

Unpaired Multi-Domain Stain Transfer for Kidney Histopathological Images

Supplementary Materials

A Loss Function

In order to better describe the Loss Function in our proposed method, we define the variables as follows:

x represents original image, and X_o represents the original staining domain.

L_t represents the target label, and L_o represents the original label.

G represents the output of the generator.

D_{adv} represents the output of Discrimination module.

D_{cls}^o represents the output of Classification module corresponding to the original style o ; D_{cls}^t is similar to D_{cls}^o .

η_{cls}^o and η_{cls}^t represent the output of auxiliary classifier 1. η_{adv} represents the output of auxiliary classifier 2.

There are three types of loss term when updating the discriminator: adversarial loss l_{adv}^D , classification loss l_{cls}^D , and auxiliary loss $l_{\eta_{adv}}^D, l_{\eta_{cls}}^D$.

a. Adversarial loss:

$$l_{adv}^D = E_{x \sim X_o} [(D_{adv}(x) - 1)^2] + E_{x \sim X_o} [(D_{adv}(G(x, L_t)))^2]. \quad (1)$$

By using adversarial loss, staining images generated by our model should be indistinguishable from real histochemical staining images.

b. Classification loss:

$$l_{cls}^D = E_{x \sim X_o} [(D_{cls}^o(x) - 1)^2]. \quad (2)$$

By using classification loss, the direction of stain transfer should be true.

c. Auxiliary loss:

$$l_{\eta_{adv}}^D = E_{x \sim X_o} [(\eta_{adv}(x) - 1)^2] + E_{x \sim X_o} [(\eta_{adv}(G(x, L_t)))^2]. \quad (3)$$

$$l_{\eta_{cls}}^D = E_{x \sim X_o} [(\eta_{cls}^o(x) - 1)^2]. \quad (4)$$

By using auxiliary loss, our stain transfer should be more accurate with the assistance of Auxiliary Classifiers.

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} .

a. Adversarial loss:

$$l_{adv}^G = E_{x \sim X_o} [(D_{adv}(G(x, L_t)) - 1)^2]. \quad (5)$$

b. Classification loss:

$$l_{cls}^G = E_{x \sim X_o} [(D_{cls}^t(G(x, L_t)) - 1)^2]. \quad (6)$$

c. Auxiliary loss:

$$l_{\eta_{adv}}^G = E_{x \sim X_o} [(\eta_{adv}(G(x, L_t)) - 1)^2]. \quad (7)$$

$$l_{\eta_{cls}}^G = E_{x \sim X_o} [(\eta_{cls}^t(G(x, L_t)) - 1)^2]. \quad (8)$$

All of a, b, and c (Eq. (5)-(8)) have the same meaning as their corresponding loss (Eq. (1)-(4)) while updating the discriminator.

d. Cycle loss:

$$l_{cyc} = E_{x \sim X_o} [| | | x - G(G(x, L_t), L_o) | | |_1]. \quad (9)$$

By using cycle loss, the content of the input images should be preserved well.

e. Identity loss:

$$l_{idt} = E_{x \sim X_o} [| | | x - G(x, L_o) | | |_1]. \quad (10)$$

By using identity loss, the content of the input images should be preserved well.

The two total loss formulations are as follows:

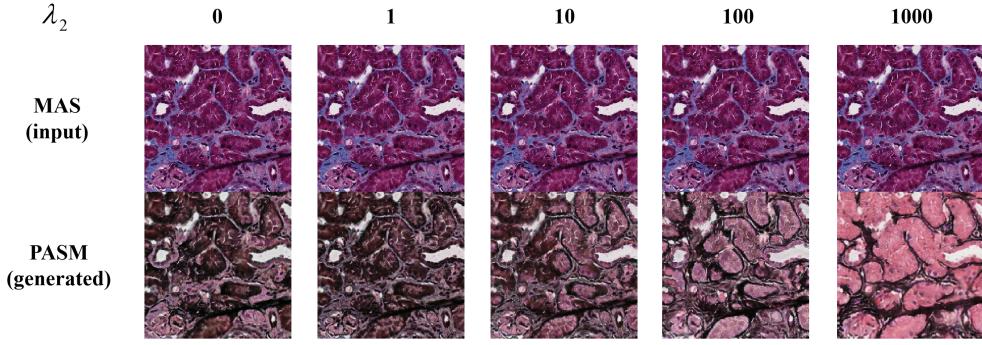
$$l_G = \lambda_1 \times (l_{adv}^G + l_{cls}^G + l_{cyc} + l_{idt}) + \lambda_2 \times (l_{\eta_{adv}}^G + l_{\eta_{cls}}^G). \quad (11)$$

$$l_D = \lambda_1 \times (l_{adv}^D + l_{cls}^D) + \lambda_2 \times (l_{\eta_{adv}}^D + l_{\eta_{cls}}^D). \quad (12)$$

