Correction of Artifacts in Formalin-Fixed Paraffin-Embedded Tissue Section Images with Contrastive Unpaired Image-to-Image Translation

by

Mohamad Kassab

B.S., in Biomedical Engineering, Bahçeşehir University, 2017

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
2022

ACKNOWLEDGMENTS

First and foremost, I would like to gratitude my parents, who have supported me tirelessly throughout my studies and during the research and writing of this master's thesis, for without them this achievement would not have been possible. I would like to thank my thesis advisor Assoc. Prof. Dr. Esin Öztürk Işık and the thesis committee for their valuable time and comments. I am grateful for the support and help of my co-advisor Dr. Mehmet Turan. I thank you for giving me the opportunity to be involved in the field of deep learning. Finally, I would like to gratitude my friends and everyone who supported me and contributed to the success of this dissertation.

ACADEMIC ETHICS AND INTEGRITY STATEMENT

I, Mohamad Kassab, 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

Correction of Artifacts in Formalin-Fixed Paraffin-Embedded Tissue Section Images with Contrastive Unpaired Image-to-Image Translation

Formalin-fixation and paraffin-embedding (FFPE) is a specimen preparation and preservation technique that has been used in histology and pathology since the late 19th century. Because the preparation of FFPE specimens is a complex, lengthy, and difficult to standardize process, and due to the complex histological and cytological characteristics of tissue, FFPE slides often contain defects. Defects arise during tissue fixation, processing, embedding, microtomy, staining, and coverslipping. These defects are referred to in images as artifacts, a term which encompasses staining inconsistencies, tissue folds, chattering, pen marks, blurring, air bubbles, and contamination, in addition to some other defects. We propose an unpaired image-to-image translation approach, FFPE++, which corrects artifacts in FFPE slides for digital pathology. Our method is a deep-learning-based approach which uses contrastive learning with spatial attention block and self-regularization loss, leading to higher quality in terms of both the visibility of textural details and cellular features. 10 board certified pathologists have performed comparative tests between our FFPE++ method and the standard FFPE sections of the ovary, thyroid, and lung, showing that our approach results in a visually coherent images for histopathological diagnosis.

Keywords: formalin-fixed paraffin-embedded, deep learning, unpaired image-to-image translation, generative adversarial networks, digital pathology, histological artifacts, histopathological image analysis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

FFPE	Formalin-Fixed Paraffin-Embedded
GAN	Generative Adversarial Network
H&E	Hematoxylin and Eosin
DP	Digital Pathology
TCGA	The Cancer Genome Atlas
OV	Ovarian Cancer
THCA	Thyroid Cancer
WSI	Whole-Slide Image
LUAD	Lung Adenocarcinoma
LUSC	Lung Squamous Cell Carcinoma
IHC	Immunohistochemistry
WSI	Whole Slide Image
ECM	Extra-Cellular Matrix
SAB	Spatial attention-based
MLP	Multi-Layer Perceptron
cGAN	Conditional GAN
CUT	Contrastive Unpaired Translation
I2I	Image to Image
DL	Deep Learning
G	Generator
D	Discriminator
AI	Artificial Intelligence
SR	Self-Regularization

1. INTRODUCTION

1.1 Motivation

Owing to the booming technology, pathologists have the privilege to examine specimens using digital whole-slide scanners that enabled rapid and reliable transfer of histopathological specimens from glass slides into a computing device that can quantitatively analyze the digitized histology images. Digital pathology may contribute to the accuracy of pathologists' diagnosis especially by incorporating machine learning techniques that improve conventional diagnosis by reducing image analysis processing time, and improving abnormality detection.

The quality of digital pathology data images is usually compromised by artifacts that arise from specimen acquisition to slide preparation and digitization. The presence of these artifacts can heavily hinder pathological examination and diagnostic procedures. Experienced pathologists can recognize these artifacts and may be able to interpret the necessary visual details despite their presence, but low-quality sections containing artifacts likely increase the time spent reviewing them and may also cause increased cognitive load, the only key for histologists and pathologists is to adjust their brains through constant practice so that artifacts and variations in staining are disregarded, and the critical visual details are extracted for best possible diagnosis [1, 2]. These limitations can be mitigated by a deep learning-based approach that addresses a common problem: artifacts in the examination routine of pathologists which may affect the accuracy of diagnosis due to the biases in the FFPE tissue sections.

Correction of artifacts can improve nucleus contours, chromatin structure, and nucleolar shape. These specific feature corrections improve workflow in studies such as cell or nucleus counting in slides of tissue and cell blocks.

In addition to improving image quality for direct examination by pathologists,

such a computational approach may also be a critical step toward better machine learning algorithms that enable partial or full automation of these studies. Last but not least, the FFPE++ approach can be used as a quality control tool in digital pathology.

1.2 Objectives

The focus of this study is to correct the following artifacts in the ovary, lung, and thyroid FFPE sections: 1) staining inconsistencies, 2) tissue folds, 3) chattering, 4) pen marks, 5) blurring, 6) air bubbles, 7) and contaminants.

To conduct comparative testing by board certified pathologists to determine if our proposed method is helpful in their pathological diagnosis.

2. BACKGROUND

2.1 Formalin-fixed paraffin-embedded (FFPE)

Formalin fixation and paraffin embedding is considered the gold standard method for preserving tissues from clinical specimens in hospitals. This technique is used all over the world in daily histopathological practice, in the diagnosis of cancer for example. The formalin fixation protocol preserves the proteins as well as the tissue morphology for later analysis [3]. The importance of this technique to pathologists is its ability to preserve rare diseases and micro-scale lesions with high-quality images for further clinical studies that span years. In addition, it is considered a cost-effective technique compared with frozen sections preservation because there is no need to worry about risks such as freezer maintenance, the need for technicians, and other additional costs to keep tissues from decaying [4, 5].

Whole slide images (WSIs) are created from slides obtained from excised tissue samples via a specific protocol: First, they are immediately subjected to a formalin fixation which aids harden and preserve the tissue to protect the proteins and morphological structure within the tissue, then dehydrated in an alcohol series, cleared in xylol agent to make sure that alcohol is absent, then embedded in paraffin blocks which allow the microtome blade to make the cutting of the tissue of desired sizes easier and finally samples are mounted on a glass slide for staining and coverslipping, then converted into digital slides using a whole slide scanner[6]. Moreover, to avoid tissue degeneration, it is important to preserve the FFPE sections under certain humidity-controlled conditions. In addition, specimens taken for molecular examination shall be cut in the nearest time before the analysis to circumvent possible antigenicity due to oxidation and hydration ramifications [7].

Typically, there are numerous applications of this technique, particularly in immunohistochemistry (IHC) [8] which is an essential tool for researchers looking for

indicators of neurodegenerative diseases and cancer. It involves coating tissue sections with a substance carrying antibodies that attach in return to specific proteins known as antigen-antibody reactions [9]. Researchers were able to extract DNA, RNA, and some other proteins from FFPE samples for biomarker identification [10]. It is also used in the field of hematology for the sake of detecting blood-related diseases caused by anomalies, and to develop drugs [11]. Also, it has been used in the field of immunological investigations, for instance, to detect diseases caused by viral infections [12]. It is very important to preserve the FFPE biospecimens under certain humidity-controlled conditions to avoid tissue degeneration. In addition, specimens taken for molecular examination shall be cut in the nearest time before the analysis to circumvent possible antigenicity caused by oxidation and hydration ramifications [7].

2.2 Artifacts in histological sections

Artifacts in histological slides are diverse and include distortions of cellular and morphological details. Figure 2.1 shows some possible artifacts that pathologists often observe in their examination routine. For instance, tissue folds veiling cellular details (**a**), chattering of rigid sections (**b**), under-staining (**c**) and over-staining (**d**) referring to staining inconsistencies, blurring due to loss of microscope focus (**e**), pen marks affecting the nuclear information (**f**), air bubbles that leave cracked areas and altered stains (**g**), and contaminants (**e**). The presence of these artifacts can cause malignant cells to become unrecognizable and benign cells to appear atypical.

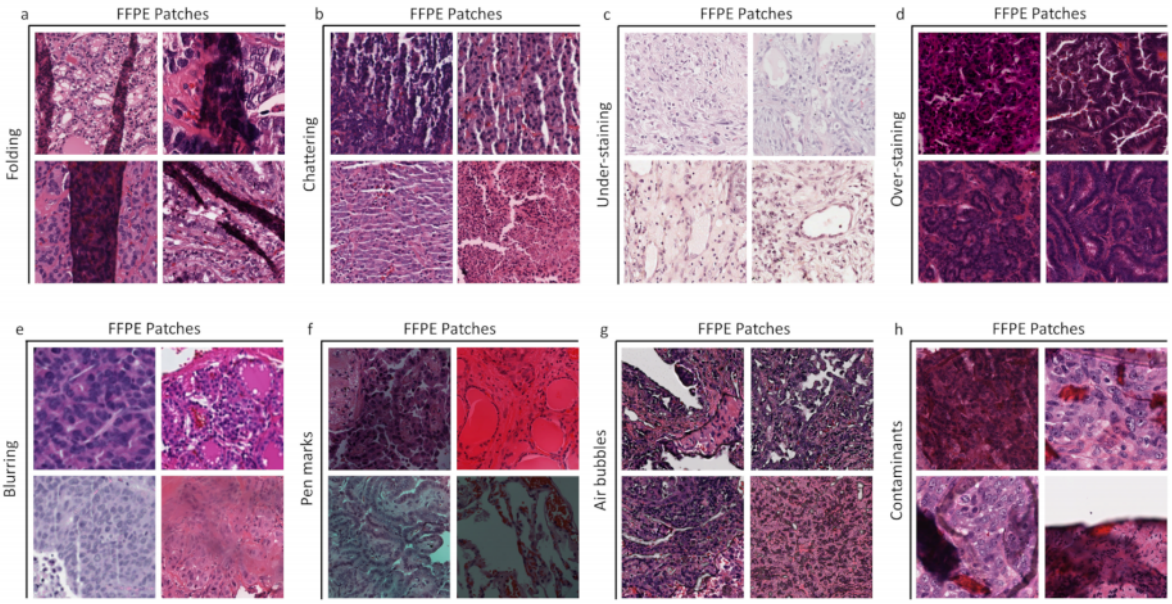

Figure 2.1: A sample of various histological artifacts representing the FFPE thyroid, lung, and ovary biospecimens.

Sometimes the presence of multiple artifacts may lead in uncommon cases to grave errors and misdiagnosis. The extent of artifact-related damage is very important in some specimens and may affect nearly the entire specimen, so it can be critical to diagnosis purposes. Any structure or a characteristic that is formed by tissue processing is known as an artifact. Some of them are easily distinguished from components of normal or diseased tissue, while others are challenging to distinguish from such entities. In histological and cytological terms, an artifact can be described as a form of tissue which is not normally present in the normal tissue that contains cellular components and extra-cellular matrix (ECM) structurally and/or visually in the specimen and prevents the microscopist from assessing the slide. The concern is to identify histological artifacts when they occur and to understand the true nature of normal tissue components or pathological changes in the whole slide image. [13] explained the various causes of artifacts. Artifacts in FFPE sections may occur 1) during tissue collection, 2) during tissue fixation, 3) during tissue processing, 4) during tissue embedding in paraffin, 5) throughout the microtomy of paraffin sectioning 6) while placing the specimen on the glass slide, 7) during staining of the tissue and coverslipping. For instance, inadequate fixation of the tissue due to thick sections, an FFPE block that is too warm or some-

times a water bath that is hotter than usual, and a mediocre flotation technique that results in tissue folds in the FFPE section. The FFPE slide is susceptible to chatter artifacts if an FFPE block is cut quickly while it is still cold. They are susceptible to contaminants, which can come from a variety of factors, such as the starch in surgical gloves. Another factor can be the use of unclean water during the flotation bath, which can lead to the presence of fungi and molds. Failure to check the glass slides for cleanliness from previous tissue residuals during the staining process will result in contaminating the tissue. Also, air bubbles can result from collapsed bubbles that are under the tissue sample that is flattened in the water bath. Over time, these bubbles can leave cracked spots as well as altered stains [14]. Blurred areas caused by loss of microscope focus. Permanent ink markers form a part of the medical records used by pathologists to label or highlight a tumor region, which hinders the data underneath the marked region after the tissue is digitized. However, pen marks can be extracted chemically but there is a high risk of deteriorating the tissue [15].

2.3 Image-to-image-translation

With the advent of Generative Adversarial Networks (GANs), many researchers have been encouraged to develop novel network architectures that overcome various limitations in computer vision, especially in the traction of image-to-image translation. The basic idea of GAN is to use two neural networks, a generator and a discriminator, trained together so that the generator synthesizes realistic images that resemble a given data image and tries to fool the discriminator, which tries to differentiate real data images from those synthesized by the generator. The aim of the image-to-image approach is to convert an image from a source domain to the content of an image from the target domain to generate an output image, while preserving the content. More specifically, content means the features of the image such as cell nuclei, chromatin, and tissue structure.

Pix2Pix is one of the most successful supervised image translation techniques that learns to map an input image to a matching output image using conditional GAN

(cGAN) and enforces an L1 loss between the generated image and its ground truth map [16]. In supervised image-to-image tasks, researchers use massive pairs of aligned images (i.e., corresponding image samples) as the source and target domains to obtain the translation model that translates the source images into the desired target images.

