DEEP LEARNING-BASED VIRTUAL SPECIAL STAINING OF H&E STAINED TISSUE SECTIONS

by

Fatma Nur Kınalı

B.S., in Biomedical Engineering, Erciyes University, 2018

Submitted to the Institute of Biomedical Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Biomedical Engineering

Boğaziçi University
2023

ACKNOWLEDGMENTS

First and foremost, I would like to gratitude my parents Gulsum Kınalı, Ahmet Kınalı, and my brother Serhat Kınalı for their constant support, motivation and patience. My achievement in completing this study would not have been possible without their tireless support of them have stood by me throughout my studies and during the research.

I would like to thank my advisor, Assist. Prof. Dr. Mehmet Turan for his invaluable guidance, support, and inspiring encouragement throughout my entire master's degree. I would like to thank the pathologists involved in this study for their valuable support, guidance in evaluation of the results and the thesis committee for their valuable time and comments.

Finally, I thank the Scientific and Technological Research Council of Turkey (TUBITAK) for the 2210 Scholarship Program, the TUBITAK 2232 International Outstanding Researcher Fellowship, and to TUBITAK Ulakbim for the Turkish National e-Science e-Infrastructure (TRUBA) cluster and data storage services.

ACADEMIC ETHICS AND INTEGRITY STATEMENT

I, Fatma Nur Kınalı, hereby certify that I am aware of the Academic Ethics and Integrity Policy issued by the Council of Higher Education (YÖK) and I fully acknowledge all the consequences due to its violation by plagiarism or any other way.

Name : ______________________________

Signature: ______________________________

Date: ______________________________

ABSTRACT

DEEP LEARNING-BASED VIRTUAL SPECIAL STAINING OF H&E STAINED TISSUE SECTIONS

Hematoxylin and eosin (H&E) staining, which is standardly applied to tissues in histopathological diagnosis, is an incredible tissue staining method that reveals the morphological features of tissues and cells and gives an idea about their biomolecular structures. However, a more accurate determination of tissue and cell structure and biomolecular components is essential for the diagnosis of diseases under the light microscope. For this reason, histochemical, immunohistochemical, immunofluorescent and genetic techniques have been applied to tissues and cells and it has been aimed to reach the closest diagnosis to the truth. However, the number of these methods is quite high and depending on the method, they can be very complex and time-consuming in terms of procurement, cost, time requirement and application. Manual preparation of each is an expensive and labor-intensive process that requires complex and difficult methods of using many chemicals, and ready-to-use kits are often costly. Here, we propose a novel virtual staining tool that transforms H&E-stained tissue images into specially stained versions in just a few minutes. Virtual staining of some of the frequently used methods in daily pathology practice with a transformation that is derived through a learning process, using H&E slide as a basis, will make the use of these methods in the histopathological diagnosis process incredibly practical, inexpensive, simple and easy to apply in a short time. As well as, we propose a novel StainKid dataset of stomach and kidney tissue samples stained with a wide collection of histological stains. The StainKid dataset can make a significant contribution to the development of computer aided diagnosis in histopathology by paving the way for new artificial intelligence-based virtual staining techniques.

Keywords: Histological Stain, Special Stain, Virtual Staining, Whole Slide Image, Generative Adversarial Network.

ÖZET

TEZİN TÜRKÇE BAŞLIĞI

Histopatolojik tanı sürecinde dokulara standart olarak uygulanan H&E boyama, doku ve hücrelerin morfolojik ozelliklerini ortaya çıkaran, biyomoleküler yapıları hakkında fikir veren kullanışlı bir boyama yontemidir. Ancak, hastalıkların teşhisi için dokunun, hücre yapısının ve biyomoleküler bileşenlerin ışık mikroskobu altında incelenerek daha doğru belirlenmesi şarttır. Bu nedenle doku ve hücrelere histokimyasal, immünohistokimyasal, immünofloresan ve genetik teknikler tek başına veya kombinasyon halinde modifiye edilmiş birçok yontem uygulanmış ve gerçeğe en yakın tanıya ulaşılması hedeflenmiştir. Ancak, bu yontemlerin sayısı oldukça fazladır ve yonteme bağlı olarak tedariği zor, maliyetli, zaman alıcı, ve uygulama açısından çok karmaşık olabilmektedir. Her birinin manuel olarak hazırlanması ve birçok kimyasalın kullanılması açısından; emek gerektiren, karmaşık, yoğun ve zor bir sureçtir. Ayrıca, kullanılan hazır kitler ve genel olarak sürecin tamamı maliyetlidir. Bu çalışmada, H&E boyalı doku kesit görüntülerini dakikalar içinde ozel histokimyasal boyalı versiyonlarına donüştüren derin oğrenme temelli bir sanal boyama aracı oneriyoruz. Günlük patoloji pratiğinde sık kullanılan bazı yontemlerin H&E boyalı doku kesitlerinden derin oğrenmeye dayalı donüşümü ile sanal olarak boyanması, bu yontemlerin histopatolojik tanı surecinde kullanımını inanılmaz derecede hızlı, pratik, ucuz, basit, kolay ve uygulanabilir hale getirecektir. Ayrıca bu çalışmada, makine oğreniminin kullanıldığı sanal boyama yaklaşımlarının geliştirilmesindeki boşluğu doldurmak için ozel olarak hazırlanmış; H&E, Massonun trikromu, Periyodik asit schiff, Jones methenamine silver, Congo kırmızısı, Toluidin mavisi dahil olmak üzere geniş bir histopatolojik boya koleksiyonuyla boyanmış goruntulerden oluşan StainKid veri setini sunuyoruz.

Anahtar Sözcükler: Histolojik Boyama, Özel Boyama, Sanal Boyama, Tüm Slayt Goruntusu, Üretken Çekişmeli Ağ.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

L	Loss
τ	Temperature
$\hat{z}$	Feature
σ_{relu}	Rectified Linear Unit Function
$\sum$	Summation

LIST OF ABBREVIATIONS

H&E	Hematoxylin and Eosin
WSI	Whole Slide Image
MT	Massons Trichrome
PAS	Periodic Acid Schiff
JMS	Jones Methenamine Silver
GAN	Generative Adversarial Network
AI	Artificial Intelligence
FFPE	Formalin-Fixed Paraffin-Embedded
CUT	Contrastive Unpaired Translation
CycleGAN	Cycle-Consistent Adversarial Network
SAB	Spatial Attention Block
ReLU	Rectified Linear Unit
MLP	Multi-Layer Perceptron
FID	Frechet Inception Distance
VTT	Visual Turing Test

1. INTRODUCTION

Histological evaluation of human tissue slides stained with hematoxylin and eosin (H&E) is considered the gold standard for the diagnosis of diseases for which tissue or cell sampling can be performed. Histologic staining is an essential part of routine histopathology as it offers insight into shape, pattern, and tumor structure, tissue, and cells by imparting contrast and color to the tissue sample, which can then be analyzed either using a bright-field microscope or on a whole slide image (WSI) of the tissue acquired with a digital WSI scanner. Tissue samples are affixed to glass slides and treated with H&E staining [6]. Hematoxylin stains the nuclei blue and eosin stains the extracellular matrix and cytoplasm pink [7]. H&E is a simple to perform and inexpensive stain that is preferred worldwide as a routine stain [8],[9],[10].

The use of H&E staining alone is not always sufficient in routine pathological practice to differentiate all tissue components and cellular structures, and showing the distribution of different chemicals [11]. Therefore, there are a variety of other histologic stains, called special stains, particularly histochemical methods that are used to identify or better highlight certain features, components, and entities that are important in the diagnosis of abnormalities or disease. Special stains are considered essential for the final diagnosis as they provide additional information or confirm the diagnosis [12],[13],[14]. For instance, Masson's trichrome (MT) highlights collagen and muscle fibers by staining them blue, while staining muscle fibers red and cytoplasm pink. Periodic Acid-Schiff (PAS) is used to assess tissue structures that carry a large amount of carbohydrate macromolecules (e.g., glycogen, glycoprotein, proteoglycans, glycolipids, and mucins) [15]. For example, connective tissue, renal tubule cells, intestinal brush borders, glycocalyx, mucus, basal laminae, and reticular fibers of connective tissue [16]. After PAS performed on tissues, the carbohydrate molecules are stained purple or magenta color [17]. Jones Methenamine Silver (JMS), stains basement membranes black by its property of oxidizing carbohydrate components to aldehydes [18]. JMS is a great agent for detecting glomerular and tubular basement membranes and is therefore

