Seam Removal for Patch-Based Ultra-High-Resolution Stain Normalization

Chi-Chen Lee*, Chi-Han Peng*
*National Yang Ming Chiao Tung University (NYCU), Taiwan
{ww888225433.ai10, pengchihan}@nycu.edu.tw

***Abstract*—Stain normalization is a key computational method in pathology that transforms histological stain images of one style to another. Modern methods are mostly based on neural image-to-image translation techniques. For very large image inputs, which are common in practice (e.g., whole slide images (WSIs)), the inferences are forced to run multiple times, each on a different subset of the image, due to GPU memory constraints. To minimize the differences between different outputs, several modifications [1], [2] of the standard instance-normalization (IN) layers have been proposed. Despite the reduced color variances, visible seams remain even with these approaches, which are disruptive to histologists when they closely examine the stitched results. These seams are also detrimental to the performance of some downstream tasks such as tumor classification. Hence, we propose a novel approach to effectively remove the seams by utilizing a Pix2Pix [3]-based neural network and an alpha blending-based post-processing step. Tested on real-world medical and natural image datasets, we found that our method performed much better than traditional Poisson image editing-based seam removal approaches. Our approaches qualitatively (in terms of the visibility of the seams) and quantitatively improved the results by prior stain normalization methods by large margins.**

I. Introduction

Histologists are experts in pathology who examine human tissues in the form of very thin layers under microscopes to study the manifestations of disease. To assist visual inspection, the tissues are stained by colorful chemical compounds, of which a common choice is hematoxylin and eosin (H&E). There are actually a large variety of histological stain images. H&E stain images alone have great color variances due to different setups of the image acquisition processes, such as tissue fixation duration, compositions of the H&E stains, or scanner settings [4]. Other stain techniques include immunohistochemistry (IHC) staining that uses enzyme or fluorophore to highlight antibody-antigen interactions. This large variety of styles of histological stain images makes human inspections and downstream computational tasks (such as tumor classification [6], [8]) more difficult to perform [5].

To mitigate the style variances problem, *stain normalization* methods are developed to transform stain images of one style to another. A typical application is to transform various stain images to a common reference color space. Another application is to transform images done by one kind of staining process to one done by another. The latter is especially useful as some stain procedures are more expensive or difficult to conduct. Modern stain normalization methods are now mostly deep learning-based that leverage neural image-to-image translation / style transfer techniques [4]–[7]. Note that the problems are mostly *unpaired* because paired stain image training samples are difficult to acquire in practice (since the staining of one tissue is irreversible).

A key challenge for neural network-based stain normalization is how to process the digitized tissue samples, come as whole slide images (WSIs), which are often of very high resolutions. WSIs can be hundreds of times bigger than the size the neural network can handle due to GPU memory constraints. The standard way is to subdivide the WSIs into smaller patches that each can be processed by the neural network, and then stitch the individual outputs together to form the final result. Unfortunately, the stitched results can have glaring "tiling artifacts" that each patches have different hues and contrasts. Recent solutions for this issue modified the instance normalization (IN) layers such that different patch's IN layers would have a common mean and variance ("Thumbnail Instance Normalization" (TIN) [2]) or spatially smoothly varied means and variances ("Kernelized Instance Normalization"(KIN) [1]). These methods significantly reduced the color variances between the patches. However, after close examinations (same as what histologists would do with histological stain images), we found that *visible seams remain at the patch boundaries* even after these methods were applied. See Figure 1 for examples. These seams can be visually disruptive to histologists' diagnosis. Moreover, we found that they are detrimental to a common downstream task of histological stain images by comparing the task's quantitative performance on regions near the patch boundaries and regions away from them.

We tried traditional seam removal techniques based on Poisson image editing but found that they could not satisfactorily remove the seams. Therefore, we propose an effective solution as follows. Our method first utilized a Pix2Pix [3]-based neural network to synthesize novel image contents on regions near the patch boundaries. As there may be new discontinuities between original images and the synthesized contents, we further applied an alpha-blending process with spatially varying weights. Note that our method is a novel application of neural image-to-image translation techniques (such as Pix2Pix) as, to our best knowledge, there were no existing papers that used such techniques for image seam removal purposes.

Tested on a real-world medical dataset and a natural image dataset (same as in KIN [1]), we found that our method worked

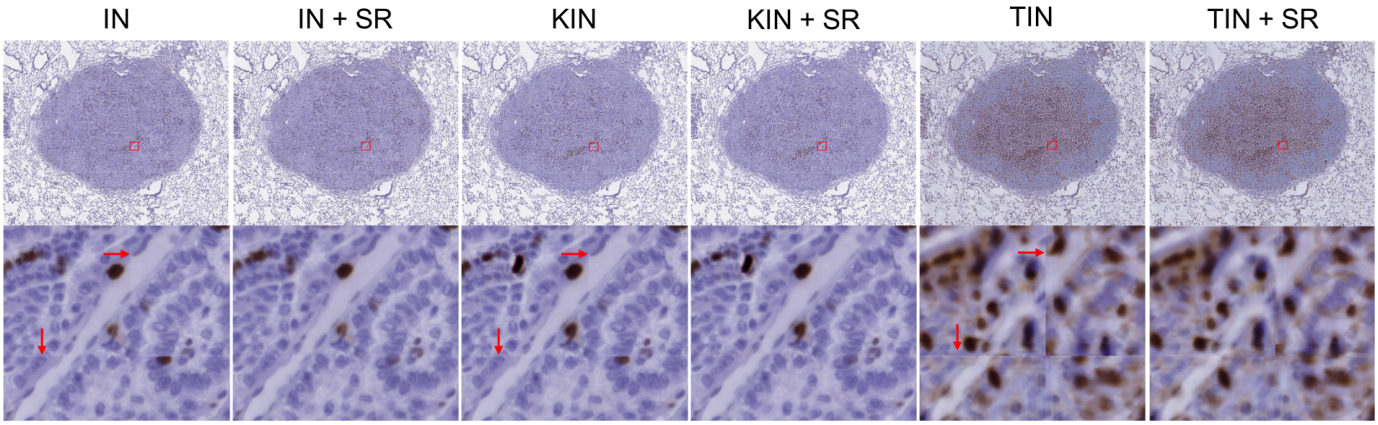

Fig. 1. From left to right, we show patch boundaries of stain normalization results by standard instance normalization (IN) of CycleGAN [17], Kernelized IN [1] (with CycleGAN), and Thumbnail IN (TIN) [2], respectively, before and after our seam removal technique is applied. Red arrows point to locations of the patch boundaries.

