2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)
April 13-16, 2021, Nice, France

2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) | 978-1-6654-1246-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/ISBI48211.2021.9433860

TEXTURE ENHANCED GENERATIVE ADVERSARIAL NETWORK FOR STAIN NORMALISATION IN HISTOPATHOLOGY IMAGES

Cong Cong[1], *Sidong Liu*[2,3], *Antonio Di Ieva*[3], *Maurice Pagnucco*[1], *Shlomo Berkovsky*[2], *Yang Song*[1]

[1]School of Computer Science and Engineering, University of New South Wales, Australia
[2]Centre for Health Informatics, Macquarie University, Australia
[3]Computational NeuroSurgery Lab, Macquarie University, Sydney, Australia

ABSTRACT

Digitised histopathology image analysis has drawn researchers' attention over recent years. However, *stain variation* due to several factors can be a significant hurdle for the diagnosis process. *Stain normalisation* can be used as an effective method to address this issue but most existing methods require careful selection of a reference image. In this work, we propose a texture enhanced pix2pix generative adversarial network (TESGAN), which takes higher contrast hematoxylin components as input and includes a novel loss function to guide the generator to produce higher quality images without the need for reference images. We implement our method as a pre-processing approach for an isocitrate dehydrogenase (IDH) mutation status classification task. Evaluated on The Cancer Genome Atlas (TCGA) glioma cohorts, the proposed model achieves Area Under Curve (AUC) of 0.967, which substantially outperforms the current state-of-the-art.

Index Terms— Stain normalisation, Conditional Generative Adversarial Networks, Content loss, IDH classification.

1. INTRODUCTION

Cell appearances and structures in tissue slides provide valuable information for diagnosing and studying diseases. To make tissue structures more recognisable, a proper staining process is required. The most widely used stain process in histopathology is the Hematoxylin and Eosin (H&E) stain [1] with which the hematoxylin stains cell nuclei blue-purple and eosion stains cell cytoplasm pink.

Traditional stain normalisation methods can be divided into two categories: global colour normalisation and stain separation. Global colour normalisation aims to match the colour distribution of the source domain to that of the target domain, based on histograms [2] or statistics of colour channels [3]. Stain separation based methods normalise images by mathematical frameworks using decomposed stain vectors from a reference image. For example, separate stain vectors have been identified by mapping RGB images to the optical density space using singular value decomposition [4, 5]. In addition, a relevance vector machine is used to classify each pixel to a stain component [6], but its supervised learning fashion increases the computation complexity.

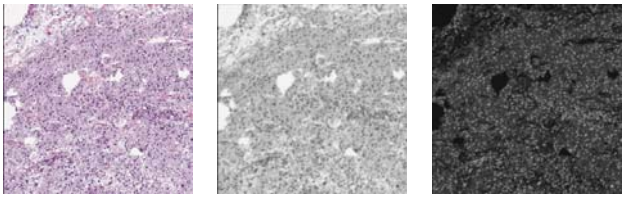


Fig. 1: Left to right: original colour image, grayscale image and hematoxylin component of the image.

Recent stain normalisation methods are increasingly using deep learning models. For instance, sparse auto-encoders are employed to decompose stain components into separate feature spaces and build tissue specific correspondence between input images and a single reference image [8]. However, the normalisation results depend heavily on the reference image's quality. Generative adversarial network (GAN) based methods using CycleGAN and pix2pix [7] have thus been proposed, which do not need a reference image and achieve excellent results. CycleGAN is used to formulate stain normalisation as an unsupervised image-to-image translation task, such as StainGAN [9] and cCGAN [10]. However, if the dissimilarity between two domains is large, especially when the tissue appearance can be vastly different in regions, CycleGAN-based methods can produce less accurate results.