Despite the better results obtained with the paired models, it is not feasible to collect a large amount of labeled data images corresponding to the input images. Therefore, the use of unpaired datasets enabled unsupervised frameworks to gain particular attention from researchers. CycleGAN [17] is an unsupervised image-to-image technique that proposes cycle consistency loss. It enforces correspondence between the source and target images so that basically an inverse translation function is learned simultaneously to minimize the pixel reconstruction loss between the source and target domains along with the adversarial loss, least square GAN. DualGAN [18] is similar to CycleGAN but uses the Wasserstein GAN loss instead, DiscoGAN [19] adds a reconstruction loss function that measures how well the original input image is reconstructed, UNIT [20] works with the shared latent space supposition, by assuming that two matching images in the opposite domain can be mapped to the same latent representation in a shared latent space, and MUNIT [21] designed a new architecture by presuming that each image can be decomposed into domain-specific features and domain-invariant content features, and these latent content and style features are exchanged to perform image translation. Moreover, the unsupervised attention-guided GAN [22] uses attention learning alongside the adversarial network, which exhibits the synthesis of more cogent images compared to classical GANs. HarmonicGAN [23] applies a smoothing parameter that yields robust similarity consistency among images in the source and target domains. U-GAT-IT [24], an unsupervised image translation model, employs an attention module that controls the geometric alterations between source and target domains, and the adaptive layer instance normalization function that assists the attention module to adjust the proportion of changes in both texture and outline. With the exception of pix2pix, all of the previously mentioned frameworks are used for training with unpaired or unaligned datasets. Most often, the unpaired image translation frameworks which utilize the cycle consistency constraint (CycleGAN, DualGAN, DiscoGAN, UNIT, MUNIT, U-GAT-IT) fail when the source and target

domains involve clutter and heterogeneity rather than minor changes.

More recently, [25] has incorporated contrastive learning by maximizing the mutual information between corresponding images of the input and output in a patch-based manner, rather than considering the entire image. [26] showed that the CUT model has impressive superiority over state-of-the-art unpaired image translation methods in both Fréchet Inception Distance (FID) and Inception Score (IS). [27] claim that unpaired models using cycle consistency loss struggle during training when shape changes, extracting objects from an image, or ignoring insignificant textures. Consequently, [28] found that CycleGAN and UNIT models are skeptical and generally insert unnecessary changes or artifacts into the output image. This encouraged us to use a different network architecture, especially in the case of histological artifact correction, where we seek stability between artifact correction and not changing the pathological required cellular information during training. Therefore, we adapt the CUT network architecture by adding the self regularization loss and the SAB modality to optimize the image translation model.

2.3.1 Image translation in histopathological imaging analysis:

Image translation frameworks have been used extensively in medical image analysis over the past decade, such as segmentation [29], super-resolution [30], and domain adaptation [31]. They have achieved compelling results, especially in the field of digital pathology, where the accessibility of whole-slide images (WSIs) of tissue samples has made digital pathology a mining ground for deep learning. For example, tasks such as nuclei segmentation [32, 33, 34], tumor detection and disease classification [35, 36], cancer grading [37, 38], detection of blur artifacts in tissue sections [39], and normalization of staining is of particular interest in histopathological image analysis for sections stained with hematoxylin and eosin (H&E) [40].

Due to the costly annotation of nuclei in H&E stained sections compared to immunohistochemistry (IHC) staining techniques, [41] used a domain adaptation tech-

nique that aims to map images of labeled nuclei stained with IHC to H&E stained sections for the purpose of data augmentation and the outcome result is later used to train a detection network on the newly generated labeled H&E stained images. Accordingly, [42] utilized an unpaired image-to-image translation technique to translate different stains for kidney biospecimens, and the resulting output is then passed to a network for semantic segmentation as a target task. [43] modified CycleGAN by adding two structural loss functions to the original network architecture [17] to improve the translation process of H&E stained tissue sections into IHC. [44] applied the pix2pix framework to convert H&E images to unstained images and contrariwise.

3. MATERIALS & METHODS

The first step in FFPE preparation is tissue acquisition, that is the patient must undergo a surgical procedure for tissue excision as shown in Figure 3.1 (**a**). Once the tissue samples are collected (**b**), they are immediately sent to the pathology laboratory for FFPE processing. Here, there are 6 main steps in preparing FFPE tissue slides: 1. fixation, 2. dehydration, 3. clearing, 4. embedding, 5. sectioning, and 6. staining, and this process usually takes about 36 hours (**c**). After the FFPE slide is ready, it is sent for pathological examination (**b**). Consequently, we show how our framework can improve the quality of FFPE slides that contain artifactual sections and are then given for pathological analysis. The FFPE++ pipeline is fed with a FFPE WSI containing artifacts as input. Then, N-by-N patches are created for each FFPE WSI (**e**). Finally, the artifact correction transformer network processes the FFPE images with artifacts one patch at a time by converting the artifactual FFPE image patches into a corrected version (**f**).

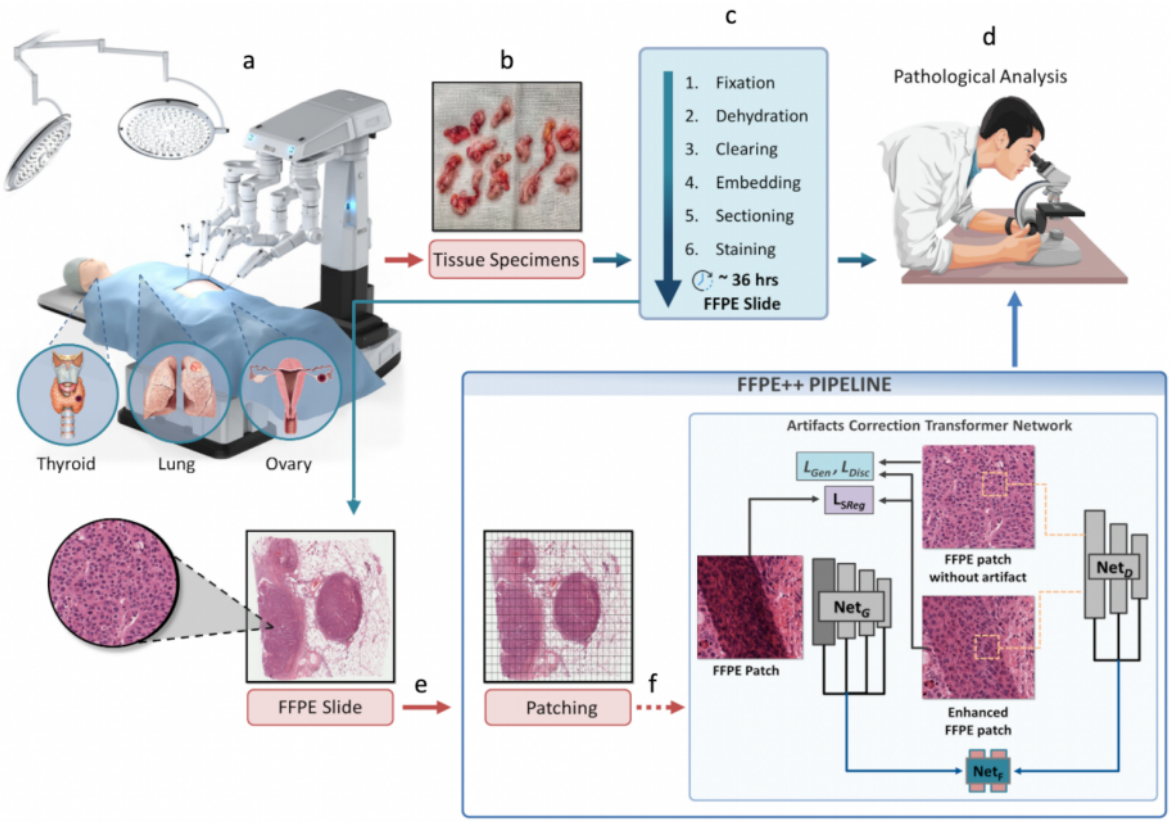

Figure 3.1: Workflow of the system. A graphical representation showing how the network pipeline fits into the histopathological evaluation of the biopsies taken.

3.1 Whole-Slide Images (WSIs) Processing:

After the WSIs are selected for each specimen, they are subjected to contour segmentation to indicate the tissue regions that need to be cropped into stacks of image patches (e.g., 512x512) without overlap, since the FFPE++ algorithm cannot be trained with the very large size of the WSIs. For this purpose, the tool CLAM [45] is used, which stores the segmented contours of the foreground at the magnification specified by the user. This results in a total number of image patches that can range from thousands (i.e., tissue sections at x20 magnification) to hundreds of thousands (i.e., tissue sections at x40 magnification). The Cancer Genome Atlas - Thyroid Cancer (TCGA-THCA), TCGA-LUAD (lung adenocarcinomas), and TCGA-LUSC (squamous cell carcinoma) a magnification of 20 x is used, and for TCGA-OV (ovarian cancer), 40x magnification is used to evaluate the performance of the FFPE++ approach at different magnifications.

3.2 Datasets:

In this master's thesis, we perform the experiments using FFPE slides from the open source Cancer Genome Atlas (TCGA). The open source TCGA datasets were used for both the training and testing phases of the FFPE++, taking into account the distribution of patients in terms of age, gender, and ethnicity. Regions that contained artifacts in the tissue sections were manually selected, and the same was done for regions of high quality.

Ovarian serous cystadenocarcinoma (TCGA-OV), which is the most common malignant form of ovarian epithelial tumors. Lung adenomas (TCGA-LUAD) and lung squamous cell carcinomas (TCGA-LUSC), these diseases are considered a leading cause of death among humans across the globe [46]. Papillary adenomas and adenocarcinomas of the thyroid (TCGA-THCA), the most common subtypes of thyroid cancer, occur in approximately 8 out of 10 cases [47]. These tissues are examined to correct for artifacts in FFPE slides. This is because tumors arising from these tissues have many problems, including fragility, cystic and air spaces, colloidal material. These neoplasias were selected based on their characteristic features. Ovarian neoplasias have unique cystic spaces. Lung neoplasias have large air spaces and complicated structure. Thyroid neoplasias have a delicate structure due to the colloidal material. Also, the nuclear characteristics of the tissue, which are important for the diagnosis of cancer.

For the ovary, we used FFPE WSIs from 38 patients from TCGA-OV. The subset consists of 11,638 artifactual FFPE patches and 10,514 high-quality patches selected for the study. Lung carcinomas are divided into small cell and non-small cell lung cancers (NSCLC). NSCLC is further subdivided into LUAD and LUSC. FFPE Whole slide images TCGA-LUAD and TCGA-LUSC projects of 177 patients are included. The subset includes 17,277 artifactual FFPE patches and 18,221 high quality images. The FFPE thyroid includes 15,351 image patches with artifacts and 23,761 FFPE patches of high quality from 230 patients.

3.3 Network Architecture:

Unlike other unpaired image translation models that require bidirectional translation due to cycle consistency constraint, they are more demanding on GPU memory and slower in training time. Therefore, we adapted the GAN network architecture from the contrastive unpaired translation (CUT) framework, which learns mapping in one direction (source to target) and simplifies training time by using a generator and a discriminator. Our objective includes four types of terms (see details in Figure 3.2**b**, and **c**): GAN loss for matching the generated images to the images in the target domain, patch noise contrastive loss, which focuses on the content correspondence between the artifactual FFPE and the FFPE++, self regularization loss is incorporated to encourage the output image to have diagnostically relevant features that are closer to the high-quality FFPE images, and SAB module is incorporated into the generator network for artifact-aware FFPE image synthesis.

As shown in the network architecture Figure 3.2, FFPE-WSIs are processed and cropped into patches of size 512x512 which are given to the "FFPE Artifacts Correction" network that converts the FFPE images with artifacts into a visually better quality patches. Then, we perform stitching of all corrected FFPE patches to create the enhanced WSI of the same tissue (**a**). The generator transforms an artifactual FFPE patch (input) 512x512 sized into corrected FFPE patch (FFPE++ output), then the output image is then given to the discriminator as well as a patch from the artifacts-free FFPE patches domain. Subsequently, patch-NCE loss is incorporated into the final objective by assisting the networks to focus on the commonalities among the source and target domains as well as the self regularization which focuses on the nucleo-cytoplasmic content preservation and mitigate the severity of artifacts correction. A small MLP network based on two layers is added to improve the content correspondence between the input image (artifactual FFPE) and the output image (FFPE++) (**b**). The SAB module is integrated to the ResNet generator for the synthesis of a visually better morphological structure and assists the model from filling vessel's lumen with an ECM (**c**).