very well qualitatively - the seams became unnoticeable in most cases, and quantitatively improved stain normalization results by KIN according to Fréchet inception distances (FID). Furthermore, we found that our method benefited a key downstream task that take histological stain images as inputs. Our contributions are summarized as follows:

- We identified a key shortcoming of current stain normalization methods for large image inputs - namely the visible seams between patches. We also quantitatively demonstrated that the visible seams are detrimental to some downstream task.
- We propose a novel method that can effectively remove the seams and consequently improved the results of current stain normalization methods quantitatively and qualitatively. To our best knowledge, our method is the first to use neural image-to-image translation techniques for seam removal applications.
- Our method is simple and easy to implement, and can be readily plugged into any existing patch-based stain normalization methods for large image inputs.

II. Related Work

A. Image Processing-Based Stain Normalization Methods

Recall that a major goal for stain normalization methods was to convert stain images to a common reference color space. Early work demonstrated that even just converting all stain images to grayscale could be beneficial to certain downstream tasks in medicine [9], [10]. Latter approaches aimed for transforming to color spaces. [13] was a classic algorithm by conducting statistical analysis in an alternative, non-RGB color space. Deconvolutional methods found a transform function that maps ground-truth source and target image pairs and used the function to transform unseen source images. The source and target images were first compressed into a small "stain vectors" for the transform function to deduct from. These vectors were chosen manually [14], by conducting SVD [11], or by segmentation and clustering [12]. Latter methods attempted to match distributions of each color channels by various approaches such as variational Bayesian Gaussian mixture models [15] and histogram normalization of the colormaps [16].

B. Deep Learning-Based Stain Normalization Methods

Stain normalization problems are well suited to be solved by unpaired image-to-image translation methods. We focus our discussion on the neural networks they used. StainGAN [5] leveraged cycle-consistent adversarial networks (CycleGAN) [17] to transfer H&E stain images from one scanner style to another (i.e., Hamamatsu Nanozoomer 2.0-HT to Aperio Scanscope XT). [4] and [18] used GANs with disentangled feature presentations (e.g., DRIT++ [20]) to perform various stain normalization tasks. In Marini et al. [19], a novel convolutional neural network (CNN)-based approach was proposed to tackle highly heterogeneous images. Some recent methods took a teacher-student model in which their neural networks were trained on *paired* images synthesized by a prior (unpaired) method. For example, Stain-Net [6] trained a simple network of 1x1 windows to conduct color normalization in a pixel-to-pixel manner based on paired examples synthesized by StainGAN. It inferences much faster than StainGAN but is less accurate (measured by FID). Lee et. al [7] later proposed a similar U-Net-based approach that is more accurate than Stain-Net while being slightly slower. In KIN [1], authors used CycleGAN, Contrastive Unpaired Translation (CUT) [21], and LSeSim [22] to synthesize virtual immunohistochemistry (IHC) stain images (which are more expensive to obtain in real world) from H&E stain images. In [23], user studies showed that synthesized stain images helped to improve the diagnosis of kidney diseases by real pathologists.

Dealing with ultra-high resolution images. To address the tiling artifacts of patch-based methods, in [24], the authors proposed using a sliding window to create left-to-right smooth transitions of the means and variances of the IN layers of each patch's neural network (they used CycleGAN). Later,

authors of TIN [2] proposed a simple solution of using a shared global mean and variance for all patches. While this approach effectively reduced inter-patch color differences and is memory efficient, the downside is that the results may be over/under-colorized. Recently in KIN [1], a convolution was used to spatially smooth the the means and variances of the patches' IN layers. Their approach shared the same advantages of TIN but avoided the over/under-coloring problem. Note that all these methods aimed to reduce the differences of color spaces of patches only, and visible seams may remain at the patch boundaries at close examinations. Therefore, our goal is to practically remove the seams.

C. Seam Removal / Image Stitching Methods

Trivial Laplacian smoothing-based approaches to remove seams between two patches, e.g., averaging out colors of adjacent pixels from different patches, would not work for our problem (it is erroneous to assume Laplacians are zero at patch boundaries, causing blurred lines).

In panoramic photography and remote sensing imaging, image stitching methods [26], [27] work to combine multiple images with overlapping fields of view. The classical methods to compute new image contents at the overlapping regions are Poisson image editing [25]-based, which compute new color values by solving Poisson equations to modify the values of one image to fit the Laplacians of another. For our problem, the patches do not overlap so these methods could not be used directly. Nevertheless, in Section IV-B, we tried an alternative way to subdivide the WSIs into overlapping patches so that Poisson-based methods can be used. We found the results to be worse than ours.

III. Method

Our method has two stages - a Pix2Pix-based step to largely remove the seams (Section III-A) followed by an alpha blending-based step (Section III-B) to mitigate the slight discontinuities caused by the previous step. We go through the details next.

A. Pix2Pix-Based Seam Removal

We begin with notations. We assume the stain normalization task transforms a WSI image of style $Stain_a$ to another style $Stain_b$ using an unpaired style transfer method such as CycleGAN [17], CUT [21], or LSeSim [22]. We denote the method as $Method$. In practice, $Method$ is run separately on each subdivided patch of the whole WSI and the inference outputs are stitched together to form the final result. We denote the input WSI as WSI_a and the stitched result as WSI_b. Note that the mechanisms in KIN [1] or TIN [2] may be used to coordinate the statistics parameters of the IN layers of each $Method$ inference to reduce the color differences. Still, visible seams are likely to remain at the patch boundaries.