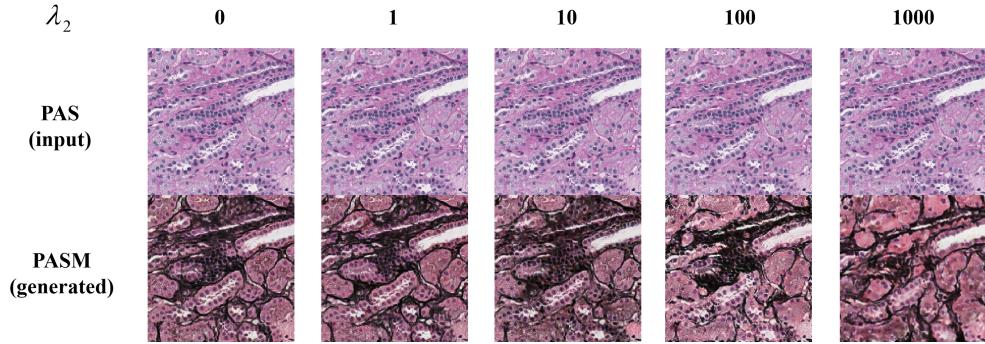
Here, we set $\lambda_1 = \lambda_2 = 10$. The details of how to choose a proper value of λ_2 are in Section B.

B Influence Analysis of Hyperparameter λ_2

In Eq. (11), (12), λ_2 is the coefficient of Auxiliary Classifiers. Here, we use the virtually generated PASM staining images from MAS staining and PAS staining to show the influence of λ_2 . If λ_2 is too small, the virtually generated images are stained mistakenly. For example, in the generated PASM images of Supplementary Fig. 1, when $\lambda_2=0$ (i.e., there are no Auxiliary Classifiers), the glomerular basement membrane, tubular basement membrane, mesangial matrix



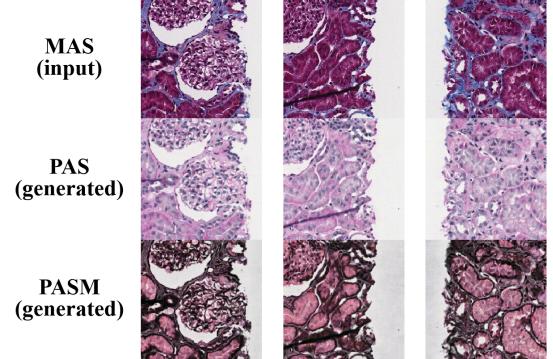
Supplementary Figure 1: The virtually generated PASM images from MAS stained images based on different values of λ_2 (Here, we set $\lambda_2 = 0, 1, 10, 100, 1000$).



Supplementary Figure 2: The virtually generated PASM images from PAS stained images based on different values of λ_2 (Here, we set $\lambda_2 = 0, 1, 10, 100, 1000$).

(GTM) are stained pink, while other parts of the tissue are stained black. Actually, PASM is used to stain GTM black. With the increase of λ_2 , this problem is gradually solved. However, when λ_2 is too large, as shown in Supplementary Fig. 2, the content of the input images cannot be preserved well (PASM and PAS staining show the same components in kidney tissue, so the transfer between PASM and PAS can be used to show our network’s ability to preserve the structure of the input images).

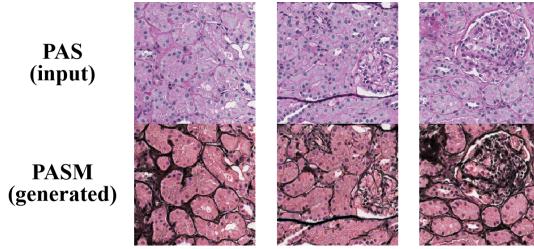
This is because when λ_2 is too small, Auxiliary Classifiers have less assistance to Classification Module and Discrimination Module, which makes the Classification Module and Discrimination Module unable to classify and judge the input images well. Therefore, the input images are poorly transferred to the new stain style, occurring the phenomenon of incorrect staining. When λ_2 is a proper value, Auxiliary Classifiers assist Classification Module and Discrimination Module well, so the results are accurate. But when λ_2 is too large, cycle loss (Eq.(9)) and identity loss (Eq. (10)) account for only a small proportion of the total loss (Eq. (11), (12)), while cycle loss and identity loss are used to retain the content of the input images. Thus, in this case, the network cannot preserve the content of the original images well. In conclusion, we need to set a proper value of λ_2 to make the images generated by the network optimal. In this paper, we set $\lambda_2=10$.