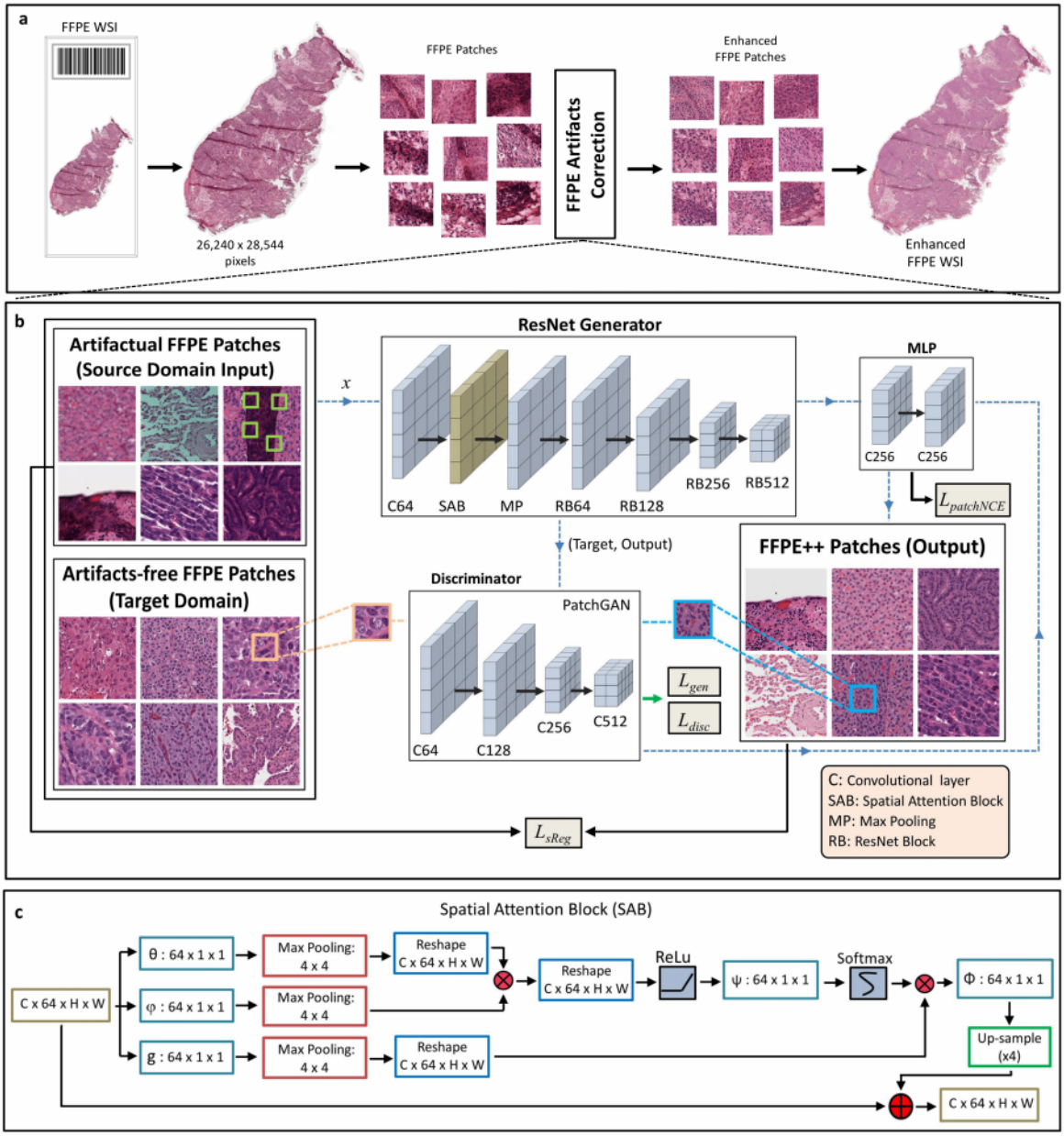

Figure 3.2: FFPE++ network architecture.

GAN loss is our first component of the total loss function:

$$\mathcal{L}_{GAN}(G, D, X, Y) = \mathbb{E}_{y \sim Y} \log D(y) + \mathbb{E}_{x \sim X} \log(1 - D(G(x))) \tag{3.1}$$

The source domain X contains x samples of FFPE patches with artifacts and the target or style domain Y) contains y samples of high quality FFPE patches without artifacts. The role of the generator G is to generate images $G(x)$ that look very similar to y in the target domain Y, and simultaneously the discriminator $D(y)$ strives to distinguish $G(x)$ from y in the target domain.

Recently, contrastive learning has been incorporated to tackle several unsupervised visual representation learning problems [48, 49, 50]. Accordingly, the task of image translation becomes more challenging when image samples from one domain comprise additional details and features in contrast to the second domain, which is the case in histopathological image analysis due to morphological complexity and tissue characteristics. Herein, the noise contrastive estimation loss denoted by InfoNCE loss has been utilized, since it guarantees content preservation [50]. It is used in a multilayer patch-wise manner to retain the similarities of the image features between the input artifactual FFPE image and the corrected FFPE output image. In other words, contrastive learning works on matching two signals, a **query**(Q) patch from the "output image" and its **positive patch** (P) from the "input image" (see Fig. 3.2**b**), whereas the other points in the image that are dissimilar, are marked as negatives (N). The Q, P, and N are mapped to L-dimension vector $v, v^+ \in \mathbb{R}^L$ and $\in \mathbb{R}^{N\times L}$, individually. $_n \in \mathbb{R}^L$ denotes the n-th negatives in L samples. We use scaling the distance between the query and the other example by a temperature τ. The cross-entropy loss is computed, showing the likelihood of prioritizing the P examples among the N.

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

The generator, which synthesises the images, has already created its feature pile z_l^s. This means that each spatial location s and each layer l in this feature pile represents a patch of the input image. Then, L-layers of interest are chosen and the generator-derived feature maps are given as input to a small two-layer network called

the Multi-Layer Perceptron (MLP) F_l, as proposed in SimCLR [48], bringing on a pile of features $\{z_l^s\}_L = \{F_l(G^l(x))\}_L$, in which G^l is the output of the $l-th$ chosen layer. We specify in layers $l \in \{1, 2, ..., L\}$ and denote $s \in \{1, ..., S_l\}$, where S_l represents the amount of spatial locations in each layer, and the corresponding image features as z_l^s $\in \mathbb{R}^{C_l}$ and the dissimilar extracted features $z_l^{S\setminus s} \in \mathbb{R}^{(S_l-1)\times C_l}$, where C_l belongs to the channel numbers in per layer, and equivalently $\{\hat{z}_l^s\}_L = \{F_l(G^l(x))\}_L$ represents the pile of features of the output image. MLP learns to estimate the input and output patches in a common embedding space.

The patch noise contrastive estimation is defined as follows:

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

Moreover, we claim that the generated and input images should share similar low-level features, such as the staining color, the outline of the nucleus, the shape of the nucleus, etc. Therefore, we enforce this constraint with a self-regularization loss added to the network architecture to restrain the generator from adding additional histopathological details during the image translation process, which could lead to misinformation.

$$\mathcal{L}_{Self-Regularization}(G, X) = ||X - G(X)||_1. \tag{3.4}$$

The final objective is as follows:

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

Here λ_{sReg} is an adjustable weight hyper-parameter to scale the distance between output and input images. The identity loss λ_{NCE} is a weight term aids to maintain the color of the input images, which is important in the case of staining variations.

SAB: details of the SAB flow-diagram is given in Figure 3.2**c**. We implemented SAB after the first convolutional block of ResNet generator. This spatial block is comprised of three convolution layers that break down the input image $\mathbf{X}$ into θ, ϕ, and g as shown in Figure 3.2**c**). The motivation behind implementing SAB mechanism is to assist the generator to focusing on the histopathological features and regions with higher diagnostic relevance in any input image $\mathbf{X} \in \mathcal{R}^{C\times 64\times H\times W}$, where C is the number of channels, 64 is the size of convolution filter, H and W indicate the height and width of the input, respectively. Thus, the non-local procedure can be described as follows:

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

Considering that f characterizes the correlation or correspondence of each pixel to another on the input $\mathbf{X}$. After reshaping $\mathbf{X}$, a dot product is performed on two convolution layers of the three components θ, ϕ, then ReLU is used as an activation function.

$$\mathbf{K} = \psi(\sigma_{ReLU}(\theta(\mathbf{X})\phi(\mathbf{X}^T)). \tag{3.7}$$

We denote by $\sigma_{(ReLU)}$ the ReLU activation function. The dot product of $\theta(\mathbf{X})\phi(\mathbf{X}^T)$ can be viewed as the covariance of the input image by measuring the degree of tendency between two feature maps on different channels. Then, we perform the ψ convolution in the $softmax$ function, accompanied by matrix multiplication between g and the output obtained from the $softmax$ function. The output resulting from the multiplication is then convolved with ϕ and up-sampled to obtain an attention map $\mathbf{P}$. Finally, an element by element sum operation between the $\mathbf{P}$ and $\mathbf{X}$ yields the output $\mathbf{F} \in \mathcal{R}^{C\times 64\times H\times W}$.

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

$$\mathbf{F} = \mathbf{P} + \mathbf{X}. \tag{3.9}$$

Here $\sigma_{softmax}$ stands for the softmax function. Besides, a short connection between $\mathbf{X}$ and $\mathbf{F}$ completes the spatial block and allows it to learn the residual mapping effectively.

3.4 Modification of Target Domain Dataset:

In the primary training phases of the ovarian cancer, the network had difficulty distinguishing filling chattering artifacts from the vessels' lumen and white regions with a stromal structure and ECM, which is problematic. Additionally, the concentration of erythrocytes in the FFPE patches in the target domain was insufficient, which was a major problem in image translation process. Therefore, blood smear images were added to the target dataset to increase the frequency of such patches so that the model would recognize the erythrocytes, besides, increasing the number of patches that contain vessel's lumen and white regions helped the model to learn more effectively as shown in Figure 3.3.

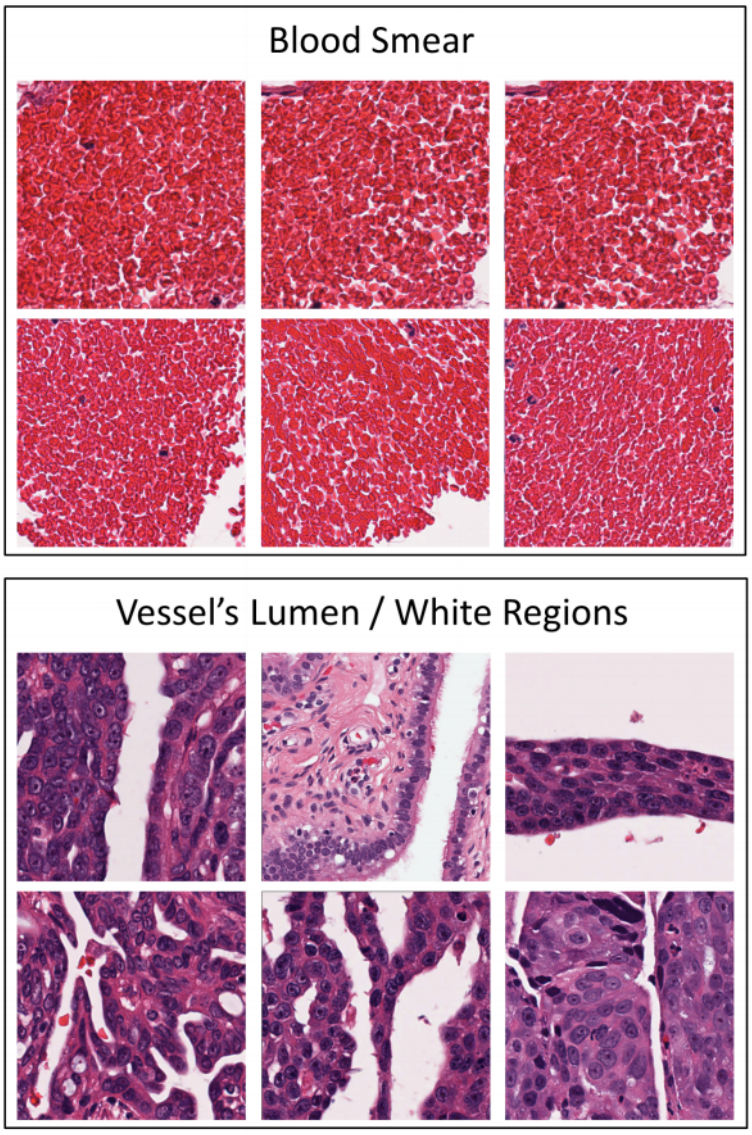

Figure 3.3: Sample images of Blood smear patches and vessel's lumen/white regions.

The outcome of the proposed implementation of blood smear patches alongside the vessel's lumen/white regions can be seen in Figures 3.4. These patches urged the model to distinguish chattering artifact and to detect RB cells among other morphological features within the FFPE tissue (see Figure 3.4, **Reconstructed FFPE Patch v2**). Similarly, this strategy was performed when training lung and thyroid FFPE tissues.

