CS-Net: A Stain Style Transfer Network for Histology Images with CS-Gate Attention

Zhengze Gong[1,2,3], Xipeng Pan[2,3,4], Chu Han[2,3], Bingjiang Qiu[2,3], Bingchao Zhao[2,3], Yu Liu[2,3], Xinyi Chen[2], Cheng Lu[2,3], Zaiyi Liu[2,3], and Gang Fang[1(✉)]

[1] Institute of Computing Science and Technology, Guangzhou University, Guangzhou 510006, China
gangf@gzhu.edu.cn

[2] Department of Radiology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, China

[3] Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Guangzhou 510080, China

[4] School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin 541004, China

Abstract. Accurate segmentation of epithelium in Hematoxylin-Eosin (HE)-stained oropharyngeal cancer (OPC) pathological images is the premise for quantitative diagnosis. However, the pathological features of OPC are highly heterogeneous in morphology, while the appearance may be similar between epithelium and other tissues (such as stroma), which increases the difficulty of the segmentation task. To solve the above problem, we propose a two-step framework for epithelium tissue segmentation, including a stain-transfer step and an image-processing step. In the first step, we propose a stain-style transfer network called CS-Net, together with our proposed attention module called CS-Gate, to transform HE-stained images into Immunohistochemistry (IHC)-stained images, with the aim of enhancing the contrast among epithelium and other tissues. In the second step, we perform a series of image processing methods on the synthesized IHC-stained images to obtain the final binary mask of the OPC epithelium. The experimental results show that CS-Net can synthesize more stable images with a higher degree of restoration than the GAN-like networks, and the accuracy of the final obtained mask is also better than the current mainstream segmentation networks, reaching 92.48%. An external validation experiment shows that CS-Net has a stronger generalization capability.

Keywords: Stain style transfer · Convolutional neural network · Attention mechanism · Oropharyngeal cancer · Tissue segmentation

1 Introduction

Oropharyngeal cancer (OPC) is the sixth most common cancer in the world, accounting for 10–15% of cancers of the head and neck [1]. Of the 5000 new cases of OPC diagnosed each year in the United States, 85–90% originated from malignancies of the epithelium

[2]. Although OPC is predominantly diagnosed in individuals over 45 years old, studies in Western Europe and the United States have shown an increasing incidence in people under 45 years old over the past 20–30 years [3]. Therefore, the precise assessment of patients with OPC can help to develop individualized treatment. Pathological examination is the gold standard for the diagnosis of OPC, while the cellular composition and spatial structure of epithelium can reflect tumor heterogeneity. Hence, accurate segmentation of cancer epithelium in OPC pathological images is significant for quantitative diagnosis and subsequent image analysis [4].

Hematoxylin-Eosin (HE) staining is a routine staining method in pathology [5]. Earlier epithelium segmentation tasks were performed by manual annotation of HE-stained pathology images by pathologists, but this process is time-consuming and inefficient [6]. In addition, since the epithelium of OPC often has similar morphological and color characteristics with other tissues (such as stroma), it is difficult to distinguish epithelium from stroma only by HE staining [7]. Using immunohistochemistry (IHC) staining for cytokeratin could specifically detect the epithelium in the OPC samples based on the principle of antigen-antibody-specific binding [8]. As shown in Fig. 1, the epithelium is not evident in the HE-stained image in Fig. 1-a, but is clearly stained brown in the section stained for cytokeratin in Fig. 1-b, while other tissues are stained light blue.

a : HE-Stained Image   b : IHC-Stained Image   c : Mask of Epithelium
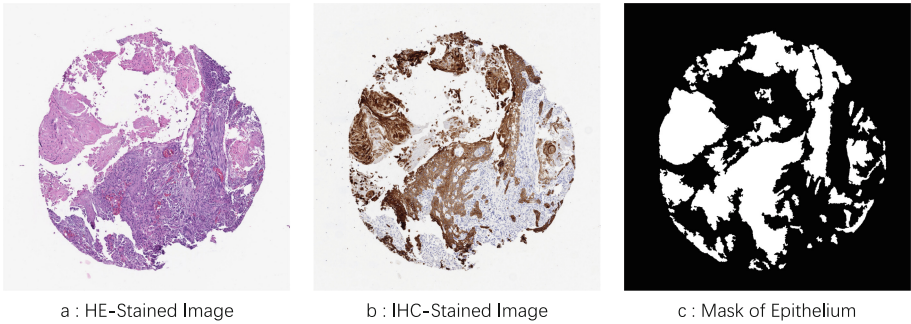
Fig. 1. A sample of the OPC dataset.

