Advanced Intelligent Systems

ADVANCED INTELLIGENT SYSTEMS Open Access

www.advintellsyst.com

RESEARCH ARTICLE OPEN ACCESS

Xstainer: A Novel Virtual Staining Tool Powered by Advanced Deep Learning Techniques

Fatma Nur Kinali[1,2] | Derya Demir[3] | Kutsev B. Ozyoruk[4,5] | Yasin Almalioglu[6] | Emre C. Kose[7] | Drew F. K. Williamson[8,9] | Gurdeniz Serin[3] | G. Evren Keles[10] | Banu S. Kumbaraci[3] | Ceren Utku[3] | Uguray P. Hacisalihoglu[11,12] | Ilknur Turkmen[13] | Figen Ozturk[14] | Kayhan Basak[15] | Yavuz Nuri Ertas[16] | Mehmet Turan[1]

[1]Department of Computer Engineering, Bogazici University, Istanbul, Türkiye | [2]Institute of Biomedical Engineering, Bogazici University, Istanbul, Türkiye | [3]Faculty of Medicine, Department of Pathology, Ege University, Izmir, Türkiye | [4]National Cancer Institute, Bethesda, Maryland, USA | [5]Department of Biomedical Engineering, Emory University School of Medicine, Atlanta, Georgia, USA | [6]Department of Computer Science, Oxford University, Oxford, UK | [7]The Institute of Cancer Research, London, UK | [8]Harvard Medical School, Boston, Massachusetts, USA | [9]Department of Pathology & Laboratory Medicine, Emory University School of Medicine, Atlanta, Georgia, USA | [10]Virasoft Corporation, New York, New York, USA | [11]Istanbul Yeni Yuzyil University Medical Faculty, Pathology Department, Gaziosmanpasa Hospital, Istanbul, Türkiye | [12]Pathology Department, Medical Faculty (English Program), Istanbul Aydın University, Istanbul, Türkiye | [13]Pathology Department, Memorial Hospitals Group, Istanbul, Türkiye | [14]Faculty of Medicine, Department of Pathology, Erciyes University, Kayseri, Türkiye | [15]Department of Pathology, Saglik Bilimleri University; Kartal Dr. Lutfi Kirdar Sehir SUAM, Istanbul, Türkiye | [16]Department of Biomedical Engineering, Erciyes University, Kayseri, Türkiye

Correspondence: Kayhan Basak (kayhan.basak@sbu.edu.tr) | Yavuz Nuri Ertas (yavuznuri@gmail.com) | Mehmet Turan (mehmet.turan@boun.edu.tr)

Received: 17 April 2026 | **Revised:** 17 April 2026 | **Accepted:** 24 April 2026

ABSTRACT

Histopathological analysis traditionally relies on hematoxylin and eosin (H&E) staining. However, comprehensive differential diagnoses often require additional histochemical stains, increasing diagnostic time and costs. To address these limitations, we introduce Xstainer, a novel virtual staining tool powered by advanced deep learning techniques. This system adeptly transforms conventional H&E-stained images into multiple histochemical visualizations, marking a significant advancement in diagnostic histopathology. To validate Xstainer's efficacy, we conducted an evaluation study involving experienced nephropathologists. Using the OmniST dataset—a carefully curated collection of 1,646 whole slide images representing diverse patient samples, including renal transplant samples, liver explants, nonmalignant renal disease, and Helicobacter pylori gastritis and paired with standard stains such as Masson's trichrome, Periodic Acid-Schiff, Jones methenamine silver, and Toluidine blue—our tool underwent intensive clinical evaluation. Our virtually stained slides enabled board-certified, experienced nephropathologists (>10) to achieve diagnostic accuracy on par with, if not superior to, traditional staining techniques. Xstainer consistently outperformed various assessment benchmarks, including patch-level visual Turing test, slide-level staining quality assessment, and showed favorable performance in the Fréchet inception distance comparison, further underscoring its transformative potential. In summary, Xstainer offers a promising solution for rapid and accurate histopathological diagnosis with considerable clinical potential.

1 | Introduction

Histological assessment of tissue slides is fundamental in disease diagnosis, particularly when tissue sampling is feasible. Hematoxylin and eosin (H&E) staining serves as the gold standard for such evaluations, providing essential information on tissue architecture and cellular morphology [1, 2]. In routine histopathology, H&E staining is widely used due to its simplicity, cost-effectiveness, and ability to provide high-contrast visualization of nuclei (blue) and extracellular matrix/cytoplasm (pink), facilitating

Fatma Nur Kinali and Derya Demir Co-first authors contributed equally to this work.

general assessments of tissue structure [3–5]. However, H&E staining alone may not suffice in distinguishing various tissue components and specific cellular structures crucial for accurate diagnoses [6]. Consequently, a range of other histologic stains, known as "special stains," is frequently employed to identify specific tissue features and to support or confirm the initial assessment made with H&E staining [7–9] (Table S1).

In renal pathology, especially for diagnosing transplant rejection, pathologists rely on various histochemical stains to evaluate tissue characteristics critical to diagnosis. For example, Masson's trichrome (MT) stain effectively visualizes collagen by staining it in shades of blue or green, while muscle fibers appear red, and cytoplasm appears pink. This coloration aids in assessing fibrosis and extracellular matrix disorders, allowing precise identification of nuclei, muscle fibers, cytoplasm, and collagen [10, 11]. MT is also used to highlight connective tissues, renal tubule cells, and basal laminae, all pertinent to renal pathology. Similarly, periodic Acid-Schiff (PAS) stain highlights carbohydrate-rich structures, including glomerular basement membranes and tubular structures, aiding in the identification of glomerular diseases and diabetic nephropathy [12]. It is particularly useful in highlighting specific components, including connective tissue, renal tubules, intestinal brush borders, glycocalyx, mucus, basal laminae, and reticular fibers of connective tissue [10]. Jones methenamine silver (JMS) stain, which oxidizes carbohydrate components to stain basement membranes black [13], is routinely used in renal biopsies for assessing glomerular and tubular basement membranes [14, 15]. Toluidine blue (TB) stain enhances visualization of DNA- and RNA-rich structures [16], aiding in the identification of mast cells, mucins, and other cellular components, and is valuable in diagnosing dysplasia, carcinoma, and other conditions associated with renal pathology [17, 18].

Despite their diagnostic value, special staining procedures are resource-intensive, requiring rigorous quality control, specialized expertise, and extended processing times. These factors lead to increased costs and a higher workload, especially in large healthcare systems [19]. Additionally, when pathologists request special stains after an initial H&E examination, the entire slide preparation process, including additional cutting and staining, must be repeated, resulting in extra expenses and prolonging the diagnostic timeline. Such delays in diagnosis are often challenging for patients, particularly when timely diagnosis is essential. Variations in staining protocols across centers and interpretive differences among clinicians further complicate the standardization of histopathological diagnoses.