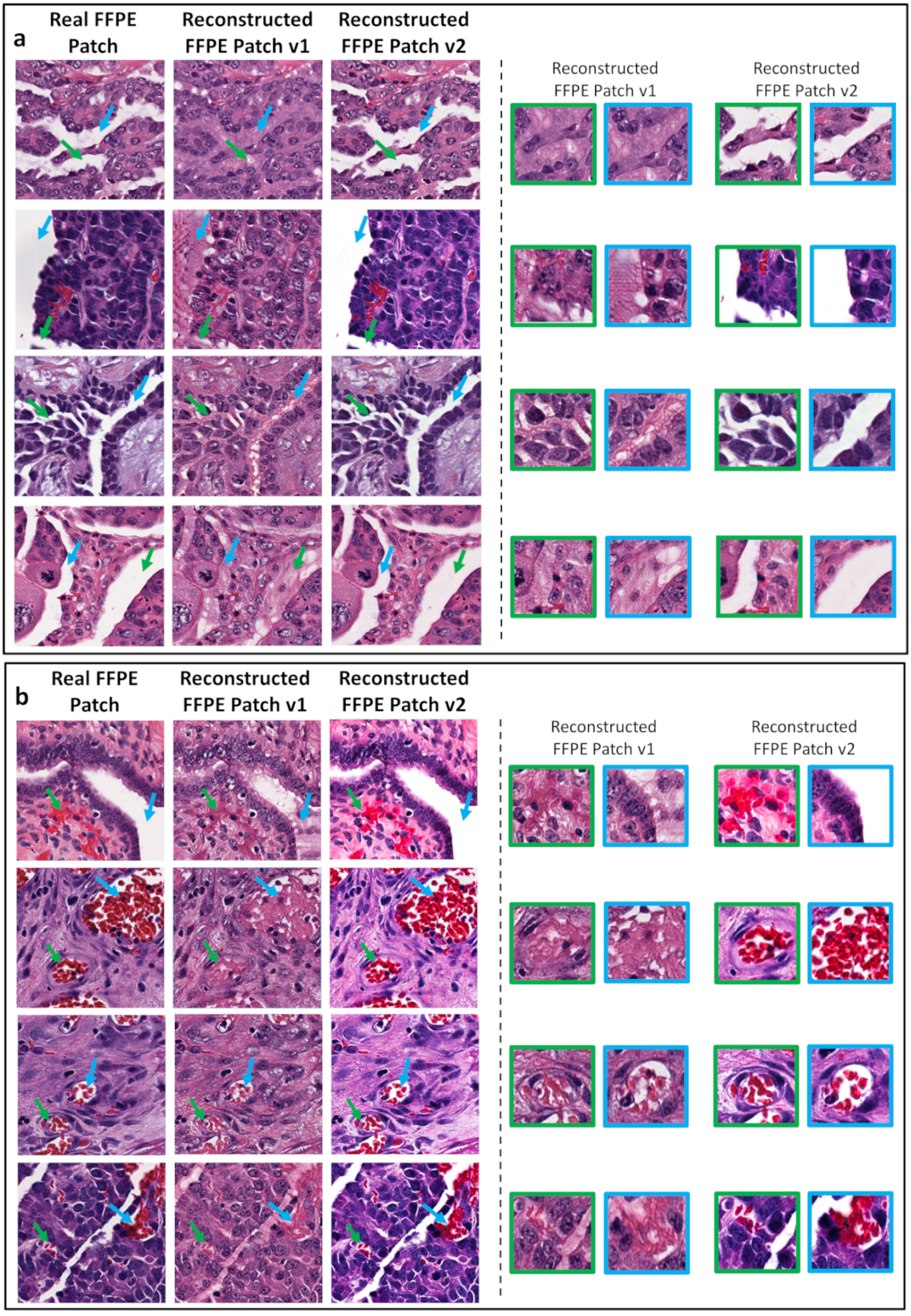

Figure 3.4: Comparison before and after the implementation of the proposed patches.

3.5 Training Details and Hyper-parameters:

To train any unpaired I2I translation model, there must be two sub-folders to host images from the source domain (artifactual FFPE patches) denoted by **trainA** and also from the target domain denoted by **trainB** (high quality FFPE patches), which means that the image translation task will be from **(AtoB)**.

We trained the FFPE++ algorithm from scratch for 200 epochs using a batch size equal to 1, we used Xavier [51] for weight initialization, instance normalization [52], Adam optimizer [53] with a learning rate equal to 0.0002.

The ResNet-based with 9 residual blocks is employed as the encoder-decoder generator [54]. An input image is down-sampled or encoded to a bottleneck layer, then the encoded layer is analyzed using a ResNet that uses skip connection layers, and finally, a series of up-sampling or decoding layers to obtain the original size of the output image. Additionally, we utilized the PatchGAN as our GAN discriminator [16], where its role is to classify N x N patches in a whole image, whether it is from the target domain or an output image, used a Least Square GAN loss [55], since it is provides stable training and visually better output image. The identity loss λ_{NCE} parameter was set as 2 for the patch noise contrastive loss, and the temperature value was set as $\tau = 0.07$. The identity loss aids to maintaining the color of the input images which is essential in the case of staining variations.

3.6 Hardware & Software:

In this master's thesis the experiments were performed using the consoles and computational hardware in the DeepMIA laboratory at the institute of BME under the supervision of Dr. Mehmet Turan as well as the web user-interface tool supported by the turkish science e-infrastucture "TRUBA" for providing their high computing services such as GPU, CPU and data storage. The ubuntu software was mainly used

as the operating system to conduct the study more efficiently in terms of performance and compatibility of deep learning packages. 64-bit 3.7 python and Anaconda3 were used to install the deep learning dependencies and packages, the curated training and testing phases of the experiments were performed using Pytorch implementation version 1.7, one NVIDIA V100 GPU with 16 GB was used along with CUDA 10.0 toolkit and cuDNN 7.5 to produce the fastest and best possible results. The FFPE WSIs processing and patching was done using the Intel Xeon multi-core CPU in the lab.

4. RESULTS

FFPE++ significantly ameliorated slide artifacts in all three tissue types and improved visualisation of cytological and nuclear detail and staining quality, as shown in Figures 4.1, 4.2, and 4.3. For instance, cellular details in the folding area, which cannot be read in any way in original sections, can be recognized very clearly after being corrected 4.1(a), 4.2(a), and 4.3(a). FFPE++ completely excluded the pen marks artifact as seen in Figure 4.1(e), Figure 4.2(e), and figure 4.3(e) and improved the cytological, nuclear, and the stromal structure of the underlying tissue. Most importantly, FFPE++ preserved tissue features, with no change in areas other than the artifact. Nuclei, nucleoli, cytoplasm, and chromatin that could not be recognized at all in the blurred region of the slide became fully readable after FFPE++ and may make it possible to reach a better diagnosis Figure 4.1(c), 4.2(c), and 4.3(c).

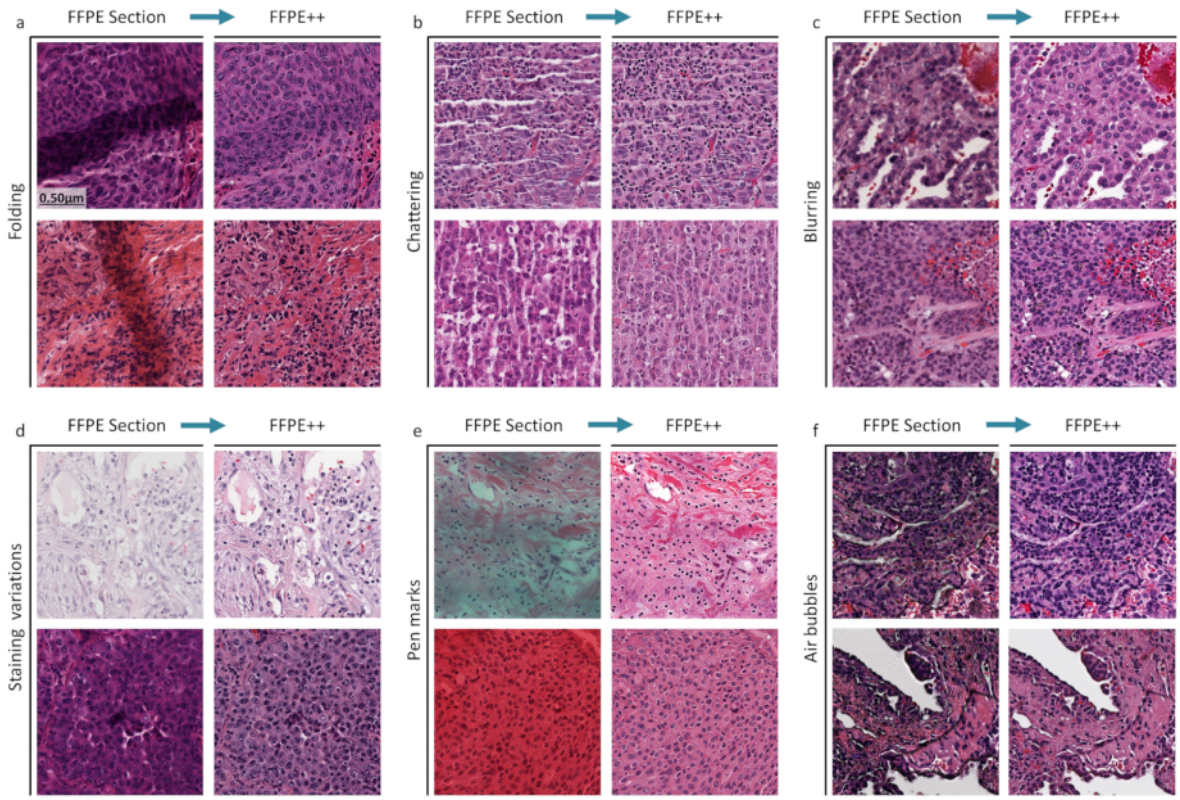

Figure 4.1: Artifacts correction in the lung FFPE sections.

In Figure 4.1, tissue folds affect the cellular density and make the tissue to seem uneven and obscure some cytological information because tissue cells are superimposed. The network adjusts the color and contrast of the folded region and allows a better identification of nuclei, nucleoli, chromatin, and cytoplasm (a). The model corrects the chattering artifact that causes to have unneeded white spaces by filling the stromal component resulting the overall tissue to look more cohesive (b). As can be seen in the blurring artifact correction examples (c), FFPE++ rectified the blurred regions and the image has improved immensely, and recovered the lost details in the extracellular matrix. Under-staining and over-staining of tissue sections is problematic for pathological diagnosis, and the quality of staining is affected by several factors that may contribute to misdiagnosis. Therefore, as seen in both images (d), the network corrected the color intensity, sharpness, and contrast of the FFPE tissue and provided a more comprehensive visualization of the tissue infiltrated with lymphocytes, nuclei, erythrocytes, and neutrophils. Correcting pen marks in FFPE sections is a major challenge for the network to make the image look visually coherent while preserving the morphological content of the tissue (e). Air bubbles inhibit pathologists' ability from properly visualizing the slide area because of the poor floatation technique, the contrast looks excessive, the cellular, nuclear, and cytoplasmic details in the bubble region cannot be comprehended, and the staining appears darker than it is. As shown by the corrected FPPE column compared to its earlier version (f), resulting in a worthwhile tissue for pathological evaluation.

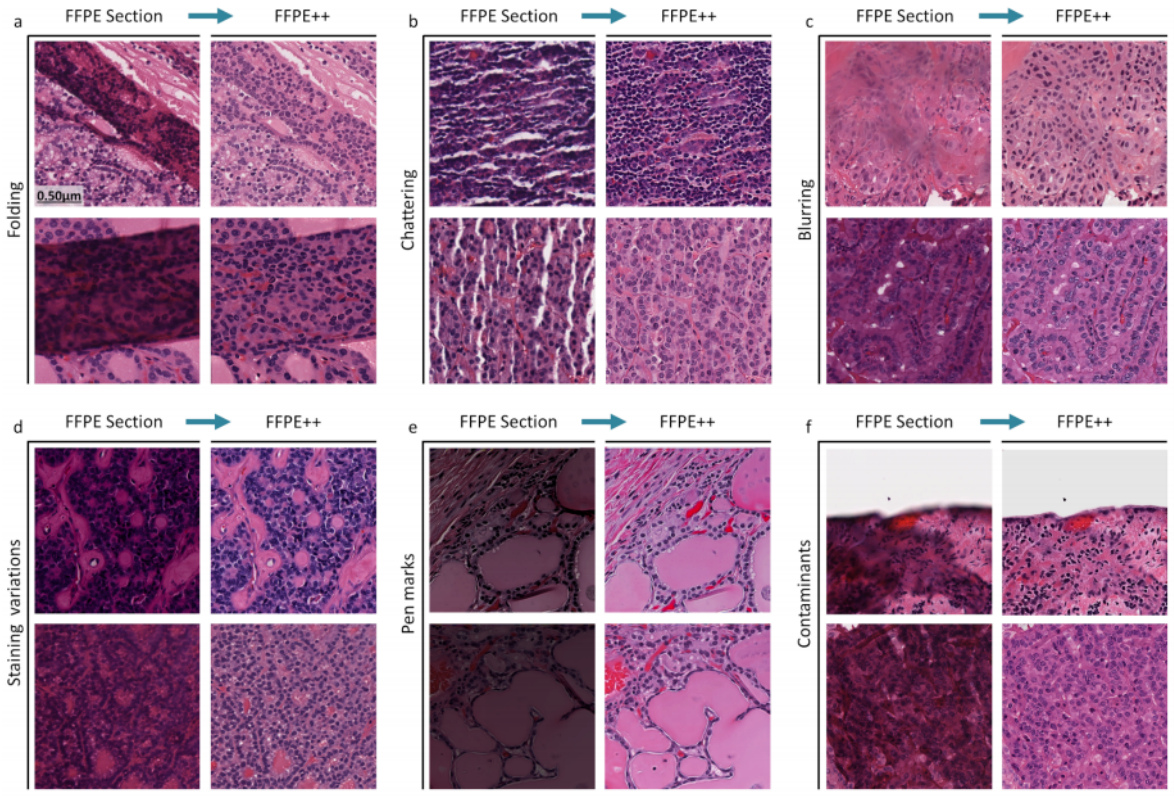

Figure 4.2: Artifacts correction in the thyroid FFPE sections.

In Figure 4.2, folding artifact veils some nuclear features due to their superimposition. The algorithm adjusts the folded region, resulting in a brighter region with smoother and more convenient to visualize cells (a). FFPE++ successfully fills the stromal component caused by the chattering artifact and renders the overall image to appear more cohesive (b). The loss of camera focus prevents pathologists to visualize cellular details, therefore increasing the sharpness of blurred images is very helpful in pathological diagnosis (c). Staining variations in FFPE tissue sections could lead to miss-leading diagnosis. The network improved the color intensity, sharpness, and contrast of the FFPE tissue and allowed a more comprehensive patches (d). Permanent ink markers are very challenging to correct, however FFPE++ is able to extract it (e). Contamination may occur if unclean water during the flotation bath contains tissue residuals from previous block, this will overlap the original section on the slide, the overlap of two tissue sections in the same region prevents the real tissue from being comprehensible. Although FFPE++ provides a complete understanding of the main section and successfully corrects the contamination in this area, it is critical not to

forget the presence of foreign tissue. Furthermore, in some cases contaminated areas are understandable, pathologists should avoid using these findings for better diagnosis (f).