Pix2pix-based methods generally produce better results than CycleGAN-based methods as they use paired images as input and are trained in a supervised fashion. The STST method [11] is based on pix2pix and treats stain normalisation as an image repainting problem. Target domain images and their converted grayscale images are used as paired images to train a pix2pix GAN in which the generator learns the colour pattern of the target domain and repaints the input grayscale image with a similar colour style. However, grayscale conversion may cause irreversible loss of stain information in the input images, leading to low contrast images and poor performance in pathology detection and classification.

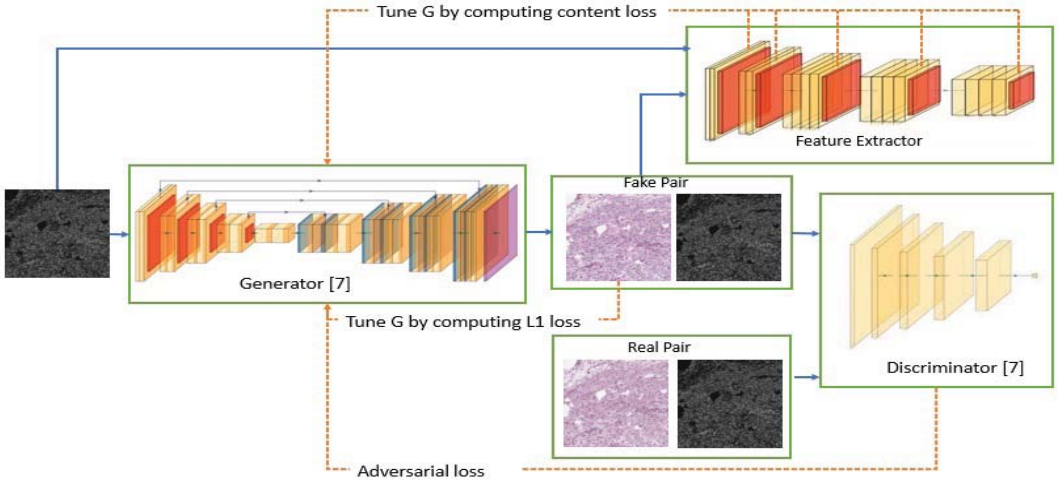


Fig. 2: Overall TESGAN model structure. Similar to [7], U-Net generator transforms hematoxylin components to images with target domain colour pattern. PatchGAN [7] discriminator judges the input image pairs as real or fake. Feature vectors produced by a pre-trained VGG19 network are used to derive the content loss.

In this paper, we propose a texture enhanced pix2pix-based method (TESGAN) to generate high quality stain normalised images without the need to select reference images. Inspired by [12], we use the hematoxylin component of the H&E stained images as the generator input since the contrast between nuclei and cytoplasm in hematoxylin components is higher than in grayscale images (Fig. 1). Moreover, we define a custom loss function which regulates the generated images to preserve both high-level and low-level features of the input images. We then apply our method to Isocitrate Dehydrogenase (IDH) classification in histopathology images. IDH is an important diagnostic, prognostic and therapeutic biomarker for glioma [13]. Our contributions include: 1) using a template-free method based on the pix2pix framework for generating texture-enhanced stain normalised images; 2) using the hematoxylin component as the model inputs to demonstrate its effectiveness in histopathology image analysis; and, 3) combining content loss and $\mathcal{L}_1$ loss to enforce the correct colour pattern and enhance the contrast in the produced images. Experiments were conducted on 921 glioma cases collected from the public TCGA dataset. Our results show promising performance improvement over the baseline [14] and state-of-the-art stain normalisation methods.

2. METHODS

We formulate the stain normalisation problem as re-painting of the hematoxylin component with the colour pattern of images from the target domain. The proposed method is illustrated in Fig. 2. The base structure of TESGAN is similar to pix2pix. We introduce an additional feature extractor component using a pre-trained VGG19 model to derive the content loss. A trained generator is used to produce stain normalised images. We then train a ResNet50 classifier for the task of IDH classification using the normalised images.

