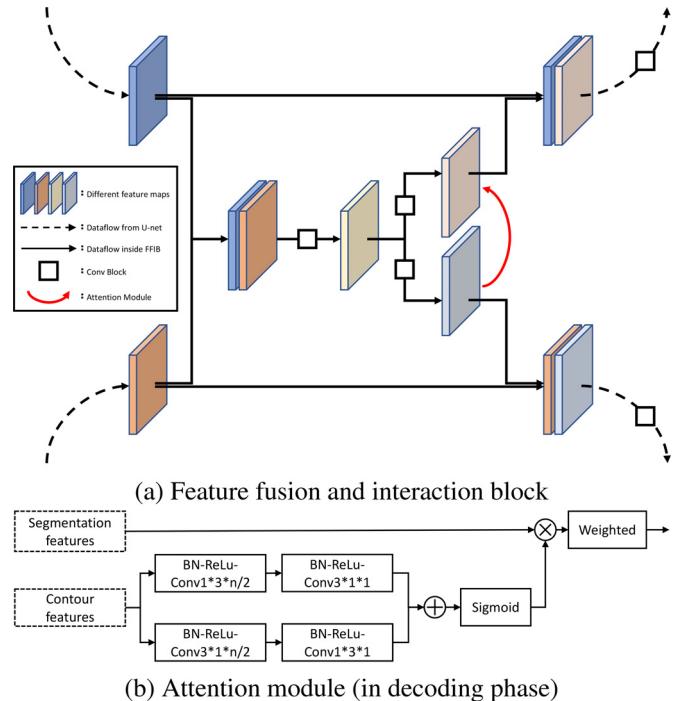


Fig. 2. Feature fusion and interaction block. The red arrow indicates an attention module, which is only applied in decoding phase. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tion task, the red arrow in Fig. 2(a). The structure of the contour attention module is demonstrated in Fig. 2(b). It encourages the segmentation task to aggregate contour information. In the encoding phase, we want two tasks to mainly focus on their own missions with feature interactions only, with no attention module. In the decoding phase, we introduce this attention module to emphasize the contour information in the segmentation task to generate segmentation results with more precise nuclei boundaries.

Table 1

Quantitative comparison with baseline models. Bold means the best performance. Ratio means the portion of the training data involved in the training phase. *T*-test is performed for the models trained by 80% samples between the metrics of each baseline model and our model. * means the result with *p*-value *p* < 0.05.

Ratio	Method	MoNuSeg			CPM-17		
		AJI	DICE	PQ	AJI	DICE	PQ
100%	(1) CNN3	0.508	0.762	–	–	–	–
	(2) Mask R-CNN	0.546	0.760	0.509	0.684	0.850	0.674
	(3) DCAN	0.525	0.793	0.492	0.561	0.828	0.545
	(4) DIST	0.560	0.786	0.443	0.616	0.826	0.504
	(5) Deep Panoptic	0.585	0.794	–	–	–	–
	(6) CIA-Net	0.620	0.818	0.577	–	–	–
	(7) HoVer-Net	0.618	0.826	0.597	0.705	0.869	0.697
	(8) TripleU	0.621	0.837	0.601	0.711	0.888	0.685
	(9) Ours	0.616	0.826	0.569	0.708	0.876	0.679
80%	(7) HoVer-Net	0.554*	0.835	0.570	0.638*	0.864	0.632
	(8) TripleU	0.567*	0.836*	0.560	0.655*	0.870	0.634
	(9) Ours	0.615	0.826	0.570	0.675	0.865	0.643

3.3.3. Loss function

The loss function of the proposed model contains two parts which correspond to the objectives of two tasks.

$$\begin{aligned}\mathcal{L}_{\text{total}} &= \mathcal{L}_{\text{seg}} + \mathcal{L}_{\text{contour}} \\ &= \mathcal{L}_{\text{seg}-\text{sd}} + \mathcal{L}_{\text{seg}-\text{st}} + \mathcal{L}_{\text{contour}-\text{sd}}\end{aligned}\quad (5)$$

Soft dice loss Since the foreground and background are usually imbalance in medical images, the model might be easily dominated by the majority area (background) which leads to a strong bias. So we adopt soft dice loss in Eq. (6) for both segmentation and contour detection tasks to balance the foreground and background.

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

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

Smooth truncated loss The huge resolution and expertise dependency make it extremely difficult for manual annotations. It is also hard to obtain perfect dense-pixel labels with precise boundaries because of the ambiguity of the blurry boundaries. Zhou et al. (2019) proposed a smooth truncated (ST) loss to tackle the imperfect label problem. The idea of ST loss is intuitive. Because the imperfect labels typically account for only a small part of the entire dataset. When training the neural network model, such imperfect labels may be the outliers in the feature space. The idea of truncated loss is to suppress such outliers (noises or imperfect labels) and let the model focus on more informative regions. Smooth truncated loss is the smoother version of it, to avoid getting hurt by a constant truncated threshold. In our model, we applied the smooth truncated loss in the segmentation task.

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

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

3.4. Implementation and training details

Our network is implemented using Tensorflow version 1.14 on a workstation equipped with an NVIDIA GeForce RTX 2080 Ti. During the process of training, we adopt several data enhancement ways such as elastic transformation, scaling, shift, rotation and flipping. Original images are cropped into patches with a resolution of 256×256 . Adam optimizer (Kingma and Ba, 2014) is applied with the batch size of 2. The learning rates of two steps in the MAML

algorithm are $1e-4$ and $1e-3$, respectively. The number of samples in the support set and the query set in the MAML algorithm are 2 and 1, respectively.

4. Datasets and evaluation metrics

We evaluate our proposed method on two public nuclei segmentation datasets, Multi-Organ Nuclei Segmentation Dataset (MoNuSeg) (Kumar et al., 2017) and CPM-17 (Vu et al., 2019).

MoNuSeg contains 30 images from 7 tissue sites, including breast, kidney, liver, prostate, bladder, colon and stomach. The resolution of each image is 1000×1000 . We split the whole dataset into a training set and a test set. The training set contains 16 images from 4 kinds of tissues, including the breast, kidney, liver, and prostate. The test set contains 8 images from the seen tissues and 6 images from the unseen tissues.

CPM-17 contains 64 images (size: 500×500 or 600×600) from 4 cancer types, including glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSCC), lower-grade glioma (LGG) and non-small cell lung cancer (NSCLC). The training set and the test set both contain 32 images. To be fair, the image split of the training set and the test set are the same as the existing methods in all the experiments.