However, there are some limitations in IHC staining. Firstly, IHC staining is time-consuming and rather not a routine clinical workflow, which requires additional costs. Secondly, the quality of IHC staining is easily affected by staining kits and conditions, which can lead to under/over-staining [6].

With the development of deep learning and the exponential growth of digitalized medical image data, the use of convolutional neural networks (CNN) for tissue/tumor lesion segmentation of digital pathology images has reached unprecedented popularity [9]. U-Net [10] was a milestone in the semantic segmentation of medical images and has been followed by many excellent segmentation networks [11, 12]. In recent years, many GAN-based methods have also been used to accomplish this task in ways of synthesis. It is worth pointing out that these methods are usually used to segment tumor lesions or necrotic regions, but when it comes to the task of segmenting epithelium regions, these methods show their shortcomings for the following reasons: firstly, the epithelium tissue

and stroma are all stained in a similar shade of purple; secondly, in terms of morphology, the epithelium does not have obvious characteristics distinguished from the stroma and lymphoid tissue, etc. [13].

To address the above challenges, inspired by the manual annotation process, we propose a two-step epithelial tissue segmentation framework and a stain-style transfer network named CS-Net. The details will be described in Sect. 3. The main contributions of this paper are as follows:

1. A new two-step framework for epithelial tissue segmentation. Compared with the single-step methods that segment epithelial tissue masks directly from HE-stained images, our two-step method overcomes the challenge of "indistinct epithelial tissue information in the HE-stained image", and achieves higher accuracy. The first step of the framework is to generate pseudo-IHC-stained images from HE-stained images; the second step is to perform threshold segmentation to obtain the binary mask from the synthetic IHC-stained image.
2. A novel stain-style transfer network and attention module. For the characteristics of pathological images, we innovatively incorporate U-Net, VGG16, and CS-Gate into U-Net to become CS-Net. CS-Gate is an integration of CBAM and attention-gate, aiming to extract features from the channel and spatial aspects simultaneously (a fusion of channel attention, spatial attention, and attention gate, hence the name CS-Gate). Compared with the mainstream segmentation networks, CS-Net can achieve higher accuracy with 92.48%; compared with the GAN-based networks, CS-Net can generate pseudo-IHC-stained images with colors and textures closer to Ground Truth, reaching SSIM with 82.02%.
3. Better generalization capability. The external generalization experiment show that CS-Net has the highest generalization performance, achieving an accuracy of 85.83% on the BCSS dataset without training on it.

2 Related Works

2.1 Tissue Segmentation

Epithelium tissue segmentation is the key step in the diagnostic analysis of digital pathology images, and the accuracy of the segmentation is crucial to the subsequent diagnosis. In recent years, due to the development of deep learning, convolutional neural networks (CNN) have shone in the field of medical image analysis, displaying superior capabilities that are not inferior to manual annotation at all. Among them, U-Net [10] is the pioneer of using CNN for medical image segmentation, which is featured by an encoder and a decoder, and the features located in the same layer are connected by skip-connection. However, as the tasks became more complex, U-Net began to show its shortcomings: it could only extract relatively simple and obvious features, while often predicting too many false-positive (FP) and false-negative (FN) regions when distinguishing epithelial tissue from stroma [14] since these regions are highly similar in morphology.

After U-Net, many excellent segmentation networks emerged, which can be divided into two categories based on their characteristics. One category is based on feature engineering, and the other is based on feature learning. For the first category, Swin-UNet [15] integrated information on the color and cellular arrangement of tissues in

texture features; SegNet [16, 32] constructed a symmetric encoding-decoding structure for capturing similarities in the overall structure and appearance of stained sections; Van [17] proposed an improved DCAN that integrates the original DCAN and the identity mapping method proposed by He [18] into the ResNet architecture. The other category is data-driven, which is highly dependent on a large number of manually annotated datasets and usually requires large data and computational resources, such as [19, 20] which focus on capturing the morphological patterns inherent within the dataset.

Despite the continuous emergence of new models, only a few have been able to significantly improve the segmentation performance of epithelial tissue in pathological images. The reason is that the staining quality of the slides varies from different labs and operators, which causes the high variability of the dataset, hence requiring stronger feature extraction capabilities.

2.2 Attention Mechanism

One way to tackle the challenge of insufficient extraction capacity is the adoption of the attention mechanism. First proposed in 2014 by Bahdanau [2, 33], the attention mechanism can help the models assign different weights to different features within an image while extracting more crucial information. In this way, networks can make more accurate judgments without imposing an extreme burden on computation and storage. In the field of computer vision, attention mechanisms can be divided into three main categories: channel attention, spatial attention, and hybrid domain attention. Channel attention generates and scores the mask of the channel, such as SE-Net [21] and ECA-Net [22]. Spatial attention can be seen as an adaptive spatial region selection mechanism, such as RAM [23], STN [9], and GE-Net [7] are the typical spatial attention modules with different emphases. The hybrid domain attention mechanism evaluates and scores both channel attention and spatial attention, emphasizing meaningful features on both dimensions, like DA-Net [5], Attention gate (AG) [24], and CBAM [25].