Recent advancements in artificial intelligence (AI), particularly deep learning techniques, have introduced promising alternatives for virtual histochemical staining, revolutionizing the field of computational pathology. One notable approach is the use of generative adversarial networks (GANs) for virtual staining, which enables the generation of artificially stained tissue images that accurately replicate the characteristics of target stains [20]. AI-based transformations of H&E into stains such as MT, JMS, PAS, PHH3, and FAP-CK [19, 21, 22] have been demonstrated in the literature. However, these studies were conducted on limited tissue and stains, and some of them were performed at the patch level rather than encompassing the entire whole slide image (WSI). Besides, these approaches are based on an early generated Cycle-Consistent Adversarial Network [23]. Zhao et al. [24] suggest that unpaired models utilizing cycle consistency loss face challenges during training when encountering variations in shape, extracting objects from images, or disregarding inconsequential textures. In line with this, Yang et al. [25] discovered that CycleGAN models exhibit skepticism and tend to introduce superfluous modifications into the resulting image. Furthermore, some studies, such as HER2 expression analysis using paired H&E and immunohistochemistry (IHC) images, remain limited to a single biomarker, while other studies are restricted to a single tissue type, which constrains their generalizability and applicability to broader pathology workflows [26, 27].

To overcome the limitations of prior models, Xstainer introduces two key innovations. First, it employs a novel self-regularizing loss function that enhances horizontal completeness and consistency across WSIs. Second, it integrates a hybrid spatial-attention block within the ResNet-9 generator architecture, effectively mitigating redundant modifications commonly introduced by earlier CycleGAN-based models. An overview of the Xstainer workflow, including dataset construction, model training, and whole-slide virtual staining, is illustrated in Figure 1.

Furthermore, the acquisition of appropriate tissue samples, their subsequent preparation, and the scanning process using a digital image scanner pose considerable challenges for researchers. To the best of our knowledge, existing histopathology datasets often have limitations, such as a limited number of slides or a sole reliance on H&E-stained images. To address these limitations and meet a critical necessity in digital pathology, we curated a new dataset, OmniST, which enabled a comprehensive evaluation of the Xstainer tool across different tissue and stain types. While not publicly available at this stage, the dataset can be accessed upon reasonable request.

Significantly, we conducted an evaluation study to diagnose kidney rejection. In collaboration with experienced nephropathologists, we rigorously assessed the diagnostic capabilities of our approach. The results of this study not only demonstrate the clinical potential of Xstainer but also underscore its relevance and applicability in digital pathology.

2 | Results

We showcased the versatility of the Xstainer tool by employing a diverse array of tissue–stain combinations (Figure 2) and applying it to WSIs from different centers (Figure 3). Figure 4 illustrates virtual patches of MT, PAS, and JMS stains and a corresponding WSI derived from H&E-stained kidney tissue. Our virtual MT staining accurately reproduces the extracellular connective tissue, as evidenced by its blue coloration, mimicking the results of conventional histochemical methods. The method effectively captures the density variations in connective tissue, resulting in darker shades in areas of higher density. Virtual PAS staining successfully highlights the basement membranes and accentuates their accumulation in the glomeruli. Furthermore, our JMS staining method selectively stains reticulin fibers between tubules and glomeruli, clearly visualizing the glomerular capillary network. These outcomes unequivocally demonstrate the efficacy of our AI-driven

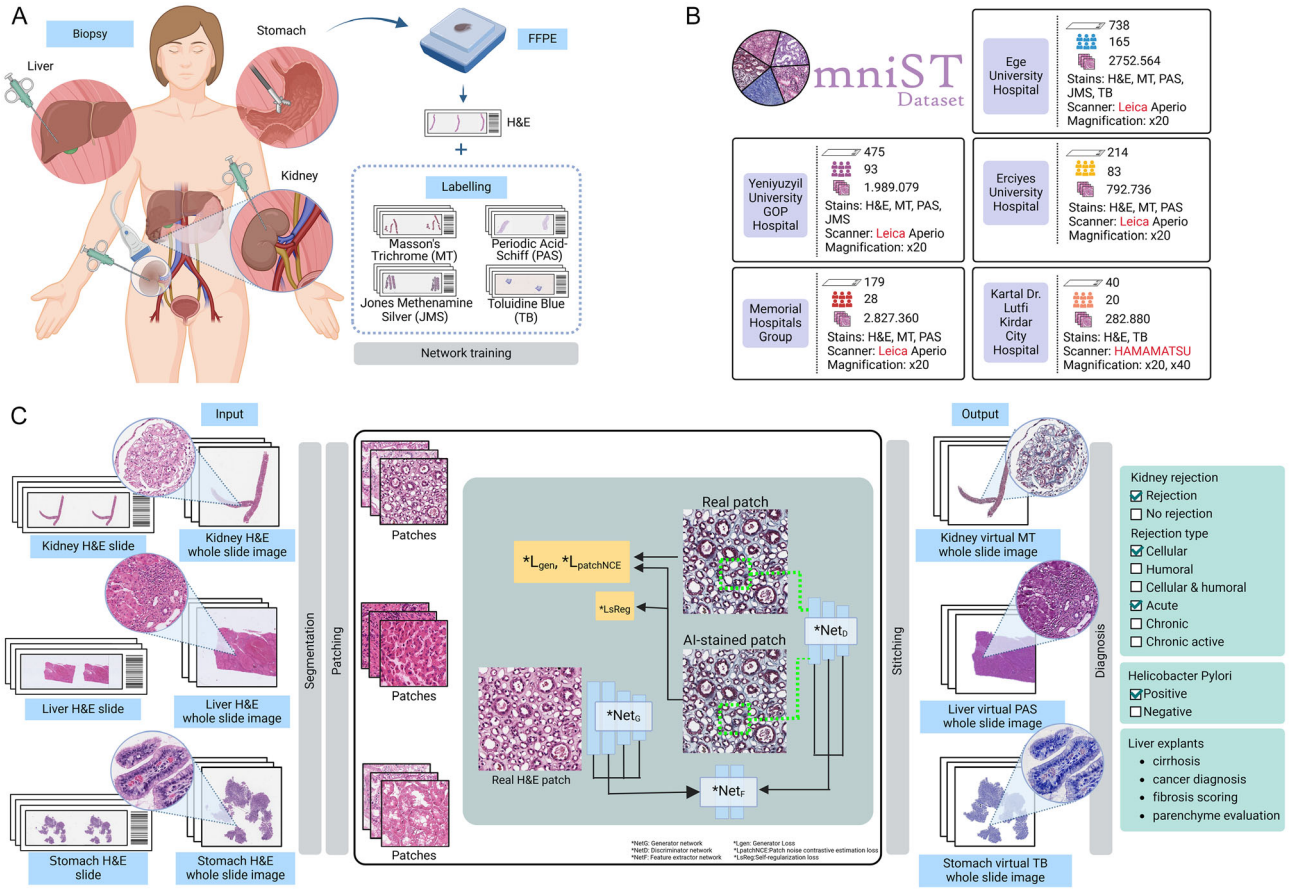

FIGURE 1 | Overview of the study. (A) Kidney, liver, and stomach tissue samples, fixed with formalin and embedded in paraffin (FFPE), are stained with H&E and various special stains. These stained images are used as training data for the model, enabling it to learn and recognize specific tissue features. (B) The OmniST dataset is a collection we gathered for our virtual staining study, which will be publicly released. (C) Separate virtual staining networks are trained for each stain. Whole slide images (WSIs) of kidney, liver, and stomach tissues are first segmented and then divided into smaller patches. Real H&E-stained patches obtained from kidney, liver, and stomach tissue samples are used as input for the virtual staining generator network. The generator network produces corresponding virtually stained patches as output. In the final step, the outputted virtually stained patches are stitched together to provide pathologists with a comprehensive view that allows for detailed examination of specific features, components, and entities to aid in the final diagnosis. This figure was created with BioRender.com.

approach in generating virtually stained images across diverse settings, highlighting its potential for diagnostic applications and the detection of specific pathological features.