In the following experiments, three common metrics are employed for the comparisons of our model against existing models to evaluate the nuclei segmentation performance at both pixel-level and instance-level.

Dice coefficient is the most common measurement for segmentation tasks. It compares the similarity between the prediction results and the ground truth, which is defined as follows:

$$\text{DICE} = \frac{2 \times |G \cap P|}{|G| + |P|} \quad (8)$$

G is the segmentation result while G denotes the ground truth.

Aggregated Jaccard Index (Kumar et al., 2017) evaluates the nuclei segmentation performance in instance-level by calculating the ratio of the aggregated intersection cardinality and the aggregated union cardinality. Comparing to Dice score, it further penalizes under- and over-segmentation. It is defined as follows:

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

where $j = \arg\max_k \frac{G_i \cap P_k}{G_i \cup P_k}$, G_i denotes the ground truth of the i th nucleus, and P_j is the segmentation result of the j th nucleus.

To handle the problem of over-penalizing the overlapping region, Panoptic Quality (PQ) (Kirillov et al., 2019) can evaluate the

Table 2

Data dependency with fewer training samples. (CPM-17).

Metrics	Ratio	HoVer-Net	TripleU	Ours
DICE	100%	0.8690	0.8880	0.8756
	80%	0.8644	0.8698	0.8651
	60%	0.8568	0.8669	0.8483
	40%	0.8522	0.8628	0.8478
	20%	0.8281	0.8228	0.8195
	AJI	0.7050	0.7110	0.7081
PQ	100%	0.6375	0.6545	0.6754
	80%	0.6303	0.6521	0.6522
	60%	0.6059	0.6355	0.6413
	40%	0.5532	0.5900	0.6027
	20%	0.6970	0.6850	0.6791
	AJI	0.6323	0.6340	0.6427
PQ	100%	0.6171	0.6212	0.6143
	80%	0.5983	0.6093	0.6073
	60%	0.5225	0.5026	0.5396
	40%			
	20%			

performance of nuclei segmentation at the instance-level with more precise quantification and interpretability, which is defined as follows:

$$PQ = \frac{|TP|}{|TP| + \frac{1}{2}|FP| + \frac{1}{2}|FN|} \times \frac{\sum_{(x,y) \in TP} IoU(x, y)}{|TP|} \quad (10)$$

x is the ground truth and y denotes the segmentation result in instance-level. Each (x, y) pair with the intersection over union IoU over 0.5 can be considered to be unique.

5. Experiments

We evaluate the effectiveness of our proposed model on data dependency (Section 5.2), domain adaptation (Section 5.3) and domain generalization (Section 5.4) by a series of quantitative comparisons. The best performances of existing models are demonstrated in Section 5.1. Ablation studies (Section 5.5) are conducted

Table 3

Data dependency with fewer training samples. (MoNuSeg).

Metrics	Ratio	HoVer-Net	TripleU	Ours
DICE	100%	0.8260	0.8370	0.8262
	80%	0.8348	0.8358	0.8258
	60%	0.8333	0.8322	0.8109
	40%	0.8269	0.8242	0.8054
	20%	0.8011	0.8094	0.7955
	AJI	0.6180	0.6210	0.6163
PQ	100%	0.5535	0.5672	0.6148
	80%	0.5503	0.5435	0.5969
	60%	0.5281	0.5347	0.5859
	40%	0.4729	0.5183	0.5448
	20%	0.4790	0.5601	0.5701
	AJI	0.5970	0.6010	0.5694
PQ	100%	0.5700	0.5359	0.5463
	80%	0.5684	0.5359	0.5369
	60%	0.5536	0.5319	0.5369
	40%	0.4839	0.4473	0.4869
	20%			

to evaluation the superiority of the *Meta-MTL* design and the *FFIB*. Qualitative comparisons and results are demonstrated in each corresponding experiment. All the statistical results and the qualitative results of the existing works are directly from the corresponding papers or generated by their released implementation.

5.1. Quantitative comparison with baseline models

We first compare our proposed model with the several baseline models, include: (1) CNN3 (Kumar et al., 2017); (2) Mask R-CNN (He et al., 2017); (3) DCAN (Chen et al., 2016); and models tailored for nuclei segmentation: (4) DIST (Naylor et al., 2018); (5) Deep Panoptic (Liu et al., 2019); (6) CIA-Net (Zhou et al., 2019), the winner of MoNuSeg challenge 2019; (7) HoVer-Net (Graham et al., 2019), multi-task nuclei segmentation model and (8) TripleU (Zhao et al., 2020), state-of-the-art model. The quantitative results with 100% training data of the existing meth-