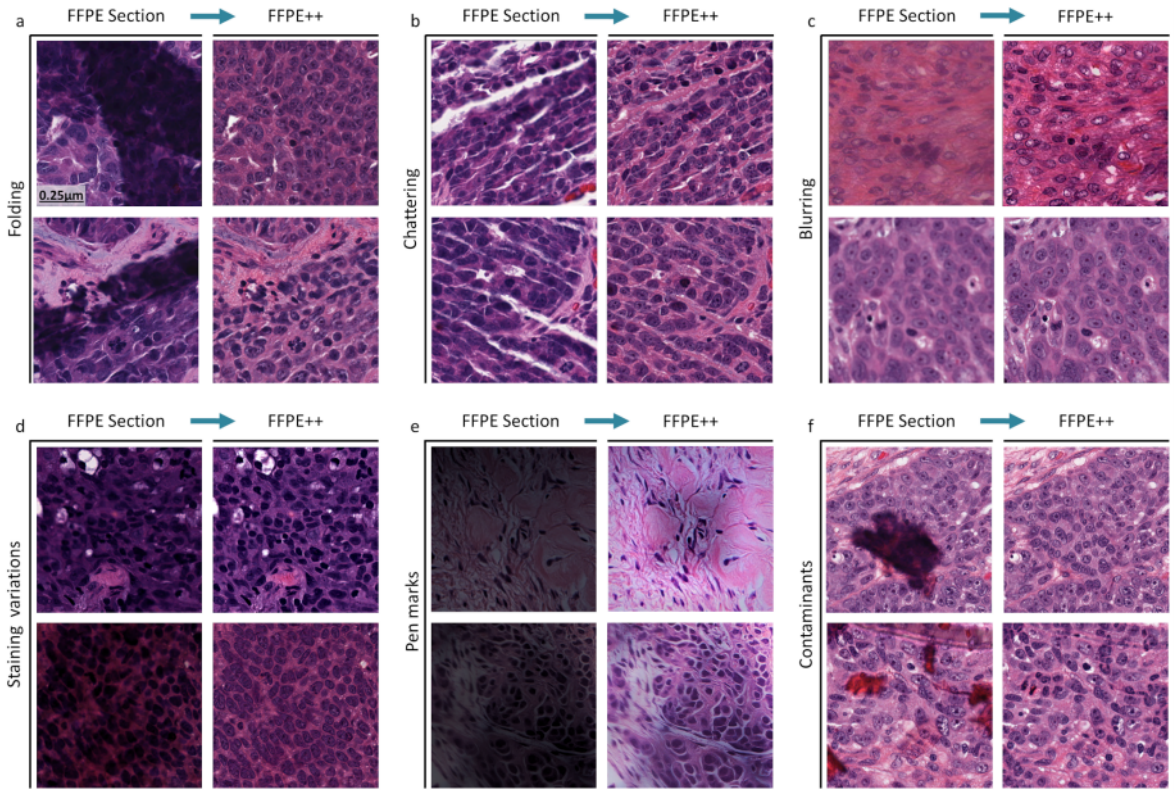

Figure 4.3: Artifacts correction in the ovary FFPE sections.

In Figure 4.3, all artifact classes are the same as in thyroid 4.2. The FFPE++ network was able to detect nuclear features that were very difficult for pathologists to see (a). White spaces which caused by the chattering artifact were successfully filled (b). As can be seen in the blurred areas (c), the sharpness of the image increased and the network was able to deblur the lost details. The color intensity, sharpness, and contrast of the FFPE tissue section of the ovary improved and allowed to have a tissue infiltrating nuclear details since staining variations are very problematic in pathological examinations (d). Correction of pen marks artifact revealed nuclear, cellular, and cytological content that were very difficult to see in the previous version (e). Contaminants in ovarian sections look smaller compared to thyroid. FFPE++ was able to rectify the contamination within the sections (f).

We also wanted to show an illustration of the network training for the thyroid, ovary, and lung tissue sections in different steps (after 10K, 100K, 200K, 400K, and 800K) of the learning process (see Figure 4.4). In thyroid sections, the tissue became more comprehensible by sharpening the nuclear details and replacing the stroma after iteration 800K. The details of the lymphocytes and nucleoli were also improved despite the low nuclear sharpness and show these corrections more clearly compared to their original version. For the ovary, the network was not only able to deal with one artifact class but it was capable of correcting multiple artifacts, including, chatter, drying and folding artifact that presented altogether in the same tissue section. At the start of the training, the cytological, nuclear details, stroma and patterns are almost incomprehensible then throughout further iterations the stromal component has been successfully restored. The correction in the fold area is very successful. The features that could not be understood at all in the ovary section became fully readable at the end of training phase. The alveolar structure in lung tissue is a major challenge for the network. However, after 800K, the visibility of cellular details and red blood cells caused by the permanent ink marker artifact was successfully corrected, and in the case of the blurred region, the network enhanced the sharpness of cellular details observed throughout the training progress.

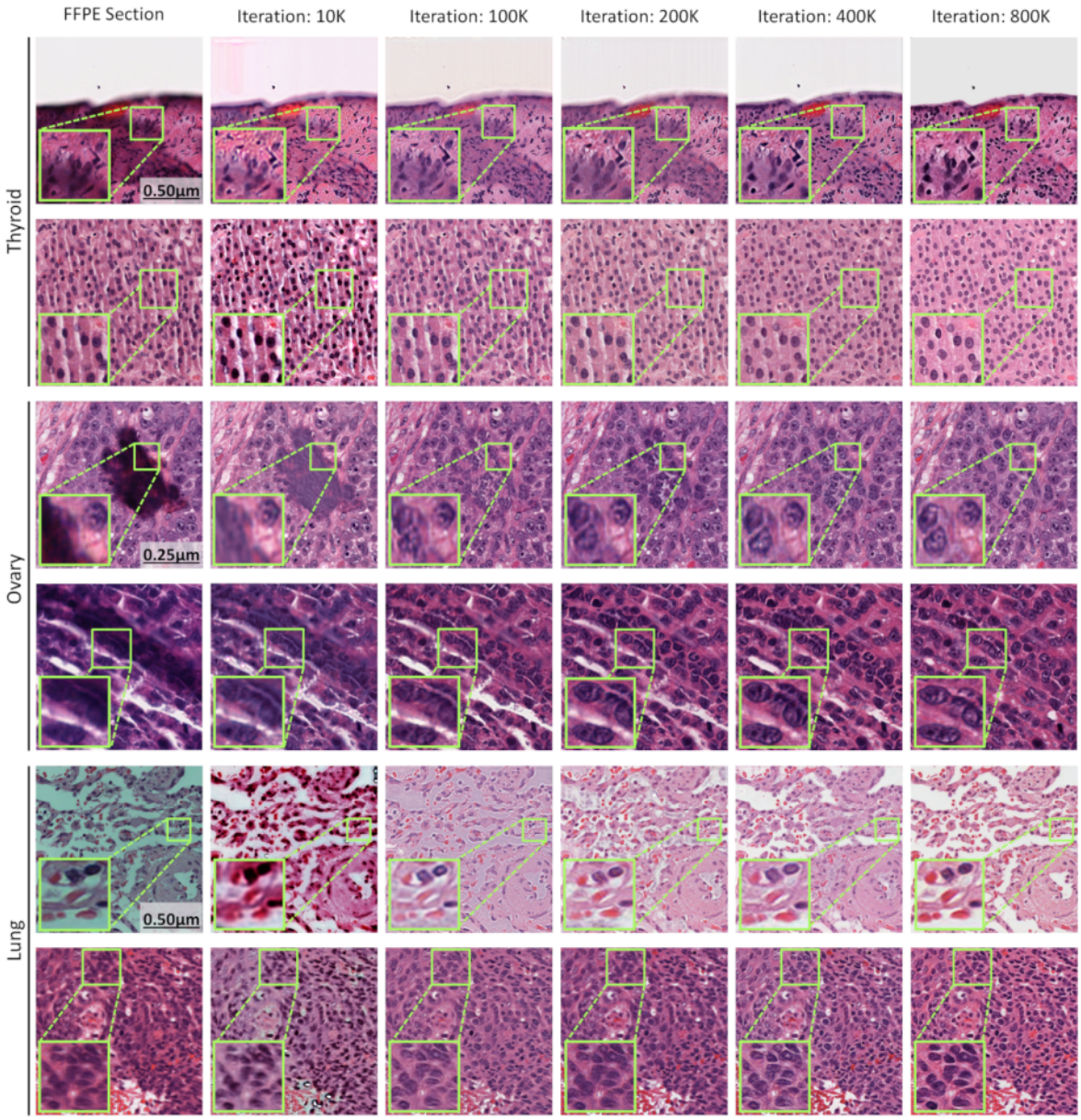

Figure 4.4: Enhancement of FFPE patch quality throughout the training process.

Moreover, we exhibit the effect of adding SAB and SR loss together in the final objective indicated as FFPE++ versus results of FFPE++ in the absence of SR (w/o SR) and SAB (w/o SAB) in Figure 4.5. For instance, SAB restrained the model from filling in stromal structures in the vessel lumen **(e)**. The results of w/o SR loss and w/o SAB are streaked and blurred as seen in **(c**,**c)**. The model struggles to replace the missing information caused by the permanent markers using only SAB module or SR loss as seen in **(a**,**a)**. For example, FFPE++ completely excluded the artifact and improved the contrast and clarity of the underlying tissue (i.e. erythrocytes, nuclei,

and cytoplasm). Therefore, the combination of SAB and SR loss shows a cumulative improvement in textural and nuclear details **(f,e,c,c)**, and enhanced staining quality as well as tissue features were preserved with no change in areas other than the artifacts (**d,b,b**). The SR loss increased the efficiency of the network to correct more effectively, white areas due to chattering artifact and filled them with stromal structure**(f)**. Also, in the FFPE++ **(b,b)** lymphocytes appear more prominent, which means they can be more easily distinguished from tumor cells.

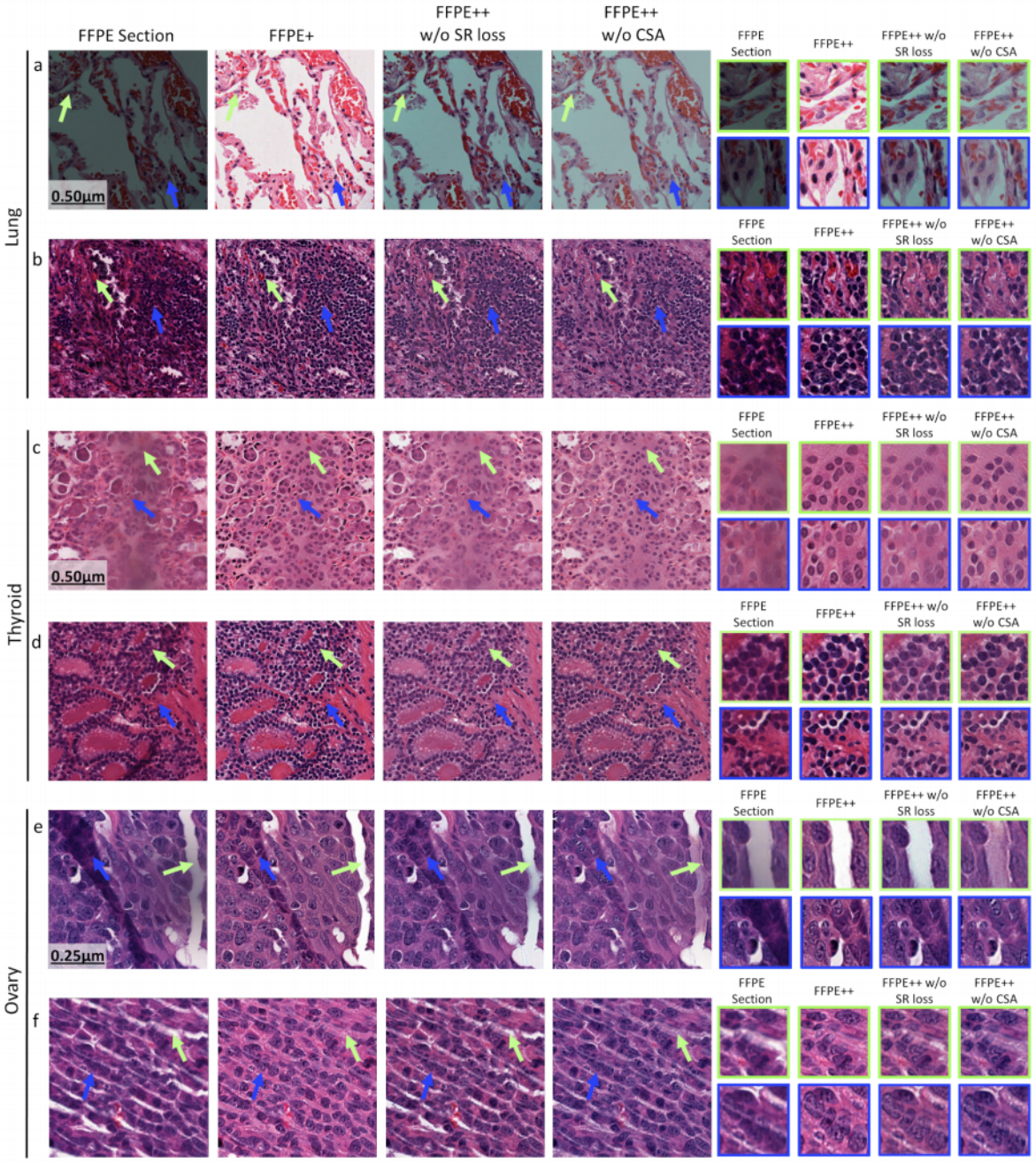

Figure 4.5: The effect of SR loss and SAB on the model performance.