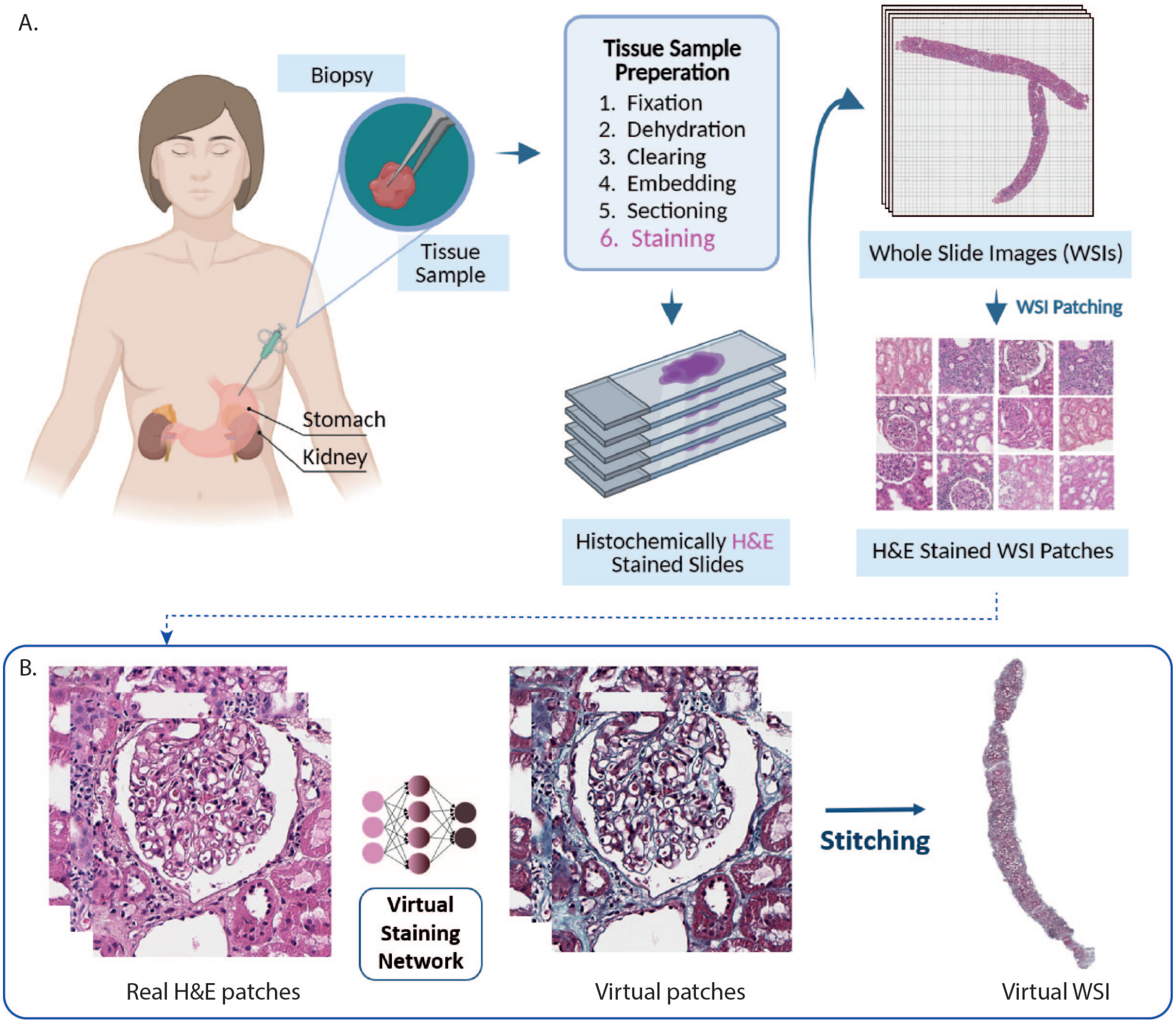

Figure 1.1 Workflow Overview. The diagram showing how the virtual staining method fits into the current histopathology workflow. **A.** A small kidney or stomach tissue specimen is taken from patient in biopsy. After the tissue preparation steps, WSIs are generated using a digital slide scanner, then NxN patches are generated from each WSI to give as input to the virtual staining networks for the different stain transformations. **B.** The real H&E stained patches generated from the kidney and stomach tissue samples given as an input to the virtual staining network, then corresponding virtually stained patches are given as output of the generator. At the final step, the outputted virtually stained patches are stitched together for review by pathologists by quickly ensuring detailed information about specific features, components, and entities for final diagnosis.

routinely used by nephropathologists for renal biopsies [19],[20]. Another histochemical stain example is toluidine blue (TB) which stains acidic tissue components and is therefore a useful agent for staining tissue components that have high DNA and RNA content [21]. Mast cell granules, mucins, and cartilage are some of the tissue types that can be stained with TB. The use of TB can be used to diagnose dysplasia, carcinoma in the oral cavity, and allergic inflammation [22],[23], [24]. Congo red (CR) stains amyloids that are caused by the aggregation of abnormal proteins in different organs, thus also known as amyloid [25]. We have summarized the details of all stains used in this study in the Table 1.1.

While special stains improve the visibility of tissue structures, making them easier to identify, the traditional histology staining process in histopathology is lengthy, labor-intensive, and costly, consuming a significant amount of time. It requires a serious histopathology laboratory organization and equipment. Tissue samples which can be limited first should be taken from patients, fixed in formalin and embedded in paraffin blocks. After the embedding, paraffinized tissue blocks are sectioned into 2-10 μm thin slides using a microtome and placed on glass slides before staining [26],[27]. If more than one stain will be used, multiple tissue sections should be sliced and separate processes must be applied for each tissue section [28]. The special staining is a more comprehensive procedure than H&E staining, it requires more effort in safety, quality control and validation of tissue steps which must be followed by a histotechnologist [29]. The fact that the procedure requires expertise, extra effort and time causes costs to rise that might be a burden on the healthcare system. Particularly, if pathologists need special staining after H&E stained tissue examination all the processes are repeated. This incurs additional costs in addition to prolonging the time for diagnosis. Prolongation of diagnosis can be burdensome to patients when a diagnosis is immediately needed. In addition, performing the staining in different centers and by different clinicians creates an obstacle to the standardization of the histopathological diagnosis process.

To address these concerns, several methods have been developed that can be used in the workflow in place of conventional histopathology staining. Early methods aimed to get rid of tissue preparation steps and provide a slide-free method for rapid imaging of intact tissue by using various microscopy techniques such as optimized light sheet microscopy [30], ultraviolet surface excitation [31], and nonlinear microscopy [32]. Recently, the application of deep learning in histopathology has gained attention. The computational pathology community has developed Generative Adversarial Networks (GANs) based methods known as virtual staining to create artificially stained tissue images that reflect the characteristics of the target stain [33]. Different imaging techniques (e.g. autofluorescence microscopy [34],[35], hyperspectral microscopy [36], quantitative phase contrast [37] and multi-photon microscopy [38]), have been used in the literature with the aim of virtually stain unstained tissue images. Approaches mentioned above are based on supervised training and require image pairs to avoid hallucinations occurs on tissue images while training neural network to align input data and corresponding output data using adjacent tissue sections [39] or restaining the same tissue sections [40],[41]. Nevertheless, collecting a perfectly aligned registered dataset in these ways is a laborious and complicated task in clinics. The second alternative demands extensive knowledge and skill in antibody responses and long-term effects on tissues, while the first alternative includes excessive usage of tissue materials. In addition, both choices may cause artifacts such as wrinkles and tissue deformations, and can interfere with neural network training. Besides, the need for a special apparatus (e.g., quantitative phase contrast microscopy, autofluorescence microscopy with DAPI attachment) prior to using virtual staining conflicts with the idea of widespread usage of virtual staining methods for all pathology laboratories.

Using deep learning to transform patches and WSIs of H&E-stained tissue sections into special stain versions is another effective approach. Pathologists usually apply H&E staining first and use it for initial diagnosis. Biopsy tissue samples may be limited, reducing the number of possible stains that can be applied to a tissue sample and limiting pathologists in analyzing tumor structure. By transforming the stain via deep learning, you can save the tedious additional tissue sections and physical staining processes that are required for other stains and extend over the next few working days. In this way, the tissue sample is preserved for further examination and no unnecessary or additional biopsy is required on the patient. Transforming stains to each other using AI reduces histopathology laboratory operating costs, by requiring fewer personnel, materials, and resources. It also reduces staining preparation time and increases staining consistency. Deep learning-based transformations of H&E into MT [42],[28], JMS and PAS, and the conversion of Ki67-CD8 to fibroblast activation protein-cytokeratin [43] have been performed in the literature. However, these approaches are developed based on CycleGANs (Cycle-Consistent Adversarial Networks) [44] or require perfectly paired image dataset.

Clearly, a decent data set is essential for virtual staining studies, but obtaining the desired and appropriate tissue samples, preparing them, and then scanning them with a digital image scanner is a daunting task for researchers. Existing datasets are unsuitable for deep learning-based virtual staining studies because they contain a limited number of slides or consist only of H&E-stained images.

Our main contributions to the literature are as follows:

- **Deep learning-based virtual staining approach for different tissue-organ combinations:** We optimized our spatial attention-based GAN model [1],[45] and utilized as a tool for virtual staining task and evaluated on five different tissue-stain combination by transforming H&E stain into MT (Kidney), PAS (Kidney), JMS (Kidney), CR (Kidney) and TB (Stomach).
- **Qualitative and quantitative study:** We calculated the Frechet Inception Distance (FID) scores for transformation performance of real stained patches to virtually stained patches for the quantitative analysis. We performed Visual Turing Tests (VTTs) involving 12 (patch-level VTT), 10 (slide-level VTT) board-certified pathologists to assess qualitatively the effectiveness of the virtual staining technique at both patch-level and slide-level. We also conducted a reader study with 3 board-certified pathologists to assess clinical performance of our virtual staining approach.
- **Publicly available StainKid dataset:** We present a comprehensive StainKid dataset collected by four centers and dedicated for development of virtual staining approaches based on machine learning. StainKid dataset comprises H&E-stained tissue images and equivalent MT (Kidney), PAS (Kidney), JMS (Kidney), CR (Kidney) and TB (Stomach) stained tissue images data as patient-based. The details of the StainKid dataset is summarized and illustrated in Figure 2.1.

Table 1.1
A comparison of the special stains with H&E in terms of utility, cost and the time required to stain per slide. The preparation time and cost information for each stain obtained from [2],[3],[4],[5].