In this paper, we integrate AG and the advantages of channel and spatial attention into a new attention module, in order to better adapt to the high variation between different pathological images and the high similarity of features within a pathological image.

2.3 Stain-Style Transfer

Another effective way to improve the accuracy of epithelium segmentation is to transfer HE-stained images into IHC-stained images, for the variability and heterogeneity would be greatly reduced. Subsequently, the epithelium mask can be obtained by image processing methods. Thus our goal can be shifted from a segmentation task to a stain-style transfer task.

Most of the style transfer models are based on Generative Adversarial Networks (GAN) [26], which consist of a Generator (G) and a Discriminator (D). Pix2pix is the pioneer of applying GAN to the field of image conversion, and Isola proposed cycleGAN [27] based on the idea of pix2pix. StainGAN [28] converted Hamamatsu staining to Aperio staining based on cycleGAN and achieved a high degree of visual similarity with the target domain. However, the GAN-based generation method actually has some drawbacks. Firstly, it is difficult to train, like cycleGAN is training two G and two D at the

same time, which requires large running memories and storage. Secondly, the loss curve oscillates substantially because different G and D are constantly passing information to each other, and the predicted output is often unstable.

There are also some methods that try to utilize CNN solely to perform stain-style transformation [29]. For example, Gatys improve the performance of stain-style transformation by adding an artistic style algorithm into CNN[30]; based on Gatys, Chen added the Sobel operator and an improved loss function to the network in order to enhance the edge information of the synthetic images [31, 34]. Compared with GAN-like methods, these single CNN models can improve the transformation accuracy by enhancing the feature extraction ability of the network itself, while effectively avoiding loss oscillation and memory consumption problems. However, there is no single CNN model for the stain-style transfer task of the pathological image so far, hence this paper attempts to propose an improved CNN model specifically for the stain-style transformation.

3   Proposed Method

We propose a two-step epithelium segmentation framework, and the key step is stain-style transfer. The two steps will be described in the following two subsections, and the pipeline of the framework is shown in Fig. 2:
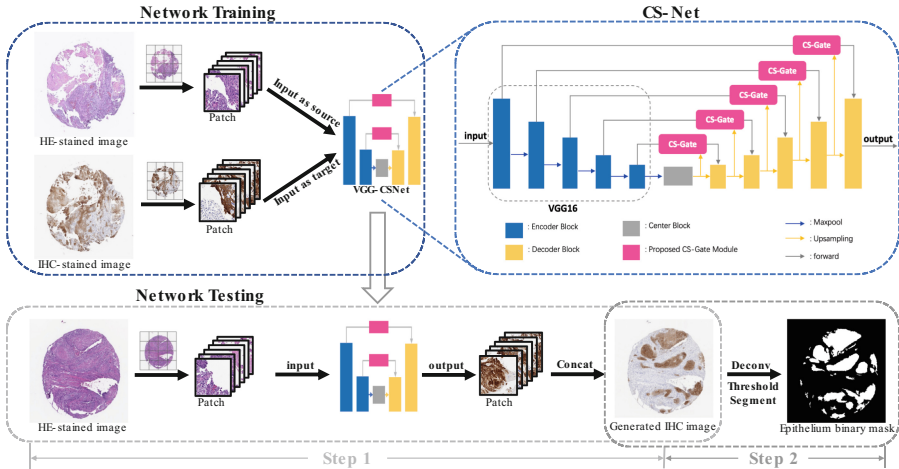
Fig. 2. The Overview of our proposed two-step OPC epithelium tissue segmentation network, and the structure of our proposed CS-Net.
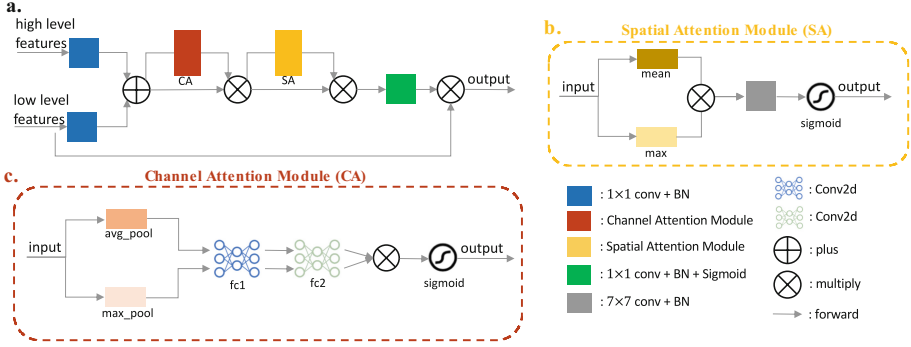
Fig. 3. The details of our proposed CS_Gate module.