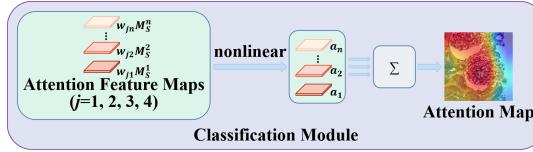
Supplementary Figure 3: The virtually generated PAS and PASM images from MAS stained images using our network.

C Stain Transfer Results Based on Non-H&E Stained Images

Supplementary Figure 3 exhibits the virtual generation of PAS and PASM stained images from MAS stained images; Supplementary Fig. 4 shows the virtual generation of PASM stained images from PAS stained images. Similarly, given an image whose staining style is in one of H&E, PAS, PASM, and MAS, our method can transfer the image to the three remainings.



Supplementary Figure 4: The virtually generated PASM images from PAS stained images using our network.



Supplementary Figure 5: The way to obtain attention map.

D Attention Maps of the Global and Local Discriminators

As shown in Supplementary Fig. 5, taking a group of attention feature maps in Classification Module as an example, the attention maps are obtained by stacking the attention feature maps across all the channels, which can reflect the area that the network pays attention to. And in our model, we use a combination of global discriminator and local discriminator. As the global discriminator contains more layers, it has a larger receptive field.

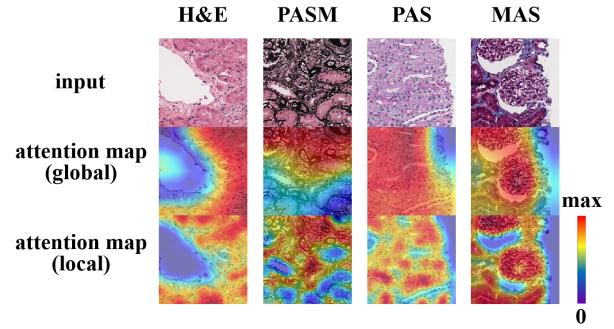
Supplementary Fig. 6 shows the attention maps corresponding to different types of staining images. Obviously, the global discriminator concentrates on the global information. For example, in all global attention maps, we can see that the global discriminator pays more attention to the tissue area and pays less attention to the background area. In contrast, the local discriminator concentrates on the local information. For example, in the local attention map of PASM staining, the network pays more attention to the black areas (GTM), while the effect of PASM staining is to stain GTM black. Based on these aspects, the area that the network focuses on is consistent with the actual meaning of the staining, so our network has strong interpretability.

E Comparison Results of Fixed φ and Our Learnable φ

In style guided normalization (SGN), the style guiding factor (SGF) φ can control the direction of virtual staining. Actually, FIN also adopted a similar idea. However, φ is fixed artificially in FIN. Thus sometimes poor results may be achieved, especially in the expanded applications, such as mixed staining.

To compare with FIN, in Table 1, we use four fixed values of φ to represent different types of staining.

The disadvantage of the fixed value of φ is that it cannot generate mixed staining between arbitrarily two types of staining. For example, if we want to generate the mixed

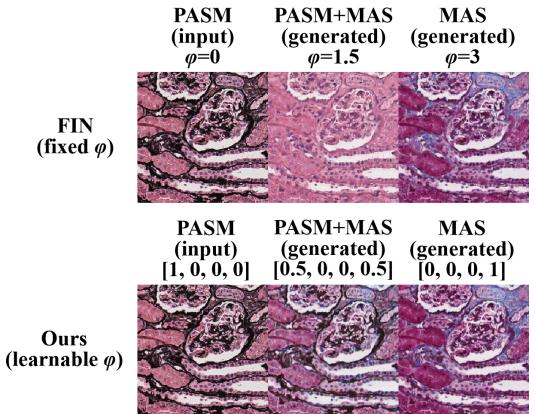


Supplementary Figure 6: Attention maps of the global and local discriminators.

