Re-staining Pathology Images by FCNN

Masayuki FUJITANI
Waseda University
masa-tani@fuji.waseda.jp

Yoshihiko MOCHIZUKI
Waseda University
motchy@aoni.waseda.jp

Satoshi Iizuka
Center for Artificial Intelligence Research, University of Tsukuba
iizuka@cs.tsukuba.ac.jp

Edgar Simo-Serra
Waseda University
esimo@aoni.waseda.jp

Hirokazu Kobayashi
Nagoya Institute of Technology
kobayasi@iu.nitech.ac.jp

Chika Iwamoto
Kyushu University
chika-i@dem.med.kyushu-u.ac.jp

Kenoki Ohuchida
Kyushu University
kenoki@surg1.med.kyushu-u.ac.jp

Makoto Hashizume
Kyushu University, Japan
mhashi@dem.med.kyushu-u.ac.jp

Hidekata Hontani
Nagoya Institute of Technology
hontani@nitech.ac.jp

Hiroshi Ishikawa
Waseda University
hfs@waseda.jp

Abstract

In histopathology, pathologic tissue samples are stained using one of various techniques according to the desired features to be observed in microscopic examination. One problem is that staining is irreversible. Once a tissue slice is stained using a technique, it cannot be re-stained using another. In this work, we propose a method for simulated re-staining using a Fully Convolutional Neural Network (FCNN). We convert a digitally scanned pathology image of a sample, stained using one technique, into another image with a different simulated stain. The challenge is that the ground truth cannot be obtained: the network needs training data, which in this case would be pairs of images of a sample stained in two different techniques. We overcome this problem by using the images of consecutive slices that are stained using the two distinct techniques, screening for morphological similarity by comparing their density components in the HSD color space. We demonstrate the effectiveness of the method in the case of converting hematoxylin and eosin-stained images into Masson's trichrome-stained images.

1 Introduction

In histopathology, pathologic tissue samples are chemically processed, or stained, using one of various techniques to enhance visible features in order to diagnose and characterize the disease based on morphology of various cell components in microscopic examination. The stain technique is chosen according to the purpose of the examination. For instance, the most basic and widely used stain is hematoxylin and eosin (H&E) stain, which is used to highlight general cellular structure. Another example is the Masson's trichrome (MT) stain, which is used to differentiate between collagen and smooth muscle in tumors.

Staining is an irreversible chemical process. Generally, once a tissue is stained using one technique, it cannot be stained again using another; thus, one sample can only be observed with a single enhancement modality. Consequently in clinical medicine, once a tissue sample is stained with one technique, features that are not highlighted by the specific stain often have to be guessed. It is true that multiple slices of a tissue can be stained differently. However, that requires significantly more time and cost. If the histopathological images can be digitally *re-stained*, more structures and features would be enhanced than possible by a chemical stain; thus such digital re-staining would be very useful as one sample can give more information than previously possible.

In a related work, Hashimoto et al. [2] used 16-band multispectral images of H&E stained liver-tissue image to enhance the fiber region. They took spectral transmittance samples from different tissue components and enhanced the fiber region according to that informa-

tion, effectively simulating the MT stain. However, this method requires a multispectral microscope, as well as careful prior calibration of spectral transmittance.

Bejnordi et al. [1] proposed a system to standardize H&E stained histopathological images to reduce the effect of variations in the color and intensity. They used the hue-saturation-density (HSD) color model [5] to align the chromatic and density distributions for each of the stain components to match the corresponding distributions from a template image. Our work similarly utilize the characterization of the stain components by chromatic and density distributions for other purposes, namely, for simulating stain by one technique based on an imaged obtained using another.

Here, we propose a method for using Fully Convolutional Neural Network (FCNN) to convert a digitally scanned pathology image of a sample stained using one technique into another image that has a simulated stain using another technique. Our method directly takes an image of a stained pathological sample and outputs an image of the same sample simulating a different stain.

An FCNN is a CNN that does not have a fully-connected layer, which allows it to process an input image of any size. To determine the color of a pixel in the output image, it uses the information from the surrounding wider region in the input image, utilizing the chromatic and density distributions. The challenge is that, as any CNN, an FCNN needs a large amount of training data, which in this case would be pairs of images showing the results of staining samples by two techniques; however, such pairs are not obtainable because chemical re-staining is not possible. In this work, we overcome this problem by using the images of consecutive slices that are stained using different techniques. We screen the pairs to be used as training data so that the slices are close enough by comparing them after a non-rigid registration, using the density component in the HSD color space, in order to compare the images with different stains.