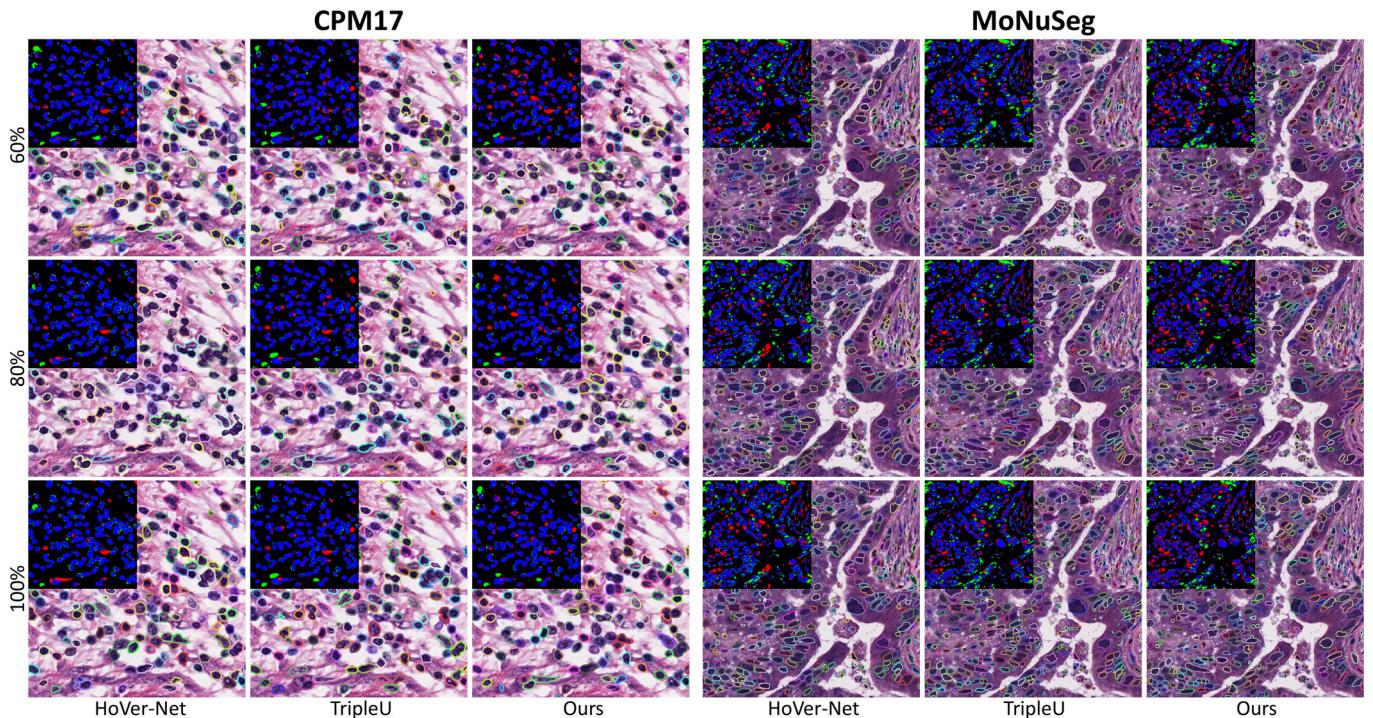


Fig. 3. Qualitative comparisons with SOTA models. We gradually deduct the training set with (a) 100%, (b) 80% and (c) 60%. In the difference maps, blue, green and red areas indicate true positive segmentation, false positive segmentation and false negative segmentation respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4

Quantitative evaluation on domain adaptation. C→M means the model was first trained on CPM-17 with a fast adaptation to MoNuSeg. M→C is the opposite.

Metrics	Ratio	C→M		M→C	
		TripleU	Ours	TripleU	Ours
DICE	15%	0.8348	0.8154	0.8206	0.8111
	10%	0.8341	0.8053	0.8194	0.7988
	5%	0.8344	0.8037	0.7936	0.7840
AJI	15%	0.5594	0.6065	0.5753	0.6013
	10%	0.5363	0.5912	0.5362	0.5796
	5%	0.5324	0.5844	0.5058	0.5563
PQ	15%	0.5390	0.5631	0.5264	0.5450
	10%	0.5254	0.5456	0.5039	0.5199
	5%	0.5243	0.5336	0.4796	0.4925

ods are directly from Zhao et al. (2020). We also show a simple comparison analysis against two state-of-the-art models, HoVer-Net and TripleU, with 80% of the training data. T-test is conducted between each baseline model and our model in all the metrics

Table 5

Domain generalization results. CPM (MoNuSeg) means using the trained model of CPM-17 to test the data in MoNuSeg and MoNuSeg(CPM) is the opposite.

Dataset	Metrics	HoVer-Net	TripleU	Ours
CPM (MoNuSeg)	DICE	0.7635	0.7801	0.7659
	AJI	0.4863	0.4769	0.5169
	PQ	0.4569	0.4436	0.4600
MoNuSeg (CPM)	DICE	0.8058	0.7660	0.8110
	AJI	0.5972	0.5331	0.5982
	PQ	0.5559	0.4578	0.5504

(80% training samples). Since the resolutions of the test images are large and the number of test images is limited, we randomly cut patches with smaller resolution (224×224) from the original images to guarantee the normal distribution of the samples. For MoNuSeg dataset, we sampled 32 patches from each test image. For CPM-17 dataset, we sampled 16 patches from each image. The metrics with p-value less than 0.05 are marked by * in

