

Unpaired Multi-Domain Stain Transfer for Kidney Histopathological Images

Yiyang Lin^{1*}, Bowei Zeng^{1*}, Yifeng Wang^{2*},
Yang Chen¹, Zijie Fang¹, Jian Zhang³, Xiangyang Ji⁴, Haoqian Wang¹, Yongbing Zhang^{2†}

¹Tsinghua Shenzhen International Graduate School, Tsinghua University

²Harbin Institute of Technology (Shenzhen)

³School of Electronic and Computer Engineering, Peking University

⁴Department of Automation, Tsinghua University

{lly20, zbw20}@mails.tsinghua.edu.com, wangyifeng@stu.hit.edu.cn, ybzhang08@hit.edu.cn

Abstract

As an essential step in the pathological diagnosis, histochemical staining can show specific tissue structure information and, consequently, assist pathologists in making accurate diagnoses. Clinical kidney histopathological analyses usually employ more than one type of staining: H&E, MAS, PAS, PASM, etc. However, due to the interference of colors among multiple stains, it is not easy to perform multiple staining simultaneously on one biological tissue. To address this problem, we propose a network based on unpaired training data to virtually generate multiple types of staining from one staining. Our method can preserve the content of input images while transferring them to multiple target styles accurately. To efficiently control the direction of stain transfer, we propose a style guided normalization (SGN). Furthermore, a multiple style encoding (MSE) is devised to represent the relationship among different staining styles dynamically. An improved one-hot label is also proposed to enhance the generalization ability and extendibility of our method. Vast experiments have demonstrated that our model can achieve superior performance on a tiny dataset. The results exhibit not only good performance but also great visualization and interpretability. Especially, our method also achieves satisfactory results over cross-tissue, cross-staining as well as cross-task. We believe that our method will significantly influence clinical stain transfer and reduce the workload greatly for pathologists. Our code and Supplementary materials are available at <https://github.com/linyiyang98/UMDST>.

Introduction

As a vital part of the histopathological analysis, which is the gold standard for diagnosis, histochemical staining can make the different components of the tissue appear in different colors, thereby facilitating pathologists to observe the tissue. The most common type of histochemical staining is hematoxylin and eosin (H&E) staining in routine pathological examinations. H&E can highlight the morphological characteristics of tissue, enabling pathologists to make preliminary diagnoses for patients. However, H&E stained

tissue cannot show all the information that pathologists need to diagnose in many diseases. Therefore, after viewing the H&E stained sections of the patient, pathologists may need other types of staining to obtain more information. For example, in kidney tissue examination, in addition to H&E staining, other types of staining like MAS staining, PAS staining, and PASM staining are also commonly used. MAS staining is mainly used to distinguish collagen fibers and muscle fibers inside the kidney tissue. PAS staining is mainly used to observe glomerular basement membrane, tubular basement membrane, and mesangial matrix (hereinafter called GTM). PASM staining and PAS staining show the same components in kidney tissue, while PASM staining can display the details more clearly (de Haan et al. 2021; Lo et al. 2021). These non-H&E stains are usually called “special stain”. They can provide additional tissue diagnostic information from different dimensions and are of great significance for the pathological diagnoses of kidney diseases.

However, compared with H&E staining, special staining requires more time and labor costs. In addition, if the pathologists need multiple types of histological staining on the patients’ tissue, it is necessary to sample the tissue numerous times for different staining. These will undoubtedly increase the cost of pathological diagnoses and increase the waiting time of patients, making the current popularizing rate of pathological diagnosis low (de Haan et al. 2021).

Recently, researchers are exploring the correlation between H&E stained images and images stained by special stains. Rivenson et al. (Rivenson et al. 2019) proposed a model to virtually generate PAS, MAS, and other types of special staining from H&E stained images, showing that there is a correlation between multiple types of stained images. However, this method requires pixel-level paired data, which is hard to obtain in clinical practice. Lo et al. (Lo et al. 2021) proposed a model to virtually generate PAS, MAS, and other types of special staining from H&E stained images based on unpaired data (Zhu et al. 2017). But in this method, a model is required for every two types of staining, so $k(k - 1)$ models are needed if we have k types of staining, which greatly wastes computing resources. Therefore, it would be much desirable to devise a model trained over

*These authors contributed equally.

†Corresponding Author: Yongbing Zhang.

Copyright © 2022, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

unpaired data to generate multiple types of virtually stained images based on any type of stained image. This process is called multi-domain stain transfer. A lot of experience has shown that, compared with the traditional histochemical staining process, the time cost of multi-domain stain transfer will be reduced by tens or even hundreds of times, which can greatly improve the diagnostic efficiency. Therefore, multi-domain stain transfer through computer-aided technology is a feasible and novel solution to the existing problems.

In recent years, Generative Adversarial Network (GAN) has developed rapidly, providing a good solution for image-to-image translation (Isola et al. 2017; Liu, Breuel, and Kautz 2017). The purpose of image-to-image translation is to transfer images from one style to another one to highlight different information in different styles. In the field of biology, image-to-image translation technology can make pathological structures that are difficult to detect in the original style clearly visible, so its application is extensive, such as stain normalization, virtual staining (Rivenson et al. 2019), graded image generation, artifacts reduction/correction, anomaly detection in brain MRI.