2.1 | Quantitative Results

2.1.1 | Fréchet Inception Distance Results

A quantitative comparison of Xstainer with baseline methods (CUT and FastCUT) using FID scores across different tissue–stain combinations is presented in Figure 5. To evaluate the effectiveness of Xstainer's virtual staining, we used the FID metric, which quantifies the distance between the distributions of real and virtually stained images. We compared Xstainer to two established baselines: CUT and FastCUT, which are commonly used in unpaired image-to-image translation tasks [28]. While CUT focuses on structural detail and FastCUT on efficiency, both have limitations in preserving global consistency and fine-grained accuracy. The FID scores for Xstainer, CUT, and FastCUT were calculated across several tissue types and stains, including kidney (MT, PAS, JMS), liver (MT, PAS), and stomach (TB). For kidney tissue stained with MT, Xstainer achieved a final FID score of 56.19, significantly lower than CUT's 72.34 and FastCUT's 78.19, highlighting its ability to generate high-quality virtual stains. Similar trends were observed across other stains and tissue types, including PAS and JMS in kidney and liver samples. For instance, Xstainer's FID score for PAS-stained kidney tissue was 48.99, outperforming CUT (61.45) and FastCUT (64.45), and for JMS-stained kidney tissue, Xstainer's score was 55.22, well below the scores of CUT (70.12) and FastCUT (73.11). Since lower FID scores indicate greater resemblance to real images, these findings highlight the effectiveness of the Xstainer model in achieving high-fidelity virtual staining, surpassing that of generic style transfer models.

2.1.2 | Qualitative Results

The experimental designs used for slide-level staining quality assessment (SSQA), patch-level visual Turing test (VTT), and the reader study are illustrated in Figure 6. The qualitative evaluation outcomes are summarized in Figure 7, including the

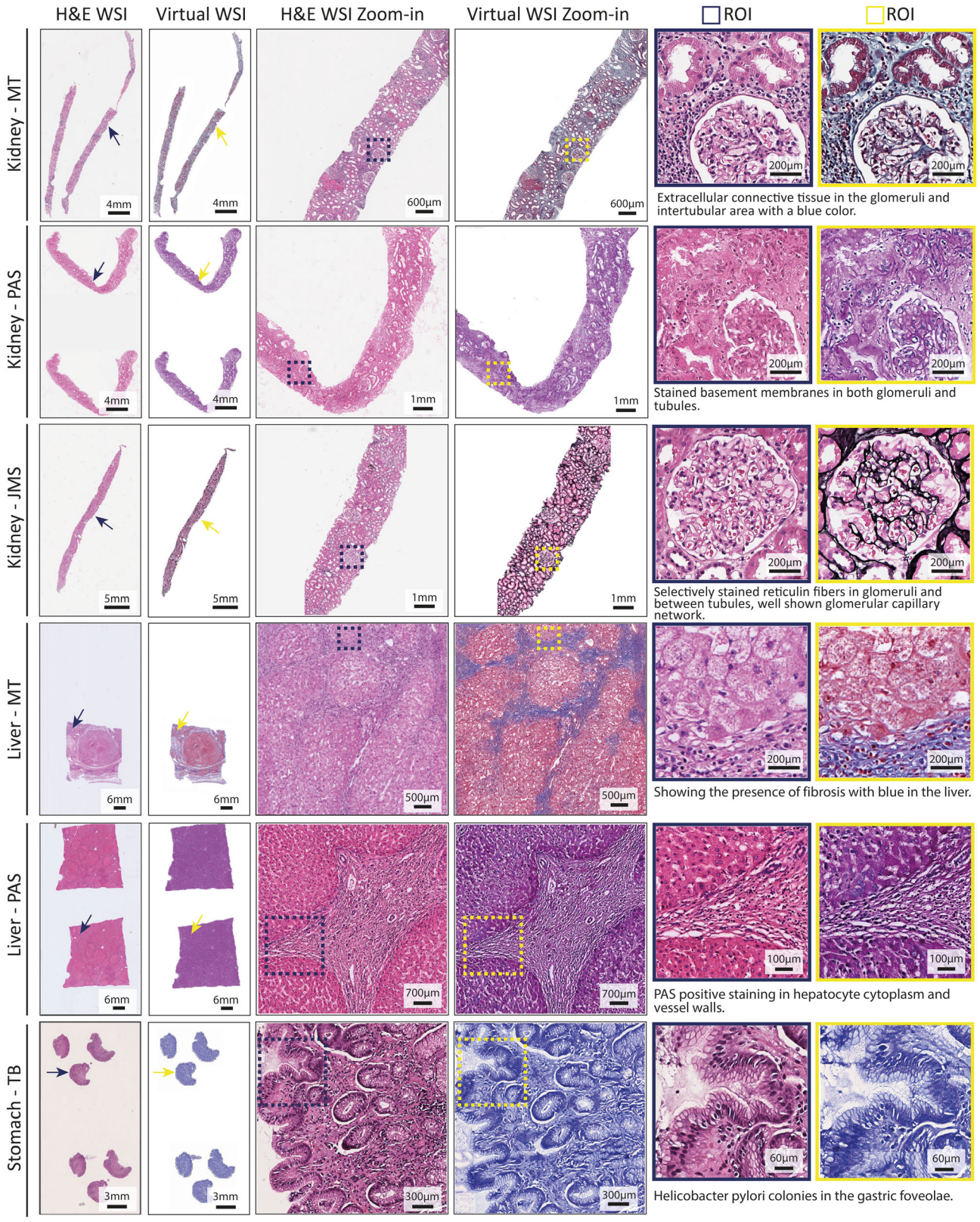

FIGURE 2 | Virtually stained WSIs derived from H&E-stained WSIs of tissues and a detailed display of target ROIs. WSIs of the kidney, liver, and stomach were virtually stained with PAS, MT, JMS, and TB stains. The WSIs of virtual stains are presented in the second column. Targeted regions of interest (ROIs) are displayed in the fourth column, and the ROI details with descriptions are shown in the last two columns, accompanied by their corresponding H&E counterparts.