2 Method

In this paper, we focus on using an H&E stained image to simulate an MT stained image. The input H&E stained image is first converted to the HSD color space [5]. The three-channel image is then input to the Fully Convolutional Neural Network. The output of the FCNN has two channels (hue and saturation in the HSD color space). We combine the density channel of the input image with this before converting back to the RGB color space for output.

2.1 HSD Color Space

The stain conversion is performed on the hue-saturation-density (HSD) color model [5]. We use the coordinate system (c_x, c_y, D), where the first two coordinates represent the hue and saturation, while D measures the density. Specifically, the conversion from the RGB space (r, g, b) to the HSD space is defined as follows:

$$c_x = \frac{D_r}{3D},$$
$$c_y = \frac{D_g - D_b}{6D} + \frac{1}{2},$$
$$D = \frac{D_r + D_g + D_b}{3},$$

where

$$D_r = -\ln(r), \quad D_g = -\ln(g), \quad D_b = -\ln(b).$$

This space is designed so that a different stain only changes the hue and saturation, while the density is only affected by the density of the stain solution. Thus, we only change the hue and saturation in order to simulate different stain.

2.2 Training Data

The FCNN is trained using input-output image pairs. The challenge here is that images showing the results of two different stain of the same samples are not obtainable because chemical re-staining is not possible. We overcome this problem by using the images of consecutive slices that are stained using different techniques. Out of a pair of large images ($\tilde{}10000 \times 10000$ pixels) of consecutive slices of a sample that are stained by the two techniques, we sample smaller subregions as training data. In the process, we screen the training pairs so that the change between the slices are not too large and they are morphologically similar. Because the pair of images are differently stained, they cannot be compared directly. As explained in §2.1, the density component in the HSD color space are relatively independent of the kind of stain. Thus we compare the slices stained differently using this component alone.

The overall density of the image, however, can be different depending on various factors such as the concentration of the stain chemical and the optical condition of the camera. Thus, we scale the density component by a constant normalization factor to match its histogram over the image. Fig. 1 shows the comparison of histograms of the D component of the two images I_1 (H&E stained) and I_2 (MT stained). It also shows the histogram obtained by scaling the D component of image I_1 by a constant normalization factor $k_{1,2}$. It can be seen that the constant normalization factor matches the histograms very well. The normalization factor $k_{1,2}$ is obtained by minimizing the squared distance of the histogram bins as vectors.

By comparing the normalized pair of density images, we screen the training pairs so that only morphologically similar pairs are used. The process is outlined in Fig. 2. The images of the consecutive slices are first matched as much as possible by non-rigid registration.

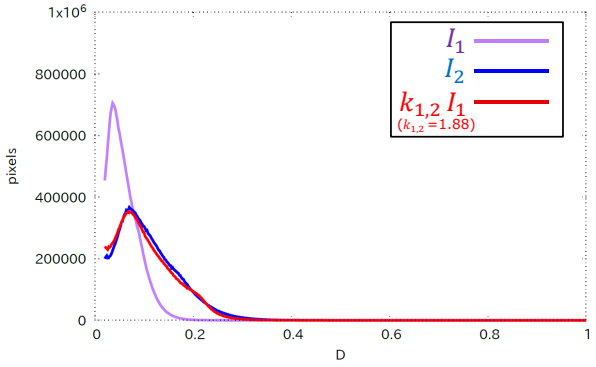

Figure 1: Histograms of the density component (D) of the H&E stained image I_1, MT stained image I_2, and I_1 with scaled density by a constant factor $k_{1,2}$.

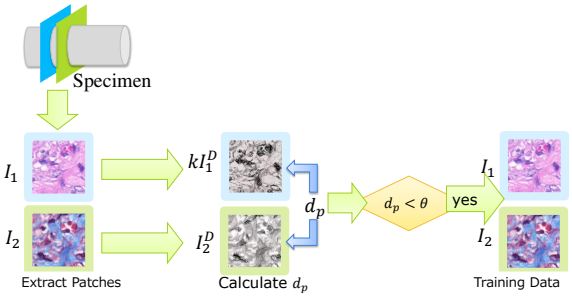

Figure 2: Training data generation. Comparing the normalized density components, we screen the training pairs to use only morphologically similar pairs.