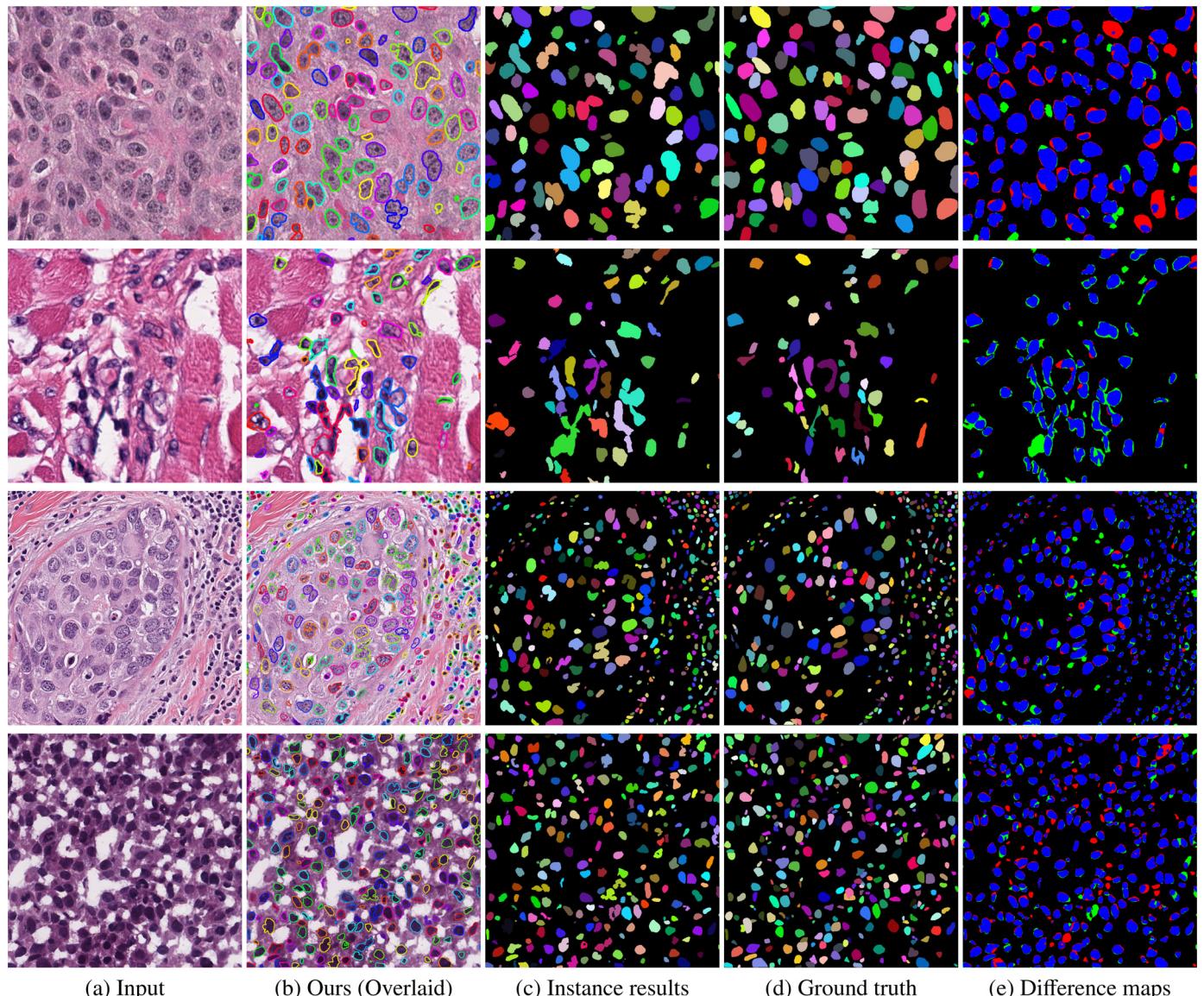


Fig. 4. Qualitative results of our model with 100% training data. Upper two rows are from CPM-17, the next two rows are from MoNuSeg. In the difference maps (e), blue, green and red areas indicate true positive segmentation, false positive segmentation and false negative segmentation respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

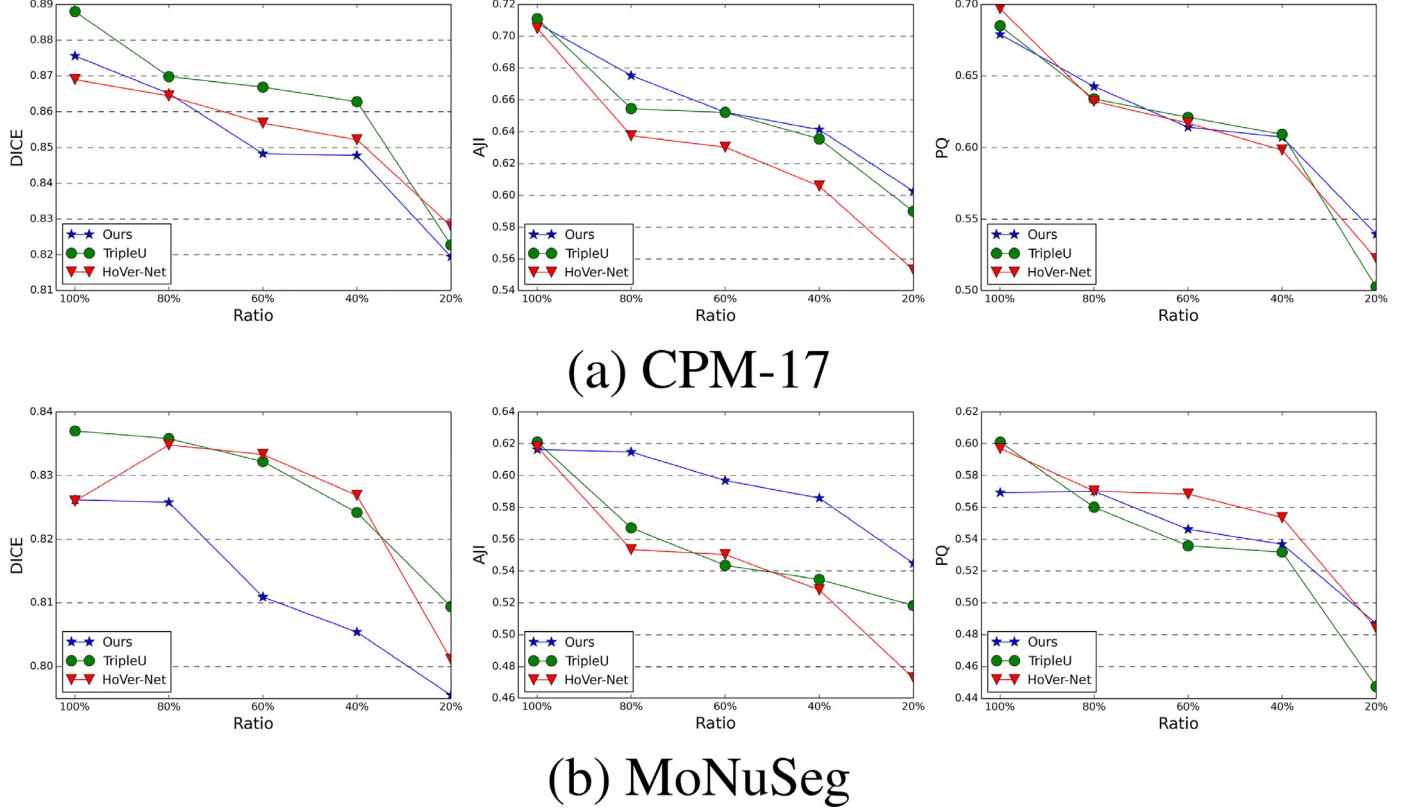


Fig. 5. Data dependency with fewer training samples. (a) Evaluation in CPM-17. (b) Evaluation in MoNuSeg. The results from left to right are DICE, ARI and PQ.

Table 1. More comprehensive comparisons on fewer training data are demonstrated in Tables 2 and 3.

Table 1 demonstrates the quantitative results. Our proposed model trained by the complete dataset outperforms most of the baseline models and achieves comparable performances with SOTA nuclei segmentation models in both datasets. In the lower part of Table 1, we can also observe that our model with 80% training data still outperforms most of the existing baseline models with even 100% training data. When comparing with two SOTA models, HoVer-Net and TripleU, the performance of the three models decreases. Our model outperforms them at the instance-level (ARI and PQ). However, the pixel-level performance (Dice) of ours is still slightly lower. Because MAML algorithm is originally designed for image-level prediction tasks and is less effective on pixel-level dense prediction. But combining MAML with multi-task learning can still improve the model generalization, especially with fewer training samples.

Fig. 3 demonstrates the qualitative comparison with two strong baseline models, HoVer-Net and TripleU. We can observe that our model with 100% training samples demonstrates comparable visual performance with SOTA models. When we deduct the training samples, the results are still promising. Here, we demonstrate more results of our proposed model (without postprocessing) on both datasets in Fig. 4.