4.1 Assessment of FFPE++ visual quality:

To further validate the FFPE++ method, we designed three surveys for each tissue sample to qualitatively assess the efficiency of our method. 10 board-certified pathologists were asked to complete three surveys for FFPE artifacts correction in the ovary, thyroid, and lung image patches. The surveys focused on determining whether pathologists preferred the corrected FFPE sections (FFPE++ images) compared with the artifactual FFPE patches in terms of pathological diagnosis. Each pathologist was shown the same survey format for each tissue sample, which referred to 15 images of uncorrected or artifactual FFPE image patches and 15 images of FFPE++ generated patches, side by side. In the first question, pathologists are asked to rate a pair of FFPE patches (raw image vs. corrected image) on a scale from 1, representing the poorest image quality, to 10, representing the best image quality, in order to make the best possible pathological diagnosis. We then asked pathologists to select which specific features (i.e., nuclei, nucleoli, cytoplasm, chromatin, stroma, erythrocytes, inflammatory cells) and which artifacts were corrected in the FFPE++ version. Finally, we asked pathologists whether any of these enhancements/changes in "FFPE++ image" could lead to any misdiagnosis or misinformation.

The results of the ovary survey are summarized in Table 4.1 and Figure 4.6, and the results of the thyroid survey are summarized in Table 4.2 and Figure 4.7, and the findings of the lung survey can be seen in Table 4.3 and 4.8. We show in tables 4.1, 4.2, and 4.3 how pathologists scored all 15 generated FFPE++ and raw FFPE image patches. The average image quality column indicates that all 10 pathologists have significantly preferred FFPE++ compared with the prior. Accordingly, the histograms in Figures 4.6, 4.7, and 4.8 show the percentage of how specifically each feature and each histological artifact class improved across all 15 FFPE++ patches. For instance, in all three histograms in Figures 4.6, 4.7, and 4.8, nuclei, nucleoli, cytoplasm, chromatin, and stroma were the most visually enhanced features in FFPE++, whereas staining quality, blurring, folding, chattering, and pen marks were most frequently selected as improved artifacts. Most importantly, they all agreed that our technique does not lead to any misdiagnosis or misinformation.

Tissue Number	Pathologist 1	Pathologist 2	Pathologist 3	Pathologist 4	Pathologist 5	Pathologist 6	Pathologist 7	Pathologist 8	Pathologist 9	Pathologist 10	Average Image quality
1 (FFPE)	2	1	4	2	3	3	1	4	2	1	2.30
1 (FFPE++)	8	7	10	10	8	7	7	7	3	7	**7.40**
2 (FFPE)	3	2	5	3	4	4	1	2	5	2	3.10
2 (FFPE++)	7	7	10	10	7	6	6	4	6	7	**7.00**
3 (FFPE)	1	1	2	1	1	1	1	2	3	1	1.40
3 (FFPE++)	8	8	10	8	8	7	8	6	6	8	**7.70**
4 (FFPE)	1	1	3	1	1	2	1	2	2	1	1.50
4 (FFPE++)	8	8	10	7	7	6	5	8	5	8	**7.20**
5 (FFPE)	3	4	3	4	3	4	3	3	4	4	3.50
5 (FFPE++)	9	9	10	10	8	7	8	9	9	9	**8.80**
6 (FFPE)	1	1	3	2	1	2	1	2	1	1	1.50
6 (FFPE++)	8	6	10	9	6	7	8	9	4	6	**7.30**
7 (FFPE)	2	1	3	2	2	1	1	2	3	1	1.80
7 (FFPE++)	7	7	10	9	7	6	4	8	4	7	**6.90**
8 (FFPE)	3	4	6	4	5	5	2	4	3	4	4.00
8 (FFPE++)	9	9	10	10	7	8	7	10	8	9	**8.70**
9 (FFPE)	2	2	5	3	3	1	1	3	1	2	2.30
9 (FFPE++)	7	7	10	9	8	6	5	6	7	7	**7.20**
10 (FFPE)	1	3	3	2	3	3	1	2	2	3	2.30
10 (FFPE++)	8	8	10	8	7	8	4	7	8	8	**7.60**
11 (FFPE)	2	4	4	3	4	2	2	4	2	4	3.10
11 (FFPE++)	9	9	10	9	7	7	4	10	4	9	**7.80**
12 (FFPE)	4	5	6	4	4	4	3	3	3	5	4.10
12 (FFPE++)	7	8	10	10	7	7	8	7	6	8	**7.80**
13 (FFPE)	1	1	3	3	3	4	1	1	2	1	2.00
13 (FFPE++)	9	9	10	10	6	7	5	9	3	9	**7.70**
14 (FFPE)	2	2	3	3	3	4	2	2	2	2	2.50
14 (FFPE++)	8	8	10	10	7	6	7	9	4	8	**7.70**
15 (FFPE)	2	1	4	4	2	3	2	2	1	1	2.20
15 (FFPE++)	8	7	10	10	6	8	9	9	5	7	**7.90**

Table 4.1: Evaluation of image quality preference for pathological diagnosis between the raw FFPE ovary section patches and FFPE++. The winner average score are in bold. FFPE, resemble to the raw FFPE ovary section patches with artifacts; FFPE++ resemble to the corrected ovary image version.

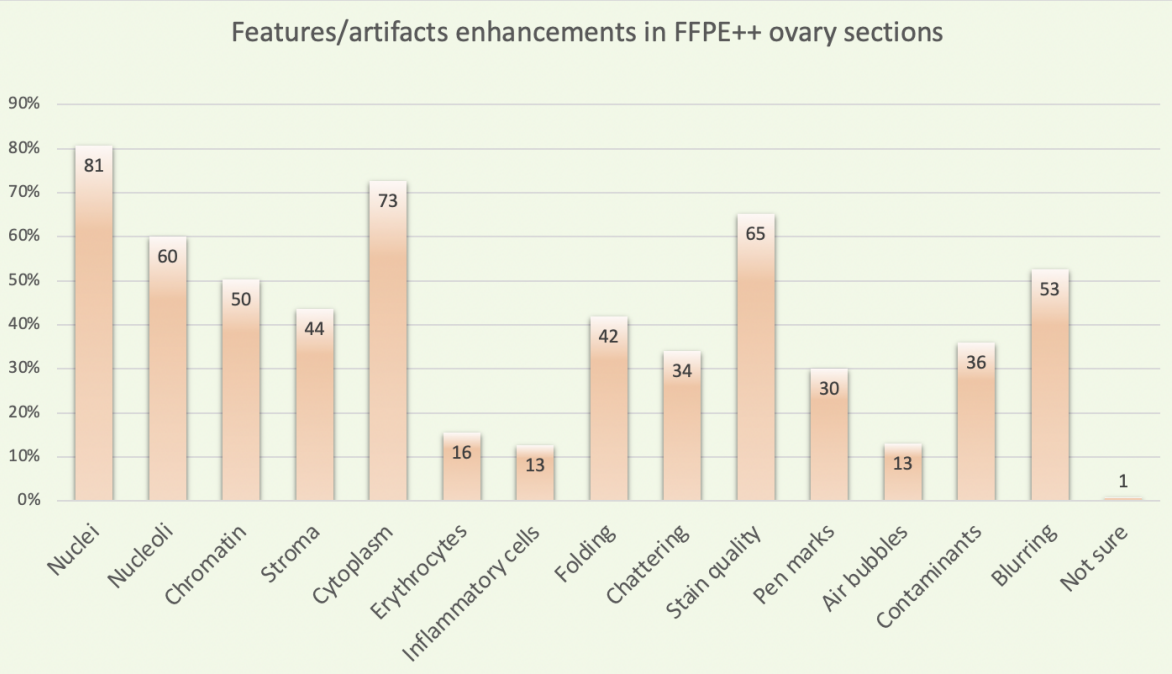

Figure 4.6: Enhanced features and artifacts in FFPE++ ovary sections. A detailed histogram showing the most highlighted histological features and artifacts as an average percentage for all 15 FFPE++ ovary section patches. Our data show that pathologists were consistently able to better identify nuclei, nucleoli, chromatin, cytoplasm, and stroma. In addition, they clearly preferred the staining quality of our technique because it successfully corrects staining inconsistencies, such as over-staining and under-staining. Blurred FFPE sections are considered the second most frequently corrected artifact which hinders pathologists from clearly identifying histopathological features. Folding, chattering, and contamination are the other corrected artifacts that pathologists agree on.

Tissue Number	Pathologist 1	Pathologist 2	Pathologist 3	Pathologist 4	Pathologist 5	Pathologist 6	Pathologist 7	Pathologist 8	Pathologist 9	Pathologist 10	Average Image quality
1 (FFPE)	1	2	2	4	3	2	6	1	3	1	2.50
1 (FFPE++)	7	10	10	8	7	9	8	8	10	7	**8.40**
2 (FFPE)	1	4	3	3	1	2	5	1	1	1	2.20
2 (FFPE++)	4	10	10	7	7	8	7	6	7	4	**7.00**
3 (FFPE)	2	3	2	2	1	1	4	2	2	2	2.10
3 (FFPE++)	7	10	10	7	7	7	7	9	7	7	**7.80**
4 (FFPE)	2	2	4	3	2	2	4	2	1	2	2.40
4 (FFPE++)	7	10	10	8	7	9	7	7	7	7	**7.90**
5 (FFPE)	4	3	4	4	5	2	5	4	2	4	3.70
5 (FFPE++)	8	10	10	8	8	9	7	9	8	8	**8.50**
6 (FFPE)	3	3	3	4	3	2	5	2	3	3	3.10
6 (FFPE++)	6	10	10	7	7	9	8	7	8	6	**7.80**
7 (FFPE)	2	2	2	2	4	1	2	2	1	2	2.00
7 (FFPE++)	7	10	10	8	7	9	7	8	7	7	**8.00**
8 (FFPE)	1	3	2	3	1	2	2	2	2	1	1.90
8 (FFPE++)	7	10	10	7	8	8	8	6	7	7	**7.80**
9 (FFPE)	2	3	4	3	3	2	4	3	2	2	2.80
9 (FFPE++)	8	10	10	7	7	8	7	6	8	8	**7.90**
10 (FFPE)	1	2	2	4	1	1	3	2	1	1	1.80
10 (FFPE++)	8	10	10	7	7	7	8	8	2	8	**7.50**
11 (FFPE)	2	3	2	3	4	2	5	3	3	2	2.90
11 (FFPE++)	8	10	10	8	7	9	8	9	8	8	**8.50**
12 (FFPE)	2	3	4	2	2	2	2	2	3	2	2.40
12 (FFPE++)	8	10	10	7	7	9	9	9	10	8	**8.70**
13 (FFPE)	1	3	2	2	1	2	2	2	1	1	1.70
13 (FFPE++)	7	10	10	7	7	9	8	10	9	7	**8.40**
14 (FFPE)	1	3	2	3	1	3	2	2	1	1	1.90
14 (FFPE++)	6	10	10	7	7	8	7	8	9	6	**7.80**
15 (FFPE)	1	3	4	3	1	2	3	1	1	1	2.00
15 (FFPE++)	7	10	10	8	7	8	7	6	9	7	**7.90**

Table 4.2: Evaluation of image quality preference for pathological diagnosis between the raw FFPE thyroid section patches and FFPE++. The winner average score are in bold. FFPE, thyroid patches with artifacts; FFPE++, corrected thyroid image patches.

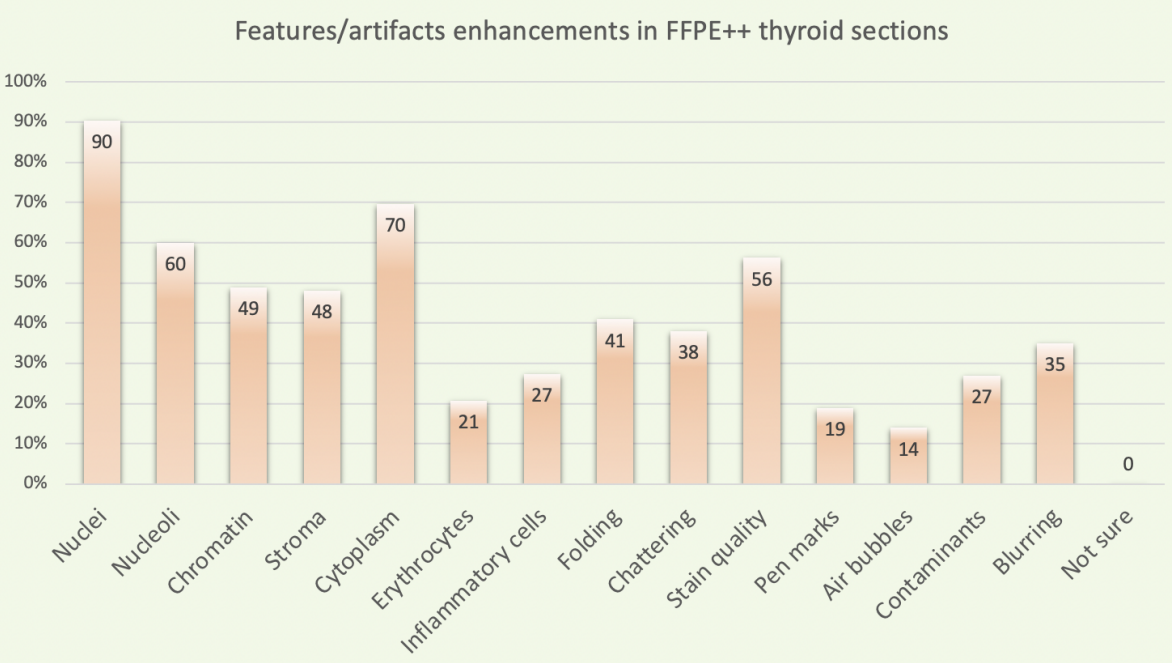

Figure 4.7: Enhanced features and artifacts in FFPE++ thyroid sections. In this histogram, we plot the average percentage for 15 FFPE++ thyroid sections. The pathologists graded the most enhanced features and artifacts in the thyroid survey. They selected nuclei, nucleoli, chromatin, cytoplasm, stroma, and inflammatory cells as the enhanced histopathological features. Similar to the ovary survey, staining quality and blurred areas are considered the most corrected artifacts. In addition, folding, chattering, pen marks, and contamination are the other corrected artifacts that pathologists agree on.