Our goal is to prepare training data to train a *paired image-to-image translation* method (we used Pix2Pix [3] and found it to work quite well) that will be used to run on regions of WSI_b near the patch boundaries, such that the synthesized image content would be similar to the original image but with the seams removed. For this goal, the training data for the Pix2Pix should be two versions of regions of WSI_b near the patch boundaries, one with the seams (source style) and one without the seams (target style). How to generate such training data? We come out the idea of running $Method$ again on WSI_a but with a horizontally and vertically shifted patch subdivision scheme such that the original patch boundaries would fall into the new patches' interiors. We denote the new result WSI_{b+}. We can then prepare the training data as the quadrangulation (i.e., subdivision into rectangles) of regions of WSI_b near the patch boundaries and their counterparts in WSI_{b+} at the same locations. Namely, the former are the source-style images with seams and the latter are the target-style images without seams. We explain the technical details in Figure 2.

Finally, after the Pix2Pix is trained, we run it on each subdivided rectangle of the regions near the patch boundaries in WSI_b and replace them by the inference outputs.

B. Alpha Blending Step

Recall that the Pix2Pix is trained on pairs of images which are subsets of stain normalization results by a particular $Method$ that one overlaps patch boundaries and another doesn't. Therefore, in principle the paired images should differ only by the existence of the visible seams. In practice, the trained Pix2Pix may still make slight changes to the whole image input. This let to new visible discontinuities between the regions processed by Pix2Pix and the rest. To minimize the discontinuity, we developed a simple alpha-blending strategy to blend the the Pix2Pix outputs and the original WSI_b contents by assigning a weight to Pix2Pix-generated content that is inversely proportional to the L1 distance to the boundaries of the Pix2Pix-applied regions. The weighting scheme is shown in Figure 3.

IV. Results

We tested our method on a Linux computer with AMD Ryzen 5 5600g CPU, NVIDIA GeForce RTX 3090 Ti GPU, and 24GB rams. We used the default settings (Adam optimizer, 200 epochs) to train the Pix2Pix. For a typical input WSI of 7895 pixels by 5586 pixels, we collected 722 paired images (550 long rectangles, 126 squares, and 46 small rectangles) as training examples to train the Pix2Pix. The training took 5000 seconds for such a WSI. The inferences of Pix2Pix to generate the new contents for regions near the patch boundaries took about 165 seconds in total. The computational cost of the alpha-blending step is neglectable. Next, we describe the datasets we tested on.

A. Datasets

Automatic Non-rigid Histological Image Registration (ANHIR) challenge [28]. This dataset consists of high-resolution whole slide images (WSIs) containing multiple types of tissue samples, including lesions, lung lobes, breast tissues, colon adenocarcinoma, etc. These images are organized in sets of consecutive tissue slices that are stained using