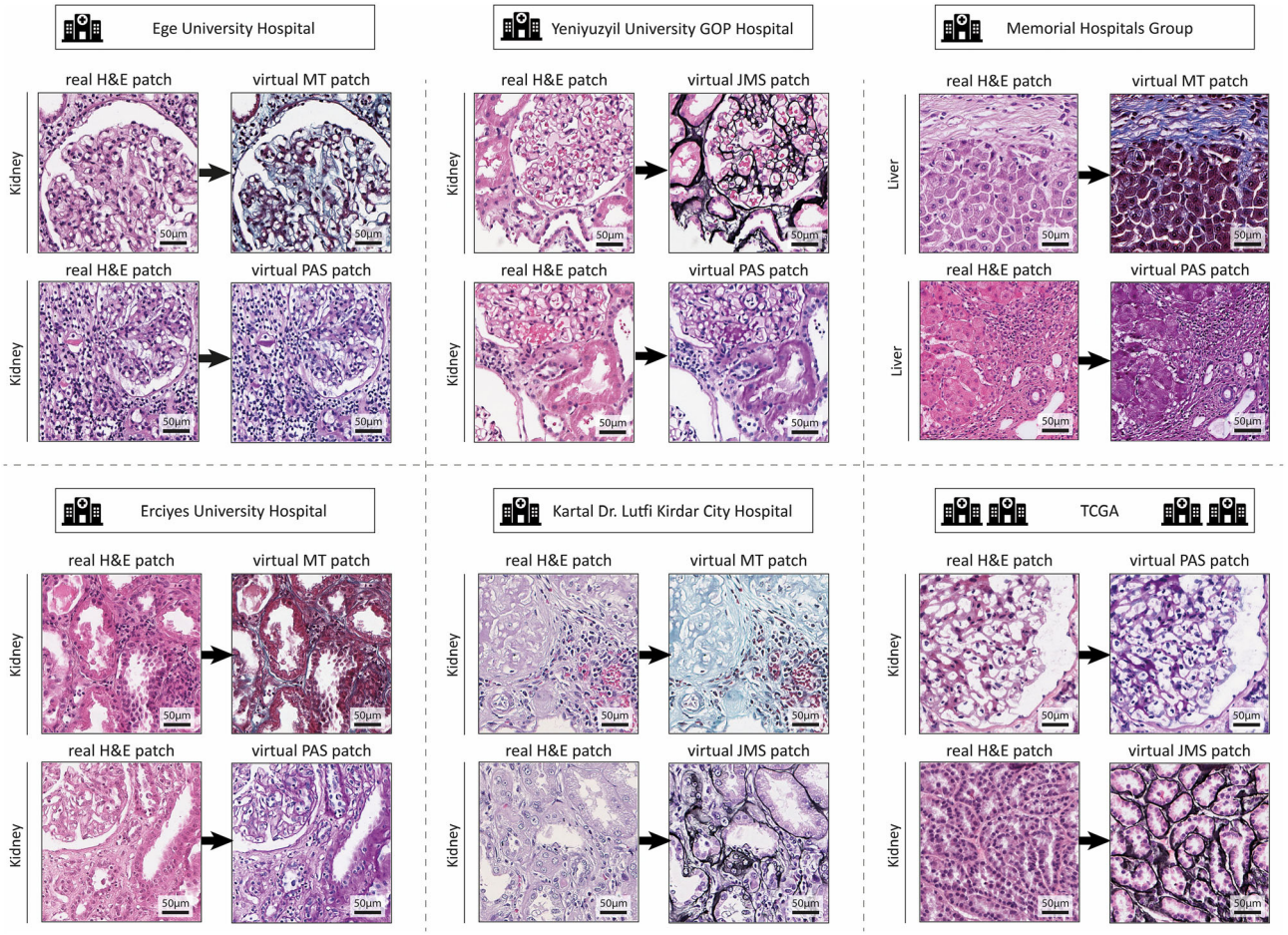

FIGURE 3 | Evaluation of the virtual staining on independent test cohorts. The virtual stains were applied to kidney Tru-Cut biopsy samples obtained from Ege University Hospital and Erciyes University Hospital. Yeniyuzyıl University GOP Hospital provided kidney Tru-Cut biopsy samples, virtually stained with JMS and PAS on H&E patches. The MT and PAS virtual stains were applied to H&E-stained liver tissue from the Memorial Hospitals Group. The JMS and MT virtual stains were applied to H&E-stained kidney tissue from Kartal Dr. Lutfi Kırdar City Hospital and TCGA.

reader study results (Figure 7A), patch-level VTT (Figure 7B), and slide-level SSQA scores (Figure 7C).

2.1.3 | Patch-Level Visual Turing Test

The VTT results revealed noteworthy trends in pathologists' ability to distinguish between real and virtually stained patches. As depicted in Figure 7B, a higher proportion of virtual images were incorrectly classified as real, compared to the real images that were accurately identified as real. For the MT stain, 27.5% of virtual images were classified as real, while only 22.5% of real images were correctly identified as real. Similarly, 26.5% of virtual images were misclassified as real for the PAS stain, compared to 23.5% of real images accurately identified. However, for the JMS stain, the results were slightly different: 24.6% of virtual images were classified as real, and 25.3% of real images were correctly identified as real.

In summary, these results indicate that pathologists demonstrated a limited ability to consistently differentiate between real and virtual images, especially for PAS and MT, underscoring the realism of the virtually stained patches.

2.1.4 | Slide-Level Staining Quality Assessment

Figure 7C provides a comprehensive report on the assessments conducted by the participating pathologists. On a four-point scale (1: unacceptable, 4: perfect), virtual stains received average scores ranging from 2.80 to 3.30, indicating "very good" performance comparable to real stains. These scores collectively suggest that the virtual staining technique achieved a "very good" staining quality, with consistent results across the participating pathologists. The survey outcomes provide valuable insights into the effectiveness of virtual staining in replicating the visual characteristics of traditional stains.

The findings highlight the virtual staining technique's ability to reliably mimic the visual attributes of traditional staining methods. This success underscores its potential as a supplementary diagnostic tool, providing a consistent and efficient alternative for pathologists in analyzing tissue samples.

2.1.5 | Reader Study (Diagnostic Test)

To assess the diagnostic utility of the virtual stains, pathologists answered histological questions using both real and virtual stains.