Tissue Number	Pathologist 1	Pathologist 2	Pathologist 3	Pathologist 4	Pathologist 5	Pathologist 6	Pathologist 7	Pathologist 8	Pathologist 9	Pathologist 10	Average Image quality
1 (FFPE)	2	2	1	1	4	2	3	6	2	1	2.40
1 (FFPE++)	9	9	8	10	8	8	10	9	10	8	**8.90**
2 (FFPE)	1	1	1	3	2	3	2	1	2	1	1.70
2 (FFPE++)	9	8	7	10	7	8	10	7	8	7	**8.10**
3 (FFPE)	4	4	2	4	4	3	2	3	2	3	3.10
3 (FFPE++)	8	8	7	10	7	5	5	7	6	9	**7.20**
4 (FFPE)	2	4	3	3	4	6	2	3	2	1	3.00
4 (FFPE++)	8	7	7	9	7	8	9	7	8	6	**7.60**
5 (FFPE)	3	3	1	2	3	4	3	4	2	1	2.60
5 (FFPE++)	8	8	5	10	7	7	6	6	8	7	**7.20**
6 (FFPE)	3	3	2	4	5	4	2	3	2	2	3.00
6 (FFPE++)	9	7	4	10	7	6	9	7	8	9	**7.60**
7 (FFPE)	1	1	1	1	1	2	3	1	2	1	1.40
7 (FFPE++)	8	6	7	10	7	7	10	6	9	8	**7.80**
8 (FFPE)	1	2	2	3	2	3	3	1	2	1	2.00
8 (FFPE++)	9	8	6	10	7	7	8	8	6	7	**7.60**
9 (FFPE)	3	3	1	3	1	6	3	1	3	1	2.50
9 (FFPE++)	7	6	6	10	7	7	9	3	7	7	**6.90**
10 (FFPE)	2	2	3	5	5	3	2	2	5	1	3.00
10 (FFPE++)	9	8	7	10	7	7	8	6	9	7	**7.80**
11 (FFPE)	2	1	2	4	1	5	1	2	3	1	2.20
11 (FFPE++)	6	5	5	10	6	7	9	6	8	7	**6.90**
12 (FFPE)	1	1	1	1	1	2	2	1	1	1	1.20
12 (FFPE++)	7	6	4	10	5	7	9	4	8	8	**6.80**
13 (FFPE)	1	2	1	1	1	3	2	1	1	1	1.40
13 (FFPE++)	7	6	4	10	5	7	8	4	6	8	**6.50**
14 (FFPE)	2	3	3	2	2	6	4	4	2	1	2.90
14 (FFPE++)	9	9	7	10	7	8	7	9	9	9	**8.40**
15 (FFPE)	4	3	2	5	3	6	3	3	7	2	3.80
15 (FFPE++)	9	9	9	10	7	8	10	8	9	8	**8.70**

Table 4.3: Evaluation of image quality preference for pathological diagnosis between the raw FFPE lung section patches and FFPE++. The winner average score are in bold. FFPE, belongs to the FFPE lung image patches with artifacts; FFPE++ belongs to the corrected lung patches.

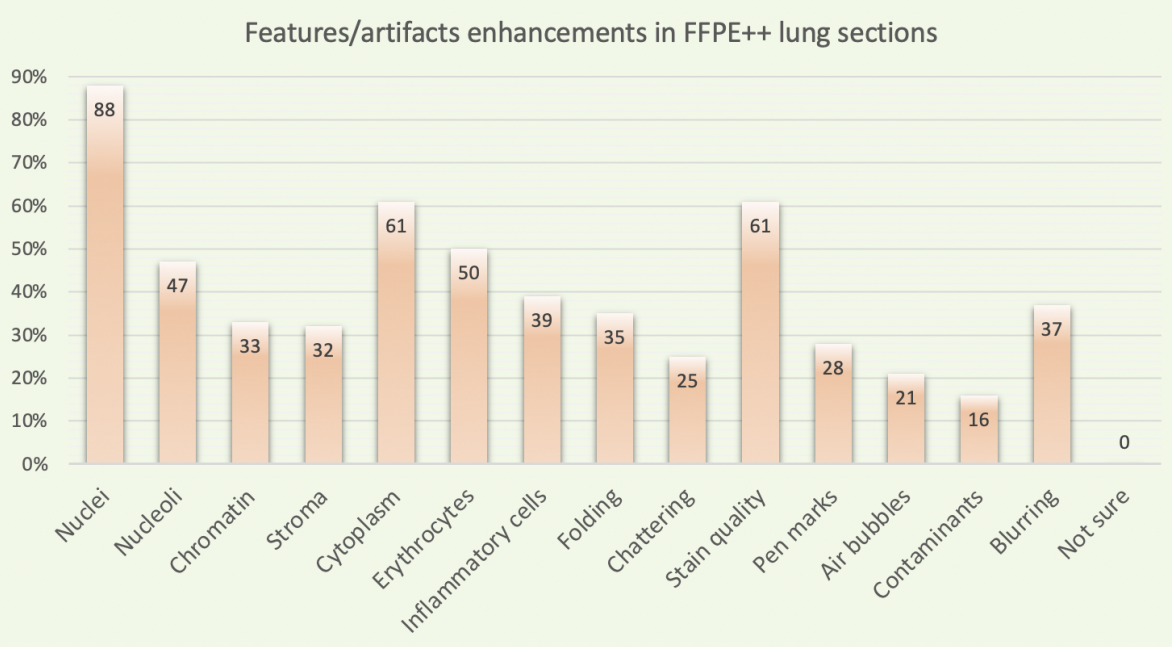

Figure 4.8: Enhanced features and artifacts in FFPE++ lung sections. Similar to the ovary and thyroid histograms, pathologists selected the same enhanced features in addition to the erythrocytes due to patches that contain haemorrhage. Once more, staining variations and blurring are considered the most successfully corrected artifacts. In addition, folding, chattering, air bubbles, and pen marks are the other corrected artifacts that pathologists agree on.

5. DISCUSSION

The discipline of AI in digital pathology has demonstrated promising results in improving the healthcare quality for patients in different medical domains. A primary goal of using AI-based algorithms in digital pathology is to reduce variations and errors. Accordingly, researchers strive to develop new tools to support pathologists in their clinical routine to improve pathological diagnoses such as the classification of cancer tumors using morphology and cellular patterns, cancer grading, and replacing chemical stains with virtual staining based-approaches, yet, approaches to correct several histological artifact classes in FFPE sections have remained an unexplored area [15, 56, 57]. We designed an unsupervised image-to-image translation algorithm that efficiently corrects various artifacts in lung, ovary and thyroid gland FFPE section patches on the textural and cellular levels and generates a better image quality for pathological diagnosis.

Due to the difficulties in standardizing the analytical process of pathological examination, technical artifacts in FFPE slides are one of the major problems in the routine practice of pathological examination. Numerous histochemical and immunohistochemical stains and molecular methods are performed on specimens to determine diagnoses and treatment options, which requires a large amount of tumor tissue. As a result, it is common for tumor tissue to be scarce or to have insufficient supply of micro scale lesions to repeat the analytical process and correct various artifacts. While the increasing applications of AI in digital histopathology slides challenge pathological diagnosis and seek solutions or assist pathologists in their critical diagnosis, FFPE++ corrects artifacts that are fundamental problems that occupy an important place in daily practice. We implemented the self-regularization loss function and SAB in the network to optimize image reconstruction and develop a robust tool for AI in digital histopathology. Incorporating SR loss function and SAB to the network improved the visualization of the tissue sections by making the pattern and cellular-nuclear details such as nucleus borders and chromatin structure become more understandable.

Hence, this combination resulted in a cumulative improvement effect in the inclusion of textural details in the ECM as seen in Figure 4.4, and 4.5.

The folding of the section on the slide makes it almost impossible to understand the cellular structures and tissue properties (Fig. 4.1a, 4.2a, 4.3a). The excess thickness in the fold area and the dark color of the stain prevent the cells from being seen. FFPE++ is very successful in correcting these artifacts and provides a clear understanding of all the details of cells in the folding area. Nucleus and nucleolus shapes and numbers, cytoplasm amounts, and cell shapes of cells in this area are recovered.

The chattering artifact complicates the understanding of normal tissues or tumors, the cells that form them, the structure of the ECM, and tumor pattern (Fig.4.1b, 4.2b, 4.3b). Correction replaces the cytoplasm and restores cell shape. This allows for the recovery and understanding of the tumor and tissue pattern. Correction does not change the nucleus structure/shape, it makes the nucleus details clear by making the nucleus border clear, the shape and number of the nucleoli visible, and correcting the chromatin structure.

The correction is most successful and effective in blurring, pen marks, and staining artifacts (Fig.4.1, 4.2, and 4.3 - a, b and c). Apart from cellular features, and tumor pattern, the presence of ECM and its organization are also critical in differentiating between thyroid papillary carcinomas (PTC) and non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP). In lesions where the nuclear details cannot be understood and the pattern cannot be read due to chatter, it will not be possible to distinguish between PTC and NIFTP by histomorphology in light microscopy [58]. FFPE++ makes the structure/contour of the nucleus, chromatin structure, number/shape and staining characteristics of the nucleolus, intracytoplasmic inclusions and pseudoinclusions understandable and helps in the differential diagnosis of PTC and NIFTP. Apart from tumor cells, FFPE++ will facilitate the understanding of inflammatory cells infiltrating the tumor and ECM. The prominence of the pleomorphism of the nucleus is critical when deciding whether ovarian serous carcinomas are high or low grade, and similarly, the nuclei details in more monotonous low-grade tumors are also

important in the differentiation of sex cord-stromal and germ cell tumors. FFPE++'s recovery of nuclear details will enable these distinctions to be made. Reading and understanding the atypia, pleomorphism, and organization of the tumor cells, which are the basis of the pattern, is a fundamental requirement of tumor typing. This is the reason why artifacts in FFPE slides are problematic. The main strength of our work is the recovery of microscopic identifiability morphology, which is the main focus of analysis of the microscopy. This will make a significant contribution to the typing, subtyping, and understanding of the tumor microenvironment of the tumors.

In lung carcinomas where the biopsy material is usually very small and the tumor cells in the biopsy are very valuable for further molecular investigations, FFPE++ will provide critical benefit in differentiating small cell carcinoma from non-small cell carcinoma. Morphologically, the distinction between small cell carcinoma and non-small cell lung carcinoma (NSCLC) has a critical role in determining the basic therapy, as well as preserving the tissue to be used for genetic alteration-related therapies. Similarly, metastatic tumors of unknown origin, undifferentiated tumors, or tumors that are difficult to morphologically evaluate will be difficult or impossible to detect in artifactual sections. Therefore, FFPE++ will make these tumors easier to understand and diagnose. The use of FFPE++ technique in sections with immunohistochemical staining other than HE sections may be the motive of further benefit from our approach. The presence of artifacts in critical immunohistochemical evaluations such as the scoring of programmed cell death protein-1 (PD-1) or programmed cell death ligand 1 (PD-L1) performed to decide on the indication for immunotherapy [59] will affect the outcome of the evaluation and cause a wrong decision in the treatment of the patient. Therefore, the FFPE++ model will provide critical benefit in these and similar situations.

The limitation of this study is related to the stitching or reconstructing a new enhanced FFPE WSI of the corrected FFPE patches. Some of the WSIs are very large in file size (i.e. 50 and 60 gigabytes) and require high RAM (random access memory) memory, in some cases we need up to 1000 gb of RAM to process the stitching. As mentioned in the Hardware & Software subsection, this study was conducted using

the laboratory infrastructure of TRUBA and deep MIA, which do not provide such a capability. Therefore, we can only stitch WSIs with small file size.

6. CONCLUSION

In this master's thesis a deep learning algorithm has been introduced that efficiently corrects various artifacts in lung, ovary and thyroid gland FFPE sections on the textural and cellular levels and generates a better image quality for pathological diagnosis during slide examination. The implementation of the self-regularization loss function and spatial attention block in the network is to optimize image reconstruction and develop a robust tool for digital histopathology.

