AGMDT: Virtual Staining of Renal Histology Images with Adjacency-Guided Multi-Domain Transfer Supplementary Materials

BMVC 2023 Submission # 409

A Loss Function

To incorporate adjacent layer information as supervision, we design the following loss function. To make it easier to explain, we define the following variables:

x refers to the input image, while $\widetilde{y}$ represents its corresponding adjacent image.

X_o represents the original staining domain, L_t represents the target label, and L_o represents the original label.

G represents the generator and D represents the discriminator. A represents the Adjacency-guided Encoder, S represents the encoder-decoder module for de-stain and re-stain, while C represents the style code generator.

The generator's loss function consists of six components, and can be divided into three parts: basic loss, auxiliary loss, and adjacency-guided loss. The basic loss, which includes adversarial loss l^G_{adv}, classification loss l^G_{cls}, cycle consistency loss l_{cyc}, and identity preserving loss l_{idt}, is used to achieve fundamental multimodal stain transfer:

- adversarial loss l^G_{adv}:

$$l_{adv}(G) = E_{x\sim X_o}[(D_{adv}(S(x,C(L_t))-1)^2] \tag{1}$$

- classification loss l^G_{cls}:

$$l_{cls}(G) = E_{x\sim X_o}[(D_{cls}(S(x,C(L_t))-1)^2] \tag{2}$$

- cycle consistency loss l_{cyc}:

$$l_{cyc}(G) = E_{x\sim X_o}[\|x - S(S(x,C(L_t),C(L_o))\|_1] \tag{3}$$

- identity preserving loss l_{idt}:

$$l_{idt}(G) = E_{x\sim X_o}[\|x - S(x,C(L_o))\|_1] \tag{4}$$

Auxiliary loss, which includes $l^G_{\eta adv}$, $l^G_{\eta cls}$, is used to improve the performance of stain transfer, and the setting is the same as that in UMDST[3].

Adjacency-guided loss l_{adj} measures the discrepancy between the generated data and its corresponding adjacency data. The loss function is formulated as follows:

$$l^G = \lambda_1 \times (l^G_{adv} + l^G_{cls} + l_{cyc} + l_{idt}) + \lambda_2 \times (l^G_{\eta adv} + l^G_{\eta cls}) + \lambda_3 \times l_{adj} \tag{5}$$

Here, we set $\lambda_1 = \lambda_2 = \lambda_3 = 10$.

As for adjacency-guided loss, we denote the Adjacency-guided Encoder as A, the encoder-decoder module for de-stain and re-stain as S and the formula for computing l_{adj} is as follows:

$$l_{adj} = E_{x,\tilde{y}}\left[\|\nabla A(S(x),\tilde{y})\|^2\right] + E_{x,\tilde{y}}[\|\tilde{y} - S(x) \circ A(S(x),\tilde{y})\|_1] \tag{6}$$

B Blind Evaluation of Staining Results

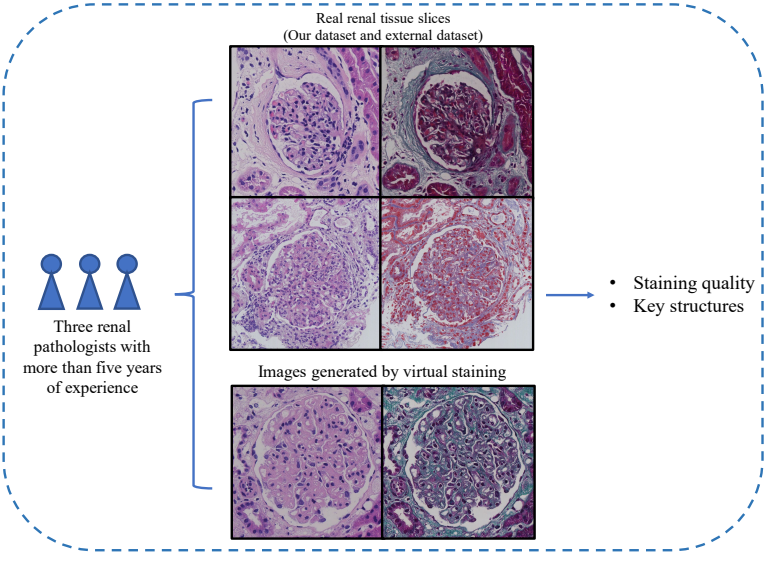

Figure 1: An overview of the blind evaluation of staining results.