From these large pair, we randomly cut out a small region. As training pairs, we use only the regions where the slices are close enough as measured by the Mean Squared Error between the matched density components.

2.3 Network Model

The structure of the network is shown in Fig. 3 and detailed in Table 1. Except for the last layer, a batch normalization layer [3] is inserted after each layer and a ReLU activation function is used. The last layer uses the sigmoid function as the activation function. As the fully convolutional network, this model can take the image of any size as the input and outputs an image of the same size. In Fig. 3, the thickness of the layers indicates the number of channels. The model first shrink the spatial dimensions as it down-samples, as well as increasing the number of channels, then it up-samples and reduces the channels. Down- and up-sampling is achieved by convolutional layers with strides.

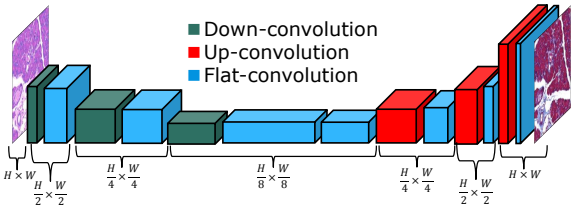

Figure 3: Netwrok model for re-staining. The thickness indicates the number of channels.

Table 1: Details of the FCNN model for digital re-staining. The size of the input image ($W \times H$) is arbitrary. The input image has the three channels (c_x, c_y, D) of the HSD color space. The output has two chromatic channels (c_x, c_y). In order not to change the output size, 0-value padding is used in all the layers. After each convolution layer other than the last is a batch normalization layer [3], after which we use the Rectified Linear Unit (ReLU) activation function. The final layer uses a sigmoid activation function.

type	kernel	stride	channel	spatial dimensions
input			3	$W \times H$
down-convolution	5×5	2×2	48	$W/2 \times H/2$
flat-convolution	3×3	1×1	128	$W/2 \times H/2$
down-convolution	3×3	2×2	256	$W/4 \times H/4$
flat-convolution	3×3	1×1	256	$W/4 \times H/4$
down-convolution	3×3	2×2	256	$W/8 \times H/8$
flat-convolution	3×3	1×1	512	$W/8 \times H/8$
flat-convolution	3×3	1×1	256	$W/8 \times H/8$
up-convolution	4×4	2×2	256	$W/4 \times H/4$
flat-convolution	3×3	1×1	128	$W/4 \times H/4$
up-convolution	4×4	2×2	128	$W/2 \times H/2$
flat-convolution	3×3	1×1	48	$W/2 \times H/2$
up-convolution	4×4	2×2	48	$W \times H$
flat-convolution	3×3	1×1	24	$W \times H$
flat-convolution	3×3	1×1	2	$W \times H$

2.4 Training

As explained in §2.2, we randomly select small region with the dissimilarity measure smaller than a threshold out of the images of consecutive slices. We use the resulting pairs of images, H&E stained input and MT stained teacher output, as training pair. As the loss function, we use a standard MSE loss. Note that the MSE is taken in the 2D chromatic components.

3 Experiments and Results

We evaluated the method by converting H&E stained sample images of mouse pancreas to MT stain.

3.1 Dataset

We prepared pathologic images as follows. The specimen was a mouse pancreas including tumor, surgi-

cally removed and treated with formalin. The paraffin-embedded tissue blocks were used to prepare sections of 4μm thickness. There were 64 pairs of consecutive slices. Each pair were H&E- and MT- stained and were scanned by Axio Scan. The image resolution was 98000×60000 pixels, each pixel roughly corresponding to 4μm. We call these the high-resolution images. We also used for some experiment down-sampled low-resolution images, which is 15% of the resolution of the original. The images of consecutive slices were non-rigid registered using the method in [4].

3.2 Training

For training, we used 58 of the 64 pairs of slices, leaving 6 for testing. We conducted two experiments using the high-resolution and low-resolution images, respectively. The training and testing scheme was the same for both. We mini-batch trained the network using AdaDelta [6] with the mini batch of 16 pairs of image patches. Each patch was taken from a random position within a randomly selected slice out of the 58 training pairs. The size of the patches were 256×256 for the high-resolution experiment and 128×128 for low resolution. We trained for 100,000 mini-batch iterations.