In addition, 10 board certified pathologists preferred FFPE++ image quality for histopathological diagnosis, and they were conveniently able to detect nucleocytoplasmic details after artifacts correction.

REFERENCES

1. Khan, S., M. Tijare, M. Jain, and A. Desai, "Artifacts in histopathology: A potential cause of misinterpretation," *Res Rev J Dent Sci*, Vol. 2, pp. 23–31, 2014.

2. Dimitriou, N., O. Arandjelović, and P. D. Caie, "Deep learning for whole slide image analysis: an overview," *Frontiers in medicine*, p. 264, 2019.

3. Becker, K.-F., C. Schott, S. Hipp, V. Metzger, P. Porschewski, R. Beck, J. Nährig, I. Becker, and H. Höfler, "Quantitative protein analysis from formalin-fixed tissues: implications for translational clinical research and nanoscale molecular diagnosis," *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, Vol. 211, no. 3, pp. 370–378, 2007.

4. Gaffney, E., P. Riegman, W. Grizzle, and P. Watson, "Factors that drive the increasing use of ffpe tissue in basic and translational cancer research," *Biotechnic & Histochemistry*, Vol. 93, no. 5, pp. 373–386, 2018.

5. Kokkat, T. J., M. S. Patel, D. McGarvey, V. A. LiVolsi, and Z. W. Baloch, "Archived formalin-fixed paraffin-embedded (ffpe) blocks: a valuable underexploited resource for extraction of dna, rna, and protein," *Biopreservation and biobanking*, Vol. 11, no. 2, pp. 101–106, 2013.

6. Becker, K.-F., C. Schott, I. Becker, and H. Höfler, "Guided protein extraction from formalin-fixed tissues for quantitative multiplex analysis avoids detrimental effects of histological stains," *PROTEOMICS–Clinical Applications*, Vol. 2, no. 5, pp. 737–743, 2008.

7. Fergenbaum, J., M. Garcia-Closas, S. Hewitt, J. Lissowska, L. Sakoda, and M. Sherman, "Loss of antigenicity in stored sections of breast cancer tissue microarrays," *Cancer Epidemiology Biomarkers and Prevention*, Vol. 13, pp. 667–672, Apr. 2004.

8. Rupprecht, C., and T. Nagarajan, *Current Laboratory Techniques in Rabies Diagnosis, Research and Prevention, Volume 2*, Vol. 2, Academic Press, 2015.

9. Ducheyne, P., *Comprehensive biomaterials*, Vol. 1, Elsevier, 2015.

10. Klopfleisch, R., A. Weiss, and A. Gruber, "Excavation of a buried treasure–dna, mrna, mirna and protein analysis in formalin fixed, paraffin embedded tissues," *Histology and histopathology, Vol. 26, nº6 (2011)*, 2011.

11. Liew, M., R. Mao, C. Wittwer, and M. Salama, "Detection of chromosomal translocations in formalin-fixed paraffin-embedded (ffpe) leukemic specimens by digital expression profiling," *International journal of laboratory hematology*, Vol. 37, no. 5, pp. 690–698, 2015.

12. Chen, C.-Y., and A. Pillay, "Protocol for the detection of treponema pallidum in paraffin-embedded specimens," in *Diagnosis of Sexually Transmitted Diseases*, pp. 295–306, Springer, 2012.

13. Taqi, S. A., S. A. Sami, L. B. Sami, and S. A. Zaki, "A review of artifacts in histopathology," *Journal of oral and maxillofacial pathology: JOMFP*, Vol. 22, no. 2, p. 279, 2018.

14. Niemann, T. H., J. G. Tranovich, and B. R. De Young, "Biopsy bag artifact," *American journal of clinical pathology*, Vol. 110, no. 2, pp. 224–226, 1998.

15. Jiang, J., N. Prodduturi, D. Chen, Q. Gu, T. Flotte, Q. Feng, and S. Hart, "Image-to-image translation for automatic ink removal in whole slide images," *Journal of Medical Imaging*, Vol. 7, no. 5, p. 057502, 2020.

16. Isola, P., J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *CVPR*, 2017.

17. Zhu, J.-Y., T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017.

18. Yi, Z., H. Zhang, P. Tan, and M. Gong, "Dualgan: Unsupervised dual learning for image-to-image translation," *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2868–2876, 2017.

19. Kim, T., M. Cha, H. Kim, J. Lee, and J. Kim, "Learning to discover cross-domain relations with generative adversarial networks," *Proc. Int. Conf. Machine Learn. (ICML)*, 03 2017.

20. Liu, M.-Y., T. Breuel, and J. Kautz, "Unsupervised image-to-image translation networks," in *NIPS*, 2017.

21. Huang, X., M.-Y. Liu, S. Belongie, and J. Kautz, "Multimodal unsupervised image-to-image translation," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 172–189, 2018.

22. Mejjati, Y. A., C. Richardt, J. Tompkin, D. Cosker, and K. I. Kim, "Unsupervised attention-guided image-to-image translation," in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, (Red Hook, NY, USA), p. 3697â3707, Curran Associates Inc., 2018.

23. Zhang, R., T. Pfister, and J. Li, "Harmonic unpaired image-to-image translation," in *ICLR*, 2019.

24. Kim, J., M. Kim, H. Kang, and K. Lee, "U-gat-it: Unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation," *ArXiv*, Vol. abs/1907.10830, 2020.

25. Park, T., A. A. Efros, R. Zhang, and J.-Y. Zhu, "Contrastive learning for unpaired image-to-image translation," in *European Conference on Computer Vision*, pp. 319–345, Springer, 2020.

26. Pang, Y., J. Lin, T. Qin, and Z. Chen, "Image-to-image translation: Methods and applications," *arXiv preprint arXiv:2101.08629*, 2021.

27. Zhao, Y., R. Wu, and H. Dong, "Unpaired image-to-image translation using adversarial consistency loss," in *European Conference on Computer Vision*, pp. 800–815, Springer, 2020.

28. Yang, C., T. Kim, R. Wang, H. Peng, and C.-C. J. Kuo, "Show, attend, and translate: Unsupervised image translation with self-regularization and attention," *IEEE Transactions on Image Processing*, Vol. 28, no. 10, pp. 4845–4856, 2019.

29. Chen, C., Q. Dou, H. Chen, J. Qin, and P. A. Heng, "Unsupervised bidirectional cross-modality adaptation via deeply synergistic image and feature alignment for medical image segmentation," *IEEE transactions on medical imaging*, Vol. 39, no. 7, pp. 2494–2505, 2020.

30. Almalioglu, Y., K. B. Ozyoruk, A. Gokce, K. Incetan, G. I. Gokceler, M. A. Simsek, K. Ararat, R. J. Chen, N. J. Durr, F. Mahmood, *et al.*, "Endol2h: Deep super-resolution for capsule endoscopy," *IEEE Transactions on Medical Imaging*, Vol. 39, no. 12, pp. 4297–4309, 2020.

31. Yang, Q., N. Li, Z. Zhao, X. Fan, I. Eric, C. Chang, and Y. Xu, "Mri cross-modality image-to-image translation," *Scientific reports*, Vol. 10, no. 1, pp. 1–18, 2020.

32. Long, F., "Microscopy cell nuclei segmentation with enhanced u-net," *BMC bioinformatics*, Vol. 21, no. 1, pp. 1–12, 2020.

33. Mahmood, F., D. Borders, R. J. Chen, G. N. McKay, K. J. Salimian, A. Baras, and N. J. Durr, "Deep adversarial training for multi-organ nuclei segmentation in histopathology images," *IEEE transactions on medical imaging*, Vol. 39, no. 11, pp. 3257–3267, 2019.

34. Naylor, P., M. Laé, F. Reyal, and T. Walter, "Segmentation of nuclei in histopathology images by deep regression of the distance map," *IEEE transactions on medical imaging*, Vol. 38, no. 2, pp. 448–459, 2018.

35. Lee, B., and K. Paeng, "A robust and effective approach towards accurate metastasis detection and pn-stage classification in breast cancer," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 841–850, Springer, 2018.

36. Xu, Y., J.-Y. Zhu, I. Eric, C. Chang, M. Lai, and Z. Tu, "Weakly supervised histopathology cancer image segmentation and classification," *Medical image analysis*, Vol. 18, no. 3, pp. 591–604, 2014.

37. Jansen, I., M. Lucas, J. Bosschieter, O. J. de Boer, S. L. Meijer, T. G. van Leeuwen, H. A. Marquering, J. A. Nieuwenhuijzen, D. M. de Bruin, and C. D. Savci-Heijink, "Automated detection and grading of non–muscle-invasive urothelial cell carcinoma of the bladder," *The American journal of pathology*, Vol. 190, no. 7, pp. 1483–1490, 2020.

38. Karimi, D., G. Nir, L. Fazli, P. C. Black, L. Goldenberg, and S. E. Salcudean, "Deep learning-based gleason grading of prostate cancer from histopathology imagesârole of multiscale decision aggregation and data augmentation," *IEEE journal of biomedical and health informatics*, Vol. 24, no. 5, pp. 1413–1426, 2019.

39. Wu, H., J. H. Phan, A. K. Bhatia, C. A. Cundiff, B. M. Shehata, and M. D. Wang, "Detection of blur artifacts in histopathological whole-slide images of endomyocardial biopsies," in *2015 37th annual international Conference of the IEEE Engineering in Medicine and biology society (EMBC)*, pp. 727–730, IEEE, 2015.

40. Janowczyk, A., A. Basavanhally, and A. Madabhushi, "Stain normalization using sparse autoencoders (stanosa): application to digital pathology," *Computerized Medical Imaging and Graphics*, Vol. 57, pp. 50–61, 2017.

41. Brieu, N., A. Meier, A. Kapil, R. Schoenmeyer, C. G. Gavriel, P. Caie, and G. Schmidt, "Domain adaptation-based augmentation for weakly supervised nuclei detection," *ArXiv*, Vol. abs/1907.04681, 2019.

42. Gadermayr, M., L. Gupta, V. Appel, P. Boor, B. M. Klinkhammer, and D. Merhof, "Generative adversarial networks for facilitating stain-independent supervised and unsupervised segmentation: a study on kidney histology," *IEEE transactions on medical imaging*, Vol. 38, no. 10, pp. 2293–2302, 2019.

43. Xu, Z., C. F. Moro, B. Bozóky, and Q. Zhang, "Gan-based virtual re-staining: a promising solution for whole slide image analysis," *arXiv preprint arXiv:1901.04059*, 2019.

44. Rana, A., G. Yauney, A. Lowe, and P. Shah, "Computational histological staining and destaining of prostate core biopsy rgb images with generative adversarial neural networks," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 828–834, IEEE, 2018.

45. Lu, M. Y., D. F. K. Williamson, T. Y. Chen, R. J. Chen, M. Barbieri, and F. Mahmood, "Data efficient and weakly supervised computational pathology on whole slide images," 2020.

46. Anusewicz, D., M. Orzechowska, and A. K. Bednarek, "Lung squamous cell carcinoma and lung adenocarcinoma differential gene expression regulation through pathways of notch, hedgehog, wnt, and erbb signalling," *Scientific reports*, Vol. 10, no. 1, pp. 1–15, 2020.

47. Nguyen, Q. T., E. J. Lee, M. G. Huang, Y. I. Park, A. Khullar, and R. A. Plodkowski, "Diagnosis and treatment of patients with thyroid cancer," *American health & drug benefits*, Vol. 8, no. 1, p. 30, 2015.

48. Chen, T., S. Kornblith, M. Norouzi, and G. E. Hinton, "A simple framework for contrastive learning of visual representations," *CoRR*, Vol. abs/2002.05709, 2020.

49. He, K., H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

50. Oord, A. v. d., Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," *arXiv preprint arXiv:1807.03748*, 2018.

51. 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.

52. Ulyanov, D., A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," *arXiv preprint arXiv:1607.08022*, 2016.

53. DP, K., and J. Ba, "Adam: A method for stochastic optimization," in *Proc. of the 3rd International Conference for Learning Representations (ICLR)*, 2015.

54. Johnson, J., A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Computer Vision – ECCV 2016* (Leibe, B., J. Matas, N. Sebe, and M. Welling, eds.), (Cham), pp. 694–711, Springer International Publishing, 2016.

55. Mao, X., Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. Paul Smolley, "Least squares generative adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, pp. 2794–2802, 2017.

56. Ali, S., N. K. Alham, C. Verrill, and J. Rittscher, "Ink removal from histopathology whole slide images by combining classification, detection and image generation models," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 928–932, IEEE, 2019.

57. Kothari, S., J. H. Phan, and M. D. Wang, "Eliminating tissue-fold artifacts in histopathological whole-slide images for improved image-based prediction of cancer grade," *Journal of pathology informatics*, Vol. 4, 2013.

58. KholovÃ¡, I., E. Haaga, J. Ludvik, D. Kalfert, and M. Ludvikova, "Noninvasive follicular thyroid neoplasm with papillary-like nuclear features (niftp): Tumour entity with a short history. a review on challenges in our microscopes, molecular and ultrasonographic profile," *Diagnostics*, Vol. 12, no. 2, 2022.

59. Paulsen, E.-E., T. K. Kilvaer, M. R. Khanehkenari, S. Al-Saad, S. M. Hald, S. Andersen, E. Richardsen, N. Ness, L.-T. Busund, R. M. Bremnes, *et al.*, "Assessing pdl-1 and pd-1 in non–small cell lung cancer: a novel immunoscore approach," *Clinical Lung Cancer*, Vol. 18, no. 2, pp. 220–233, 2017.