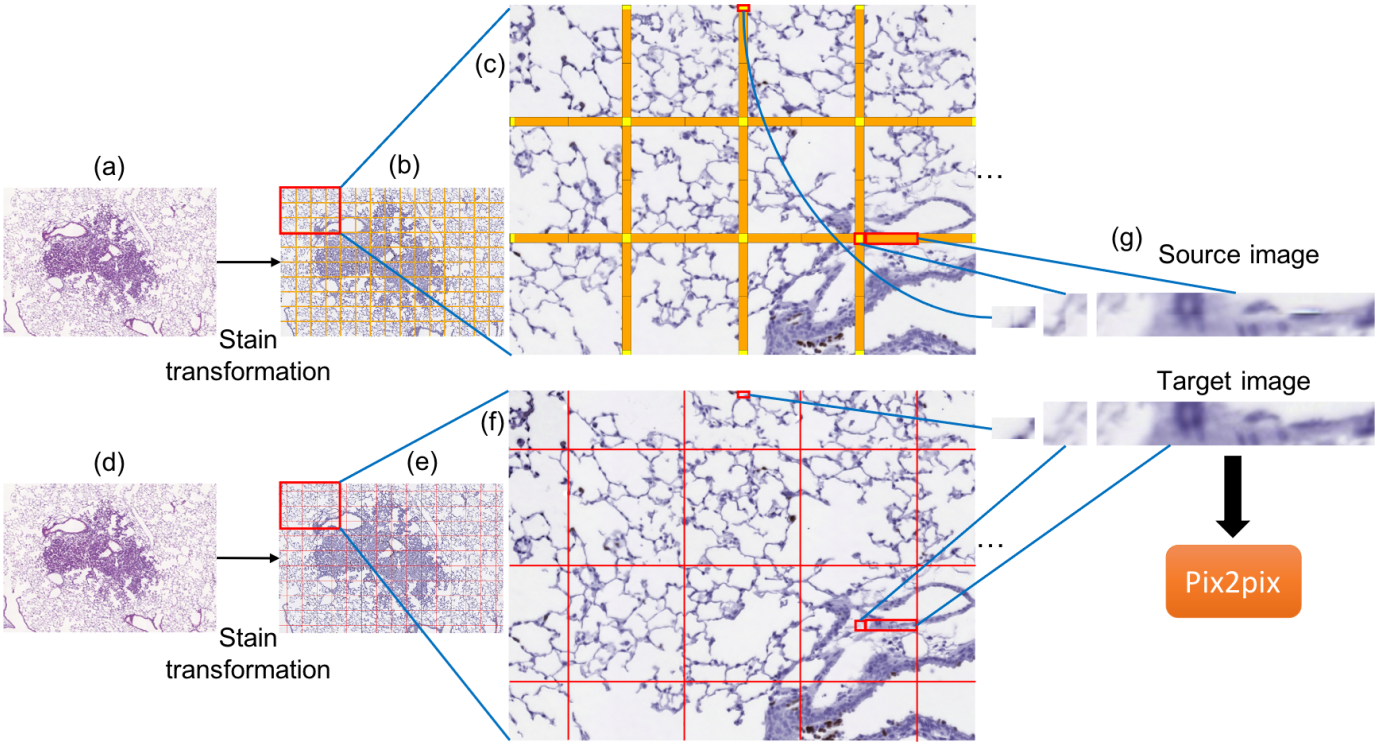

Fig. 2. (a to c) shows the original patch-based stain normalization pipeline to transform a given WSI (a) to another style (b) by running a style transfer method (e.g., CycleGAN, CUT, or LSeSim) on each of the subdivided patch. (c) shows a close-up view of the patches and the quadrangulation of the regions near the patch boundaries. The regions are defined by a "dilation size" along the patch boundaries. By default we use a dilation size of 20 pixels for patches of 512 by 512 pixels (we conducted ablation studies of other sizes in Section IV-F). The quadrangulation turns the regions into long rectangles in either horizontal or vertical directions (along the patch edges), squares (at the 4-way junctions between patches), and small vertical or horizontal rectangles (at the WSI boundaries). Note that all these quadrangulated polygons overlap some original patch boundaries. The long rectangles are further cut into two for the reason we see next. (d to f) shows a second run of the patch-based stain normalization pipeline but with a new patch subdivision scheme which is the original scheme shifted horizontally and vertically by half of the patch size. Now, for every quadrangulated polygon in the original stain normalization result, we can find a counterpart in the second run's result at the same location. (g) We collect the pairs of images to train a Pix2Pix.

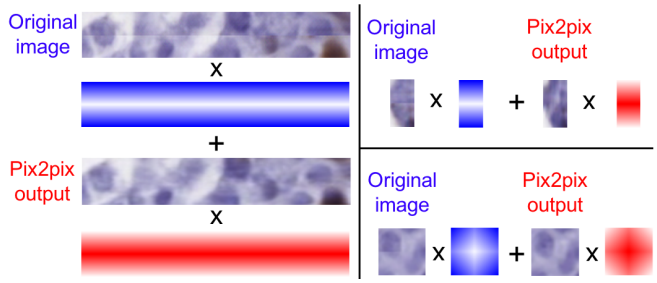

Fig. 3. For each of the three types of quadrangulated polygons, we show the weightings of the original image contents and the Pix2Pix outputs in blue and red, respectively.

various dyes, such as H&E, Ki-67, ER/PR, CD4/CD8/CD68, among others. Following the testing procedures in KIN [1], we selected lung lesion WSIs as our dataset. There are two H&E stain images with resolutions of 8899x7328 and 7895x5586 pixels and two corresponding Ki-67 stain images with resolutions of 8915x7336 and 7898x5573 pixels in this category. Note that the H&E and corresponding Ki-67 images are not pixel-to-pixel matched. For both pairs of H&E / Ki-67 images, we ran existing stain normalization pipelines (i.e., IN, KIN, TIN) to get a synthetic Ki-67 image. We then trained a Pix2Pix on the first image pair only and tested the trained Pix2Pix on the second (for quantitative comparisons) and both (for qualitative observations).

Kyoto summer2autumn dataset. To verify the effectiveness of our method, we utilized an additional dataset of natural images provided by the authors of KIN [1]. The dataset consists of two unpaired sets of high resolution (3456x5184 pixels) images, namely 17 and 20 photos taken in Kyoto during the summer and autumn seasons, respectively. The stain normalization pipeline in KIN was then used to transfer the summer images to winter. For our experiments, we split the dataset into a testing set of 4 summer images and the rest as the training set (to train the Pix2Pix).

B. Poisson Image Editing-Based Method

For comparison, we implemented a Poisson image editing [25]-based approach for seam removal. As the traditional method worked on overlapping areas only (recall that it optimizes for new pixel values that are based on the values of the first image but shall have the Laplacians of the second image), we run the stain normalization pipeline again but with a shifted patch subdivision scheme so that the original

	ANHIR_CycleGAN		ANHIR_CUT		ANHIR_LSeSim	
	FID ↓	Corr. ↑	FID ↓	Corr. ↑	FID ↓	Corr. ↑
IN	89.71	0.9814	76.09	0.9931	87.46	**0.9988**
IN + PB	97.41	0.9807	99.73	0.9900	89.26	0.9871
IN + SR	**85.09**	**0.9833**	**73.74**	**0.9932**	**86.72**	**0.9988**
KIN	123.50	0.9914	102.12	0.9969	93.71	**0.9990**
KIN + PB	123.49	0.9905	103.25	**0.9971**	103.25	0.9984
KIN + SR	**114.83**	**0.9916**	**97.18**	**0.9971**	**90.53**	0.9989
TIN	**200.48**	**0.9865**	**174.62**	0.9989	228.52	**0.9891**
TIN + SR	208.12	0.9864	177.63	0.9989	**225.80**	0.989

TABLE I
QUANTITATIVE RESULTS FOR THE ANHIR DATASET USING VARIOUS STYLE TRANSFER METHODS AND WITH THE STANDARD IN LAYERS (IN), KIN [1], OR TIN [2], TO SYNCHRONIZE THE NEURAL NETWORK INFERENCES FOR EACH PATCHES. +PB MEANS USING POISSON BLENDING AND +SR MEANS USING OUR METHOD FOR SEAM REMOVAL

Kyoto_CycleGAN	FID↓
IN	152.17
IN + SR	**150.33**
KIN	177.48
KIN + SR	**176.61**
TIN	192.37
TIN + SR	**184.64**

TABLE II
QUANTITATIVE RESULTS ON THE KYOTO DATASET.

patches and the new patches are overlapping. In this way, the standard Poisson-based blending scheme can be applied on the overlapping areas. As reported in the next sub-sections, the results are worse than our method.

C. Quantitative Comparisons

We quantitatively evaluated our method's impact on the quality of patch-based stain normalization results with the standard IN layers, the kernelized IN layers in KIN [1], and the mean/variance-averaged IN layers in TIN [2]. Following the experiment designs in previous work (TIN and KIN), we chose *Fréchet Inception Distance (FID)* and *histogram correlation* for testing. These are commonly used metrics to evaluate the performance of generative models by measuring the similarity between two sets of images (e.g., gt and synthesized) by either utilizing a pre-trained neural network (Inception) or by comparing the distributions of histograms. To sum up, our method significantly improved the results by standard IN layers and the kernelized IN layers. Our method did not quantitatively affect the results by TIN much. This may be because TIN introduced significant blurring artifacts to the original style transfer results, making the Pix2Pix harder to train. The results are shown in Table II. Histogram correlations could not be calculated on the Kyoto dataset because the source and target images have no correspondences.