2.1. Stain Normalisation

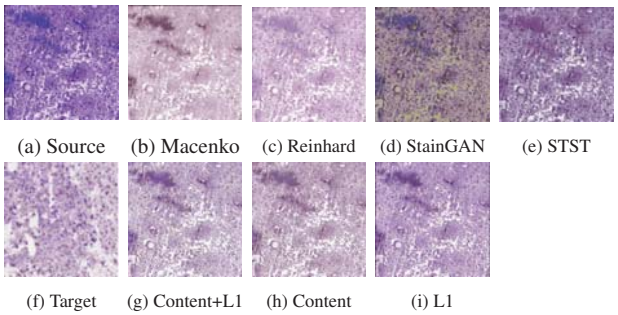
Our TESGAN framework is designed based on the pix2pix architecture, which takes paired images (x, y) to train the generator, with x as the input, y as the target output and $G(x, z)$ as the generated image (z is the noise vector). To train the generator network, both adversarial loss and $\mathcal{L}_1$ loss are used. The adversarial loss encourages the generator to produce images that fool the discriminator and the $\mathcal{L}_1$ loss forces the generator to produce images that are close to the target image y.

In our approach, we do not have a reference or target image to normalise an input image. However, to utilise the supervised learning capability of pix2pix, we use the colour image from the target domain as the target y and the corresponding hematoxylin component h of the target image as the input x to train the network. In this way, our approach is essentially to repaint the hematoxylin component with the colour pattern of the target domain. Specifically, U-Net is used as the generator, which takes the hematoxylin component h of the target domain RGB images y as input and produces stain normalised images $G(h, z)$. PatchGAN [7] is used as the discriminator, which takes the paired input of h and stained images, and classifies whether the given stained images are real or fake.

Furthermore, unlike the objective function in pix2pix, we add a content loss to enhance texture features in the generator output. The overall loss function is defined as follows:

$$\begin{aligned}\mathcal{L}_{TESGAN}(G, D) = \mathbf{E}_{h,y}[\log(D(h, y))]+ \\ \mathbf{E}_h[\log(1 - D(h, G(h, z))] + \mathcal{L}_{Content\_L1}.\end{aligned} \quad (1)$$

The first part of the loss function is a similar adversarial loss



(a) Source (b) Macenko (c) Reinhard (d) StainGAN (e) STST

(f) Target (g) Content+L1 (h) Content (i) L1

Fig. 3: Comparison of stain normalised results.

to most GANs and the second part is a $\mathcal{L}_{Content\_L1}$ loss that combines $\mathcal{L}_1$ loss and content loss. $\mathcal{L}_1$ loss encourages the produced image to be less blurry with richer low-level features, such as texture and colour. Content loss helps preserve high-level features, such as structural information [15]. The formula for $\mathcal{L}_{Content\_L1}$ is as follows:

$$\mathcal{L}_{content\_L1}(G) = \lambda_{L1}\mathcal{L}_1(G) + \lambda_{content}\mathcal{L}_{Content}(G), \quad (2)$$

$$\mathcal{L}_1(G) = \mathbf{E}_{h,y,z}[||y - G(h,z)||], \quad (3)$$

$$\mathcal{L}_{content}(G) = \sum_{j}^{n} \omega_j \frac{1}{C_j H_j W_j} ||\phi_j(h) - \phi_j(G(h,z))||. \quad (4)$$

To calculate the content loss, a feature extractor is added to the model to extract feature vectors from $G(h,z)$ and h. We use a pre-trained VGG19 network as a feature extractor. Here we use the feature map ϕ_j produced from the convolution layer j before each max pooling layer; C_j, H_j, W_j correspond to the number of channels, height and width of ϕ_j, while ω_j is the weight factor of ϕ_j.

2.2. Hematoxylin Component Extraction

Classifiers usually produce better results with images that have enhanced texture features. As can be observed from Fig. 1, the hematoxylin component of the image enhances the texture features with a better contrast compared to RGB and grayscale images. Thus, we extract the hematoxylin component by estimating the optical density of each RGB channel [16]. With the help of the proposed loss function, the content information of the input hematoxylin component is preserved, which enables us to produce texture-enhanced images.