As a novel image-to-image translation method, UGATIT (Kim et al. 2020) can transfer images with high accuracy based on unpaired images. However, UGATIT could only be applied between two domains. On the contrary, StarGAN (Choi et al. 2018) can generate images of multiple types using a single network. Consequently, StarGAN provides a good idea for multi-domain stain transfer, but sometimes poor results may be achieved. Moreover, the task extensibility of StarGAN may not be so well, and it cannot be utilized in the expanded applications like mixed staining and staining contrast tuning.

In this paper, to address these problems, we propose a method using unpaired data to accurately generate multiple types of virtually stained images via a single network. Our main contributions are as follows:

- We propose a new method of multi-domain stain transfer, which can use a single network to generate multiple types of virtually stained images;
- We propose a style guided normalization (SGN) method, which can control the direction of stain transfer and significantly improve the quality of the virtually stained images;
- We propose a multiple style encoding (MSE) method, which can dynamically represent the relationship among different types of staining style and strengthen the network's cross-task generalization ability by the improved one-hot label;
- Compared with traditional methods, the proposed method has strong cross-tissue and cross-staining generalization ability. Our method has a powerful ability to extract staining invariant features from the images.

Related Work

Virtual Staining in Histopathological Analysis

In recent years, image-to-image translation has developed rapidly in computer vision, and some work has been successfully applied in the field of virtual staining. Rivenson

et al. (Rivenson et al. 2019) and Zhang et al. (Zhang et al. 2020) proposed a GAN-based virtual staining method from unstained images to H&E stained images and extended it to MAS staining, PAS staining, etc. But this method requires pixel-level paired data, which is difficult to obtain in histopathology. Using unpaired data, Li et al. (Li et al. 2021) proposed a virtual staining method from unstained images to H&E stained images, and Liu et al. (Liu et al. 2021) completed virtual staining from H&E stained images to Ki67 stained images (Xu et al. 2020). In addition, Lo et al. (Lo et al. 2021) accomplished virtual staining from H&E images to PAS images and extended it to MAS, PASM, etc. But these methods cannot achieve the goal of generating multiple types of stained images using a single network. This paper proposes a multi-domain stain transfer method based on a single network, which does not require paired data.

Normalization in Image-to-Image Translation

Recent research on image-to-image translation shows that instance normalization (IN) (Ulyanov, Vedaldi, and Lempitsky 2016) can directly normalize the feature statistics of an image to remove its original style, so it is commonly applied in image-to-image translation tasks. In recent studies, based on IN, adaptive instance normalization (AdaIN) (Huang et al. 2018), conditional instance normalization (CIN) (Dumoulin, Shlens, and Kudlur 2017), and batch-instance normalization (BIN) (Nam and Kim 2018) have achieved better results than IN. In addition to IN, layer normalization (LN) (Ba, Kiros, and Hinton 2016) is also proposed to pay more attention to the features of the target domain in image-to-image translation tasks. Combining the advantages of IN and LN, Kim et al. (Kim et al. 2020) proposed adaptive layer instance normalization (AdaLIN) to select the appropriate ratio between IN and LN, making the task of image style transfer more thorough. However, AdaLIN cannot achieve the transfer among multiple styles. To realize the transfer among multiple styles, Pizzati et al. (Pizzati, Cerri, and de Charette 2021) proposed FIN, in which the direction of transfer is controlled by labels artificially added. But this method cannot dynamically represent the relationship among different styles. In our work, we devise a new method of style guided normalization (SGN), in which labels are learned through the network so that the labels can accomplish the task of dynamical representation.

Method

From the perspective of staining style transfer, a tissue image is composed of its content (principal structure and morphological content) and style (the types of staining, including H&E staining, PAS staining). The essence of stain transfer is to preserve the content of the original image while transferring its style to the target one.

In this work, we propose a GAN-based multi-domain stain transfer method. We first introduce the network structure of the generator and discriminator. In the generator, we devise a style guided normalization (SGN) method and a multiple style encoding (MSE) method based on an improved one-hot label. In the discriminator, we introduce two