D. Qualitative Comparisons

In Figure 4 and Figure 5, we show visual comparisons of regions near the patch boundaries before and after our seam removal method, and by the Poisson blending-based (PB) method, for stain normalization results by various kinds of pipelines (IN, KIN, TIN) based on three different style transfer techniques (CycleGAN, CUT, and LSeSim). Our findings are three folds. First, we found that visible seams *really* are a problem for all kinds of pipelines we tested. Second, we found our method worked effectively to remove the seams and meanwhile did not introduce noticeable discontinuities between processed regions and the rest. Third, we found the Poisson-based method were less effective at removing the seams. In Figure 6, we show results on the Kyoto dataset.

	Interior	Boundary	Boundary + SR
Accuracy:	0.813±0.006	0.803±0.007	0.806±0.007

TABLE III
PERFORMANCE OF A TUMOR CLASSIFICATION METHOD [6] ON REGIONS OF STAIN NORMALIZATION RESULTS (IN+CYCLEGAN) AWAY FROM THE PATCH BOUNDARIES ("INTERIOR"), NEAR THE PATCH BOUNDARIES ("BOUNDARY"), AND BOUNDARY REGIONS AFTER OUR SEAM REMOVAL METHOD WAS APPLIED.

E. Downstream Task Experiment

Following the testing procedures in StainNet [6], we leveraged the public Camelyon16 dataset [8] to train and test a tumor classification neural network. We used the classification neural network trained in StainNet (using a SqueezeNet [29], which was trained on 40000 256x256 patches subtracted from 170 WSIs from the Radboud University Medical Center and tested on 10000 256x256 patches subtracted from 50 WSIs from the University Medical Center Utrecht. They labeled the patches as normal or abnormal (i.e., containing tumors) for training and testing. We further separate the testing patches into "interior" and "boundary" (i.e., touching the boundaries of the 512x512 subdivided patches to run stain normalization) ones. In other words, the boundary patches would contain visible seams while the interior ones would not. We then tested the classification method on whole-WSI stain normalization results (generated using per-patch CycleGANs). Our findings are two folds. First, we found that the classification neural network did performed worse on the boundary patches then on the interior ones. This indicated that the visible seams are indeed detrimental to the classification neural network. Second, we found that the classification performance got slightly improved after our seam removal method was applied. See Table III for testing results. In Figure 7, we show examples of normal and abnormal (have tumors) patches and their seam-removal results in stain normalization results.

F. Ablation Studies

In Table IV (a-c), we show results of using different sizes (dilation width) for the regions processed by Pix2Pix. Note that by default we use a dilation size of 20 (width is 20+20). We found that using a smaller size (10) would led to less effective seam removal, while using sizes larger than 20 led to little differences but came with a higher computational cost. In Table IV (d), we show results of only doing the Pix2Pix step (skip the alpha-blending step). We found that there are slight discontinuities at the boundaries of regions processed by Pix2Pix and this is reflected in the worse performance. In Figure 8, we show qualitative examples for ablation studies.