3. EXPERIMENTS AND RESULTS

3.1. Dataset and Implementation

In this study, the publicly available whole-slide image (WSI) dataset from The Cancer Genome Atlas (TCGA) [17] Lower Grade Glioma (LGG) and Glioblastoma Multiforme (GBM) cohorts, which consist of 921 glioma patients, is used to develop and test the model. All the TCGA WSIs have been labelled as either IDH wildtype or mutant based on immunohistochemistry and/or genetic sequencing. We randomly split the dataset into training, validation and test sets in a 80:10:10 ratio. We further partition each slide into 1024 $\times$ 1024-pixel patches at 10$\times$ magnification and remove the patches that have less than 50% tissue content. A more detailed description of the dataset and patch selection can be found in [14]. In total, there are 17,686 patch images in the training set, 2,233 in the validation set and 2,310 in the test set, respectively.

To train our method, we need to specify the source domain and target domain. Here, we assume that slides from the same tissue source site (TSS) are uniform in staining. We thus group our dataset into various TSS groups and select the one with the largest number of images as our target domain and the rest as the source domain. Under this setting, we have 3,424 images in the target domain.

Both the stain normalisation model and ResNet50 classifier are trained for 30 epochs with the ADAM optimiser. The ResNet50 classifier is initialised with the ImageNet pre-trained model. The initial learning rate is 0.0002, which is reduced by half if the loss of validation set does not decrease for 5 epochs. We set λ_{L1} as 0.25, $\lambda_{content}$ as 0.75 and ω_j as 0.2. The models are developed in PyTorch and trained using NVIDIA Tesla P100 GPUs.

Table 1: Classification results with different stain normalisation methods.

	Reinhard [3]	Macenko [4]	StainGAN [9]	STST [11]	TESGAN
R	0.825	0.800	0.826	0.865	**0.885**
PR	0.825	0.791	0.876	0.918	**0.939**
F1	0.815	0.838	0.878	0.891	**0.911**
AUC	0.912	0.914	0.917	0.918	**0.967**

3.2. Experimental Results

Classification performance is tested using the same classifier architecture trained with images normalised by different methods and results are shown in Table 1. The results are evaluated based on recall (R), precision (PR), F1-score (F1) and AUC. TESGAN outperforms all the benchmarked baselines. Moreover, sample stain normalised images of our method and other stain normalisation approaches are shown in Figs. 3b–3e. The quality of images produced by traditional methods, such as Macenko [4] and Reinhard [3], depends on the choice of reference images. Our method requires no reference images and produces images with colour pattern visually closer to the target domain. It can also be seen that STST [11] and TESGAN outperform StainGAN [9], possibly due to the fact that pix2pix-based methods use paired inputs whereas CycleGAN-based methods use unpaired inputs;

Table 2: Classification results with different loss functions.

	Non-normalised	L1	Content	Content+L1
R	0.720	0.816	0.857	**0.885**
PR	0.892	0.918	0.889	**0.939**
F1	0.797	0.857	0.792	**0.911**
AUC	0.908	0.918	0.878	**0.967**

therefore CycleGAN results degrade if the two domains differ substantially. In addition, compared with STST [11] which uses grayscale images as generator input, the normalised images produced by TESGAN demonstrate better contrast and higher colour similarity with the target images, hence indicating the advantages of using the hematoxylin component.

We further evaluate the classification results of stain normalised data using the TESGAN model with different loss functions. From Table 2, it can be seen that using stain normalisation is effective and all metrics are superior to non-normalised data. Figs. 3g–3i show that $\mathcal{L}_1$ loss alone helps preserve the colour pattern of target images and content loss alone helps increase the contrast but the colour feature is not fully preserved. By combining these two losses, the produced images not only preserve the colour pattern of the target image but also show clearer tissue morphology.

The current state-of-the-art in IDH classification [14] has an AUC of 0.938, which focuses on GAN-based data augmentation. In our experiments, we used the same dataset as [14] with the same patch samples and achieved an improved AUC of 0.967 without excessive data augmentation. This further demonstrates the advantage of our method and importance of texture features for histopathology image analysis.