3.1 Step 1: Stain-Style Transfer

The first step is stain-style transfer, which is to transfer HE-stained images into IHC-stained images to highlight the epithelium region. We propose a stain-style transfer model named CS-Net (based on U-Net and CS-Gate module, hence the name CS-Net), and the structure is shown in the blue box in Fig. 2. We have made two improvements based on U-Net.

The first improvement is to replace the encoder of the U-Net with VGG16, which possesses stronger feature extraction capabilities. It was observed that the pseudo-IHC stained images generated by U-Net (as a generative network) were blurred and faint with severe FP phenomenon. The reason is that U-Net does not extract deeper, essential features at the encoding stage. This problem can be effectively solved by replacing the encoder with VGG16, which has a more reasonable convolutional layer design and convolutional depth.

The second improvement is the insertion of the attention module. Inspired by the attention gate and CBAM module, we innovatively combine these two modules into one and named CS-Gate. Our approach is to combine the advantages of AG and CBAM together: the low and high-level features are encoded separately, then combine together and extracted by the channel and spatial attention to be encoded into a feature weight. In this way, CS-Gate allows the network to receive multi-dimensional information, as shown in Fig. 3. Moreover, we utilize the smooth L1Loss function to better fit the stain-style transfer task for pathology images, and it can be described as:

$$smooth\ L1\ Loss = \frac{1}{n}\sum_i z_i \tag{1}$$

where:

$$z_i = \begin{cases} 0.5(x_i - y_i)^2, & |x_i - y_i| < 1 \\ |x_i - y_i| - 0.5, & otherwise \end{cases} \tag{2}$$

x_i and y_i represent the predicted value and actual value.

3.2   Step 2: Image Processing

The second step is image processing described in Algorithm 1. The corresponding mask would be obtained from the synthesized pseudo-IHC-stained images vis this algorithm.

Algorithm 1 Image Processing Method

Input: Synthesized IHC-stained Image I_{ihc}
Output: Correspond epithelium binary mask I_{bin}

1: Given an input I_{ihc}, select its blue channel as the grayscale map, $I_{gray} = I_{ihc}[:, :, 3]$

2: Set the value of threshold $T_{max} = 210$, to convert I_{gray} into I_{mask}, $I_{mask} = where(0 < I_{gray} < T_{max})$

3: Fill the regions of small holes with area<5000, $I_{bin} = morphology.remove\_small\_holes(I_{mask}, 5000)$

4: Remove the isolated regions with area<800, $I_{bin} = morphology.remove\_small\_object(I_{bin}, 800)$

5: Use the close operation on the I_{bin} with an elliptical structure of size 5×5 as the structural element, $I_{bin} = cv2.morphologyEx(I_{bin}, MORPH\_CLOSE, 5×5)$

6: Fill the small pore regions of area<5000 again to get the Ground Truth mask, $I_{bin} = morphology.remove\_small\_holes(I_{mask}, 5000)$

7: Return I_{bin}

It is worth pointing out that the values of 5000 and 800 in steps 3 and 4 are empirical thresholds obtained from experimental observations on a portion of the dataset. It is observed that the small holes with an area < 5000 inside the tissue and the spots with an area > 800 outside the tissue are generally noisy rather than mistakes due to segmentation, so the epithelial tissue mask obtained by Algorithm 1 is closer to the real label.

4   Experiments

4.1   Dataset Description

The dataset used in this paper is the tissue microarray (TMA) pathology images from the OPC of Guangdong Provincial People's Hospital, with 208 samples of 2048×2048 sizes, each with a HE-stained image, a paired IHC-stained image, and an epithelial tissue mask (as shown in Fig. 1). All samples are divided in the ratio of train set: validation set: test set $= 6:2:2$. We use a sliding window of 256×256 sizes from left to right and then from top to bottom with overlap $= 50\%$ to take the image patches. All the predicted patches are stitched back together in the order of segmentation, and the maximum value of the overlapping parts is taken as the final pixel value of the region.