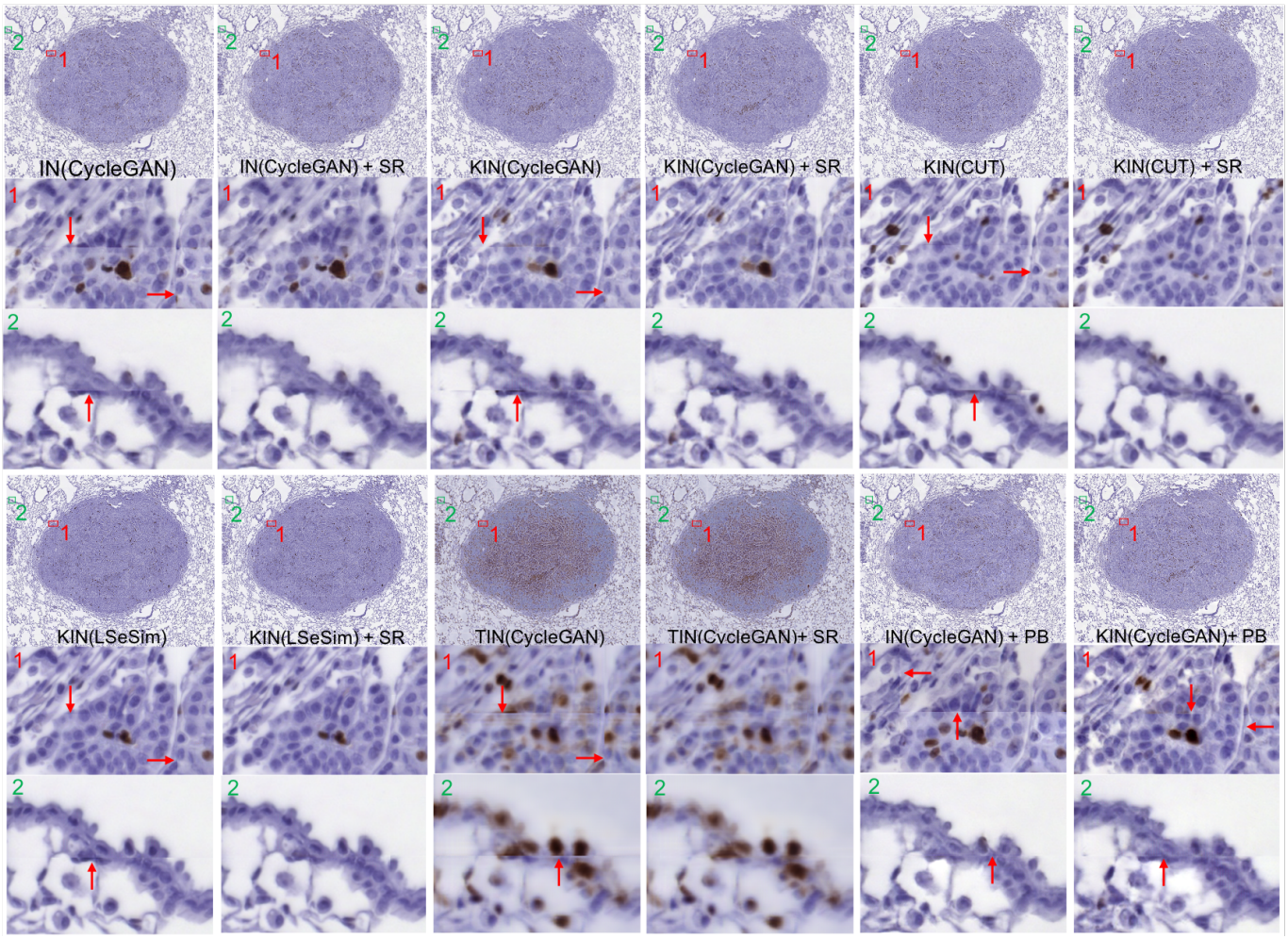

Fig. 4. Qualitative results for the ANHIR dataset. SR means our seam removal method. PB means Poisson blending.

	FID↓	Corr.↑
No SR	89.71	0.9814
(a) SR (width 10+10)	92.40	0.9822
(b) SR (width 20+20)	85.09	0.9833
(c) SR (width 40+40)	85.26	0.9846
(d) SR (Pix2Pix only)	87.66	0.9839

TABLE IV
ABLATION-STUDY RESULTS. WE SHOW PERFORMANCES OF USING DIFFERENT SIZES FOR THE REGIONS PROCESSED BY PIX2PIX (A-C) AND PERFORMANCE OF DOING ONLY THE PIX2PIX STEP (D).

V. CONCLUSION

To sum up, our method can be understood as a modern take of the traditional Poisson image editing-based approach to blend color contents of overlapping images: instead of relying on solving Laplacian or gradient-based optimization problems, we leveraged powerful image-to-image translation neural networks (e.g., Pix2Pix) to learn the non-trivial conversion of regions near visible seams to the seamless counterparts. This is also a novel use of such neural techniques for seam removal applications. We found our method worked well qualitatively - not only the seams were removed, but also no additional discontinuities were introduced between the processed parts and the original images thanks to the alpha-blending post-processing step. The quantitative benefits of our method are verified in two ways. First, the FID and histogram correlation of stain normalization results both got improved after being processed by our method. Second, for a key downstream task (tumor classification), the worse performance on regions near patch boundaries (compared to the other parts) got improved after our method. We consider the main disadvantage to be the relatively high inference costs of the Pix2Pix. This is because we ran Pix2Pix on each of the rectangles separately. For future work, we would like to try to improve the performance by minimizing the numbers of inference calls.

Acknowledgement This work is funded by the National Science and Technology Council of Taiwan (project number 111R10286C).

REFERENCES

[1] Ming-Yang Ho, Min-Sheng Wu, and Che-Ming Wu, "URUST: Ultra-high-resolution unpaired stain transformation via Kernelized Instance