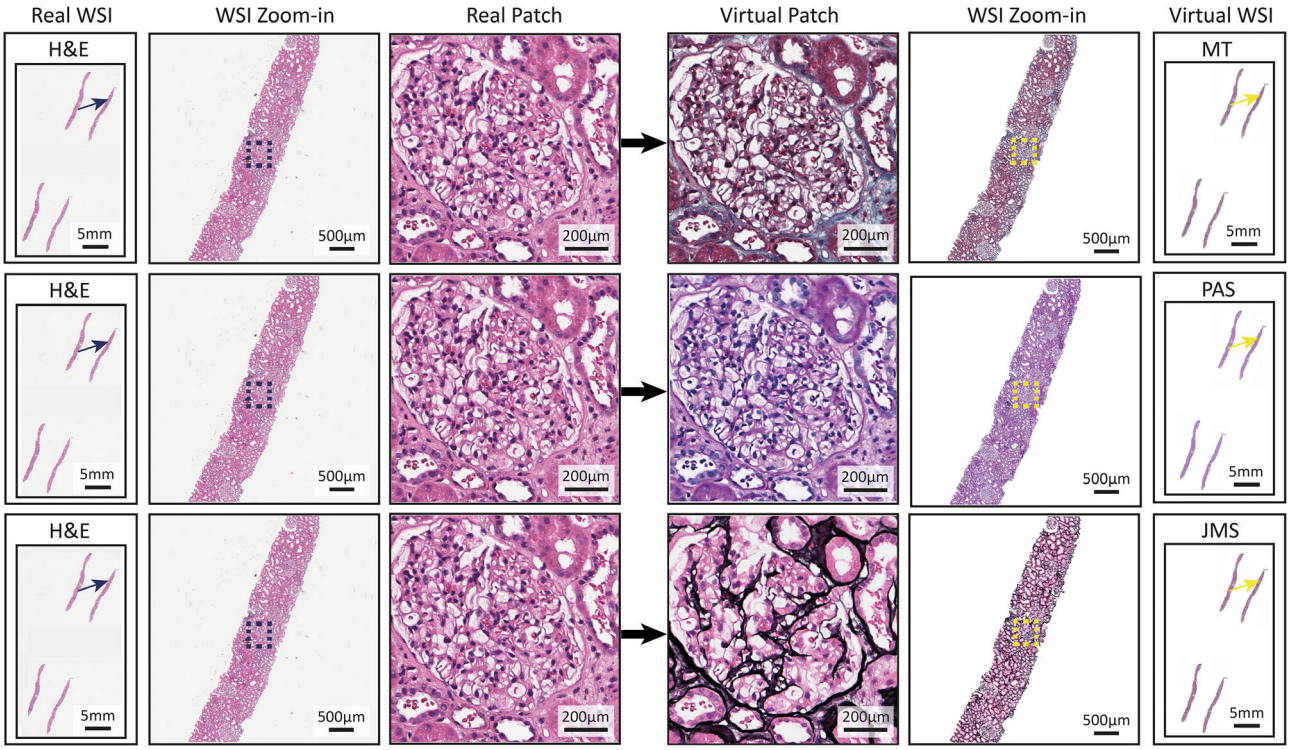

FIGURE 4 | Illustration of the transformation of H&E image into MT, PAS, and JMS patch and WSI images using virtual special staining in kidney Tru-Cut biopsy tissue sections.

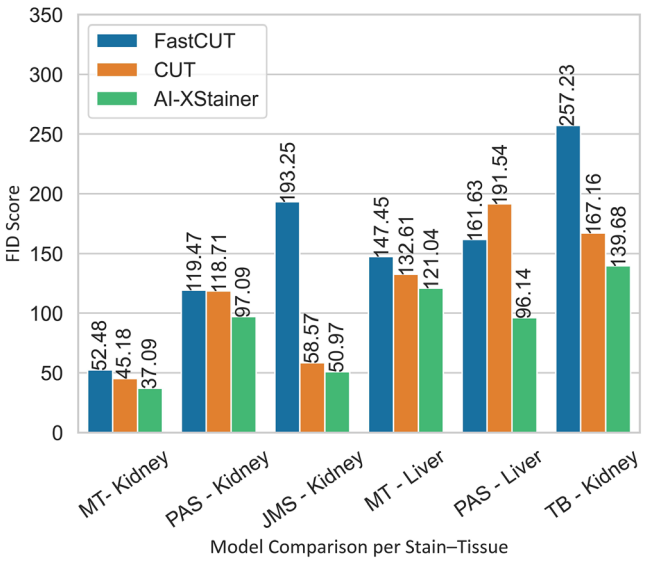

FIGURE 5 | Quantitative results. The bar plots show the performance of Xstainer compared to alternative style transfer methods for transforming H&E images to other stains, with lower FID scores indicating greater resemblance.

Diagnostic accuracy was comparable between Phase 1 (virtual images) and Phase 2 (real images) across the three diagnostic questions (Figure 7A). These questions focused on identifying key diagnostic features, such as fibrosis, basement membrane integrity, and cellular morphology. The analysis yielded a test statistic of 3.5 with a *p*-value of 0.58, indicating no statistically significant difference in the accuracy of rejection diagnosis between the two groups. This result shows that images produced by Xstainer may provide the same information as real stained images in clinical decision-making.

To further evaluate the clinical relevance of this finding, an effect size was calculated using the standard formula for nonparametric tests: $r = Z / \sqrt{N}$, where Z is the standardized test statistic, and N is the total number of observations, resulting in an effect size of: $r = -0.0926$ (reported as absolute value: $r = 0.07$). According to Cohen's guidelines ($r < 0.1$), this corresponds to a small effect size, suggesting that the lack of statistical significance may be due to the limited sample size and the resulting low statistical power.

3 | Materials and Methods

3.1 | Study Design

The study included tissue samples from the kidney, liver, and stomach sourced from six centers to ensure diverse staining variations and tissue characteristics. Only patches derived from WSIs were included, and artifacts or staining irregularities were excluded. The evaluation process combined quantitative metrics, such as FID scores, pathologists' expert assessments, and an analysis of the diagnostic impact, explicitly focusing on renal transplant rejection. This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki and was approved by the institutional ethics committees of Istanbul Yeni Yuzyil University (Approval No: 2022/11-939) and Memorial Healthcare Group (Approval No: 95 – 11/04/2023).

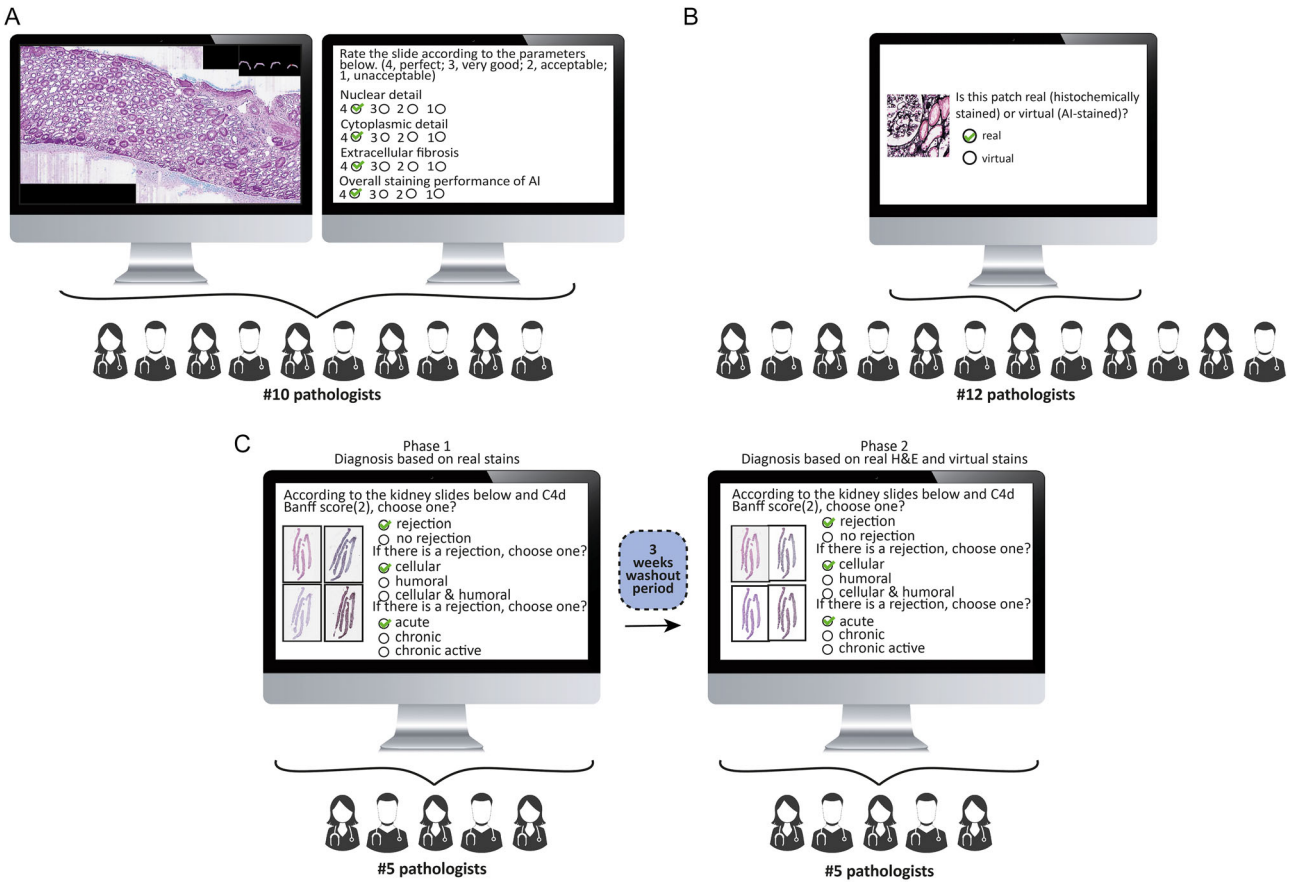

FIGURE 6 | Experimental designs to evaluate the virtual staining performance. (A) SSQA design. (B) Patch-level visual Turing test (VTT) design. (C) Reader study design.