5.2. Data dependency: fewer training samples

One of the major purposes of this paper is to minimize the data dependency of the model to reduce the labeled data requirement. In this experiment, we test the data dependency of our proposed model comparing with two SOTA nuclei segmentation models, HoVer-Net (Graham et al., 2019) and TripleU (Zhao et al., 2020). We set up this experiment by progressively reducing the amount of

training samples randomly with a 20% descending proportion until 20% of the training samples are left. To be fair, we train three models using the same configurations, e.g. same dataset and same image split. The models of two competitors were trained using their released codes.

Tables 2 and 3 demonstrate the quantitative results in MoNuSeg and CPM-17 respectively. For better visualization, we draw these two tables in Fig. 5(a) and (b). As can be seen, in both datasets, the performances (in the blue line) of our proposed model decrease more smoothly than in the other two models with the reduction of training samples. Our proposed model maintains a more stable performance than the other two SOTA methods when gradually decreasing the ratio of the training samples. For HoVer-Net, it can perform nuclei segmentation and classification simultaneously. But associating three branches (semantic segmentation, HoVer and classification) will also increase the data dependency which might lead to unstable results when with fewer training samples. For TripleU, it is effective in pixel-level prediction (DICE) but less effective in instance-level prediction (ARI and PQ). For our proposed model, we achieve the best ARI and stable PQ values but worse DICE scores.

Fig. 9 demonstrates the direct outputs of the results from each model without postprocessing, including overlaid results and the difference maps. Here, all the results are from the same local position for better comparison. The results from top to bottom are generated from the models trained with 20%, 40%, 60%, 80% and 100% of the training data. When increasing the training data, the segmentation results generated by the three models are becoming more precise. Compared with two other competitors, our model demonstrates visually more stable segmentation results.

Table 6
Ablation study of Meta-MTL design. (CPM-17).

Metrics	Ratio	Ours (w/o MTL)	Ours (w/o MTL*)	Ours (w/o MAML)	TripleU (with MAML)	Ours
DICE	100%	0.8601	0.8585	0.8629	0.8791	0.8756
	80%	0.8265	0.8533	0.8598	0.8562	0.8651
	60%	0.8236	0.8446	0.8454	0.8320	0.8483
	40%	0.7445	0.8400	0.8372	0.8154	0.8478
	20%	0.7210	0.7874	0.7529	0.7828	0.8195
AJI	100%	0.6773	0.6677	0.6852	0.6757	0.7081
	80%	0.6232	0.6602	0.6712	0.6751	0.6754
	60%	0.5898	0.6424	0.6473	0.6223	0.6522
	40%	0.4924	0.6368	0.6300	0.6205	0.6413
	20%	0.4691	0.5101	0.5227	0.5511	0.6027
PQ	100%	0.6419	0.6433	0.6572	0.7042	0.6791
	80%	0.5664	0.6326	0.6389	0.6262	0.6427
	60%	0.5340	0.6106	0.6090	0.5844	0.6143
	40%	0.4082	0.5574	0.5917	0.5752	0.6073
	20%	0.3479	0.4486	0.4148	0.4946	0.5396

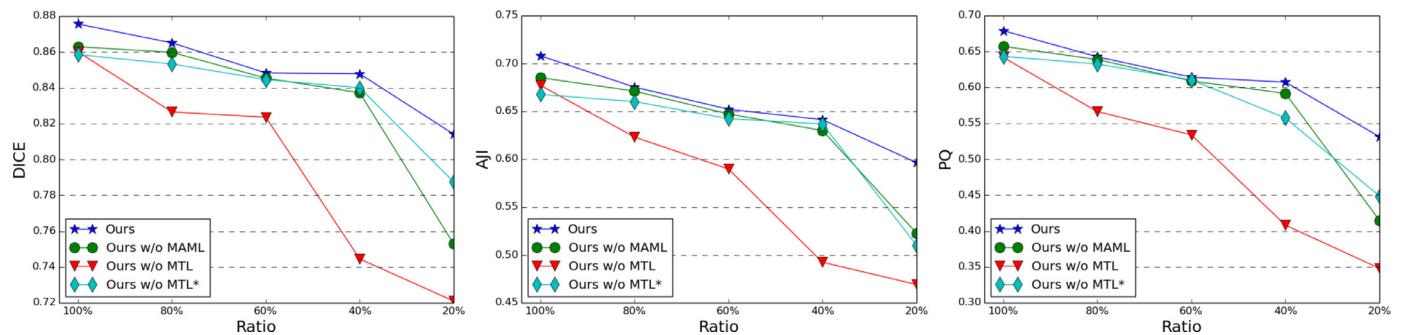


Fig. 6. Ablation study of Meta-MTL design. (CPM-17).

Table 7
Ablation study on FFIB. (CPM-17).

Metrics	Ours w/o FFIB	Ours with FFIB(w/o attention)	Ours
DICE	0.8221	0.8552	0.8756
AJI	0.5970	0.6693	0.7081
PQ	0.5209	0.6284	0.6791

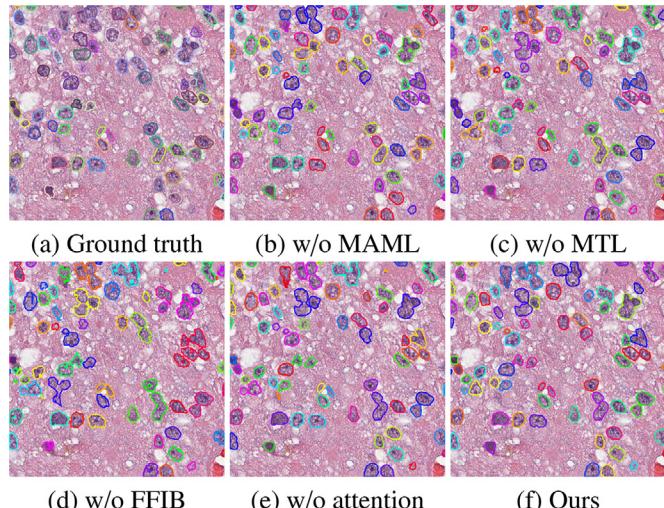


Fig. 7. Qualitative results of two ablation studies. (CPM-17).