Stain	**Utility**	**Colors**	**Prep. Time (approx.)**	**Cost per slide (approx.)**
H&E	general staining [46],[47],[48]	purple-nuclei pink-extracellular matrix	15 min	$6
MT	collagen [50],[51] connective tissue [49]	red-keratin, muscle fibers pink-cytoplasm blue-collagen,bone black-nucleus	1 hour	$20
PAS	glycogen glycoproteins glycolipids mucins [52],[53]	magenta-glycogen, mucin	30 min	$17
JMS	basement membrane [12]	black-basement membranes pink-cytoplasm	1 hour	$15
CR	amyloid [25]	purple to red-mucin blue - background	20 min	$20
TB	sulfates [21] carboxylates glycolipids phosphate radicals	purple to red - mucin blue - background	20 min	$17

2. DATASET

2.1 StainKid Dataset Collection

We collected tissue samples from four different centers in Turkey: Ege University Faculty of Medicine, Yeni Yuzyil University Medical Faculty Gaziosmanpasa (GOP) Hospital, Saglik Bilimleri University Kartal Dr. Lutfi Kirdar City Hospital, Erciyes University Faculty of Medicine Hospitals. In total, the dataset consists of 1645 WSIs from 417 biopsy cases collected between 2002 and 2022.

The StainKid dataset comprises a valuable collection of biopsy samples derived from renal and gastric tissues, with the purpose of performing virtual staining of different tissue types. The dataset includes a diverse range of biopsy samples, such as transplant kidney biopsies, diagnostic kidney biopsies, and stomach biopsies after detection of Helicobacter pylori. The kidney cases are stained with MT, PAS, JMS, and CR and have been collected over the years for clinical purposes, thereby representing daily practice.

To create an independent test cohort for CR staining, 40 amyloid-positive kidney slices stained with H&E and CR were collected from Kartal Dr. Lutfi Kirdar City Hospital. In addition, 40 tissue slices from Helicobacter pylori-positive gastric tissue specimens were collected and stained with H&E and TB at the same hospital.

The StainKid dataset contains a large amount of data, with a total of 778 WSIs and 2752564 patches from Ege University Faculty of Medicine, 563 WSIs and 1989079 patches from Yeni Yuzyil University Medical Faculty GOP Hospital, 224 WSIs and 792736 patches from Erciyes University Faculty of Medicine Hospitals, and 80 WSIs and 282880 patches from Saglik Bilimleri University Kartal Dr. Lutfi Kirdar City Hospital. The dataset is patient-based and includes information on patients' age, gender, biopsy date for all kidney cases, banff score, and additionally transplantation date for transplantation cases.

The publication of the StainKid dataset is expected to inspire further research in digital pathology, and contribute significantly to the existing literature. This dataset is highly valuable for creating training and independent test cohorts, and is expected to motivate the development of accurate and efficient methods for virtual staining in the future. The overview of the StainKid dataset can be seen in Figure 2.1.

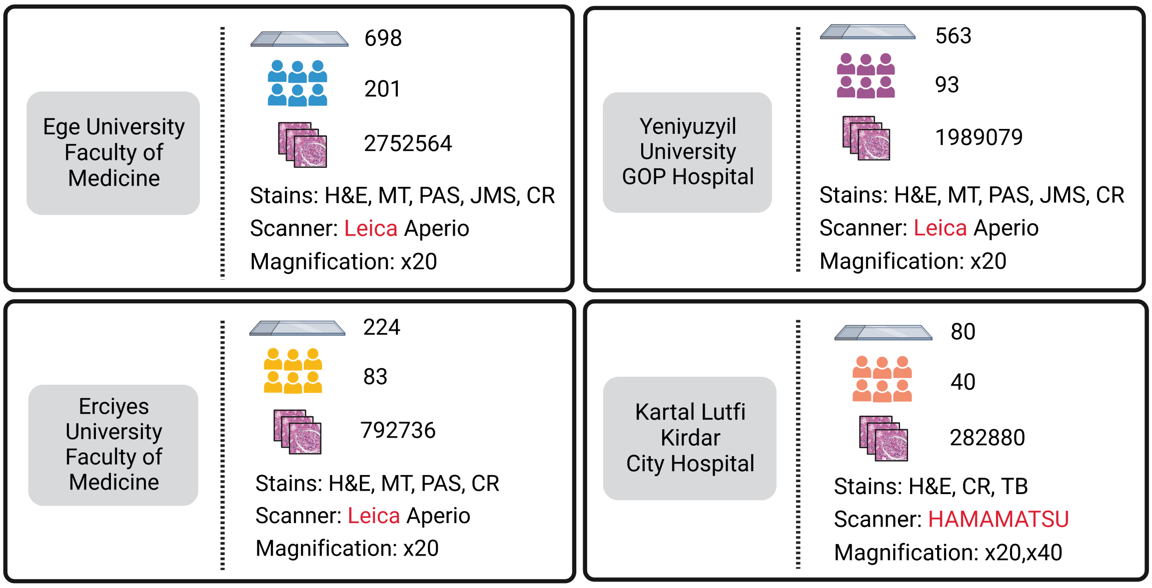

Figure 2.1 Overview of the StainKid Dataset.

2.2 Training Set

We have trained the model for each tissue-stain combination, individually. For the kidney, 144 WSIs, 150869 patches for H&E → MT, 120 WSIs, 124635 patches for H&E → PAS, 105 WSIs, 53629 patches for H&E → JMS from Ege University Faculty of Medicine, 40 WSIs, 523784 patches for H&E → CR, and 40 WSIs, 12465 patches for H&E → TB from Kartal Dr. Lutfi Kirdar City Hospital were utilized for training.

2.3 Test Set

We divided the data as a training set and one or two independent test cohorts for all tissue-stain combinations, and we tested the model on patient-based data, where each patient having WSI from both the source domain (H&E stained WSI) and the target domain (special stained WSI). For the kidney, we used 36 WSIs, 98232 patches for H&E → MT, 20 WSIs, 57918 patches for H&E → PAS, 34 WSIs, 99722 patches for H&E → JMS from Yeni Yuzyil University Medical Faculty and Erciyes University Faculty of Medicine Hospital for testing. 20 WSIs, 118436 patches, consisting of those not used in the training set, for H&E → CR from Kartal Dr. Lutfi Kirdar City Hospital, and 20 WSIs, 64687 patches for H&E → TB from Ege University Faculty of Medicine. To ensure the integrity and accuracy of our data, we carefully curated our datasets by eliminating patches exhibiting poor quality, faded staining, or clinically misleading features. This curation process enabled us to provide the model with high-quality data during the learning process, thus improving the robustness and reliability of our findings.

2.4 WSI Processing

We utilized state-of-the-art digital slide scanners, including the Leica Aperio AT2 and the in-house Hamamatsu Nanozoomer s60, to analyze glass slides of kidney and stomach tissues, respectively. The kidney slides were scanned at 20x magnification, while the stomach slides were scanned at 40x magnification to enable differentiation of Helicobacter pylori.

To perform image analysis, we employed the CLAM algorithm [54], which enabled us to accurately segment tissue regions within each whole slide image (WSI) while excluding any holes present in the tissue. We further extracted 512 x 512 image patches from each WSI, with no overlap between patches, for subsequent analysis. We carefully curated our train and test sets to ensure optimal performance and reliability of our results. The details of the train and test data are provided in Table 2.1.

Table 2.1
Data characteristics of train and test data for each stain transformation.

Stain Transformation	**Tissue**	**Number of Patches**		**Patch size**	**Resolution**
		Training	**Test**		
H&E → MT	Kidney	150869	98232	512x512	20x
H&E → PAS	Kidney	124635	57918	512x512	20x
H&E → JMS	Kidney	53629	99722	512x512	20x
H&E → CR	Kidney	523784	118436	512x512	20x
H&E → TB	Stomach	50036	64687	512x512	40x

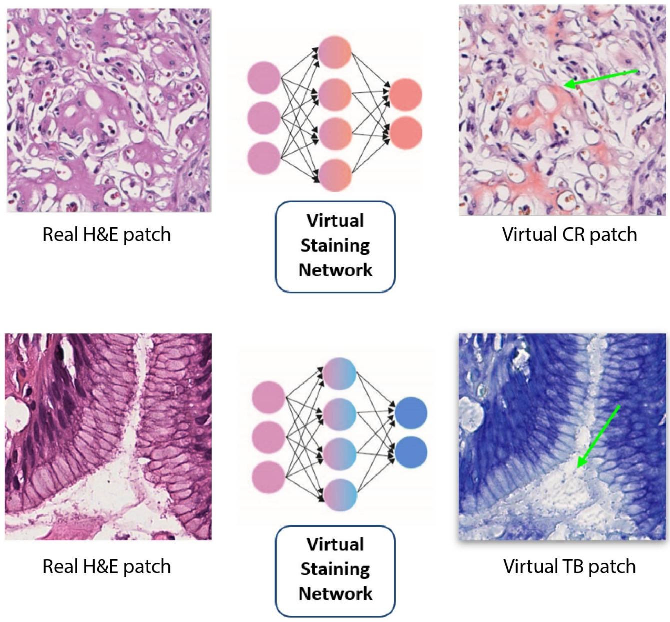

Figure 2.2 Initial patch-level results for CR (kidney) and TB (stomach).