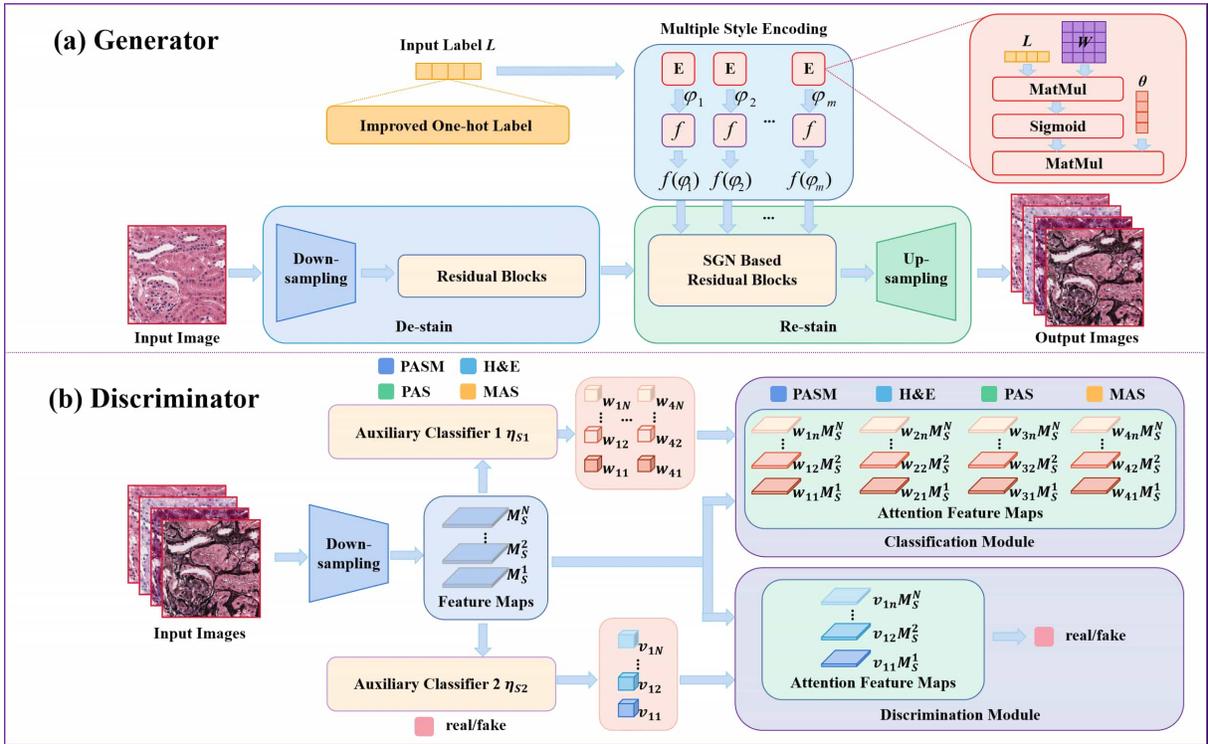


Figure 1: The network structure of our method. (a) The network structure of generator; (b) The network structure of discriminator. In the Multiple Style Encoding (MSE), m represents the times of MSE operation, and $f(\varphi_i)(i = 1, 2, \dots, m)$ contains the corresponding affine transformation factors $f_\gamma(\varphi_i)$ and $f_\beta(\varphi_i)$.

auxiliary classifiers, which distinguish the real/fake as well as the differences among various types of staining. Next, the loss function is provided. Finally, we describe our training process.

Structure of Generator

The network structure of the generator is shown in Fig. 1(a), which is composed of De-stain, Re-stain, and Multiple Style Encoding (MSE). De-stain consists of Down-sampling and Residual Blocks, whose goal is to remove the style of the input image and retain its content. Re-stain includes Up-sampling and Residual Blocks, whose purpose is to generate images of the target staining style according to the content of the input image and the output $f(\varphi)$ of MSE. The function of MSE is to represent the relationship among different types of staining style according to the input label L .

In addition, in the Residual Blocks of De-stain, we employ the normalization method of IN; and in the Residual Blocks of the Re-stain, we utilize the normalization method of SGN.

Style Guided Normalization (SGN) In the Re-stain (see Fig. 1(a)), we need to balance the content of the input image and the style of the target staining domain, which directly affects the quality of the virtually stained image. To achieve better balance, we propose a style guided normalization, which is formulated as:

$$SGN(\varphi) = f_\gamma(\varphi) \cdot (\rho \cdot \hat{a}_I + (1 - \rho) \cdot \hat{a}_L) + f_\beta(\varphi), \quad (1)$$

where \hat{a}_I and \hat{a}_L represent the channel-wise and layer-wise normalizations of the input feature maps, ρ determines the ratio between \hat{a}_I and \hat{a}_L . Here, the content of the input image is represented by $\rho \cdot \hat{a}_I + (1 - \rho) \cdot \hat{a}_L$. And the style of the target staining domain is represented by affine transformation factors $f_\gamma(\varphi)$ and $f_\beta(\varphi)$.

Actually, LIN (Kim et al. 2020) also adopted the similar idea. However, $f_\gamma(\varphi)$ and $f_\beta(\varphi)$ are fixed values in LIN, thus LIN cannot be adapted to multiple staining domains. In contrast, we devise a MSE module to overcome this shortcoming in the following.

Multiple Style Encoding (MSE) To better represent the style of different staining domains, we define a style guiding factor (SGF) φ , which is represented as:

$$\varphi = \text{sigmoid}(LW)\theta, \quad (2)$$

where W and θ are the learnable style coding matrix and learnable style coding vector, respectively. Here, W is a matrix of $n \times n$, and θ is a $n \times 1$ vector, with n being the number of the staining domains. L represents the label of the target staining domain. Once φ is obtained, the scaling factor $f_\gamma(\varphi)$ and $f_\beta(\varphi)$ in Eq.(1) can be computed as:

$$\begin{aligned} f_\gamma(\varphi) &= \varphi \cdot \gamma_1 + \gamma_2 \\ f_\beta(\varphi) &= \varphi \cdot \beta_1 + \beta_2, \end{aligned} \quad (3)$$

where $\gamma_1, \gamma_2, \beta_1$, and β_2 are learnable parameters.

Here, different L determines different staining domains. Obviously, one straightforward way to represent L is the traditional one-hot label, by which we can “take out” the style coding matrix W row by row. However, the network’s task extendibility would be greatly limited by one-hot label when there exist differences between training labels and testing labels. To address this problem, we propose an improved one-hot label, which is defined as:

$$L = [\delta, \delta, 1, \delta, \dots, \delta] \in \mathbb{R}^n \quad (4)$$

where δ is a small value, and it is set to 0.05 in our experiment. It should be noted δ can be arbitrarily set according to actual needs.