5.3. Effectiveness on domain adaptation

Our model can also be easily adapted to a new domain with few labeled samples. We conduct this experiment to evaluate the effectiveness of our model on domain adaptation compared with

the SOTA nuclei segmentation model TripleU (Zhao et al., 2020). Given two datasets MoNuSeg and CPM-17, for each model, we first train it on one dataset using 100% of training data, and then use only a small fraction (5%, 10%, 15%) of the data from the other dataset to fine-tune the model.

Table 4 demonstrates the fine-tuned model from CPM-17 to MoNuSeg ($C \rightarrow M$) and from MoNuSeg to CPM-17 ($M \rightarrow C$). We can find that our proposed model can keep relatively high AJI and PQ scores in three training data ratios, even we only use 5% of training data to fine-tune the model. DICE scores of our model are weaker than TripleU. The same scenario as the previous experiment, it means that our proposed network achieves better instance-level performances but worse pixel-level accuracy, resulting in less accurate nuclei boundary but better nuclei detection rate. This observation actually matches the specialty of our model design that focuses on more general features. And that is the reason why we achieve better instance-level performance in domain adaptation than TripleU with only a few training samples.

5.4. Effectiveness on domain generalization

Due to the difficulties in labeling manual annotations for histopathology images, domain generalization capability is also an inevitable property for histopathology image segmentation. So we set up an experiment to evaluate the model generalization capability for segmenting nuclei to the unseen domains. To test the domain generalization, we use the well-trained model of one dataset with 100% training data and use another dataset (unseen domain) to test it. Since there is no overlapping organ between two datasets, the test data of these two datasets are completely unseen by each other. The results of HoVer-Net and TripleU are generated by their provided implementation and the model. We first use the trained model of CPM-17 to test MoNuSeg, named CPM (MoNuSeg). Then we swap these two datasets.

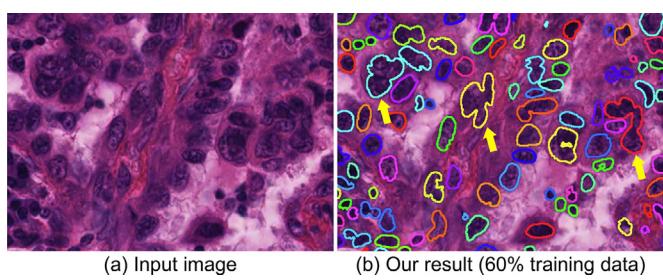


Fig. 8. Limitation of our proposed model. Our model is not able to separate the attached nuclei (yellow arrows) in the direct output segmentation map. It relies on an additional postprocessing step like watershed algorithm to achieve instance-level segmentation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5 demonstrates the comparisons of the domain generalization capability. As can be seen, our proposed model generally outperforms SOTA models in both unseen datasets. Interestingly, all three models achieve much better performance in MoNuSeg (CPM) than CPM (MoNuSeg). Because the data in CPM-17 is less difficult than that in MoNuSeg. This observation can be proved in **Table 1** that all the statistical results in CPM-17 are higher than those in MoNuSeg. Comparing the three models, our model achieves better domain generalization capability than the other two SOTA models.

5.5. Ablation studies

In this subsection, we conduct two ablation studies to evaluate the superiority of our proposed model. In the first ablation study,