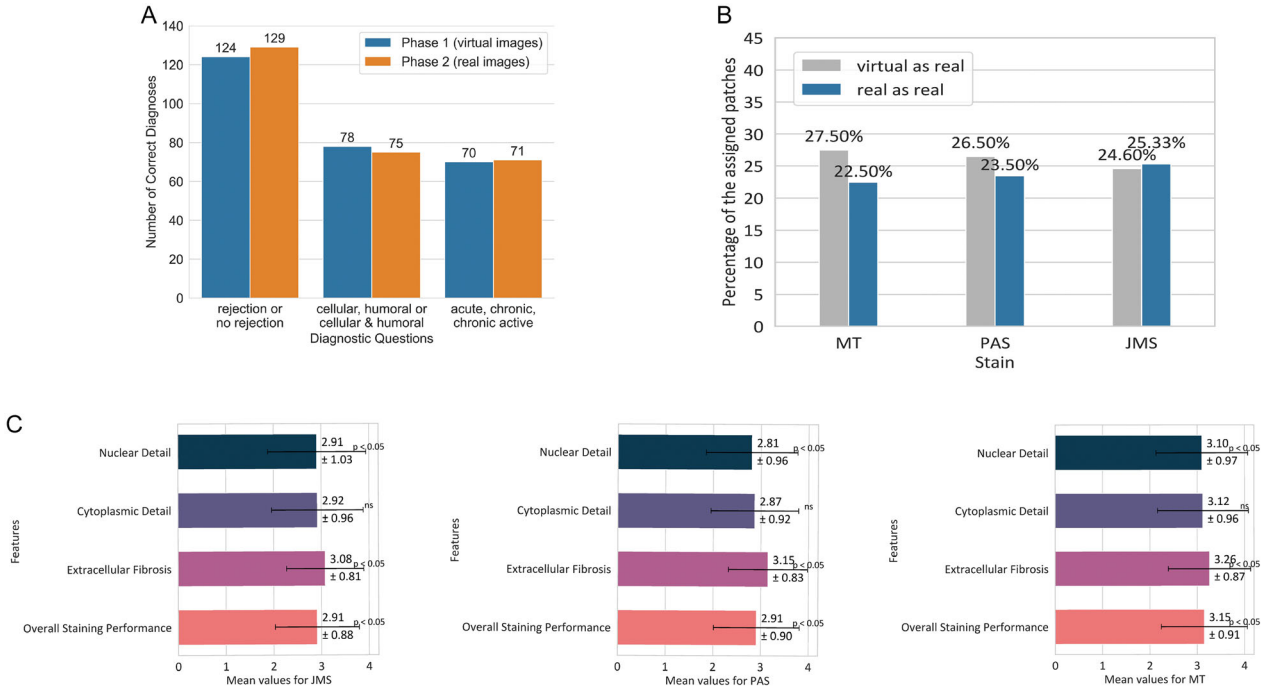

FIGURE 7 | Qualitative results. (A) Bar plots comparing the correct diagnosis of the first phase (virtual images) and second phase (real images) of a three-question reading study for diagnosing kidney rejection. (B) Pathologists' assessment of differentiating between real and virtual images at the patch level in the visual Turing test (VTT). (C) Mean values and standard deviations of pathologists' evaluations, using a rating scale of 1–4 (1 = unacceptable, 2 = acceptable, 3 = very good, 4 = perfect), were computed for the features in virtual MT, PAS, and JMS WSIs at the SSQA.

3.2 | Dataset

3.2.1 | OmniST Dataset Collection

Tissue samples were collected from five different centers in Türkiye: Ege University Hospital, Yeni Yuzyil University Gaziosmanpasa (GOP) Hospital, Kartal Dr. Lutfi Kirdar City Hospital, Erciyes University Hospital, and Memorial Hospitals Group, to create the OmniST dataset. The dataset comprises a total of 1,646 WSIs obtained from 389 biopsy cases collected between 2002 and 2022.

The OmniST dataset represents a valuable collection of biopsy samples derived from renal, liver, and gastric tissues, which were utilized for virtual staining of different tissue types. The dataset includes a diverse range of biopsies, including transplant kidney, explant liver, and stomach after Helicobacter pylori infection was detected. The kidney and liver cases were stained with H&E, MT, PAS, and JMS (except for liver) over the course of several years, reflecting routine clinical practice. For virtual TB staining, a total of 80 tissue slices (40 from each) were collected from Helicobacter pylori-positive gastric tissue samples at Kartal Dr. Lutfi Kirdar City Hospital and Ege University Hospital, then stained with H&E and TB. The dataset was curated to include stains relevant to the diagnosis of specific diseases, ensuring alignment with the staining practices used in disease diagnosis.

Comprising a substantial amount of data, the OmniST dataset consists of 738 WSIs and 2,752, 564 patches from Ege University Hospital, 475 WSIs and 1,989,079 patches from Yeni Yuzyil University GOP Hospital, 214 WSIs and 792,736 patches from Erciyes University Hospital, 40 WSIs and 282,880 patches from Kartal Dr. Lutfi Kirdar City Hospital, and 179 WSIs and 2,827,360 patches from Memorial Hospitals Group. This patient-based dataset includes information such as patients' age, gender, biopsy date for all kidney cases, Banff score, and additional transplantation dates for transplantation cases. The publication of the OmniST dataset is expected to encourage further research in digital pathology and contribute significantly to the field. This dataset is highly valuable for creating training and independent test cohorts, and it is expected to stimulate the development of accurate and efficient methods for virtual staining in the future.

3.2.2 | Training Set

The model was trained individually for each tissue–stain combination. For the kidney, training utilized 144 WSIs and 150,869 patches for the H&E → MT combination, 120 WSIs and 124,635 patches for the H&E → PAS combination, 105 WSIs and 53,629 patches for the H&E → JMS combination, all obtained from Ege University Hospital. Additionally, 40 WSIs and 50,036 patches for the H&E → TB combination were used, sourced from Kartal Dr. Lutfi Kirdar City Hospital. For the liver, 42 WSIs and 10,000 patches for MT staining, and 39 WSIs and 10,000 patches for PAS staining from Memorial Hospitals Group were employed. To ensure the integrity and accuracy of the data, a meticulous curation process was conducted to eliminate patches exhibiting poor quality, faded staining, or clinically misleading features. This curation process ensured that the model was trained with high-quality data, enhancing the robustness and reliability of the findings. The detailed characteristics of the training and test datasets for each tissue–stain transformation, including the number of WSIs, patches, patch size, and resolution, are summarized in Table 1.