4. CONCLUSIONS

In this paper, we propose an unsupervised stain normalisation model based on the pix2pix framework. The proposed TESGAN model takes the hematoxylin component of the image as the paired input and repaints it to match the stain style in the target domain. Our evaluation results show that our model produces higher quality images that have high colour similarity with the target domain. Our method is applied to IDH mutation status classification and shows improved performance over the prior art.

5. COMPLIANCE WITH ETHICAL STANDARDS

A public dataset was used and no ethical approval was required.

6. ACKNOWLEDGEMENTS

This study is supported in part by Australian Research Council (FT190100623) and Australian National Health and Medical Research Council (1160760). The authors have no other relevant financial or non-financial interests to disclose.

7. REFERENCES

[1] A. H. Fischer, J. K. A. *et al.*, "Hematoxylin and eosin staining of tissue and cell sections," *CSH protocols*, vol. 2008, 2008.

[2] D. Coltuc, P. Bolon *et al.*, "Exact histogram specification," *IEEE TIP*, vol. 15, no. 5, pp. 1143–1152, 2006.

[3] E. Reinhard, M. Adhikhmin *et al.*, "Color transfer between images," *IEEE Trans. Comp. Graph. App.*, vol. 21, no. 5, pp. 34–41, 2001.

[4] M. Macenko, M. Niethammer *et al.*, "A method for normalizing histology slides for quantitative analysis," in *ISBI*, 2009, pp. 1107–1110.

[5] M. T. McCann, J. Majumdar *et al.*, "Algorithm and benchmark dataset for stain separation in histology images," in *ICIP*, 2014, pp. 3953–3957.

[6] A. M. Khan, H. El-Daly *et al.*, "A gamma-gaussian mixture model for detection of mitotic cells in breast cancer histopathology images," in *ICPR*, 2012, pp. 149–152.

[7] P. Isola, J. Y. Zhu *et al.*, "Image-to-image translation with conditional adversarial networks," in *CVPR*, 2017, pp. 1125–1134.

[8] A. Janowczyk, A. Basavanhally *et al.*, "Stain normalization using sparse autoencoders (StaNoSA): application to digital pathology," *CMIG*, vol. 57, pp. 50–61, 2017.

[9] M. T. Shaban, C. Baur *et al.*, "StainGAN: Stain style transfer for digital histological images," in *ISBI*, 2019, pp. 953–956.

[10] Z. Xu, C. F. Moro *et al.*, "GAN-based virtual re-staining: a promising solution for whole slide image analysis," *arXiv preprint arXiv:1901.04059*, 2019.

[11] P. Salehi and A. Chalechale, "Pix2pix-based stain-to-stain translation: A solution for robust stain normalization in histopathology images analysis," *arXiv preprint arXiv:2002.00647*, 2020.

[12] B. Zhao, X. Chen *et al.*, "Triple U-net: Hematoxylin-aware nuclei segmentation with progressive dense feature aggregation," *MedIA*, vol. 65, p. 101786, 2020.

[13] H. Yan, D. W. Parsons *et al.*, "IDH1 and IDH2 mutations in gliomas," *New England journal of medicine*, vol. 360, no. 8, pp. 765–773, 2009.

[14] S. Liu, Z. Shah *et al.*, "Isocitrate dehydrogenase (IDH) status prediction in histopathology images of gliomas using deep learning," *Scientific reports*, vol. 10, no. 1, pp. 1–11, 2020.

[15] J. Johnson, A. Alahi *et al.*, "Perceptual losses for real-time style transfer and super-resolution," in *ECCV*, 2016, pp. 694–711.

[16] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution," *AQCH*, vol. 23, no. 4, pp. 291–299, 2001.

[17] K. Clark, B. Vendt *et al.*, "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," *Journal of digital imaging*, vol. 26, no. 6, pp. 1045–1057, 2013.