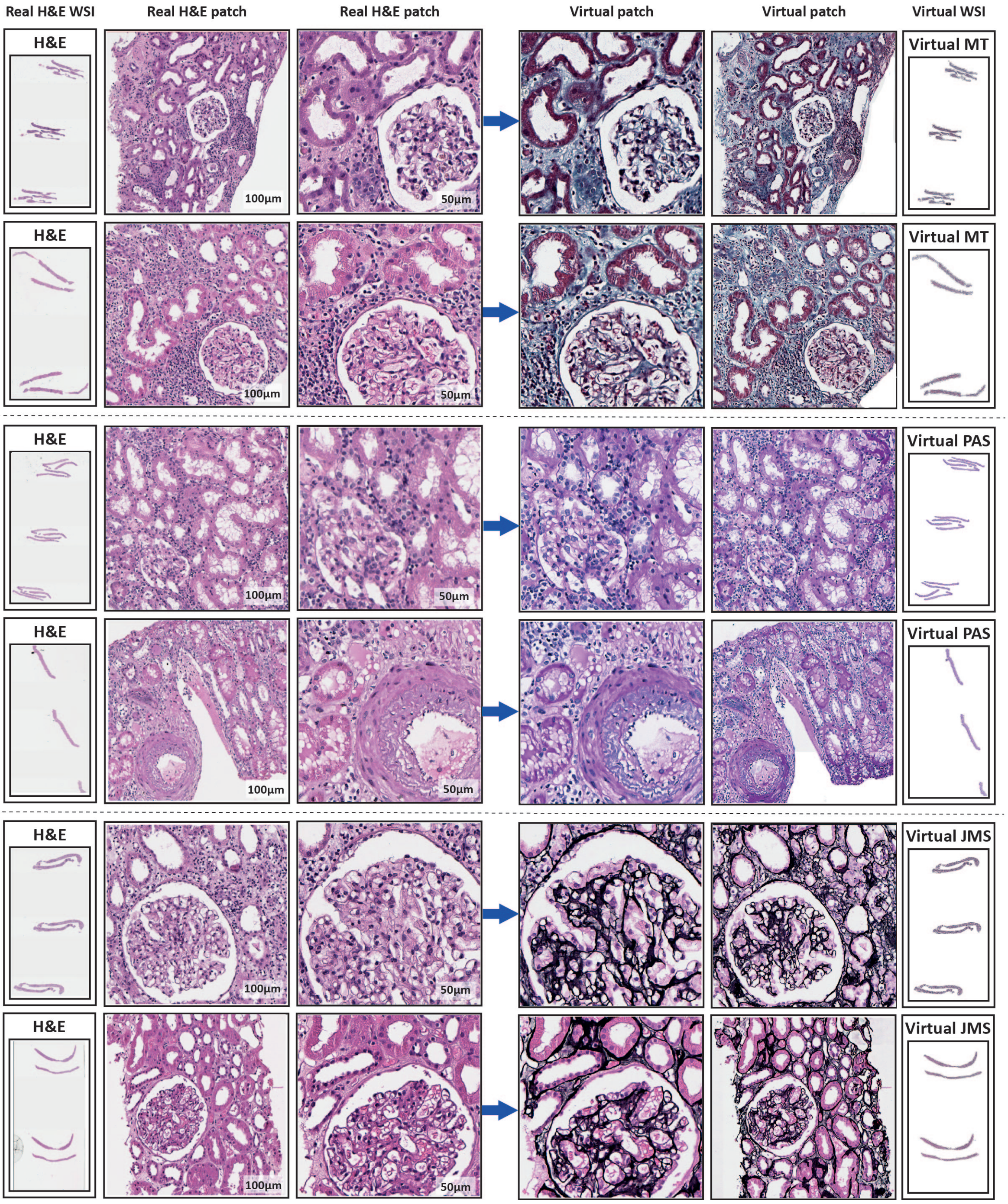

Figure 2.3 Virtual MT, PAS and JMS patches and WSIs obtained from H&E-stained kidney tru-cut biopsy tissue slides. The left columns of the blue arrow displays real H&E WSIs and their corresponding patches created from them given as an iput to the generator. The right side of the blue arrows displays virtually stained MT, PAS and JMS stained patches output of the generator, and their corresponding stitched WSIs. Each row of images in the binary format corresponds to a different center, showing that the generative model performs effectively on data from different centers.

3. METHOD

3.1 Network Architecture

To conduct stain transformation between unpaired images, we utilized our GAN based style transfer architecture [1] (see Figure 4.3). To supervise GAN training, it combines adversarial, contrastive, and self-regularization losses. The typical adversarial loss:

$$\begin{aligned}\mathcal{L}_{GAN}(G, D, X, \mathbf{Y}) = {} & E_{y \sim \mathbf{Y}} \log D(y) + \\ & E_{x \sim \mathbf{X}} \log(1 - D(G(x))),\end{aligned} \tag{3.1}$$

The generator G(x) creates virtually stained image patches from the real H&E stained patches. The discriminator D(y) endeavors to discriminate whether the input patch is a virtually stained patch generated by G(x) or real stained patch, y.

Transformation in histopathology images is a demanding task due to its cellular morphology and tissue structure. Adversarial loss tends to learn and banish the differences between the real stained patch and the virtually stained patch. However, at the same time, diagnostically meaningful content needs to be preserved in the transformation. To preserve content, a noise contrastive loss function is utilized. Correspondence between real H&E-stained patch and virtually stained patch is found by taking a query patch from a virtually stained sample patch in the target domain compared to a matching real H&E-stained patch in the source domain at the same position. This coparison is done in such a way that similar patches create positive pairings and dissimilar stains produce negative pairs. K negative samples and the query sample, positive are being mapped to L-dimension vector $v^{-} \in R^{M \times L}$ and $v, v^{+} \in R^{L}$ severally. The m-th negatives in L samples are indicated by the symbol $v_k^{-} \in R^{L}$. The distance of scaling between the sample patches and query is calculated by using a temperature $\tau = 0.07$. The cross-entropy loss is computed, demonstrating the likelihood of prioritising the positives between the negatives.

$$\mathcal{L}(v, v^{+}, v^{-}) = -\log\left[\frac{\exp(v.\ v^{+}/\tau)}{\exp(v.v^{+}/\tau) + \sum_{n=1}^{N}\exp(v.v_{n}^{-}/\tau)}\right]. \tag{3.2}$$

The generator network creates its feature piles which has spatial position s, each layer l in this feature pile depicts a region of the input patch. Later, the feature piles are given as an input to a two layered network referred as Multi Layer Perceptron (MLP) as proposed in [55]. The final features formed by the corresponding image features $\hat{z}_l^s$, final features in the output domain z_l^s, and different extracted features $z_l^s, z_l^{S\backslash s}$ characterize patch noise contrastive estimation loss:

$$\mathcal{L}_{PatchNCE}(G, \mathbf{F}, \mathbf{X}) = E_{x\sim\mathbf{X}}\sum_{l=1}^{L}\sum_{s=1}^{S_l}\mathcal{L}(\hat{z}_l^s, z_l^s, z_l^{S\backslash s}). \tag{3.3}$$

Finally, self-regularization loss to hinders the network from any clinically illusive features occurred on the patches:

$$\mathcal{L}_{sReg}(G, X) = ||\mathbf{X} - G(\mathbf{X})||_1 \tag{3.4}$$

The hybrid loss function:

$$\begin{aligned}\mathcal{L}_{AI-FFPE} = \mathcal{L}_{GAN}(G, D, \mathbf{X}, \mathbf{Y}) + \\ \lambda_{sReg}\mathcal{L}_{sReg}(G, X) + \\ \lambda_{\mathbf{X}}\mathcal{L}_{patchNCE}(G, \mathbf{F}, \mathbf{X}) + \\ \lambda_{\mathbf{Y}}\mathcal{L}_{patchNCE}(G, \mathbf{F}, \mathbf{Y}).\end{aligned} \tag{3.5}$$

Spatial Attention Block (SAB). With integration of SAB into the generator, it was aimed to focus on clinically significant features in the learning process. The SAB performs a non-local convolution operation for any input, $\mathbf{X}$. In the SAB, there are three convolutional layers that divide the input into three elements θ, ϕ, and g [1].

$$\mathbf{Z} = f(\mathbf{X}, \mathbf{X}^{\top})g(\mathbf{X}), \tag{3.6}$$

f denotes the correlation between the pixels in the input patch $\mathbf{X}$. A dot product operation is utilized on θ and ϕ, after the convolutional operations. Then, the Rectified Linear Unit (ReLU) function, σ_{relu} was used for activation.

$$\mathbf{P} = \psi(\sigma_{relu}(\theta(\mathbf{X})\phi(\mathbf{X})^{\top})). \tag{3.7}$$

$\theta(\mathbf{X})\varphi(\mathbf{X})^{\top}$, yields a measure of input co-variance can be explained as the degree of slope between two feature maps on different channels. A normalization is applied to output P with softmax function, $\sigma_{softmax}$. Then, a matrix multiplication between the output of g and softmax function is applied. After that, we convolved and up-sampled the outcome of ϕ multiplication to generate the attention map S. Eventually, an element sum operation between the S and the X produces the output $\mathbf{F} \in R^{N\times 64\times H\times W}$.

$$\mathbf{S} = \phi(\sigma_{softmax}(\mathbf{K})g(\mathbf{X})) \tag{3.8}$$

$$\mathbf{F} = \mathbf{S} + \mathbf{X} \tag{3.9}$$

3.2 Hyperparameters

Unlike traditional image-to-image translation models, our network architecture works in a single direction between the source (real H&E images) and target domains (virtually stained images), allowing faster training and lower memory requirements. We trained our model for each tissue-stain combination, individually: MT (kidney), JMS (kidney), PAS (kidney), CR (kidney), TB (stomach). In the training, we performed Adam optimizer with an initial learning rate of 0.0002, and $\beta_1 = 0.5$ and $\beta_2 = 0.999$ moementums for 8 epochs. We utilized 1 batch size, Xavier weight initialization [56] and instance normalization [57]. Besides, λ_{NCE} parameter is set to equal 2 for identity loss, and the temperature value (τ) is set to equal 0.07.

3.3 Computational Hardware and Software

WSI processing (patching and stitching) were performed on Xeon multi-core CPUs, Nvidia V100 GPUs using CLAM framework [54] in Python (version 3.6.13). Deep virtual staining networks were trained and tested on Nvidia V100 GPU for each tissue-stain combination using Pytorch (version 1.9.0).