In addition, we use BCSS (Breast Cancer Cases) Dataset as external verification to prove the generalization of the proposed method, which contains 151 HE-stained whole-slide images (WSIs), corresponding to 151 tissue classification labels (Fig. 4).
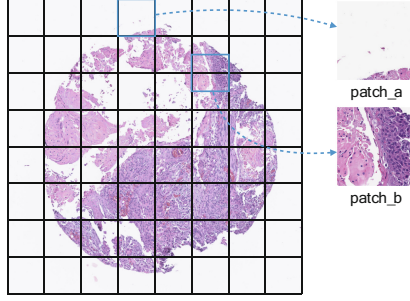
Fig. 4. The method of dividing a 2048 × 2048 image into several 256 × 256 patches.

4.2 Evaluation Metrics

In this paper, Structural Similarity (SSIM) is used to evaluate the similarity between the generated IHC-stained images and Ground Truth (GT):

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{3}$$

where x stands for the generated IHC-stained image; y stands for the GT; μ_x and μ_y represent the mean value of x and y; σ_x and σ_y represent the standard deviation of x and y; σ_{xy} represents the covariance of x and y; c_1 and c_2 are constants respectively.

In addition, five metrics (Accuracy, Precision, Specificity, Recall, and F1-score) are used to evaluate the results of the final epithelium binary masks obtained via the proposed framework compared with the common segmentation networks:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Specificity = \frac{TN}{FP + TN} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$F_1 - score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{8}$$

The meanings of TP, TN, FP, and FN are shown in the following Table 1:

Table 1. Confusion Matrix.

| Predict\Actual	Positive	Negative
Positive	TP	FP
Negative	FN	TN

4.3 Implementation Details

The device is Dell Precision 3640, with Intel Core i9-10900K CPU@3.70 GHz × 20 and NVIDIA GeForce RTX 3090. Keep the hyper-parameters the same for all experiments, Loss = L1Loss, Batch_size = 16, epoch = 20, Learning_rate = 0.0002, and the Learning_rate decay strategy is to halve every two epochs. To avoid the influence of experimental errors, the results of each network are the average values of training three times ± standard deviation.

 Firstly, we conduct experiments on U-Net as the baseline and then combine U-Net with attention gate, CBAM, and our CS-Gate module respectively. Secondly, we replace the encoder of U-Net with VGG16 (which turns into VGG16UNet), then plug the above three attention modules into it. Finally, we compare our method with the two most commonly used generative networks (cycleGAN and pix2pix). The results are shown in Table 2:

Table 2. Comparison of other networks and our CS-Net.

| Models	Accuracy	Precision	Recall	Specificity	F1-score	SSIM
U-Net (U1)	0.8486	0.6618	0.8148	0.8383	0.7303	0.7559
AG + U1	0.8622	0.6506	0.8202	0.8569	0.7256	0.7753
CBAM + U1	0.8649	0.6573	0.8307	0.8596	0.7338	0.7712
CS-Gate + U1	0.8371	0.6367	0.8360	0.8511	0.7228	0.7690
VGG16UNet (U2)	0.8833	0.6987	0.8280	0.8982	0.7578	0.7702
AG + U2	0.9082	0.7742	0.8543	0.9054	0.8122	0.8032
CBAM + U2	0.9018	**0.7933**	0.8443	0.9001	0.8180	0.8092
pix2pix	0.7731	0.6888	0.6769	0.8186	0.6827	0.6746
cycleGAN	0.6851	0.6072	0.6851	0.7233	0.6438	0.7397
Ours	**0.9248**	0.7855	**0.8845**	**0.9255**	**0.8320**	**0.8202**
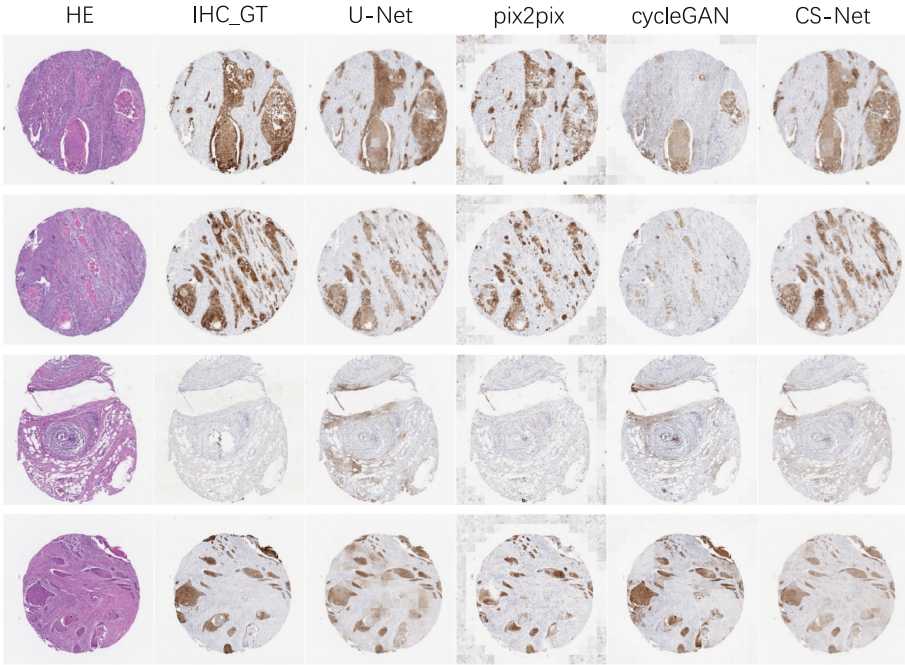
Fig. 5. Pseudo-IHC-stained images generated by different models.