In this paper, n is set to 4. Here, we use $[1, \delta, \delta, \delta]$ for PASM staining, $[\delta, 1, \delta, \delta]$ for H&E staining, $[\delta, \delta, 1, \delta]$ for PAS staining, and $[\delta, \delta, \delta, 1]$ for MAS staining (In fact, the relationship between different labels and different types of staining can be changed).

Based on the improved one-hot label, our network can adapt to the testing labels which are not included during training. For example, our network can adapt to the labels where more than one positions are not 0, and the labels where the sum of each position is not 1. Therefore, the task extendibility of our network can be significantly enhanced, which can be validated by the following two expanded applications in the clinic:

- Mixed staining (Zhang et al. 2020): In practical applications, pathologists expect to observe multiple types of staining in the same image for commonly expressed information. However, the majority of available datasets contain images with only one type of staining. To solve this problem, for example, we can input the label like $[1 - \xi, 0, 0, \xi]$ (ξ determines the proportion of the two types of staining) to our network to obtain the results of the co-staining of PASM and MAS;
- Staining contrast tuning: Some of the images in the training set are stained insufficiently, and the staining contrast of these images is low. If such data is utilized for training, the network will generate images of low staining contrast during testing, causing inconvenience to the pathologists’ diagnoses. To solve this problem, for example, our label can be $[0, 0, 0, 1 + \tau]$ (the larger the τ is in a certain range, the larger the staining contrast will be) for tuning the staining contrast among different colors in the generated results of MAS staining.

Structure of Discriminator

The network structure of the discriminator is shown in Fig. 1(b), which includes Down-sampling, Auxiliary Classifiers, Classification Module, and Discrimination Module. Down-sampling extracts the critical features of the input image, and then the extracted features are fed into the Auxiliary Classifiers. Auxiliary Classifiers, whose goals are similar to Classification Module and Discrimination Module, respectively, are designed to assist the corresponding module in paying more attention to significant positions in the input images. Thus, the quality of the virtually stained images can be improved, and strong interpretability can be achieved. Classification Module distinguishes the staining types of the input

Algorithm 1: The training process of our method.

- 1: **for** number of training iterations **do**
 - 2: Select an original image x , and the corresponding original label is L_o .
 - 3: Select a target label L_t .
 - 4: **Stage 1. Updating the discriminator**
 - 5: Obtain the target image
 - a. Input x to De-stain and obtain $G_{De}(x)$;
 - b. Input L_t to MSE and obtain the affine transformation factors $f(\varphi)$ of SGN by Eq. (2), (3);
 - c. Feed $G_{De}(x)$ and $f(\varphi)$ to Re-stain to obtain the target image $G(x, L_t) = G_{Re}(G_{De}(x), f(\varphi))$.
 - 6: Calculate l_D by Eq. (6) and update the discriminator based on back propagation.
 - 7: **Stage 2. Updating the generator**
 - 8: Obtain the reconstructed image $G(G(x, L_t), L_o)$ by the same way as **Stage 1. 5**.
 - 9: Obtain the identity image $G(x, L_o)$ by the same way as **Stage 1. 5**.
 - 10: Calculate l_G by Eq. (5) and update the generator based on back propagation.
 - 11: **end for**
-

images, and Discrimination Module judges whether the input images are real or fake.

Moreover, to improve the accuracy of our network, we use a combination of global discriminator and local discriminator (Kim et al. 2020).

Loss Function

There are five types of loss term when updating the generator: adversarial loss l_{adv}^G , classification loss l_{cls}^G , auxiliary loss $l_{\eta adv}^G, l_{\eta cls}^G$, cycle loss l_{cyc} , and identity loss l_{idt} . And there are three types of loss term when updating the discriminator: adversarial loss l_{adv}^D (Mao et al. 2017), classification loss l_{cls}^D , and auxiliary loss $l_{\eta adv}^D, l_{\eta cls}^D$. The two total loss formulations are as follows:

$$l_G = \lambda_1 \times (l_{adv}^G + l_{cls}^G + l_{cyc} + l_{idt}) + \lambda_2 \times (l_{\eta adv}^G + l_{\eta cls}^G). \quad (5)$$

$$l_D = \lambda_1 \times (l_{adv}^D + l_{cls}^D) + \lambda_2 \times (l_{\eta adv}^D + l_{\eta cls}^D). \quad (6)$$

Here, we set $\lambda_1 = \lambda_2 = 10$. The details of all the loss terms and the influence of the value of λ_2 are in Supplementary materials.

Training Process

In order to better describe the training process in our proposed method, we define the output of De-stain in the generator as G_{De} and the output of Re-stain as G_{Re} . The final output of the generator is defined as G . And the training process of our method is shown in Algorithm 1.

Experiments and Results

We have evaluated our proposed model over ANHIR dataset (Borovec et al. 2020). The results show that our method can

realize multi-domain stain transfer with high quality. Our method has strong generalization ability, and it still has superior performance on images whose staining types have not been seen during the training process. In addition, our model can be expanded to many tasks in the clinic, such as mixed staining and staining contrast tuning.