4. RESULTS

Figure 2.3 shows Virtual MT, PAS and JMS patches and WSIs obtained from H&E-stained kidney WSIs. In MT virtual staining, the method marks the extracellular connective tissue in blue. Consistent with the real histochemical method, the areas where the connective tissue is denser are virtually painted in a darker color. In PAS virtual staining, the method successfully selects and stains the basement membranes and can show the accumulation in the glomeruli. The method selectively stains reticulin fibers between tubules and glomeruli for JMS. It shows the glomerular capillary network well. Notably, each row of images in the binary format corresponds to a different center, indicating that the generative model performs effectively on data from different centers. These findings prove the capability of our AI-based approach to produce virtually stained images in diverse settings and suggest its broad applicability for downstream analysis.

Figure 2.2 shows initial patch-level results for CR (kidney) and TB (stomach). The output patches generated by the virtual stain generator network show visually recognizable patterns, indicating promising results for the detection of amyloid on the virtual CR and of Helicobacter pylori in TB. The observed patterns in the generated patches highlight the ability of the model to accurately detect and differentiate specific biomarkers such as amyloid and Helicobacter pylori, which are important indicators of various pathological conditions.

4.1 Quantitative Results

PyTorch standard framework to compute the FID [58] measure is utilized in order to objectively examine the performance of virtual staining at the patch-level. A deep network space's Gaussian distribution of real and virtually stained images is determined, together with the difference between the distributions' means and standard deviations. The FID score continually declines as the iterations go on, but the virtually stained images get more realistic and higher quality. The following is the FID formulation:

$$\mathcal{I}_{FID} = ||\vec{m_1} - \vec{m_2}||_2^2 + Tr(C_1 + C_2 - 2 * \sqrt{(C_1 * C_2)}) \tag{4.1}$$

where $\vec{m_1}$ and $\vec{m_2}$ represent mean values of the real and virtually stained patches, C_1 and C_2 are covariance matrices. Tr denotes trace function, the sum of the diagonal elements of the matrix.

Similarity between the real stained and virtually stained patches are measured with FID, and compared on virtual staining tasks with generic style transfer models: CUT and FastCUT. For each virtual staining task, the AI-FFPE, FastCUT [45], and CUT [45] models were trained with the same training data (see Table 2.1) and then tested separately with the same real stained patches. FID Scores were calculated using the test results of each model (virtually stained patches) and real patches. Folders including real patches and virtual patches patches for each stain for the FID calculation contains: 24086 real patches, 73011 virtual patches for MT, 1592 real patches, 3666 virtual patches for PAS, 16944 real patches, 79379 virtual patches for JMS, 161897 real patches, 205135 virtual patches for CR, and 66 real patches, 41240 virtual patches for TB. The results showed that the AI-FFPE model proved to be more successful than other models in the virtual staining task for each stain, as indicated by its low FID score, which signifies high similarity. The results can be seen in Table 4.1.

4.2 Qualitative Results

4.2.1 Slide-level VTT

A comprehensive survey was conducted to evaluate the performance of virtual staining as applied to WSIs from the perspective of pathologists. The study utilized 90 WSIs, with 30 WSIs each for MT, PAS, JMS staining techniques that had successfully undergone virtual slide-level staining. Ten pathologists from different countries, including Turkey and the United Kingdom, participated in the survey.

The board-certified pathologists were requested to rate the overall staining performance of the WSIs presented to them, using a ranking system where scores of 1, 2, 3, 4 are respectively represented as unacceptable, acceptable, very good, and perfect. [28]. The pathologists were also asked to evaluate the success in visualizing key features, including nuclear detail, cytoplasmic detail, and extracellular fibrosis, which are essential for examination in diagnosis.

The results of the survey were analyzed using the average and standard deviation metrics. The results indicates that the performance of virtual staining is considered "very good" (3 on the scale of 1 to 4) for all three stains according to the pathologists point of view (see Table 4.2).

Table 4.1
Comparative Frechet ID results of AI-FFPE method with other image translation models on virtual staining task for each tissue-stain combination.

	FID↓ (MT)	**FID↓ (PAS)**	**FID↓ (JMS)**	**FID↓ (CR)**	**FID↓ (TB)**
FastCUT	52.485	119.471	193.249	158.47	257.23
CUT	45.178	118.712	58.571	129.632	167.156
AI-FFPE	37.088	97.093	50.968	62.125	139.683

Table 4.2
Results of the slide-level VTT performed with 10 pathologists.

	Nuclear Detail		**Cytoplasmic Detail**		**Ext. Fibrosis**		**Overall S. P.**	
	Mean	**STD**	**Mean**	**STD**	**Mean**	**STD**	**Mean**	**STD**
PAS	3.05	1.02	2.5	0.7	3.28	0.87	3.12	0.94
MT	2.87	0.96	2.95	0.7	3.19	0.78	2.96	0.87
JMS	3.03	0.99	3.07	1.41	3.04	0.8	2.96	0.86

4.2.2 Patch-level VTT

A patch-level survey was designed with real and virtual patches. In this survey, 12 pathologists were randomly shown 150 patches, including 25 real and 25 virtual patches for each of the three stains. The aim of the VTT was to determine if the pathologists could accurately distinguish between the real patches and virtual patches.

The results were analyzed using confusion matrices [59], shown in 4.2.2, and various performance metrics were calculated based on True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP) values. These metrics included recall, precision, accuracy, and F1-score. Low values of accuracy, precision, recall, and F1-score indicates that the pathologists could not differentiate real patches and virtual patches with a high degree of certainty.

The recall scores ranged from 0.45 to 0.50, the precision scores ranged from 0.43 to 0.51, the accuracy scores ranged from 0.45 to 0.51, and the F1-scores ranged from 0.45 to 0.51. These results of this patch-level survey demonstrate that the pathologists were not able to accurately differentiate between real and virtual images, as evidenced by the low values for recall, precision, accuracy, and F1-score, which are represented in Table 4.3. The results have important implications for the use of virtual images in medical diagnosis and underscore the need for further research in this area.

Table 4.3
Pathologists' performance metrics in patch-level VTT for distinguishing real and virtual images: low Precision, accuracy, and F1 scores.

	MT	PAS	JMS
recall	0.45	0.47	0.506
precision	0.451	0.436	0.518
accuracy	0.451	0.431	0.518
F1	0.45	0.452	0.512

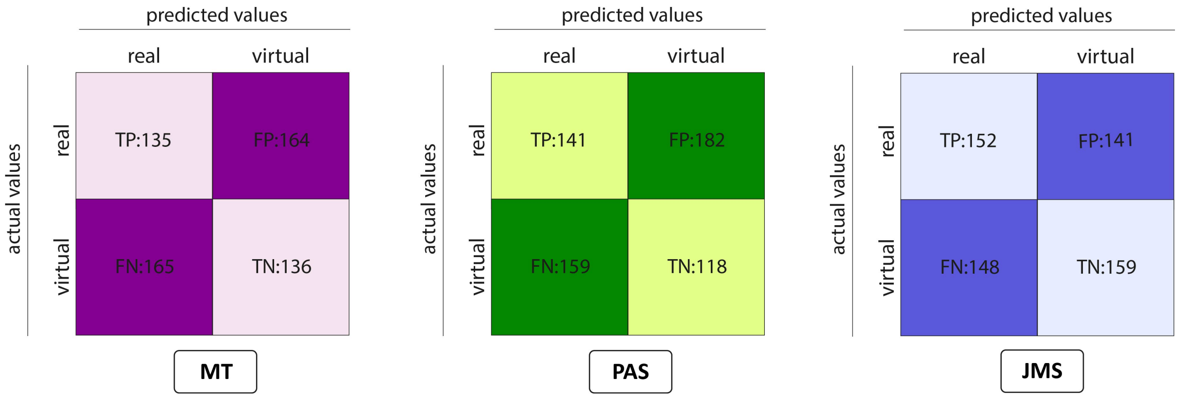

Figure 4.1 Confusion matrices for results of the patch-level VTT.

4.2.3 Reader Study(Diagnostic Test)

A survey was conducted to evaluate the impact of the virtual staining performance in the diagnostic process. In this study, the survey was designed for the diagnosis of renal transplant rejection at slide-level. Three experienced nephropathologists participated in the study. They were shown the virtually stained JMS, MT, PAS, real H&E stained WSIs, as well as the c4d Banff score of the same patients. The pathologists were first asked to determine the presence of rejection and then to identify the type of rejection (acute, chronic, chronic active, cellular, humoral, or cellular & humoral).

The results showed that the pathologists achieved an accuracy of %87 in determining the presence of rejection and an accuracy of %32 in determining the type of rejection. Given that pathologists take into account various factors such as clinical findings [60] and Banff classification [61] along with images when making their diagnosis, this accuracy level is considered acceptable. The reader study provides evidence that virtual staining can be used effectively to aid in the diagnosis of renal transplant rejection at the WSI level. The results suggest that virtual staining can provide valuable information to support the diagnostic process, while the pathologists still consider other important factors in the diagnostic decision-making process.