3.3 Results and Analysis

For testing, we input the 6 test images into the trained network. Fig. 4 shows some close-ups of the results in the high-resolution experiment. It shows the H&E stained input, the output (simulated MT stain), the output with the density matched with the neighboring slice, and the neighboring slice (MT stained) for comparison. Note that the network predicts the chromatic components (c_x, c_y) only; the constant factor on the density is merely for the convenience of comparison. Fig. 5 shows two entire images in the low-resolution experiment.

Quantitative comparison of the results is rather difficult since we don't have the ground truth and also there is no prior work on re-staining. We attempt the following two evaluations.

In [5], the authors showed that in the HSD color space only the 2D chromatic components are needed for all possible distinctions. Thus, for the Example 1 of Fig. 4, we compare the distribution of the chromatic components of the input, the output, and the neighboring slice. It is shown in the left plot in Fig. 6. It can be seen that the chromatic distribution of the output lies centrally on the real MT stained sample.

Also, we compare the difference between various MT stained slices and simulated MT stained slices in the high-resolution experiment. It is shown in the right plot in Fig. 6. We fix a reference MT stained slice. The red dots show the Mean Squared Error between the reference slice and the simulated MT stains, i.e., the output from H&E slices removed from the reference slice by the distance (number of slices) seen on the horizontal axis. The green dots indicate the MSE between the reference slice and other MT stained real slices, removed from the reference slice by the distance (number of slices) seen on the horizontal axis. To see the influence of registration error, we also show the magenta dots that are the same error as red. This shows the output simulated MT stain is close enough as compared to the real neighboring slices.

4 Conclusion

In this paper, we have proposed a method for using Fully Convolutional Neural Network (FCNN) to simulate a re-staining of pathology images. To overcome the challenge of obtaining training pairs we used consecutive slices that are stained using different techniques, screening for morphological similarity by comparing their density components in the HSD color space. The experiments showed that the proposed method produces simulated re-stain that is close to real stain.

References

[1] B. E. Bejnordi, G. Litjens, N. Timofeeva, I. Otte-Höller, A. Homeyer, N. Karssemeijer, and v. d. L. Jeroen A.W.M. Stain specific standardization of whole-slide histopathological images. *IEEE Transactions on Medical Imaging*, 35(2):404–415, Feb. 2016.

[2] N. Hashimoto, Y. Murakami, P. A. Bautista, M. Yamaguchi, T. Obi, N. Ohyama, K. Uto, and Y. Kosugi. Multispectral image enhancement for effective visualization. *Optics Express*, 19(10):9315–9329, 2011.

[3] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *Proceedings of the 32nd International Conference on Machine Learning*, pages 448–456, 2015.

[4] N. Kawamura, H. Kobayashi, T. Yokota, H. Hontani, C. Iwamoto, K. Ohuchida, and M. Hashizume. Landmark-based reconstruction of 3d smooth structures from serial histological sections. In *Proc. SPIE 10581*, pages 1–7. International Society for Optics and Photonics, 2018.

[5] J. A. W. M. van der Laak, M. M. M. Pahlplatz, A. G. J. M. Hanselaar, and P. C. M. de Wilde. Hue-saturation-density (HSD) model for stain recognition in digital images from transmitted light microscopy. *Cytometry*, 39(4):275–284, Apr. 2000.

[6] M. D. Zeiler. ADADELTA: An adaptive learning rate method. arXiv:1212.5701v1 [cs.LG], 2012.