Dataset

In the ANHIR dataset, there are five sets of high-resolution tissue slides of the human kidney. In each set, there are four slides of consecutive tissue stained with different types of stain (H&E, MAS, PAS, and PASM staining), and the slides are basically similar in tissue structure but not pixel-level paired. The magnification of the slides is $40\times$, and the resolution is $0.2528\mu\text{m}/\text{pixel}$. The slides are downsampled to 25% of the original size. Additionally, the ANHIR dataset also contains mouse kidney tissue slides, which are stained with PAS, CD31, and other stains.

In training, we only use the sets of human kidney slides. We use four sets (kidney_1, kidney_2, kidney_3, kidney_4) as the training set, and one set (kidney_5) as the testing set. In addition, since the color of the H&E stained slide in kidney_1 is quite different from those in other sets, this H&E stained slide is removed during training. We cut these slides into 256×256 images with an overlap of 192. There are 39764 images in the training set (7500 for H&E, 12008 for MAS, 9616 for PAS, and 10640 for PASM), and 8387 images in the testing set (2023 for H&E, 2241 for MAS, 2004 for PAS, and 2119 for PASM).

Experimental Details

Our model is implemented with Python based on PyTorch on a computer with Intel(R) Core(TM) i5-10400 CPU, 16 GB RAM, and one NVidia RTX 3090 GPU.

Before being input to the network, each image is preprocessed by four data augmentation strategies: random crop, random flip, random rotation, and random resize. Our model is trained for 300000 iterations. During training, we use the Adam (DP and Ba 2015) optimizer with $\beta_1=0.5$ and $\beta_2=0.999$. The learning rate is set as 0.0001 initially and decreases using linear decay after 150000 iterations. Meanwhile, the batch size of the training dataset is set to 1.

Stain Transfer Results

Figure 2 exhibits the virtual generation of MAS, PAS, and PASM stained images from H&E stained images, and it can be seen that our network can realize multi-domain stain transfer. For example, it can be observed from the 2nd column of Fig. 2, in the generated MAS images, the collagen fibers are stained blue, and the muscle fibers are stained red. And as shown in the 2nd, 3rd, and 4th columns of Fig. 2, the structures of the input H&E images are well preserved. Moreover, the network can also generate other types of stained images based on non-H&E stained images, which are provided in Supplementary materials.

Generalization Ability Analysis We give the generalization ability analysis in the following three aspects: cross-tissue, cross-staining, and cross-task.

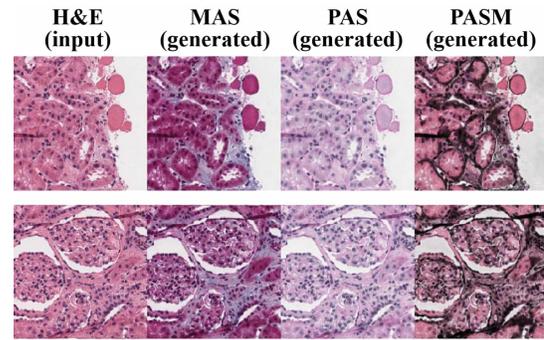


Figure 2: The virtual generation of MAS, PAS, and PASM stained images from H&E stained images using our network.

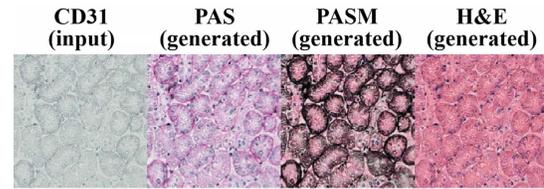


Figure 3: The virtual generation of PAS, PASM, and H&E stained images from CD31 stained images using our network.

(1) We choose the mouse kidney tissue stained by CD31 as an example to analyze cross-tissue and cross-staining generalization ability. It should be noted that CD31 stained images (Sullivan et al. 2015; Leong and Wright 1987) are not included in the training set. Similar to (Vasiljević et al. 2021), we input the CD31 stained images to our network, and PAS, PASM, and H&E stained results are generated. As shown in Fig. 3, the network can still learn the main features of CD31 stained images, and the staining is basically accurate. For example, in the generated PASM images, GTM can be stained black accurately. This is because various types of staining images are fed into our network, which enables our model to extract staining invariant features (i.e. content) from images of different staining styles.

(2) Similar to (Zhang et al. 2020), we use two extended tasks (mixed staining and staining contrast tuning) to verify the cross-task generalization ability of our network. In Fig. 4, we can see that fed a label like $[0.5, 0, 0, 0.5]$, our network can generate the mixed staining images of MAS and PASM. In clinical practice, if PASM staining and MAS staining are not performed on the same kidney tissue, it is not conducive for pathologists to judge whether the information expressed by the two types of staining exists in the same position of the tissue. However, as shown in Fig. 4, in the virtual PASM-MAS mixed staining, the glomerular basement membrane, tubular basement membrane, mesangial matrix (GTM), and the collagen fibers can be exhibited at the same time. The GTM are black, and the collagen fibers are blue. This method can show the relative position and the correlations of GTM and collagen fibers accurately (DONG et al. 2010).