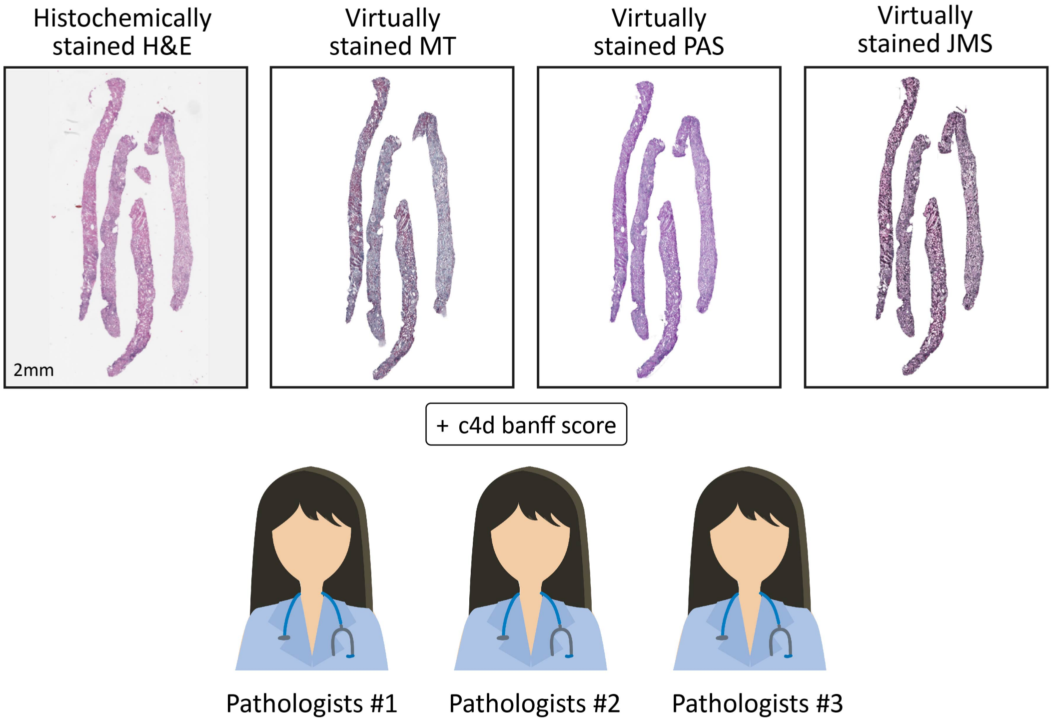

Figure 4.2 Schema of the survey that pathologists diagnose kidney rejection using real H&E stained WSI together with virtually MT, PAS, and JMS stained WSIs.

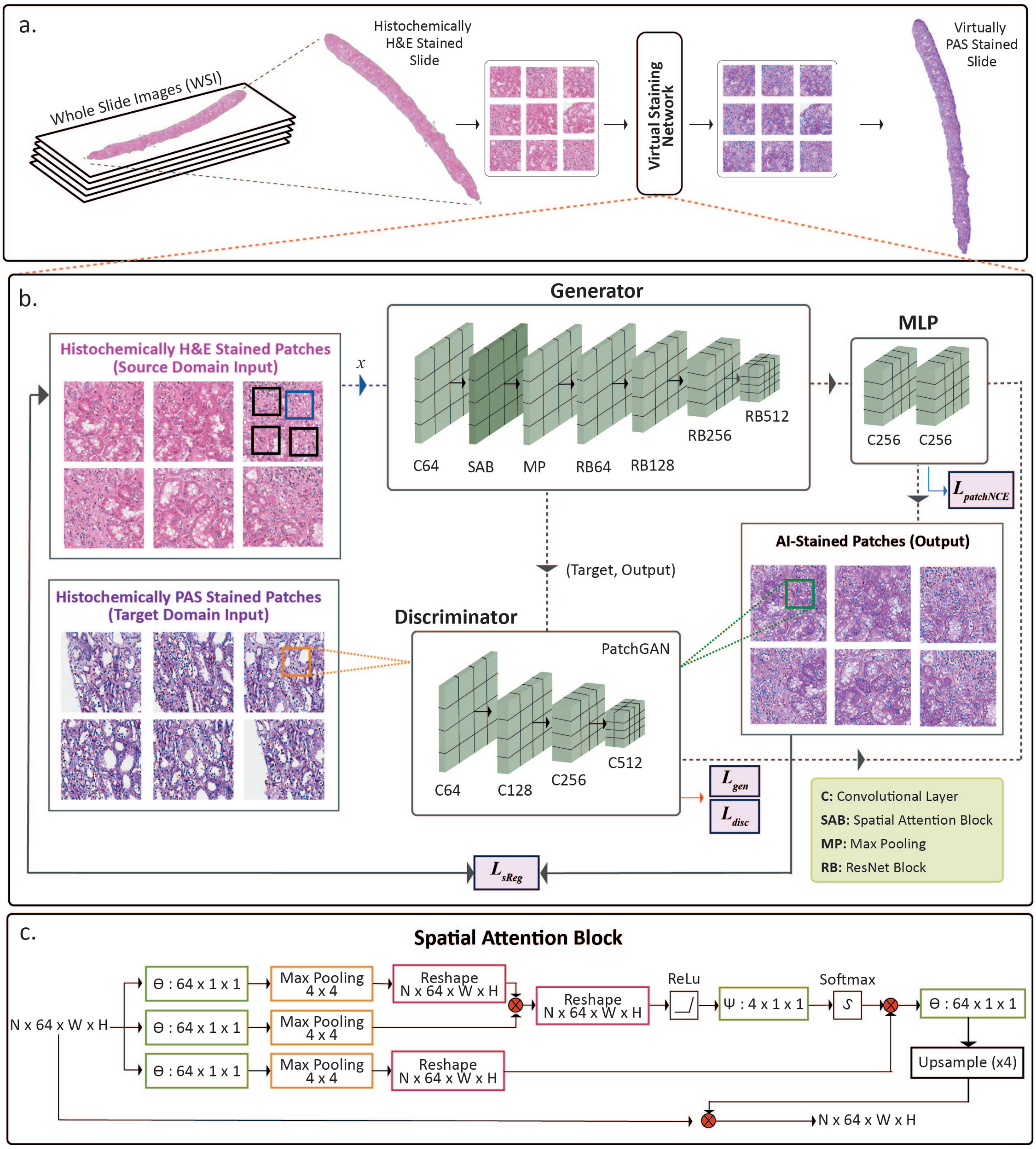

Figure 4.3 Network Architecture [1].

5. DISCUSSION AND CONCLUSION

In this thesis, an AI model is utilized to perform virtual staining of H&E-stained kidney and stomach biopsies. A novel large dataset of paired WSIs of H&E-stained non-malignant kidney biopsies and paired WSIs of H&E-stained rejection kidney transplant biopsies, as well as a smaller sample of paired WSIs of H&E and TB-stained, helicobacter pylori-positive stomach biopsies were collected for training and testing of the model. We used a GAN with a spatial attention block as the AI model and employed FID, VTTs, and reader study as the metrics for evaluation. Our results showed that the AI model can generate virtually stained patches and WSIs that are very similar to their traditionally stained counterparts and, the virtual staining does not affect the diagnostic accuracy of pathologists. Figure 1.1 shows how the virtual staining method fits into the current histopathology workflow.

Our study contributes to the growing field of virtual staining, which aims to provide faster, cheaper, better, and more expressive alternatives to traditional staining methods. Virtual staining has several advantages over traditional staining, such as reducing the need for multiple biopsies, preserving tissue integrity, saving time and resources, enhancing the contrast and resolution of the images, and enabling the visualization of multiple biomarkers, simultaneously [33],[62],[63],[64]. Virtual staining can also overcome some of the limitations and challenges of traditional staining, such as variability, reproducibility, standardization, quality control, and interpretation [65].

Our study is one of the first to apply virtual staining to kidney biopsies, which are commonly used for the diagnosis and prognosis of various kidney diseases, for example acute kidney injury, chronic kidney disease, glomerulonephritis, and transplant rejection [12],[66],[67]. Kidney biopsies are usually stained with H&E, which provides a general overview of the tissue morphology but may not be sufficient to reveal the subtle changes and specific features of some kidney [19]. Therefore, additional stains and methods, such as PAS, MT, JMS, CR, immunofluorescence, and electron microscopy, are often required to complement the H&E staining and to provide more information on the glomerular, tubular, interstitial, and vascular structures and lesions [12],[68]. However, these additional stains are time-consuming and costly and may consume or damage the precious tissue samples. Moreover, some of these stains have low contrast, poor resolution, or high background noise, which may hamper the image quality and interpretation. Therefore, virtual staining of kidney biopsies could offer a valuable alternative to the traditional staining methods by generating high quality patches and WSIs of different stains from a single H&E-stained biopsy without the need for additional biopsies or stains.

Our study is also one of the first to use a GAN with a spatial attention block as the AI model for virtual staining. GANs are a type of deep neural network comprised of two opposing components: a generator and a discriminator. The primary objective of the generator is to generate images that appear authentic enough to deceive the discriminator. In contrast, the discriminator's main goal is to differentiate between genuine and synthetic images. GANs have been widely used for various image syntheses tasks, such as image-to-image translation [69], super-resolution [70], data augmentation [71] and style transfer [72]. However, GANs also have some drawbacks, such as mode collapse, instability, and low diversity [73]. To address these issues, we incorporated a spatial attention block into the generator of our GAN, which can learn to focus on the most pertinent areas of the input image and generate more realistic and diverse output images. Figure 2.3 that the generator successfully produces realistic virtually stained patches. Our spatial attention block makes it so that the neural network can capture the long-range dependencies and the inter-channel relationships of the input image [1].

Four different metrics were utilized to assess the performance of the AI model: FID, patch-level VTT, slide-level VTT and reader study. The FID is a quantitative metric that measures the distance between the feature distributions of the real and generated images using a pre-trained inception network. The FID metric is commonly employed to assess both the excellence and variety of the synthetic images produced, and a lower FID score indicates a higher similarity between the real and virtually stained images. Patch-level VTT is a qualitative metric, where pathologists were asked to identify which patch is real and which is virtual, from a pair of patches of the same tissue region of a WSI with different stains. Slide-level VTT is an another qualitative metric, where pathologists are asked to rate given virtual WSIs according to the features of nuclear detail, cytoplasmic detail, and extracellular fibrosis and general staining performance of the AI model for MT, PAS and JMS stains. Reader study is another qualitative metric that involves pathologists evaluating the diagnostic features and outcomes of the real and generated images, and comparing their agreement and consistency. The reader study was designed and performed to assess the clinical utility and validity of the virtual WSIs, and the virtual staining method in general.