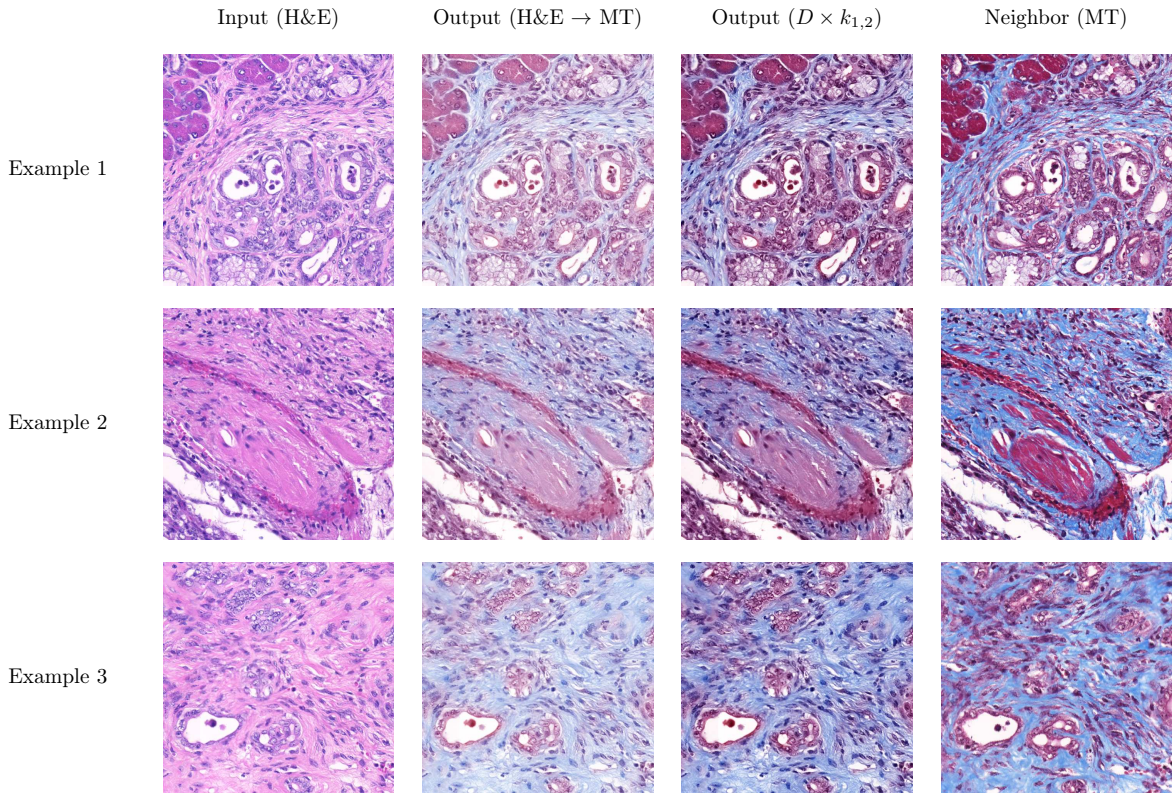

Figure 4: Close-up of the results with the high-resolution images. First column: input (H&E). The output simulates the MT stain but it only has the chromatic components. Second column: output with the density component of the input. Third column: output with density multiplied by a constant factor $k_{1,2}$. Last column: real MT stain of the neighboring slice for comparison.