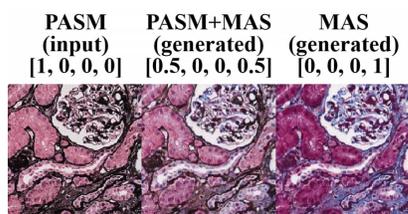


Figure 4: The result of mixed staining of MAS and PASM.

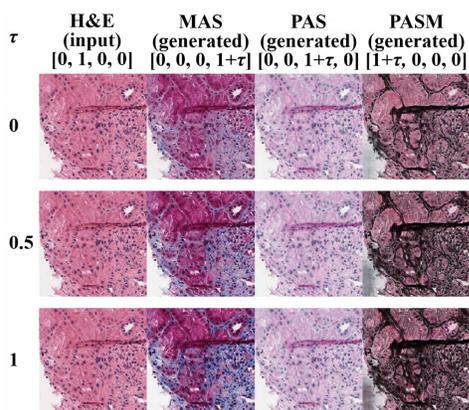


Figure 5: The results of staining contrast tuning

Figure 5 shows the results of tuning the staining contrast. We can see that given a label like $[0, 0, 0, 1+\tau]$ (τ is in a certain range), the contrast of the stained image can be tuned. For example, as τ increases, the color difference between the red areas and the blue areas stained by MAS becomes more obvious.

This is because the application of the improved one-hot label enables our network the appealing ability to adapt to different extended tasks.

Comparison Results

Here, we give the stained results of our method as well as other competing methods. The baseline methods include UGATIT which can achieve stain transfer based on unpaired data and StarGAN which can realize the transfer among multiple domains. We also compare similar frameworks like MUNIT and FUNIT.

Figure 6 shows the stain transfer results of the same H&E image using different networks. It can be seen that both MUNIT (Huang et al. 2018) and FUNIT (Liu et al. 2019) perform well in style, but they perform poorly in content, failing to preserve the original structure well. StarGAN (Choi et al. 2018) performs well in content, but it performs poorly in style, occurring mistakes in staining. For example, in the generation results of PASM, GTM are stained pink, while nearly all of the remaining parts of the tissue are stained black. UGATIT (Kim et al. 2020) and our network perform well in both content and style, but UGATIT cannot realize multi-domain stain transfer.

Further, in stain transfer, extracting the morphological content of input images is a crucial step, as it ensures the

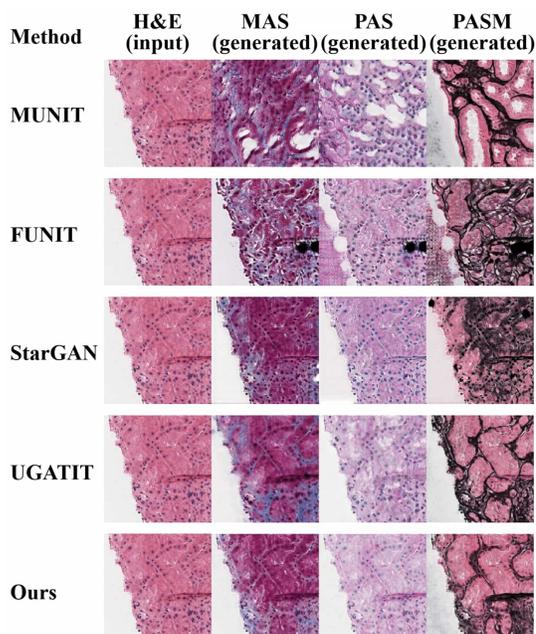


Figure 6: The comparison of the results of stain transfer of the same H&E image using MUNIT, FUNIT, StarGAN, UGATIT, and our method.

	MAS	PAS	PASM
MUNIT	0.145±0.073	0.115±0.019	0.110±0.055
FUNIT	0.332±0.052	0.318±0.034	0.246±0.018
StarGAN	0.520±0.046	0.543±0.051	0.491±0.057
UGATIT	0.584±0.043	0.510±0.038	0.399±0.064
Ours	0.682±0.039	0.674±0.040	0.600±0.025

Table 1: The CSS of different methods (higher is better).

	MUNIT	FUNIT	StarGAN	UGATIT	Ours
ER	1.2	2.8	3.2	4.4	4.4

Table 2: The evaluation results (ER) of different models (The mean results of the 5 pathologists, 1-lowest, 5-highest).

correctness of the structure. The evaluation metric Contrast-Structure Similarity (CSS) can express how much structural information is preserved from the original image, so we use CSS to measure the network’s ability to extract content. Table 1 shows that our method achieves the best among these methods. This is because, compared with the network trained with only one type of staining images (such as UGATIT), our network is trained with multiple types of stained images, so our network has a stronger ability to extract the structure of various types of staining images. Moreover, this phenomenon can be more evident in the types of staining not included in the training set. In Fig. 7, taking the PASM images as examples, we can see compared to our results, UGATIT ignores some of GTM and cannot stain it black, and the image generated by UGATIT appears blurry.

Furthermore, we invite five pathologists to evaluate the re-

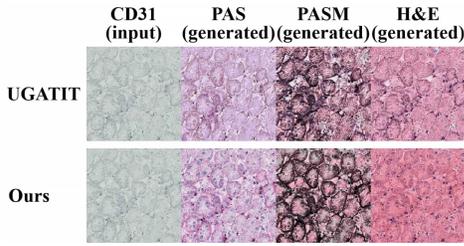


Figure 7: The comparison of the results of stain transfer of the same CD31 image using UGATIT and our method.