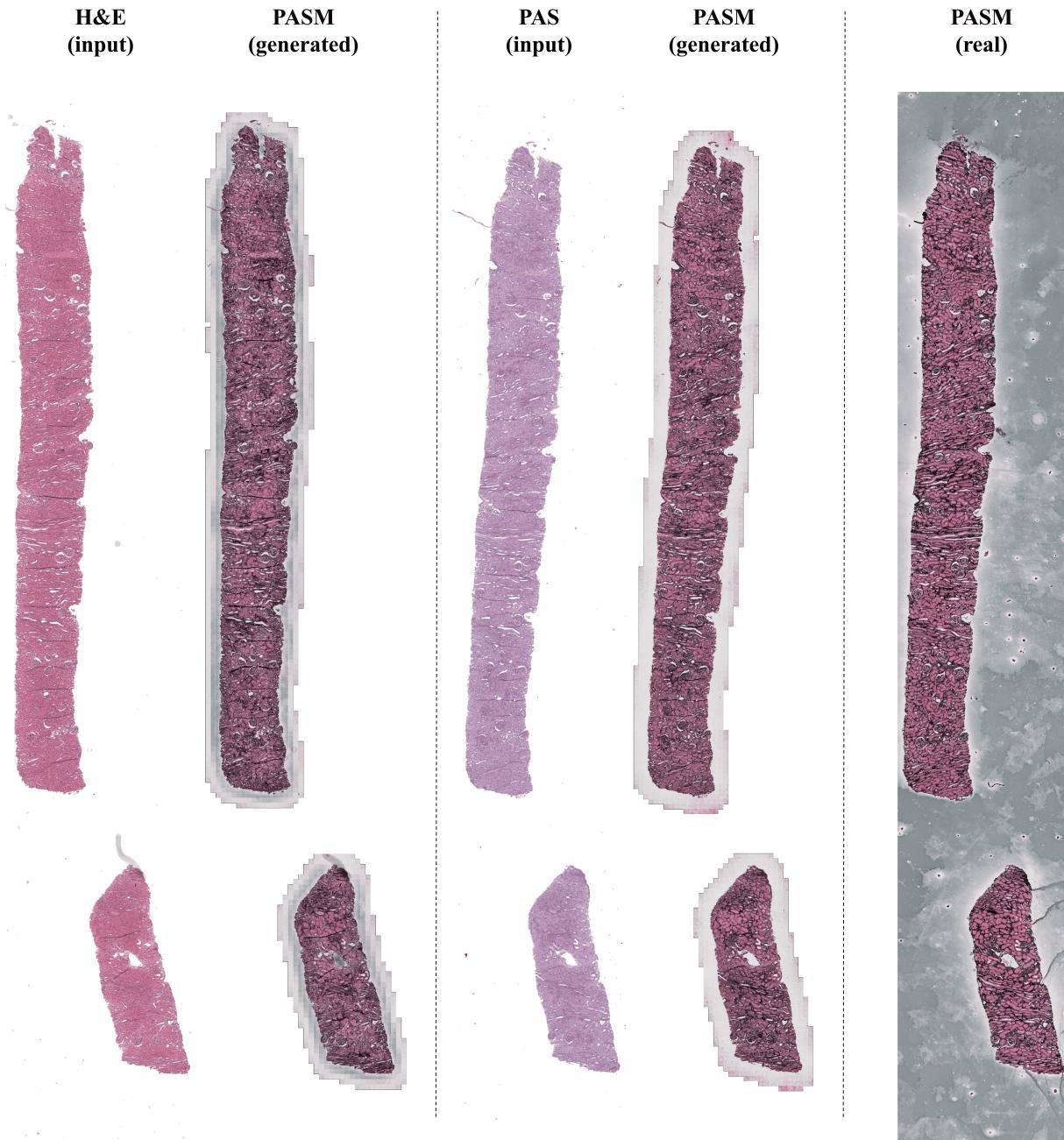
Table 1: The values of φ representing different staining domains.

Staining domains	Values of φ
PASM	0
H&E	1
PAS	2
MAS	3

staining of PASM and MAS, we can input $\varphi = \frac{0+3}{2} = 1.5$ to the network. And if we want to generate the mixed staining of H&E and PAS, we can also feed the network $\varphi = \frac{1+2}{2} = 1.5$. Therefore, $\varphi = 1.5$ is used to represent both the mixed staining of H&E-PAS and PASM-MAS. Obviously, such a fixed value of φ is highly ambiguous, indicating that FIN cannot dynamically represent the relationship among different types of staining styles. Supplementary Fig. 7 shows the virtual staining results when $\varphi=1.5$ is fed to the network. It can be seen that the φ with fixed values does not express any of the information of PASM staining and MAS staining (PASM can stain GTM black, while MAS can stain muscle fibers red and collagen fibers blue). However, in our method, the common expressed information between PASM and MAS can be obtained.



Supplementary Figure 7: The comparison results.



Supplementary Figure 8: Slide-level results using our network.

F Slide-Level Results

As shown in Supplementary Fig. 8, we input all the tiles of H&E and PAS slides in lung-lesion_5 into the network to generate PASM stained images and then splice these generated images into slides. Comparing our results with the real PASM stained slide of its consecutive tissue, we can see that our model is very accurate.