The results showed that the AI model achieved the lowest FID scores of 37.088, 97.093, 50.968, 62.125, 139.683 for compared to the other models, which indicates a high similarity between the real and virtual images (see Table 4.1). The results also showed that the AI model achieved low recall (0.45,MT;0.57,PAS;0.5,JMS), precision (0.45,MT;0.43,PAS;0.51,JMS), accuracy (0.45,MT;0.43,PAS;0.51,JMS) rates and F1-scores (0.45,MT;0.45,PAS;0.51,JMS) in the patch-level VTT, which indicates that the pathologists could not discern between the real and virtual images (see Table 4.3). The slide-level VTT results, which are shown in 4.2, proved that the virtual staining performance of our model was also successful at the slide-level, rated as "very good" (3 on a 1 to 4 scale) for pathologists. Moreover, our results showed that the AI model achieved a high accuracy (87%, presence of rejection) and acceptable accuracy (32%, type of rejection) in the reader study. These results showed that the AI model is able to produce high-quality, realistic, and clinically valid images of different stains from a single H&E stained biopsy.

In conclusion, we employed and evaluated an AI model that can perform virtual staining of H&E stained kidney biopsies, as well as other tissue types. We collected, used, and proposed a novel, large, patient-based, publicly available StainKid dataset of paired whole slide images of H&E-stained non-malignant and transplant kidney biopsies and a smaller sample of paired WSIs of H&E and TB-stained, helicobacter pylori-positive stomach biopsies. We employed a GAN with a spatial attention block as the AI model [1] and used FID, VTT, and reader study as the metrics for evaluation.

Our results showed that the AI model can generate virtually stained patches and WSIs that are quite similar to their traditionally stained counterparts, and that the virtual staining does not affect the diagnostic accuracy of pathologists. Our study contributes to the field of virtual staining, which offers a valuable alternative to the traditional staining methods by providing faster, cheaper, better, and more expressive images of different stains from a single H&E stained biopsy. Our study also has some limitations that should be addressed in Chapter 6, such as collecting more data for the StainKid dataset, extending the AI model, and conducting a larger reader study. We believe that use of the AI model has a great ability to enhance the quality and efficiency of pathology practice and to facilitate the diagnosis and prognosis of various kidney diseases.

6. FUTURE WORKS

Our study has some limitations that should be acknowledged and addressed in future work. First, our dataset of paired WSIs of H&E and TB-stained, helicobacter pylori-positive stomach biopsies was relatively small, which may limit the generalization competence of the AI model for other tissue types and stains. Therefore, we plan to collect and annotate more data from different tissue types and stains and to evaluate the AI model on a larger and more diverse dataset. Second, the AI model was only evaluated on non-malignant or transplant kidney biopsies, which may not reflect the complexity and variability of malignant kidney biopsies. Therefore, we plan to extend the AI model to handle malignant kidney biopsies and to evaluate its performance on different types and stages of kidney cancer. Third, the AI model was evaluated by 12 pathologists from different centers around the world, yet still, these pathologists opinions may not represent the opinions and preferences of the wider pathology community. Therefore, we plan to conduct an even larger and even more comprehensive reader study involving more pathologists from different backgrounds and experiences and to collect more feedback and suggestions from them.

REFERENCES

1. Ozyoruk, K. B., S. Can, B. Darbaz, and *et al.*, “A deep-learning model for transforming the style of tissue images from cryosectioned to formalin-fixed and paraffin-embedded,” *Nature Biomedical Engineering*, pp. 1–13, 2022.
2. USask, “Histology core facility cost recovery price list.” https://healthsciences.usask.ca/facility-services/Histology/histology-core-facility-price-list-june-2022.pdf.
3. YSM, “Comparative pathology research core-price list.” https://medicine.yale.edu/compmed/mrp/pricelist/.
4. LSU, “Histology fee schedule.” https://www.lsu.edu/vetmed/laddl/histology$_f$*ees.pdf*.
5. WUSM, “Special stains.” https://therisingsea.org/notes/FoundationsForCategoryTheory .pdf.
6. Titford, M., “The long history of hematoxylin,” *Biotechnic & Histochemistry*, Vol. 80, no. 2, pp. 73–78, 2005.
7. Bancroft, J., and C. Layton, “The hematoxylins and eosin. bancroftâs theory and practice of histological techniques,” *Elsevier*, Vol. 173, p. 186, 2013.
8. King, D. F., and L. A. King, “A brief historical note on staining by hematoxylin and eosin,” *The American Journal of Dermatopathology*, Vol. 8, no. 2, p. 168, 1986.
9. Dapson, R., and R. Horobin, “Dyes from a twenty-first century perspective,” *Biotechnic & Histochemistry*, Vol. 84, no. 4, pp. 135–137, 2009.
10. Wittekind, D., “Traditional staining for routine diagnostic pathology including the role of tannic acid. 1. value and limitations of the hematoxylin-eosin stain,” *Biotechnic & Histochemistry*, Vol. 78, no. 5, pp. 261–270, 2003.
11. Larson, K., H. H. Ho, P. L. Anumolu, and T. M. Chen, “Hematoxylin and eosin tissue stain in mohs micrographic surgery: a review,” *Dermatologic Surgery*, Vol. 37, no. 8, pp. 1089–1099, 2011.
12. Cathro, H. P., S. S. Shen, and L. D. Truong, “Diagnostic histochemistry in medical diseases of the kidney,” in *Seminars in Diagnostic Pathology*, Vol. 35, pp. 360–369, Elsevier, 2018.
13. Iezzoni, J. C., “Diagnostic histochemistry in hepatic pathology,” in *Seminars in Diagnostic Pathology*, Vol. 35, pp. 381–389, Elsevier, 2018.
14. Orah, N., O. Rotimi, and F. Abdulkareem, “The use of special stains in liver biopsy interpretation: Implications for the management of liver disease in nigeria,” *Nigerian Journal of Clinical Practice*, Vol. 19, no. 4, pp. 523–529, 2016.
15. Gurina, T. S., and L. Simms, “Histology, staining,” 2020.
16. Al Drees, A., M. S. Khalil, and M. Soliman, “Histological and immunohistochemical basis of the effect of aminoguanidine on renal changes associated with hemorrhagic shock in a rat model,” *Acta Histochemica et Cytochemica*, p. 16025, 2017.
17. Baum, S., “The pas reaction for staining cell walls,” *Cold Spring Harbor Protocols*, Vol. 2008, no. 8, pp. pdb–prot4956, 2008.

18. “Sağlam a. jones methenamine silver. pathologyoutlines.com website..” https://www.pathologyoutlines.com/topic/stainsJMS.html. Accessed: August 21st, 2022.

19. Walker, P. D., T. Cavallo, and S. M. Bonsib, “Practice guidelines for the renal biopsy,” *Modern Pathology*, Vol. 17, no. 12, pp. 1555–1563, 2004.

20. Walker, P. D., “The renal biopsy,” *Archives of Pathology & Laboratory Medicine*, Vol. 133, no. 2, pp. 181–188, 2009.

21. Sridharan, G., and A. A. Shankar, “Toluidine blue: A review of its chemistry and clinical utility,” *Journal of 0ral and Maxillofacial Pathology: JOMFP*, Vol. 16, no. 2, p. 251, 2012.

22. Epstein, J. B., C. Scully, and J. Spinelli, “Toluidine blue and lugol’s iodine application in the assessment of oral malignant disease and lesions at risk of malignancy,” *Journal of Oral Pathology & Medicine*, Vol. 21, no. 4, pp. 160–163, 1992.

23. Gandolfo, S., M. Pentenero, R. Broccoletti, M. Pagano, M. Carrozzo, and C. Scully, “Toluidine blue uptake in potentially malignant oral lesions in vivo: clinical and histological assessment,” *Oral Oncology*, Vol. 42, no. 1, pp. 88–94, 2006.

24. Miller, R., B. Simms, and A. Gould, “Toluidine blue staining for detection of oral premalignant lesions and carcinomas,” *Journal of Oral Pathology & Medicine*, Vol. 17, no. 2, pp. 73–78, 1988.

25. Yakupova, E. I., L. G. Bobyleva, I. M. Vikhlyantsev, and A. G. Bobylev, “Congo red and amyloids: history and relationship,” *Bioscience Reports*, Vol. 39, no. 1, 2019.

26. Akki, A., Q. Liu, S. M. Chung, K. E. Tanaka, and N. C. Panarelli, “Submit the entire gallbladder,”

27. Scanziani, E., “Immunohistochemical staining of fixed tissues,” *Mycoplasma Protocols*, pp. 133–140, 1998.

28. de Haan, K., 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 h&e stained tissues into special stains,” *Nature Communications*, Vol. 12, no. 1, pp. 1–13, 2021.

29. Adyanthaya, S., and M. Jose, “Quality and safety aspects in histopathology laboratory,” *Journal of Oral and Maxillofacial Pathology: JOMFP*, Vol. 17, no. 3, p. 402, 2013.

30. Glaser, A. K., N. P. Reder, Y. Chen, E. F. McCarty, C. Yin, L. Wei, Y. Wang, L. D. True, and J. T. Liu, “Light-sheet microscopy for slide-free non-destructive pathology of large clinical specimens,” *Nature Biomedical Engineering*, Vol. 1, no. 7, pp. 1–10, 2017.