It can be seen that for U-Net, whether it is plugged with AG, CBAM, or CS-Gate, the results have not improved even slightly decreased. However, after replacing the encoder with VGG16 to become VGG16UNet, the results have been greatly improved (from 84.86% to 88.33%). Then three attention modules were plugged into it subsequently, the accuracy and SSIM were improved in different aspects. Among them, our proposed CS-Gate improves the most (from 88.33% to 92.48%). From these results, it is reasonable to deduce that the original encoder of U-Net is not capable of extracting effective features enough from HE-stained images. But once replacing the encoder with VGG16, which has stronger feature extraction capabilities, the accuracy has been significantly improved, and the three attention modules finally work when they are added into VGG16UNet.

4.4 Qualitative Analysis Results

Figure 5 shows the pseudo-IHC-stained images generated by different generative networks. Observing the first and the fourth row, the checkerboard artifact is greatly reduced by CS-Net; in the second row, CS-Net generates the most similar IHC-stained image, while restoring color and texture features as much as possible; in the third row, CS-Net can reduce the FP phenomenon most. It can be seen that CS-Net has synthesized the most similar image to the GT.

Figure 6 shows the mask obtained from the pseudo-IHC-stained images via Algorithm 1. It can be seen that CS-Net can generate the most accurate mask. U-Net has the second-best performance after CS-Net but is still showing a high false positive in the
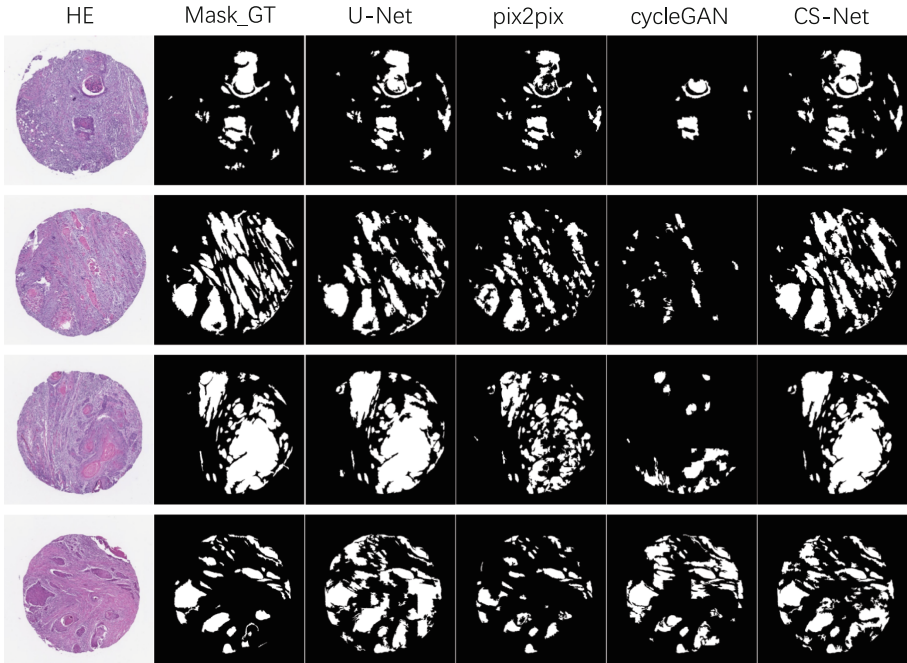
Fig. 6. The epithelium binary masks obtained from the pseudo-IHC-stained images via Algorithm 1.

fourth row. Pix2pix and cycleGAN have the worst performance, showing not only a high chessboard artifact but also high FP and FN phenomena in all samples.