The goal of virtual staining is to enable pathologists to use the generated images directly for diagnosis, just like conventional stained sections. Therefore, we designed a comprehensive experiment in which pathologists were invited to evaluate the quality of images generated by virtual staining methods.

We invited three pathologists with more than five years of clinical experience to participate in the experiment. The evaluation index of the experiment is divided into two items, namely the staining quality and the key structure. The pathologists scored each item from 0 to 5. The higher the score, the better the image quality in that item, and the average score of the two items is taken as the overall quality score. In principle, the clinical diagnostic requirements are met with a score of 4 or above. The corresponding relationship between the score and the degree of staining is as follows:

- 4-5: The staining style and clarity of the key structure basically meet the diagnostic needs.
- 3-4: The staining style is not good, the structure is not well displayed, or although the structure is displayed, the relationship between cells and the glomerular basement membrane is not well displayed.
- 2-3: The internal structure is incorrect and the style is not good.
- Below 2: The structure is completely wrong.

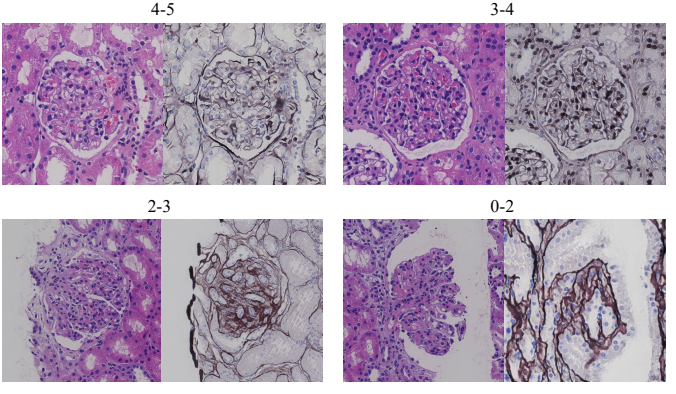

Figure 2: The corresponding relationship between the score and the degree of staining.

In order to facilitate the judgment of the pathologists, the dataset we provide consists of glomerular images, including three parts: the images generated by different virtual staining methods, the images from our glomerulus-aligned renal histological dataset, and the images of real renal tissue from external dataset. We randomly divide the data to be tested into three groups and ensure that each group of data is scored by two pathologists, i.e., each pathologist scores two groups of data, and the final score of each group of data is the average of the scores of the two pathologists.

Table 1: Evaluation of staining results by pathologists

	Masson	**PASM**	**PAS**
Real(our part)	4.25	**4.92**	4.33
Real(external part)	3.48	4.77	**4.46**
FUNIT[2]	0.69	1.42	2.25
MUNIT[1]	2.67	0.75	2.42
UGATIT[2]	2.67	3.67	3.08
UMDST[3]	4.10	4.09	4.17
AGMDT	**4.59**	4.17	4.25

The experimental results are shown in Table 1. Among the virtual staining methods, AGMDT has the best performance. The images generated by AGMDT have achieved more

than 4 points in all three types of staining, and the staining level is close to that of real renal tissue slices. Especially in Masson, it has achieved the highest score.

References

[1] Xun Huang, Ming-Yu Liu, Serge Belongie, and Jan Kautz. Multimodal unsupervised image-to-image translation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 172–189, 2018.

[2] Junho Kim, Minjae Kim, Hyeonwoo Kang, and Kwanghee Lee. U-gat-it: Unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation. *arXiv preprint arXiv:1907.10830*, 2019.

[3] Yiyang Lin, Bowei Zeng, Yifeng Wang, Yang Chen, Zijie Fang, Jian Zhang, Xiangyang Ji, Haoqian Wang, and Yongbing Zhang. Unpaired multi-domain stain transfer for kidney histopathological images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 1630–1637, 2022.

[4] Ming-Yu Liu, Xun Huang, Arun Mallya, Tero Karras, Timo Aila, Jaakko Lehtinen, and Jan Kautz. Few-shot unsupervised image-to-image translation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10551–10560, 2019.