3.2.3 | Test Set

The data were divided into a training set and one or two independent test cohorts for all tissue–stain combinations. The model was tested using patient-based data, where each patient had WSIs from both the source domain (H&E-stained WSI) and the target domain (specially stained WSI). For the kidney, 36 WSIs and 37,725 patches for the H&E → MT combination, 20 WSIs and 31,158 patches for the H&E → PAS combination, and 34 WSIs and 17,365 patches for the H&E → JMS combination were used for testing, sourced from Yeni Yuzyil University GOP Hospital and Erciyes University Hospital. For stomach tissues, 20 WSIs and 25,012 patches for the H&E → TB combination were obtained from Ege University Hospital. Regarding liver tissues, 23 WSIs and 5,234 patches, along with 20 WSIs and 5,415 patches (consisting of those not used in the training set), were utilized from Memorial Hospitals Group for testing the virtual staining process.

3.2.4 | WSI Processing

State-of-the-art digital slide scanners, including the Leica Aperio AT2 and in-house Hamamatsu Nanozoomer s60, were used to scan the glass slides. The kidney and liver slides were scanned at 20x magnification, while the stomach slides were scanned at 40x magnification to facilitate the differentiation of Helicobacter pylori. The CLAM [29] algorithm was used for accurate image analysis, enabling precise segmentation of tissue regions within each WSI while excluding holes in the tissue. Subsequently, 512×512 image patches were extracted from each WSI without

TABLE 1 | Data characteristics of the train and test data for each stain transformation.

Stain transformation	Tissue	Number of WSIs		Number of patches		Patch size	Resolution
		Training	Test	Training	Test		
H&E → MT	Kidney	144	36	150 869	37 725	512 × 512	20x
H& E→ PAS	Kidney	120	20	124 635	31 158	512 × 512	20x
H&E → JMS	Kidney	105	34	53 629	17 365	512 × 512	20x
H&E → MT	Liver	42	20	10 000	5415	512 × 512	20x
H&E → PAS	Liver	39	23	10 000	5234	512 × 512	20x
H&E → TB	Stomach	40	20	50 036	25 012	512 × 512	40x

overlap, facilitating further analysis. The train and test sets were meticulously curated to ensure optimal performance and reliable results. Table 1 provides detailed information about the train and test data.

3.2.5 | Network Architecture

For stain transformation between unpaired images, we employed a GAN-based style transfer architecture, as described in our previous work [30] (Figure 8). To effectively train the GAN, we utilized a combination of adversarial, contrastive, and self-regularization losses. The typical adversarial loss:

$$L_{\mathrm{GAN}}(\mathrm{G}, \mathrm{D}, \mathrm{X}, \mathrm{Y}) = E_{\mathrm{y}\tilde{\mathrm{Y}}} \log \mathrm{D}(\mathrm{y}) + E_{\mathrm{x}\tilde{\mathrm{X}}} \log(1 - \mathrm{D}(\mathrm{G}(\mathrm{x}))) \quad (1)$$

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 a real stained patch, y.

The transformation of histopathology images poses significant challenges due to the complex cellular morphology and tissue structure involved. While adversarial loss plays a crucial role in minimizing the differences between real stained patches and virtually stained patches, it is essential to ensure that diagnostically relevant information is preserved during the transformation process.

To address this, we incorporated a noise-contrastive loss function that focuses on preserving the content of the transformed patches. This loss function establishes correspondence between a query patch from a virtually stained sample patch in the target domain and a matching real H&E-stained patch in the source domain, both located at the same position. The comparison is performed in a manner that generates positive pairings for similar patches and negative pairings for dissimilar stains. By using this noise-contrastive loss function, we encourage the preservation of critical diagnostic features while effectively transforming the stains in histopathology images. This ensures that the transformed images maintain their diagnostically meaningful content, contributing to the accuracy and reliability of our stain transformation process. K negative samples and the query sample (positive) are mapped to L-dimensional vectors v− ∈ RM ×L and v, v+ ∈ RL, respectively. The m-th negatives in L samples are indicated by the symbol $\mathrm{v}^{-}{}_{\mathrm{k}} \in \mathrm{R}^{\mathrm{L}}$. The distance of scaling between the sample patches and the query is calculated by using a temperature $\tau = 0.07$. The cross-entropy loss is computed, indicating a tendency to prioritize positives over negatives.

$$\mathrm{L}(\mathrm{v}, \mathrm{v}+, \mathrm{v}-) = -\log[\exp(\mathrm{v}.\mathrm{v}+/\tau)/\exp(\mathrm{v}.\mathrm{v}+/\tau) + \sum_{\mathrm{n}=1}^{\mathrm{N}} \exp(\mathrm{v}.\mathrm{v}_{\mathrm{n}}^{-}/\tau)] \quad (2)$$

The generator network creates feature piles ž, each with spatial positions; each layer l in a feature pile depicts a region of the input patch. Later, the feature piles are given as input to a two-layered network called multilayer perceptron (MLP) [31]. The final features formed by the corresponding image features $\mathrm{z}^{\mathrm{s}}{}_{1}$, final features in the output domain $\mathrm{z}^{\mathrm{s}}{}_{1}$, and different extracted features $\check{\mathrm{z}}^{\mathrm{s}}{}_{1}$, $\mathrm{z}^{\mathrm{S}\backslash\mathrm{s}}{}_{1}$ characterize patch noise contrastive estimation loss:

$$L_{\mathrm{PatchNCE}}(\mathrm{G}, \mathrm{F}, \mathrm{X}) = E_{x\tilde{X}} \sum_{l=1}^{L} \sum_{s=1}^{S1} \left(\check{\mathrm{z}}_1^{s} z z_1^{s} z_1^{S/s} \right). \quad (3)$$

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

$$L_{\mathrm{sReg}}(\mathrm{G}, \mathrm{X}) = ||\mathrm{X} - \mathrm{G}(\mathrm{X})||_1 \quad (4)$$

The hybrid loss function:

$$\begin{aligned} L_{\mathrm{Xstainer}} = {} & L_{\mathrm{GAN}}(\mathrm{G}, \mathrm{D}, \mathrm{X}, \mathrm{Y}) + \lambda_{\mathrm{sReg}} L_{\mathrm{sReg}}(\mathrm{G}, \mathrm{X}) \\ & + \lambda_{\mathrm{X}} L_{\mathrm{patchNCE}}(\mathrm{G}, \mathrm{F}, \mathrm{X}) + \lambda_{\mathrm{Y}} L_{\mathrm{patchNCE}}(\mathrm{G}, \mathrm{F}, \mathrm{Y}) \end{aligned} \quad (5)$$

Spatial attention block (SAB). With the integration of SAB into the generator, the focus was on clinically significant features in the learning process. The SAB performs a nonlocal convolution operation for any input, X. In the SAB, three convolutional layers divide the input into three elements θ, φ, and g.

$$Z = f(\mathrm{X}, \mathrm{X}\top)\mathrm{g}(\mathrm{X}) \quad (6)$$

f denotes the correlation between the pixels in the input patch X. A dot product operation is utilized on θ and φ, after the convolutional operations. Then, the rectified linear unit (ReLU) function, σrelu, was used as the activation function.