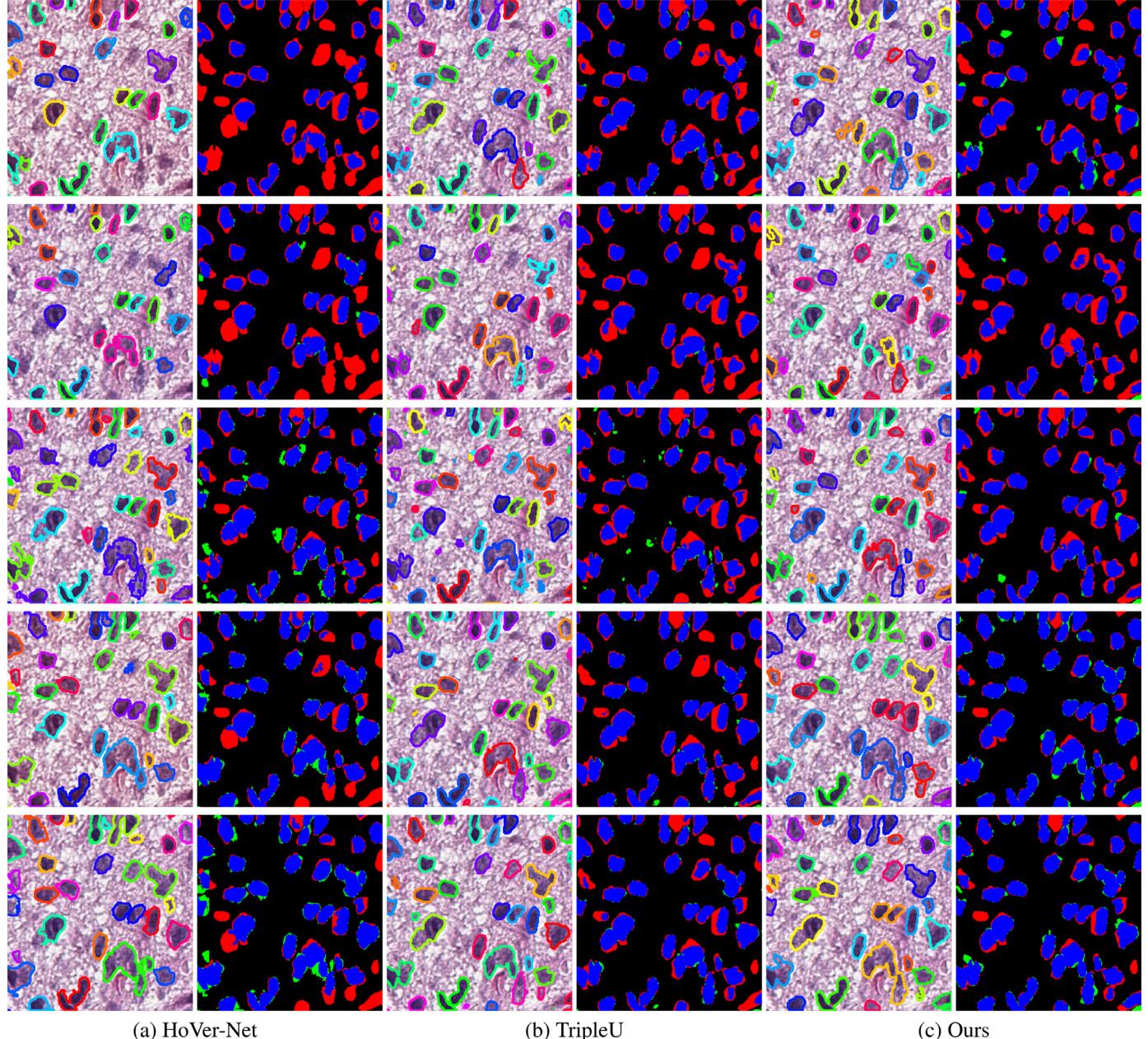


Fig. 9. Qualitative comparisons of data dependency (CPM-17). From top to bottom are the results and difference maps generated by the models trained with 20%, 40%, 60%, 80% and 100% of training data, respectively. In the difference maps, blue, green and red areas indicate true positive segmentation, false positive segmentation and false negative segmentation, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

we validate the effectiveness of Meta-MTL. We compare Meta-MTL model to three variants: (1) Ours (w/o MAML): Multi-task learning without MAML. (2) Ours (w/o MTL): Single-task learning with MAML (the identical segmentation task of the proposed model). (3) Ours (w/o MTL^{*}): Single-task learning with MAML (the segmentation task with a comparable network size to the proposed model). Similar to previous experiments, we train the network using CPM-17 dataset with five deduction rates of the training data (100%, 80%, 60%, 40%, 20%). Table 6 demonstrates the quantitative comparison which is also visualized in Fig. 6. When simply disabled the contour detection task and FFIB block, the performance drastically decreases because losing one task and one feature fusion and interaction branch means losing more than a half network parameters, resulting in reducing the network capability. When increasing the network size to a comparable level to the proposed model, the segmentation performance returns to a reasonable level but is still less effective due to the lack of the auxiliary task. Without MAML but with multi-task learning, the quantitative results decrease rapidly when with only 20% training samples. Besides the ablation studies of our proposed model, we are also curious about whether the existing multi-task baseline model TripleU is suitable for MAML. Therefore, we introduce an additional experiment by training TripleU with MAML using the same configuration with our proposed model in Table 6. The quantitative results show that when with 100% training samples, TripleU outperforms our proposed model. When reducing the training samples, our model achieves more promising results than TripleU with MAML. The experimental results prove the superiority of our Meta-MTL design.

In the second ablation study, we validate the effectiveness of FFIB. We compare our model to two variants: (1) Meta multi-task learning without FFIB. We directly concatenate the features in the contour task to the corresponding levels of the segmentation task. (2) Meta multi-task learning with FFIB without contour attention. Table 7 demonstrates the quantitative results. We conduct this experiment using the complete training data in CPM-17 dataset. Without FFIB, three metrics decrease as expected. When applied FFIB but without the attention module, the model demonstrates better results at pixel-level and instance-level than without FFIB. But without the contour attention module, the segmentation results are with less precise boundaries compared with the complete model, as demonstrated in Fig. 7. Visualization results show that our proposed model can achieve the best performance with all the equipped components.

6. Conclusion

In this paper, we propose a Meta-MTL model for nuclei segmentation. We aim to establish a general model with fewer training samples and less data dependency by associating an optimization-based meta-learning technique and a multi-task learning strategy. The proposed model achieves comparable performances with existing SOTA nuclei segmentation models with the complete training data. The superiority of our model appears when we gradually deduct the training samples. And it can still achieve promising results compared with existing models when fewer training data are involved. In the meanwhile, our proposed model achieves better domain adaptation and generalization capabilities which are essential properties in medical image segmentation. With such properties, our proposed model could be an initial segmentation model for the pathologist-in-the-loop labeling strategy with only very limited start-up pixel-level labels, to save the annotation efforts.

However, there are a few limitations in this work. First, the meta-learning model tends to focus on learning the internal characteristics of the task itself. Therefore, it generally increases the model generalizability and adaptability. It can be more easily adapted to the unseen domain or transferred to another domain

with fewer training samples. However, such properties may lead to slightly lower pixel-level precision than SOTA models because the model is not specifically fit for one specific dataset. There are two potential solutions to overcome this limitation. For example, a self-supervised learning task can be introduced to pre-train the model and learn a better initial feature representation for histopathology images. Another solution is that we can introduce more data augmentations to further enrich the training samples, such as stain augmentation, cutout, cutmix and etc. The second limitation is that the proposed model is not able to directly separate the attached or overlapping nuclei (pointed by the yellow arrows in Fig 8) due to the lack of specific design like HoVer branch in HoVer-Net. However, by leveraging the contour detection task output and the watershed algorithm, we can also achieve promising instance segmentation performance, as shown in the previously presented experiments.

In addition, our proposed approach aiming for few-shot learning may have the potential to be applied to other medical imaging modalities. For example, automatic detection or segmentation of lesions on CT and MRI is still a challenging and tedious task for disease diagnosis, treatment response and etc in clinical practice. Hence, the proposed learning strategy could be used to handle the challenge of few curated data available as a potential solution.

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.

Acknowledgments

This work was supported by the Key-Area Research and Development Program of Guangdong Province (No. 2021B0101420006), the National Science Fund for Distinguished Young Scholars (No. 81925023), the National Key R&D Program of China (No. 2021YFF1201003), Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application (No. 2022B1212010011), the National Natural Science Foundation of China (Nos. 82072090, 81771912, 62172120 and 82071892), the National Science Foundation for Young Scientists of China (Nos. 62102103, 62002082, 82102034 and 82102019), High-level Hospital Construction Project (Nos. DFJHBF202105, DFJH201805), Guangdong Provincial Key Laboratory of Cyber-Physical Systems (No. 2020B1212060069) and National&Local Joint Engineering Research Center of Intelligent Manufacturing Cyber-Physical Systems, Guangxi Science and Technology Project (No. 2019GXNSFFA245014), China Postdoctoral Science Foundation (Nos. 2021M690753 and 2020M682643).