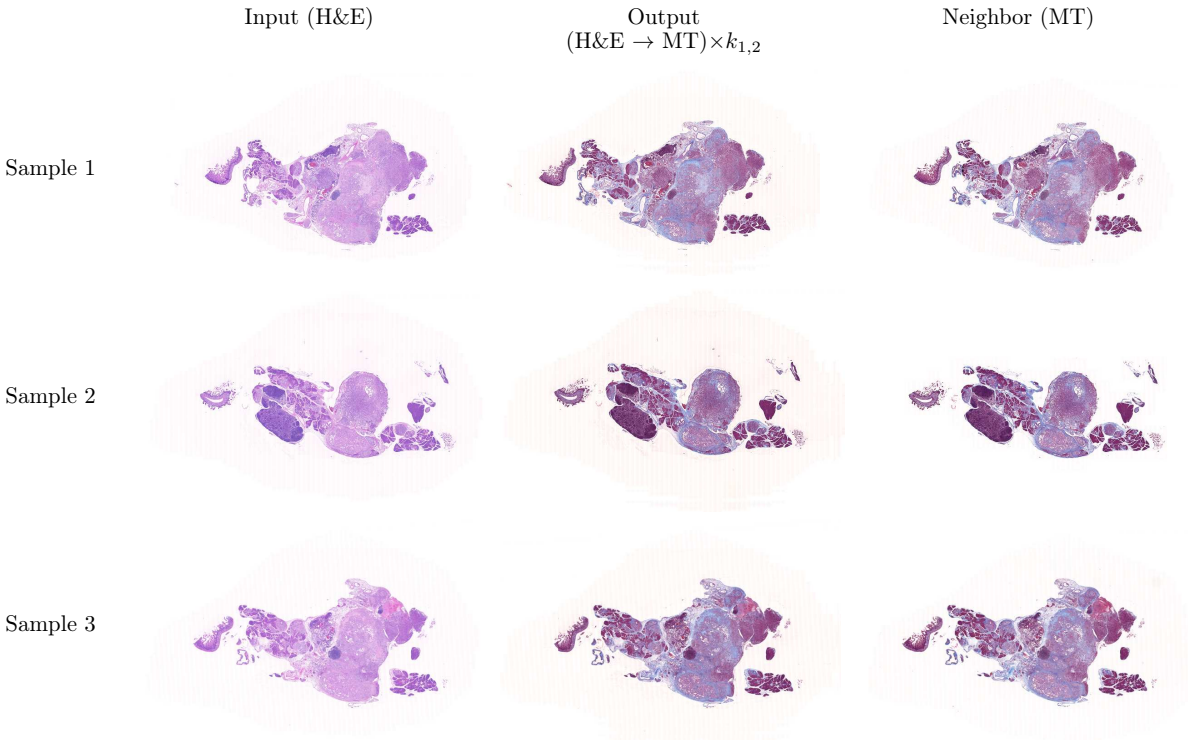

Figure 5: The whole specimen results with the low-resolution images. Sample 3 is a different specimen from Sample 1 and 2. First column: input (H&E). Second column: output (simulated MT) with the density component of the input multiplied by a constant factor $k_{1,2}$. Last column: real MT neighboring slice for comparison.

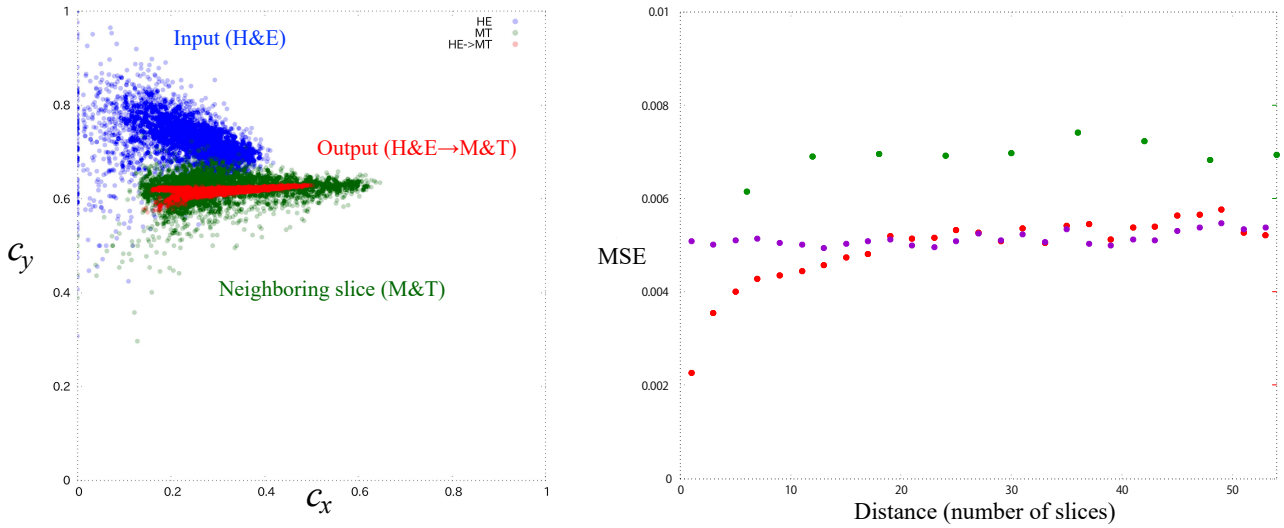

Figure 6: **Left:** Choromatic distributions of the input, the output, and the neighboring slice of the Example 1 in Fig. 4. **Right:** Mean Squared Error in (c_x, c_y) between the output simulated MT stain and real MT stained nearby slices. Red: between the reference MT slice and the simulated MT stains. Green: between the reference and other real MT slices. Magenta: between the reference MT slice and the simulated MT stains..