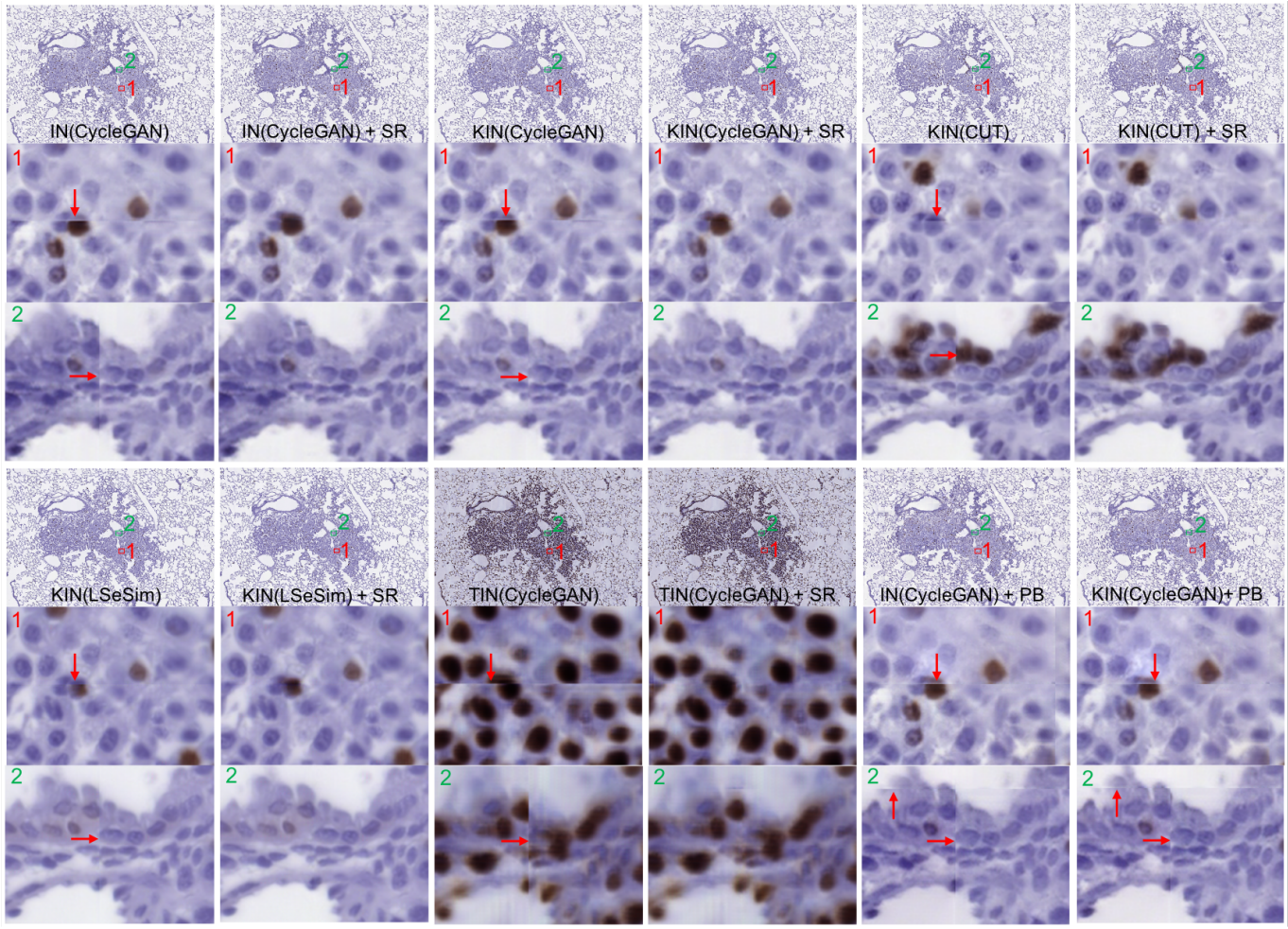

Fig. 5. More qualitative results for the ANHIR dataset.

Normalization," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.

[2] Chen, Zhe, Wang, Wenhai, Xie, Enze, Lu, Tong, and Luo, Ping, "Towards Ultra-Resolution Neural Style Transfer via Thumbnail Instance Normalization," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022.

[3] Isola, Phillip, Zhu, Jun-Yan, Zhou, Tinghui, and Efros, Alexei A, "Image-to-Image Translation with Conditional Adversarial Networks," in *CVPR*, 2017.

[4] Sophia J. Wagner, Nadieh Khalili, Raghav Sharma, Melanie Boxberg, Carsten Marr, Walter de Back, and Tingying Peng, "Structure-Preserving Multi-domain Stain Color Augmentation Using Style-Transfer with Disentangled Representations," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, 2021, pp. 257-266.

[5] M. Tarek Shaban, Christoph Baur, Nassir Navab, and Shadi Albarqoun, "Staingan: Stain Style Transfer for Digital Histological Images," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2017.

[6] Hongtao Kang, Die Luo, Weihua Feng, Shaoqun Zeng, Tingwei Quan, Junbo Hu, and Xiuli Liu, "StainNet: A Fast and Robust Stain Normalization Network," in *Frontiers in Medicine*, 2021.

[7] C. Lee, P. Kuo, and C. Peng, "HE Stain Normalization using U-Net," in *2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering (BIBE)*, 2022, pp. 29-32.

[8] Ehteshami Bejnordi, Babak, Veta, Mitko, Johannes van Diest, Paul, van Ginneken, Bram, Karssemeijer, Nico, Litjens, Geert, van der Laak, Jeroen A. W. M., and the CAMELYON16 Consortium, "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer," in *JAMA*, vol. 318, no. 22, 2017, pp. 2199-2210.

[9] P.W. Hamilton, P.H. Bartels, D. Thompson, N.H. Anderson, R. Montironi, and J.M. Sloan, "Automated Location of Dysplastic Fields in Colorectal Histology using Image Texture Analysis," in *Journal of Pathology*, 1997.

[10] Antonio Ruiz, Olcay Sertel, Manuel Ujaldon, Umit Catalyurek, Joel Saltz, and Metin Gurcan, "Pathological Image Analysis Using the GPU: Stroma Classification for Neuroblastoma," in *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2007.

[11] Marc Macenko, Marc Niethammer, J. S. Marron, David Borland, John T. Woosley, Xiaojun Guan, Charles Schmitt, and Nancy E. Thomas, "A Method for Normalizing Histology Slides for Quantitative Analysis," in *IEEE International Symposium on Biomedical Imaging*, 2009.

[12] Massimo Salvi, Nicola Michielli, and Filippo Molinari, "Stain Color Adaptive Normalization (SCAN) algorithm: Separation and standardization of histological stains in digital pathology," in *Computer Methods and Programs in Biomedicine*, 2020.

[13] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley, "Color Transfer between Images," in *IEEE Engineering in Medicine and Biology Magazine*, 2001.

[14] Arnout Ruifrok and Dennis A Johnston, "Quantification of Histochemical Staining by Color Deconvolution," in *Analytical and Quantitative Cytology and Histology*, 2001.

[15] Derek Magee, Darren Treanor, Doreen Crellin, Michael Shires, Katherine Smith, Kevin Mohee, and Philip Quirke, "Colour Normalisation in Digital Histopathology Images," in *Proc Optical Tissue Image analysis in Microscopy, Histopathology and Endoscopy*, 2009.

[16] Sonal Kothari, John H. Phan, Richard A. Moffitt, Todd H. Stokes, Shelby

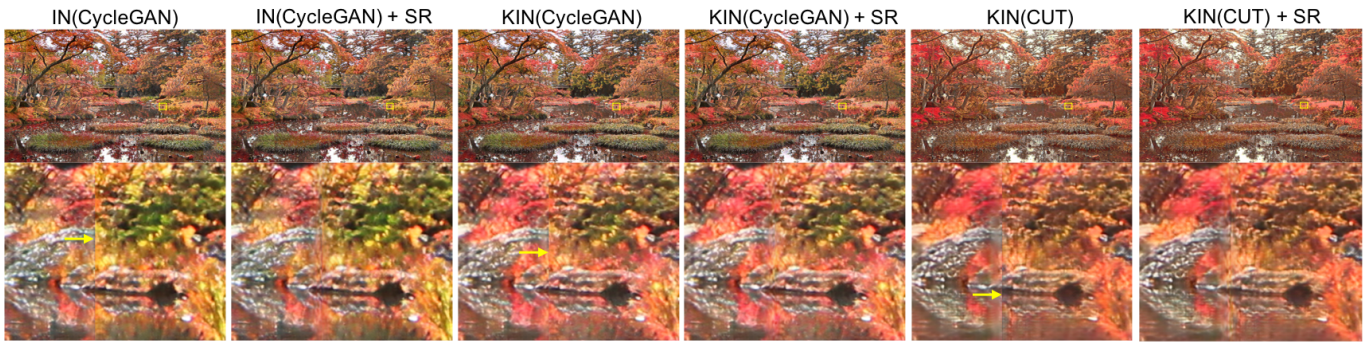

Fig. 6. Qualitative results for the Kyoto dataset.