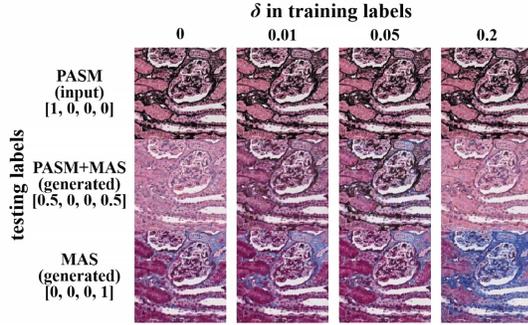


Figure 8: The results of mixed staining based on different values of δ .

sults generated by all the models, and the evaluation results (ER) can be seen in Table. 2.

Ablation Study

In the improved one-hot label, the small value δ enables the network to adapt to the problem caused by the differences between training labels and testing labels. Here, we discuss the impact of the improved one-hot label and the value of hyperparameter δ by two extended tasks.

Mixed Staining of MAS and PASM PASM can stain GTM black, while MAS can stain muscle fibers red and collagen fibers blue (de Haan et al. 2021; Lo et al. 2021).

In Fig. 8, it can be observed that with the increase of δ , the quality of mixed staining becomes better at first and then gets worse, which is in line with our expectations. When $\delta=0$ (i.e., the traditional one-hot label), the mixed staining image can neither show the black areas stained by PASM nor the red areas and the blue areas stained by MAS. When δ is given a proper value (for example, $\delta=0.05$), a better result can be obtained since the black areas, the red areas, and the blue areas can be displayed in the same image. When δ is too large (for example, $\delta=0.2$), the performance of mixed staining becomes worse again. In addition, if δ is too large, the virtually stained MAS image also becomes worse since some areas are stained blue mistakenly.

This is because when $\delta=0$, the network does not adapt to the input mode where more than one positions in the label are not 0. In contrast, when δ is a proper value, the network can adapt to the input mode well. However, the improved one-hot label is a double-edged sword: when δ is too large,

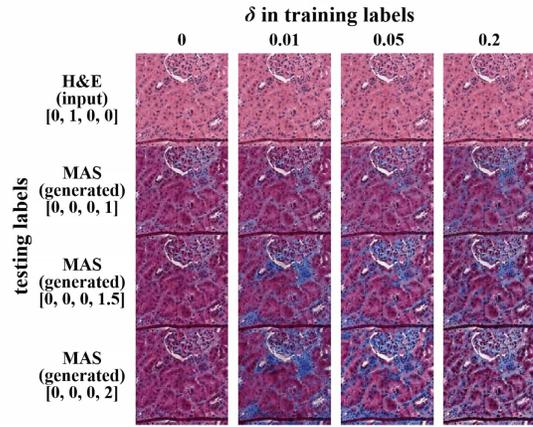


Figure 9: The results of staining contrast tuning based on different values of δ .

the performance is also poor. Based on the above observations, in this paper, we take $\delta=0.05$.

Staining Contrast Tuning In Fig. 9, we provide the staining contrast tuning results of MAS staining, which can stain the collagen fibers blue. During testing, our network is fed the label $[0, 0, 0, 1 + \tau]$ (the value of τ is 0, 0.5, 1, respectively, as shown in the leftmost column of Fig. 9). Similarly, when δ is given a proper value, the staining contrast changes obviously with the increase of τ . And if δ is too small or too large, the staining contrast is basically unchanged as τ increases.

Conclusion

In this paper, we propose a multi-domain stain transfer method based on unpaired data, enabling virtual staining of multiple types from only one staining type. We also devise style guided normalization (SGN) to control stain transfer direction efficiently. Moreover, multiple style encoding (MSE) is proposed to represent the relationship among different types of staining styles dynamically. Besides, the improved one-hot label in MSE enables the network to have a strong task extendibility, which can meet many clinical needs, such as mixed staining, staining contrast tuning. In addition, compared to the network that only transfers between two types of stain, such as UGATIT, our network has the superior ability to extract the staining invariant features and can adapt to the images whose staining types are not included in the training set.

Based on the above improvements, in clinical practice, our network can bring great convenience to pathologists and save a lot of time and money for patients. In addition, in scientific research, our method can be used for data augmentation and then improves the accuracy of downstream tasks (such as segmentation, detection).

In the future, we hope to extend our proposed method to immunohistochemical (IHC) staining technology, which is of great significance for the study of protein expression at the tissue and cell levels.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (61922048&62031023), in part by Guangdong Special Support (2019TX05X187), and in part by the Shenzhen Science and Technology Project (JCYJ20200109142808034&JCYJ20180508152042002).