Figure 7 illustrates that CS-Net has higher stability and lower loss than GAN-based generative networks (such as cycleGAN and pix2pix). Table 1 also shows that CS-Net can synthesize more similar images (about 10% higher in SSIM).
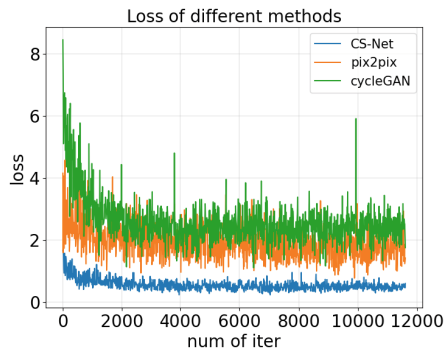
Fig. 7. The comparison of loss curves of 3 methods.

4.5 Generalization Performance

Figure 8 shows the obtained epithelium binary masks of BCSS patches predicted by different methods, aiming to validate their generalization ability. It can be seen that U-Net predicts accurately in some cases, but does not perform well overall. Pix2pix and cycleGAN both exhibit poor generalization ability. CS-Net shows the best generalization capability and robustness among these methods.
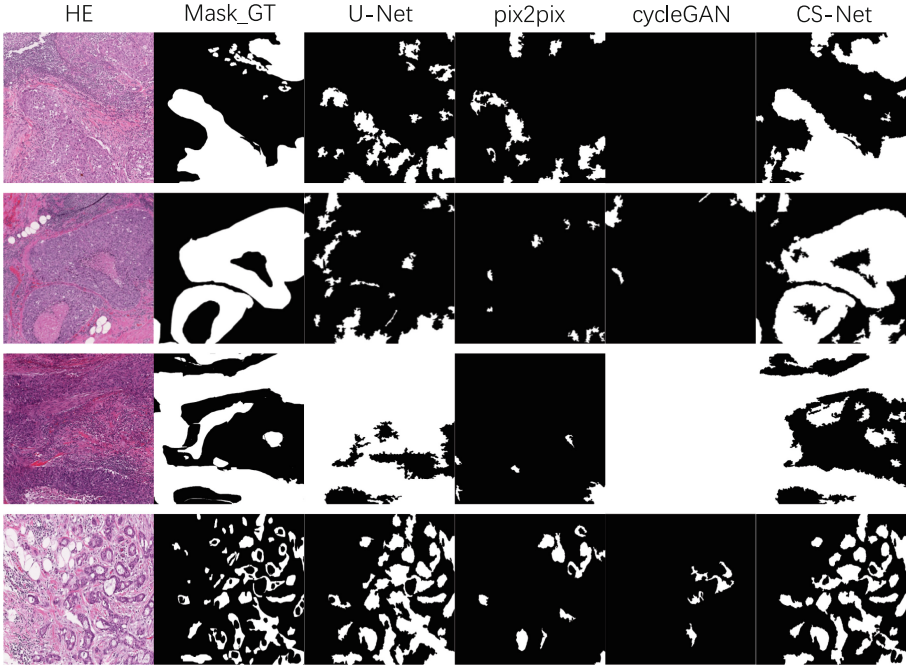

Fig. 8. The epithelium binary masks of BCSS images predicted by different methods.

5 Conclusion

This paper proposes a stain-style transfer network named CS-Net, which has strong feature extraction capability and can synthesize more stable and higher-quality images with better color and texture details. CS-Net outperforms the UNet-based network when performing segmentation task, and outperforms some GAN-based networks when performing style-transfer task, with an outstanding accuracy of 92.48%. CS-Net also shows a higher generalization ability than the U-Net-based and GAN-based networks.

CS-Net is the key step of our presented two-step framework to perform epithelium tissue region segmentation, which is to first generate a psuedo-IHC-stained image and then perform image processing to obtain the mask. Compared with the one-step methods which solely uses HE-stained images to do the segmentation task, our method shows higher superiority.

However, there are also some shortcomings in this work, and more researches could be done in the future. Firstly, we found that the more epochs are trained, the better the results would be, so better results would be obtained theoretically if time allowed. Secondly, a larger batch size would yield better results, but also take up more running memory. If the computing power is increased, such as calling multiple GPUs or providing more running memory to run a larger batch size.

