UNPAIRED H&E TO PR STAIN TRANSFER WITH SELF-SUPERVISED AUXILIARY SEGMENTATION

Learning from Data
(Fall 2022)

Yiyang Lin 2020214323   Sen Zeng 2022214125

INTRODUCTION

With the rapid development of deep neural networks, some work has been successfully applied in virtual staining, However, without expert annotations, achieving accurate virtual staining of IHC based on unpaired data is a challenging task. In this project, we propose a model to accurately transfer H&E images to IHC images by using a self-supervised auxiliary segmentation task to help extract the features necessary for generating IHC images. The self-supervised segmentation proxy task is based on the observation that the positive areas in IHC images are easy to be segmented.

METHOD

Our model consists of IHC-H&E-IHC and H&E-IHC-H&E processes.

1. The IHC-H&E-IHC and H&E-IHC-H&E processes

As shown in Fig.1(a), we first feed a real IHC image into GIHC2H&E to generate a virtual H&E image, and a reconstructed IHC image is generated after feeding the virtual H&E image to GH&E2IHC. By employing cycle loss, the reconstructed IHC image is constrained to be consistent with the real IHC image. In the process of generating the reconstructed IHC image, we simultaneously input the virtual H&E image to SH&E and obtain the positive areas of the virtual H&E image. Moreover, the label obtained by inputting the real IHC image to SIHC is used to constrain SH&E by using segmentation loss. we also input the real H&E image into SH&E and use the label obtained by inputting the virtual IHC image into SIHC as ground truth to constrain SH&E, as shown in Fig.1(b).

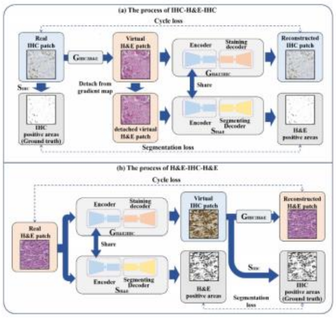


Fig. 1. The overall structure of our model. (a): The process of IHC-H&E-IHC; (b): The process of H&E-IHC-H&E.

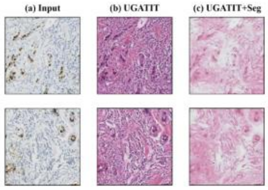


Fig. 2. Real IHC images and virtual H&E results under different conditions.

2. The impact of segmentation on virtual H&E images

As shown in Fig.2(a)&(b), the model without the auxiliary segmenter SH&E (UGATIT) maps the brown areas in IHC to dark purple areas in H&E, and maps the blue areas in IHC to light purple areas in H&E, which constitutes many color differences that exist rarely in real H&E images. The color differences show that UGATIT mainly learns the color mapping, and the real mapping of H&E images and IHC images is not obtained. As shown in Fig.1(a), we detach SH&E from the current gradient map. And since SH&E and GH&E2IHC share the encoder which can extract both the color features and morphological features well, the virtual H&E images are much closer to the real H&E images.

EXPERIMENTS

We have evaluated our proposed model over a breast cancer dataset. The results show our model can transfer H&E images into PR (a typical type of IHC) images with unpaired patches efficiently, indicating that the self-supervised auxiliary segmentation task improves the accuracy greatly.

Figure 3 exhibits the virtual generation of PR stained images from H&E stained images. In Fig.3(a), the slide-level results of a positive slide and a negative slide are provided.

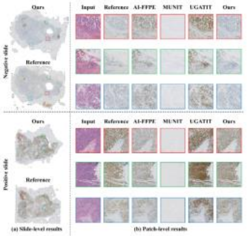
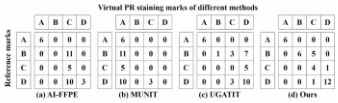


Fig. 3. The slide-level and patch-level results of our model and competing models.


Virtual PR staining marks of different methods

Reference marks

(a) AI-FFPE

	A	B	C	D
A	6	0	0	0
B	0	0	11	0
C	0	0	5	0
D	0	0	10	3

(b) MUNIT

	A	B	C	D
A	6	0	0	0
B	11	0	0	0
C	5	0	0	0
D	10	0	3	0

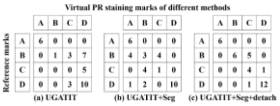
(c) UGATIT

	A	B	C	D
A	6	0	0	0
B	0	1	3	7
C	0	0	0	5
D	0	0	3	10

(d) Ours

	A	B	C	D
A	6	0	0	0
B	0	6	5	0
C	0	0	4	1
D	0	0	1	12

Fig. 4. Confusion matrices for the judgment results of Fiji IHC Toolbox. A-paracancerous tissue area, B-negative area, C-weakly positive area, D-positive area.


Virtual PR staining marks of different methods

Reference marks

(a) UGATIT

	A	B	C	D
A	6	0	0	0
B	0	1	3	7
C	0	0	0	5
D	0	0	3	10

(b) UGATIT+Seg

	A	B	C	D
A	6	0	0	0
B	4	3	4	0
C	0	4	1	0
D	1	2	0	10

(c) UGATIT+Seg+detach

	A	B	C	D
A	6	0	0	0
B	0	6	5	0
C	0	0	4	1
D	0	0	1	12

Fig. 5. Confusion matrices for the judgment results of Fiji IHC Toolbox. A-paracancerous tissue area, B-negative area, C-weakly positive area, D-positive area.

We employ Fiji IHC Toolbox to evaluate the results generated by all the models. The confusion matrices of the judged results are shown in Fig.4. It can be seen that our model achieves the highest accuracy compared to all the unsupervised competing models.

We separate SH&E from the current gradient map to ensure the virtual H&E images are much closer to the real H&E images. As shown in Fig.5(c), we can see by introducing and detaching SH&E, the accuracy of IHC generation improves a lot.

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

Models	AI-FFPE	MUNIT	UGATIT	Ours
CSS	0.629±0.169	0.384±0.207	0.613±0.174	**0.672±0.158**

Table 1 shows that our model achieves the best among these models, as the proposed self-supervised auxiliary segmenter can improve our model's focus on morphological information.

CONTRIBUTIONS

1. We propose a stain transfer model for the virtual generation of IHC images from H&E images, and apply it to PR virtual staining for the first time.
2. We propose a self-supervised image style transfer model. With the observation of the easy segmentation property of PR images, we introduce a segmentation proxy task to improve the accuracy of style transfer.
3. We enhance the model's focus on morphological information by introducing the auxiliary segmentation task. Moreover, the gap between real H&E images and virtual H&E images can be reduced, and the features extracted in the virtual H&E images can be closer to those extracted in the real ones. Ultimately, the accuracy of IHC virtual staining can be improved.

CONCLUSIONS

we propose an H&E to PR stain transfer model with self-supervised auxiliary segmentation, which can be trained with unpaired patches. The ground truth of the self-supervised auxiliary segmentation task is obtained by employing the property that PR images are easily segmented by the thresholding method. We introduce a detaching operation in the training process to ensure the high quality of virtual H&E required for generating PR results accurately.

Experiments on PR images show that our model is superior to traditional unsupervised models, and our approach can also achieve pleasing results on the external dataset.