References

- Ba, J. L.; Kiros, J. R.; and Hinton, G. E. 2016. Layer normalization. *arXiv preprint arXiv:1607.06450*.
- Borovec, J.; Kybic, J.; Arganda-Carreras, I.; Sorokin, D. V.; Bueno, G.; Khvostikov, A. V.; Bakas, S.; Eric, I.; Chang, C.; Heldmann, S.; et al. 2020. ANHIR: automatic non-rigid histological image registration challenge. *IEEE transactions on medical imaging*, 39(10): 3042–3052.
- Choi, Y.; Choi, M.; Kim, M.; Ha, J.-W.; Kim, S.; and Choo, J. 2018. StarGAN: Unified Generative Adversarial Networks for Multi-domain Image-to-Image Translation. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8789–8797.
- de Haan, K.; Zhang, Y.; Zuckerman, J. E.; Liu, T.; Sisk, A. E.; Diaz, M. F.; Jen, K.-Y.; Nobori, A.; Liou, S.; Zhang, S.; et al. 2021. Deep learning-based transformation of H&E stained tissues into special stains. *Nature Communications*, 12(1): 1–13.
- DONG, Y.; YANG, S.-c.; CHEN, W.-f.; et al. 2010. Applying double stain of PASM and Masson trichrom to pathological diagnosis of glomerulonephritis and glomerulopathy. *China Practical Medicine*, 2010: 22.
- DP, K.; and Ba, J. 2015. Adam: A method for stochastic optimization. In *Proc. of the 3rd International Conference for Learning Representations (ICLR)*.
- Dumoulin, V.; Shlens, J.; and Kudlur, M. 2017. A learned representation for artistic style. *ICLR (2017)*.
- Huang, X.; Liu, M.-Y.; Belongie, S.; and Kautz, J. 2018. Multimodal Unsupervised Image-to-image Translation. In *Proceedings of the European Conference on Computer Vision (ECCV)*.
- Isola, P.; Zhu, J.-Y.; Zhou, T.; and Efros, A. A. 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1125–1134.
- Kim, J.; Kim, M.; Kang, H.; and wang Hee Lee, K. 2020. U-GAT-IT: Unsupervised Generative Attentional Networks with Adaptive Layer-Instance Normalization for Image-to-Image Translation. In *International Conference on Learning Representations*.
- Leong, A.-Y.; and Wright, J. 1987. The contribution of immunohistochemical staining in tumour diagnosis. *Histopathology*, 11(12): 1295–1305.
- Li, X.; Zhang, G.; Qiao, H.; Bao, F.; Deng, Y.; Wu, J.; He, Y.; Yun, J.; Lin, X.; Xie, H.; et al. 2021. Unsupervised content-preserving transformation for optical microscopy. *Light: Science & Applications*, 10(1): 1–11.
- Liu, M.-Y.; Breuel, T.; and Kautz, J. 2017. Unsupervised image-to-image translation networks. In *Advances in neural information processing systems*, 700–708.
- Liu, M.-Y.; Huang, X.; Mallya, A.; Karras, T.; Aila, T.; Lehtinen, J.; and Kautz, J. 2019. Few-Shot Unsupervised Image-to-Image Translation. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 10550–10559.
- Liu, S.; Zhang, B.; Liu, Y.; Han, A.; Shi, H.; Guan, T.; and He, Y. 2021. Unpaired Stain Transfer Using Pathology-Consistent Constrained Generative Adversarial Networks. *IEEE Transactions on Medical Imaging*, 40(8): 1977–1989.
- Lo, Y.-C.; Chung, I.-F.; Guo, S.-N.; Wen, M.-C.; and Juang, C.-F. 2021. Cycle-consistent GAN-based stain translation of renal pathology images with glomerulus detection application. *Applied Soft Computing*, 98: 106822.
- Mao, X.; Li, Q.; Xie, H.; Lau, R. Y.; Wang, Z.; and Paul Smolley, S. 2017. Least squares generative adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, 2794–2802.
- Nam, H.; and Kim, H.-E. 2018. Batch-instance normalization for adaptively style-invariant neural networks. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, 2563–2572.
- Pizzati, F.; Cerri, P.; and de Charette, R. 2021. CoMoGAN: continuous model-guided image-to-image translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14288–14298.
- Rivenson, Y.; Wang, H.; Wei, Z.; de Haan, K.; Zhang, Y.; Wu, Y.; Günaydin, H.; Zuckerman, J. E.; Chong, T.; Sisk, A. E.; et al. 2019. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nature biomedical engineering*, 3(6): 466–477.
- Sullivan, H. C.; Edgar, M. A.; Cohen, C.; Kovach, C. K.; HooKim, K.; and Reid, M. D. 2015. The utility of ERG, CD31 and CD34 in the cytological diagnosis of angiosarcoma: an analysis of 25 cases. *Journal of clinical pathology*, 68(1): 44–50.
- Ulyanov, D.; Vedaldi, A.; and Lempitsky, V. 2016. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*.
- Vasiljević, J.; Feuerhake, F.; Wemmert, C.; and Lampert, T. 2021. Towards histopathological stain invariance by Unsupervised Domain Augmentation using generative adversarial networks. *Neurocomputing*, 460: 277–291.
- Xu, Z.; Li, X.; Zhu, X.; Chen, L.; He, Y.; and Chen, Y. 2020. Effective Immunohistochemistry Pathology Microscopy Image Generation Using CycleGAN. *Frontiers in Molecular Biosciences*, 7.
- Zhang, Y.; de Haan, K.; Rivenson, Y.; Li, J.; Delis, A.; and Ozcan, A. 2020. Digital synthesis of histological stains using micro-structured and multiplexed virtual staining of label-free tissue. *Light: Science & Applications*, 9(1): 1–13.
- Zhu, J.-Y.; Park, T.; Isola, P.; and Efros, A. A. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, 2223–2232.