References

- Behl, H. S., Najafi, M., Arnab, A., Torr, P. H., 2018. Meta learning deep visual words for fast video object segmentation. *arXiv preprint arXiv:1812.01397*
- Cao, Z., Zhang, T., Diao, W., Zhang, Y., Lyu, X., Fu, K., Sun, X., 2019. Meta-Seg: a generalized meta-learning framework for multi-class few-shot semantic segmentation. *IEEE Access* 7, 166109–166121.
- Caruana, R., 1997. Multitask learning. *Mach. Learn.* 28 (1), 41–75.
- Chamanzar, A., Nie, Y., 2020. Weakly supervised multi-task learning for cell detection and segmentation. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 513–516.
- 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.
- Cheng, Z., Qu, A., 2020. A fast and accurate algorithm for nuclei instance segmentation in microscopy images. *IEEE Access* 8, 158679–158689.

- Crawshaw, M., 2020. Multi-task learning with deep neural networks: a survey. *arXiv preprint arXiv:2009.09796*
- Finn, C., Abbeel, P., Levine, S., 2017. Model-agnostic meta-learning for fast adaptation of deep networks. *arXiv preprint arXiv:1703.03400*
- Graham, S., Jahanifar, M., Azam, A., Nimir, M., Tsang, Y.-W., Dodd, K., Hero, E., Sahota, H., Tank, A., Benes, K., et al., 2021. Lizard: a large-scale dataset for colonic nuclear instance segmentation and classification. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 684–693.
- 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.
- 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.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.
- Hospedales, T., Antoniou, A., Micaelli, P., Storkey, A., 2020. Meta-learning in neural networks: a survey. *arXiv preprint arXiv:2004.05439*
- Hou, L., Nguyen, V., Kanevsky, A.B., Samaras, D., Kurc, T.M., Zhao, T., Gupta, R.R., Gao, Y., Chen, W., Foran, D., et al., 2019. Sparse autoencoder for unsupervised nucleus detection and representation in histopathology images. *Pattern Recognit.* 86, 188–200.
- 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.
- Javed, S., Mahmood, A., Fraz, M.M., Koohbanani, N.A., Benes, K., Tsang, Y.-W., Hewitt, K., Epstein, D., Snead, D., Rajpoot, N., 2020. Cellular community detection for tissue phenotyping in colorectal cancer histology images. *Med. Image Anal.* 63, 101696.
- Kingma, D. P., Ba, J., 2014. Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980*
- Kirillov, A., He, K., Girshick, R., Rother, C., Dollár, P., 2019. Panoptic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9404–9413.
- Koch, G., Zemel, R., Salakhutdinov, R., et al., 2015. Siamese neural networks for one-shot image recognition. ICML Deep Learning Workshop, Lille, vol. 2.
- 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.
- Lin, R., Sun, L., Liu, Z., Lu, H., Pang, C., Luo, X., 2020. MADNet: a fast and lightweight network for single-image super resolution. *IEEE Trans. Cybern.* 51 (3), 1443–1453.
- 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: IJCAI, pp. 861–868.
- Liu, Q., Dou, Q., Heng, P.-A., 2020. Shape-aware meta-learning for generalizing prostate MRI segmentation to unseen domains. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 475–485.
- Lu, C., Koyuncu, C., Corredor, G., Prasanna, P., Leo, P., Wang, X., Janowczyk, A., Bera, K., Lewis, J., Velcheti, V., et al., 2020. Feature-driven local cell graph (flock): new computational pathology-based descriptors for prognosis of lung cancer and HPV status of oropharyngeal cancers. *Med. Image Anal.* 68, 101903.
- Moeskops, P., Wolterink, J.M., van der Velden, B.H., Gilhuijs, K.G., Leiner, T., Viergever, M.A., Işgum, I., 2016. Deep learning for multi-task medical image segmentation in multiple modalities. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 478–486.
- Naik, D.K., Mammone, R.J., 1992. Meta-neural networks that learn by learning. In: [Proceedings 1992] IJCNN International Joint Conference on Neural Networks, Vol. 1. IEEE, pp. 437–442.
- 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.
- Nichol, A., Schulman, J., 2018. Reptile: a scalable metalearning algorithm. *arXiv preprint arXiv:1803.02999* 2 (3), 4.
- Oda, H., Roth, H.R., Chiba, K., Sokolić, J., Kitasaka, T., Oda, M., Hinoki, A., Uchida, H., Schnabel, J.A., Mori, K., 2018. Besnet: boundary-enhanced segmentation of cells in histopathological images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 228–236.
- Pan, S.J., Yang, Q., 2009. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* 22 (10), 1345–1359.
- 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.
- Ruder, S., 2017. An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*
- Skrede, O.-J., De Raedt, S., Kleppe, A., Hveem, T.S., Liestol, K., Maddison, J., Askautrud, H.A., Pradhan, M., Nesheim, J.A., Albregtsen, F., et al., 2020. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet* 395 (10221), 350–360.
- Snell, J., Swersky, K., Zemel, R. S., 2017. Prototypical networks for few-shot learning. *arXiv preprint arXiv:1703.05175*
- Sun, M., Fang, H., Zhang, G., Yao, Z., Zhou, X., Qing, C., 2019. Efficient nucleus detection in digital pathology images using multi-task and multi-scale instance segmentation network. In: 2019 International Conference on Advanced Mechatronic Systems (ICAMechS). IEEE, pp. 1–6.
- Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H., Hospedales, T.M., 2018. Learning to compare: relation network for few-shot learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1199–1208.
- Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D., et al., 2016. Matching networks for one shot learning. *Adv. Neural Inf. Process. Syst.* 29, 3630–3638.
- 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, 53.
- Wan, Y., Yang, B., Wong, D. F., Zhou, Y., Chao, L. S., Zhang, H., Chen, B., 2020. Self-paced learning for neural machine translation. *arXiv preprint arXiv:2010.04505*
- Xiao, H., Kang, B., Liu, Y., Zhang, M., Feng, J., 2019. Online meta adaptation for fast video object segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 42 (5), 1205–1217.
- Yamashita, R., Long, J., Longacre, T., Peng, L., Berry, G., Martin, B., Higgins, J., Rubin, D.L., Shen, J., 2020. Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *Lancet Oncol.* 22 (1), 132–141.
- Zhang, Y., Yang, Q., 2017. A survey on multi-task learning. *arXiv preprint arXiv:1707.08114*
- Zhao, B., Chen, X., Li, Z., Yu, Z., Yao, S., Yan, L., Wang, Y., Liu, Z., Liang, C., Han, C., 2020. Triple U-Net: hematoxylin-aware nuclei segmentation with progressive dense feature aggregation. *Med. Image Anal.* 65, 101786.
- 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.