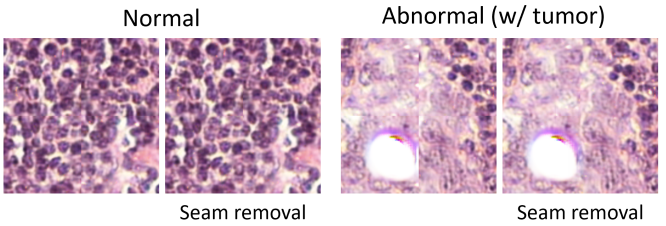

Fig. 7. . The left two are normal (no tumors) 256x256 patches near patch boundaries before and after our seam-removal method. The right two are abnormal (w/ tumors) ones. They are taken from a stain normalization result by CycleGAN.

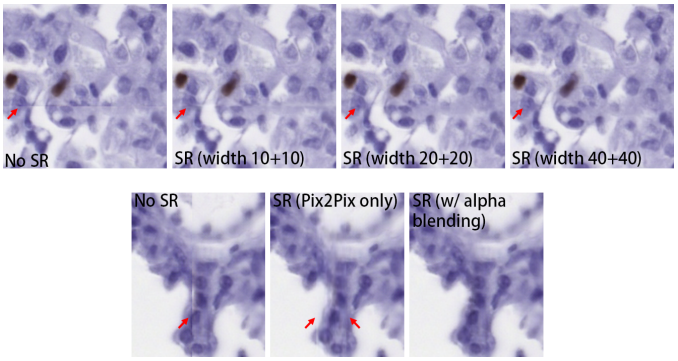

Fig. 8. Ablation-study comparisons of using different sizes for the regions processed by Pix2Pix (top) and doing only the Pix2Pix step (bottom).

E. Hassberger, Qaiser Chaudry, Andrew N. Young, and May D. Wang, "Automatic Batch-Invariant Color Segmentation of Histological Cancer Images," in *IEEE Int. Symposium on Biomedical Imaging*, 2011.

[17] Zhu, Jun-Yan, Park, Taesung, Isola, Phillip, and Efros, Alexei A, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," in *IEEE International Conference on Computer Vision (ICCV)*, 2017.

[18] Atefeh Ziaei Moghadam, Hamed Azarnoush, Seyyed Ali Seyyedsalehi, and Mohammad Havaei, "Stain transfer using Generative Adversarial Networks and disentangled features," in *Computers in Biology and Medicine*, 2022.

[19] Niccol'o Marini, Manfredo Atzori, Sebastian Ot´alora, Stephane Marchand-Maillet, and Henning M¨uller, "H&E-Adversarial Network: a Convolutional Neural Network to Learn Stain-Invariant Features through Hematoxylin and Eosin Regression," in *IEEE International Conference on Computer Vision (ICCV)*, 2021.

[20] Hsin-Ying Lee, Hung-Yu Tseng, Jia-Bin Huang, Maneesh Kumar Singh, and Ming-Hsuan Yang, "Diverse Image-to-Image Translation via Disentangled Representations," in *European Conference on Computer Vision (ECCV)*, 2018.

[21] Taesung Park, Alexei A. Efros, Richard Zhang, and Jun-Yan Zhu, "Contrastive Learning for Unpaired Image-to-Image Translation," in *European Conference on Computer Vision (ECCV)*, 2020.

[22] Zheng, Chuanxia, Cham, Tat-Jen, and Cai, Jianfei, "The Spatially-Correlative Loss for Various Image Translation Tasks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2021.

[23] K. de Haan, Y. Zhang, J.E. Zuckerman, T. Liu, A.E. Sisk, M.F. Diaz, K.Y. Jen, A. Nobori, S. Liou, S. Zhang, et al., "Deep learning-based transformation of HE stained tissues into special stains," in *Nature communications*, 2021.

[24] Amal Lahiani, Jacob Gildenblat, Irina Klaman, Shadi Albarqouni, Nassir Navab, and Eldad Klaiman, "Virtualization of tissue staining in digital pathology using an unsupervised deep learning approach," in *European Congress on Digital Pathology*, 2019.

[25] Patrick Pérez, Michel Gangnet, and Andrew Blake, "Poisson Image Editing," in *ACM SIGGRAPH 2003 Papers*, 2003, pp. 313–318.

[26] Wei LYU, Zhong ZHOU, Lang CHEN, and Yi ZHOU, "A survey on image and video stitching," in *Virtual Reality Intelligent Hardware*, 2019.

[27] Shiqi Liu, Jie Lian, Xuchen Zhan, Cong Liu, Yuze Tian, and Hongwei Duan, "Automatically eliminating seam lines with Poisson editing in complex relative radiometric normalization mosaicking scenarios," in *Arxiv*, 2021.

[28] Jiří Borovec, Jan Kybic, Ignacio Arganda-Carreras, Dmitry V. Sorokin, Gloria Bueno, Alexander V. Khvostikov, Spyridon Bakas, Eric I-Chao Chang, Stefan Heldmann, Kimmo Kartasalo, Leena Latonen, Johannes Lotz, Michelle Noga, Sarthak Pati, Kumaradevan Punithakumar, Pekka Ruusuvuori, Andrzej Skalski, Nazanin Tahmasebi, Masi Valkonen, Ludovic Venet, Yizhe Wang, Nick Weiss, Marek Wodzinski, Yu Xiang, Yan Xu, Yan Yan, Paul Yushkevich, Shengyu Zhao, and Arrate Muñoz-Barrutia, "ANHIR: Automatic Non-Rigid Histological Image Registration Challenge," in *IEEE Transactions on Medical Imaging*, vol. 39, no. 10, 2020, pp. 3042-3052.

[29] Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, and Kurt Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and ¡0.5MB model size," in *arXiv*, 2016.