References

1. Amgad, M., et al.: Structured crowdsourcing enables convolutional segmentation of histology images. Bioinformatics **35**(18), 3461–3467 (2019)
2. Bahdanau, D., Kyunghyun, C., Yoshua, B.: Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473 (2014)
3. Chang, H., Alexander, B., Paul, S., Bahram, P.: Classification of tumor histology via morphometric context. In: Proceedings of the IEEE conference on computer vision and pattern Recognition, pp. 2203–2210 (2013)
4. Chen, H., Qi, X., Yu, L., Dou, Q., Qin, J., Heng, P.A.: DCAN: deep contour-aware networks for object instance segmentation from histology images. Med. Image Anal. **36**, 135–146 (2017)
5. Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., Lu, H.: Dual attention network for scene segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3146–3154 (2019)
6. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
7. Hu, J., Shen, L., Albanie, S., Sun, G., Vedaldi, A.: Gather-excite: Exploiting feature context in convolutional neural networks. In: Advances in neural information processing systems, vol. 31 (2018)
8. Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.1125–1134 (2017)
9. Jaderberg, M., Karen, S., Andrew, Z., et al.: Spatial transformer networks. In: Advances in Neural Information Processing Systems, vol. 28 (2015)
10. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III, pp. 234–241. Springer International Publishing, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
11. Levy, J.J., et al.: A large-scale internal validation study of unsupervised virtual trichrome staining technologies on nonalcoholic steatohepatitis liver biopsies. Mod. Pathol. **34**(4), 808–822 (2021)
12. Xu, R., et al.: Histopathological tissue segmentation of lung cancer with bilinear CNN and soft attention. BioMed. Res. Int. **2022**, 1–10 (2022). https://doi.org/10.1155/2022/7966553
13. Salehi, P., Chalechale, A.: Pix2pix-based stain-to-stain translation: a solution for robust stain normalization in histopathology images analysis. In: 2020 International Conference on Machine Vision and Image Processing (MVIP), pp. 1–7. IEEE (2020)
14. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
15. Cao, H., et al.: Swin-unet: Unet-like pure transformer for medical image segmentation. arXiv preprint arXiv:2105.05537 (2021)

16. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**(12), 2481–2495 (2017)

17. Van Eycke, Y.R., Balsat, C., Verset, L., Debeir, O., Salmon, I., Decaestecker, C.: Segmentation of glandular epithelium in colorectal tumours to automatically compartmentalise IHC biomarker quantification: A deep learning approach. Med. Image Anal. **49**, 35–45 (2018)

18. He, K., Zhang, X., Ren, S., Sun, J.: Identity mappings in deep residual networks. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9908, pp. 630–645. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46493-0_38

19. Sirinukunwattana, K., Ahmed, S.E., Raza, Y.-W.T., Snead, D., Cree, I., Rajpoot, N.: A spatially constrained deep learning framework for detection of epithelial tumor nuclei in cancer histology images. In: Guorong, W., Coupé, P., Zhan, Y., Munsell, B., Rueckert, D. (eds.) Patch-MI 2015. LNCS, vol. 9467, pp. 154–162. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-28194-0_19

20. Zhong, C., Han, J., Borowsky, A., Parvin, B., Wang, Y., Chang, H.: When machine vision meets histology: a comparative evaluation of model architecture for classification of histology sections. Med. Image Anal. **35**, 530–543 (2017)

21. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7132–7141 (2018)

22. Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu. Q.: ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks (2019)

23. Mnih, V., Heess, N., Graves, A., et al.: Recurrent models of visual attention. In: Advances in neural information processing systems, vol. 27 (2014)

24. Oktay, O., Jo, S., Loic, L.F., Matthew, L., Mattias, H., Kazunari, M., Kensaku, M, et al.: Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999 (2018)

25. Woo, S., Park, J., Lee, J.-Y., Kweon, I.S.: Cbam: Convolutional block attention module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 3–19. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_1

26. Goodfellow, I., et al.: 2020. Generative adversarial networks. *Communications of the ACM* 63. ACM New York, NY, USA: 139–144

27. Mirza, M., Osindero, S.: Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784 (2014)

28. Shaban, M.T., Baur, C., Navab, N., Albarqouni, S.: Staingan: stain style transfer for digital histological images. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (Isbi 2019), pp. 953–956. IEEE (2019)

29. Li, X., Liu, S., Kautz, J., Yang, M.H.: Learning Linear Transformations for Fast Arbitrary Style Transfer (2018)

30. Gatys, L.A., Alexander, S.E., Bethge, M.: Image style transfer using convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2414–2423 (2016)

31. 陈志鹏, 郑文秀, and 黄琼丹. 基于Sobel滤波器的图像风格转换算法 (2021)

32. Zheng, Q., Zhu, J., Tang, H., et al.: Generalized label enhancement with sample correlations. IEEE Trans. Knowl. Data Eng. **35**(1), 482–495 (2021)

33. Lu, H., Li, Y., Chen, M., et al.: Brain intelligence: go beyond artificial intelligence. Mobile Netw. Appl. **23**, 368–375 (2018)

34. Li, H., Pun, C.M., Xu, F., et al.: A hybrid feature selection algorithm based on a discrete artificial bee colony for Parkinson's diagnosis. ACM Trans. Internet Technol. **21**(3), 1–22 (2021)