31. Fereidouni, F., Z. T. Harmany, M. Tian, A. Todd, J. A. Kintner, J. D. McPherson, A. D. Borowsky, J. Bishop, M. Lechpammer, S. G. Demos, *et al.*, “Microscopy with ultraviolet surface excitation for rapid slide-free histology,” *Nature Biomedical Engineering*, Vol. 1, no. 12, pp. 957–966, 2017.

32. Tao, Y. K., D. Shen, Y. Sheikine, O. O. Ahsen, H. H. Wang, D. B. Schmolze, N. B. Johnson, J. S. Brooker, A. E. Cable, J. L. Connolly, *et al.*, “Assessment of breast pathologies using nonlinear microscopy,” *Proceedings of the National Academy of Sciences*, Vol. 111, no. 43, pp. 15304–15309, 2014.

33. Rivenson, Y., K. de Haan, W. D. Wallace, and A. Ozcan, “Emerging advances to transform histopathology using virtual staining,” *BME Frontiers*, Vol. 2020, 2020.

34. Rivenson, Y., H. Wang, Z. Wei, K. de Haan, Y. Zhang, Y. Wu, H. Günaydın, J. E. Zuckerman, T. Chong, A. E. Sisk, *et al.*, “Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning,” *Nature Biomedical Engineering*, Vol. 3, no. 6, pp. 466–477, 2019.

35. Zhang, Y., K. de Haan, Y. Rivenson, J. Li, A. Delis, and A. Ozcan, “Digital synthesis of histological stains using micro-structured and multiplexed virtual staining of label-free tissue,” *Light: Science & Applications*, Vol. 9, no. 1, pp. 1–13, 2020.

36. Bayramoglu, N., M. Kaakinen, L. Eklund, and J. Heikkila, “Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks,” in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 64–71, 2017.

37. Rivenson, Y., T. Liu, Z. Wei, Y. Zhang, K. de Haan, and A. Ozcan, “Phasestain: the digital staining of label-free quantitative phase microscopy images using deep learning,” *Light: Science & Applications*, Vol. 8, no. 1, pp. 1–11, 2019.

38. Borhani, N., A. J. Bower, S. A. Boppart, and D. Psaltis, “Digital staining through the application of deep neural networks to multi-modal multi-photon microscopy,” *Biomedical Optics Express*, Vol. 10, no. 3, pp. 1339–1350, 2019.

39. Fujitani, M., Y. Mochizuki, S. Iizuka, E. Simo-Serra, H. Kobayashi, C. Iwamoto, K. Ohuchida, M. Hashizume, H. Hontani, and H. Ishikawa, “Re-staining pathology images by fcnn,” in *2019 16th International Conference on Machine Vision Applications (MVA)*, pp. 1–6, IEEE, 2019.

40. Mercan, C., G. Mooij, D. Tellez, J. Lotz, N. Weiss, M. van Gerven, and F. Ciompi, “Virtual staining for mitosis detection in breast histopathology,” in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 1770–1774, IEEE, 2020.

41. Hong, Y., Y. J. Heo, B. Kim, D. Lee, S. Ahn, S. Y. Ha, I. Sohn, and K.-M. Kim, “Deep learning-based virtual cytokeratin staining of gastric carcinomas to measure tumor–stroma ratio,” *Scientific Reports*, Vol. 11, no. 1, pp. 1–12, 2021.

42. Levy, J. J., C. R. Jackson, A. Sriharan, B. C. Christensen, and L. J. Vaickus, “Preliminary evaluation of the utility of deep generative histopathology image translation at a mid-sized nci cancer center,” *BioRxiv*, 2020.

43. Lahiani, A., J. Gildenblat, I. Klaman, S. Albarqouni, N. Navab, and E. Klaiman, “Virtualization of tissue staining in digital pathology using an unsupervised deep learning approach,” in *European Congress on Digital Pathology*, pp. 47–55, Springer, 2019.

44. Zhu, J.-Y., T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2223–2232, 2017.

45. Park, T., A. A. Efros, R. Zhang, and J.-Y. Zhu, “Contrastive learning for unpaired image-to-image translation,” in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16*, pp. 319–345, Springer, 2020.

46. Rosai, J., “Why microscopy will remain a cornerstone of surgical pathology,” *Laboratory Investigation*, Vol. 87, no. 5, pp. 403–408, 2007.

47. Fischer, A. H., K. A. Jacobson, J. Rose, and R. Zeller, “Hematoxylin and eosin staining of tissue and cell sections,” *Cold Spring Harbor Protocols*, Vol. 2008, no. 5, pp. pdb–prot4986, 2008.

48. Bancroft, J. D., and M. Gamble, *Theory and practice of histological techniques*, Elsevier Health Sciences, 2008.

49. Alturkistani, H. A., F. M. Tashkandi, and Z. M. Mohammedsaleh, “Histological stains: a literature review and case study,” *Global Journal of Health Science*, Vol. 8, no. 3, p. 72, 2016.

50. Foot, N. C., “The masson trichrome staining methods in routine laboratory use,” *Stain Technology*, Vol. 8, no. 3, pp. 101–110, 1933.

51. Leonard, A. K., E. A. Loughran, Y. Klymenko, Y. Liu, O. Kim, M. Asem, K. McAbee, M. J. Ravosa, and M. S. Stack, “Methods for the visualization and analysis of extracellular matrix protein structure and degradation,” in *Methods in Cell Biology*, Vol. 143, pp. 79–95, Elsevier, 2018.

52. “2 - histological and histochemical stains and reactions,” in *Muscle Biopsy (Fifth Edition)* (Dubowitz, V., C. A. Sewry, and A. Oldfors, eds.), pp. 14–23, London: Elsevier, fifth edition ed., 2020.

53. Kain, R., *Histopathology*, pp. 185–193. New York, NY: Springer New York, 2017.

54. Lu, M. Y., D. F. Williamson, T. Y. Chen, R. J. Chen, M. Barbieri, and F. Mahmood, “Data-efficient and weakly supervised computational pathology on whole-slide images,” *Nature Biomedical Engineering*, Vol. 5, no. 6, pp. 555–570, 2021.

55. Chen, T., S. Kornblith, M. Norouzi, and G. Hinton, “A simple framework for contrastive learning of visual representations,” in *International Conference on Machine Learning*, pp. 1597–1607, PMLR, 2020.

56. Glorot, X., and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pp. 249–256, JMLR Workshop and Conference Proceedings, 2010.

57. Ulyanov, D., A. Vedaldi, and V. Lempitsky, “Instance normalization: The missing ingredient for fast stylization,” *ArXiv Preprint arXiv:1607.08022*, 2016.

58. Seitzer, M., “pytorch-fid: FID Score for PyTorch.” https://github.com/mseitzer/pytorch-fid, August 2020. Version 0.2.1.

59. Ting, K. M., *Confusion Matrix*, pp. 260–260. Boston, MA: Springer US, 2017.

60. Brennan, D. C., T. Alhamad, F. A. Malone, and J. Vella, “Kidney transplantation in adults: Clinical features and diagnosis of acute renal allograft rejection,”

61. Jeong, H. J., “Diagnosis of renal transplant rejection: Banff classification and beyond,” *Kidney Research and Clinical Practice*, Vol. 39, no. 1, p. 17, 2020.

62. Javaeed, A., S. Qamar, S. Ali, M. A. T. Mustafa, A. Nusrat, and S. K. Ghauri, “Histological stains in the past, present, and future,” *Cureus*, Vol. 13, no. 10, 2021.

63. Rao, R. S., S. Patil, B. Majumdar, and R. G. Oswal, “Comparison of special stains for keratin with routine hematoxylin and eosin stain,” *Journal of international oral health: JIOH*, Vol. 7, no. 3, p. 1, 2015.

64. Black, J. G., and L. J. Black, *Microbiology: principles and explorations*, John Wiley & Sons, 2018.

65. Bai, B., X. Yang, Y. Li, Y. Zhang, N. Pillar, and A. Ozcan, “Deep learning-enabled virtual histological staining of biological samples,” *arXiv preprint arXiv:2211.06822*, 2022.

66. Hogan, J. J., M. Mocanu, and J. S. Berns, “The native kidney biopsy: update and evidence for best practice,” *Clinical Journal of the American Society of Nephrology*, Vol. 11, no. 2, pp. 354–362, 2016.

67. Fogo, A. B., A. H. Cohen, R. B. Colvin, J. C. Jennette, C. E. Alpers, *et al.*, *Fundamentals of renal pathology*, Vol. 1, Springer, 2014.

68. Herrera, G. A., and R. L. Lott, “Silver stains in diagnostic renal pathology,” *Journal of Histotechnology*, Vol. 19, no. 3, pp. 219–223, 1996.

69. Alotaibi, A., “Deep generative adversarial networks for image-to-image translation: A review,” *Symmetry*, Vol. 12, no. 10, p. 1705, 2020.

70. Tian, C., X. Zhang, J. C.-W. Lin, W. Zuo, Y. Zhang, and C.-W. Lin, “Generative adversarial networks for image super-resolution: A survey,” *arXiv preprint arXiv:2204.13620*, 2022.

71. Sandfort, V., K. Yan, P. J. Pickhardt, and R. M. Summers, “Data augmentation using generative adversarial networks (cyclegan) to improve generalizability in ct segmentation tasks,” *Scientific Reports*, Vol. 9, no. 1, p. 16884, 2019.

72. Yang, M., and J. He, “Image style transfer based on dpn-cyclegan,” in *2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, pp. 141–145, IEEE, 2021.

73. Saxena, D., and J. Cao, “Generative adversarial networks (gans) challenges, solutions, and future directions,” *ACM Computing Surveys (CSUR)*, Vol. 54, no. 3, pp. 1–42, 2021.