$$P = \psi(\sigma_{\mathrm{relu}}(\theta(X)\varphi(X)^{\top})) \quad (7)$$

$\theta(X)\varphi(X)^{\top}$ yields a measure of input co-variance that can be explained as the degree of slope between two feature maps on different channels. Normalization is applied to output P with softmax function, $\varphi_{\mathrm{softmax}}$. Then, a matrix multiplication between the output of g and the softmax function is applied. After that, we convolved and upsampled 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 F ∈ RN × 64 × H × W.

$$S = \varphi(\sigma_{\mathrm{softmax}}(K)g(X)) \quad (8)$$

$$F = S + X \quad (9)$$

3.2.6 | Hyperparameters

Unlike traditional image-to-image translation models, our network architecture operates in a single direction between the source (real H&E images) and target (virtually stained images) domains, enabling faster training and lower memory requirements. We trained our model for each tissue–stain combination individually: MT (Kidney), PAS (Kidney), JMS (Kidney), MT (Liver), PAS (Liver), and TB (Stomach). In the training, we used the Adam optimizer with an initial learning rate of 0.0002 and momentum parameters $\beta 1 = 0.5$ and $\beta 2 = 0.999$ for eight epochs. We utilized 1 batch size, Xavier weight initialization [32], and instance normalization [33]. The loss weights in the hybrid loss function were empirically set as follows: the weight for L_{sReg} was set to 0.03, and λ_X and λ_Y were both set to 1.

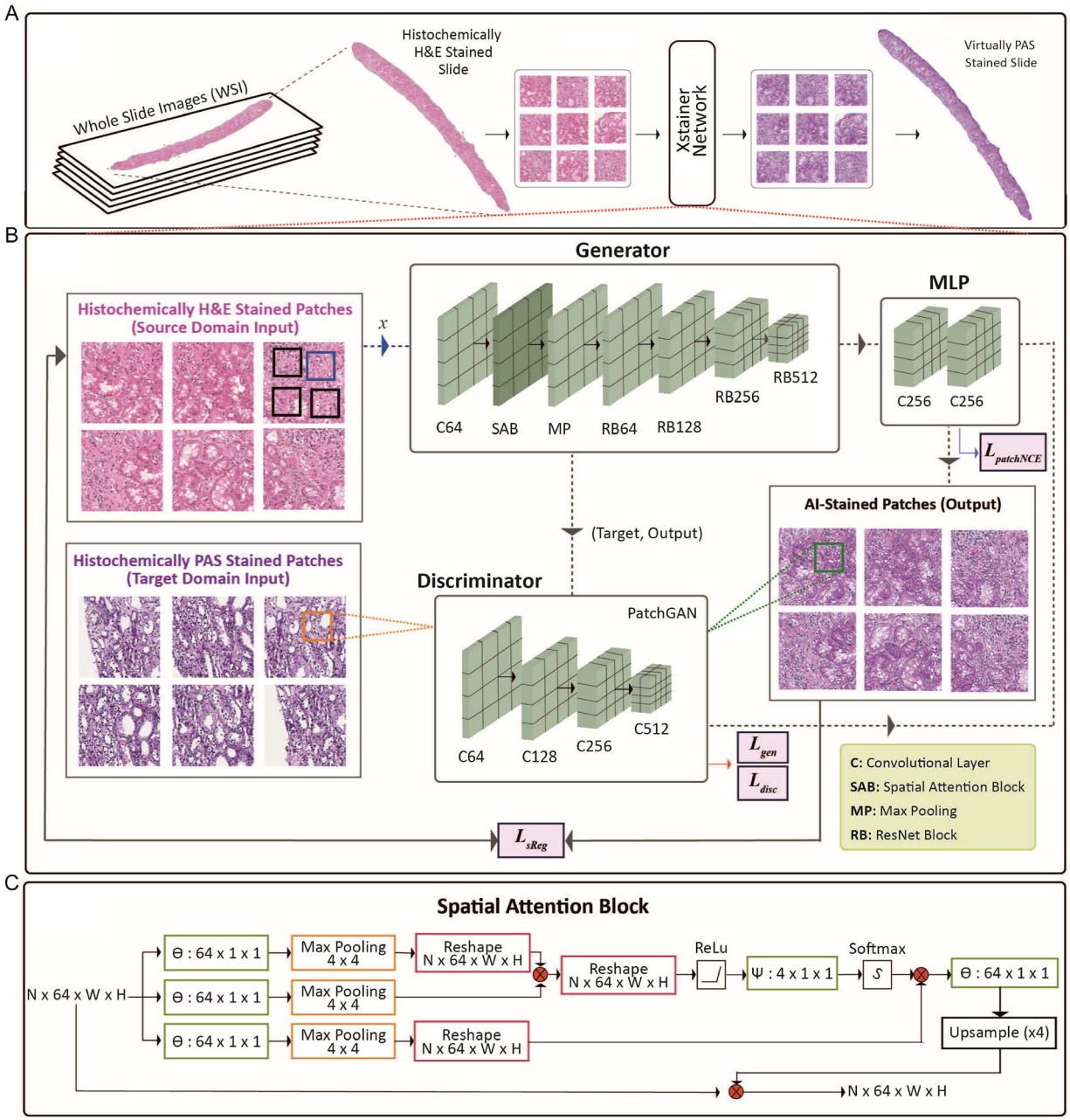

FIGURE 8 | Xstainer network architecture [24, 25, 30]. (A) Overview of the process. Initially, WSIs histochemically stained with H&E are cropped into 512×512 pixel square patches. These patches serve as input to the Xstainer networks, transforming them into virtually stained patches. These virtually stained patches are then stitched together to create final WSIs for pathologists to examine. (B) Detailed diagram of Xstainer networks. The SAB embedded in the ResNet-9 generator plays a notable role in the network architecture. The generator consists of ResNet blocks (RB) with different filter sizes. Each RB's filter size is indicated at the end of the block. The SAB output is then passed to the discriminator, which comprises convolutional layers (C) with filter sizes specified by numbers at the end of each layer (e.g., C128 corresponds to a convolutional block with a filter size of 128). The discriminator's output is concatenated with the image from the target domain (virtual patches). The adversarial loss, which combines L_{gen} and L_{disc}, encourages the generated images to resemble the target domain. To emphasize content preservation and focus on the relationship between the two domains, the overall objective function includes SR (super-resolution) and a patch-wise noise-contrastive estimation loss. This approach has been designed to capture long-range dependencies and inter-channel relationships within the input images. Additionally, a two-layer MLP network is incorporated alongside the generator to enhance the association between the input and the generated data. (C) Flow diagram of SAB in the ResNet-9 generator for virtually stained patch synthesis. This diagram illustrates the sequential steps involved in the SAB embedded in the ResNet-9 generator. It depicts the process of synthesizing virtually stained patches.

3.2.7 | FID

To objectively assess the performance of virtual staining at the patch level, we employed the PyTorch standard framework [34] to compute the FID measure. This metric quantifies the statistical similarity between the virtual stains and real counterparts by comparing feature distributions in a learned latent space. Low FID scores across all stain types indicated high visual similarity

between the virtual and real stains, establishing a strong foundation for interpretability by pathologists. This measure quantifies the dissimilarity between the Gaussian distributions of real and virtually stained images, accounting for differences in mean and standard deviation. As iterations progress, the FID score steadily decreases, while the virtually stained images become more realistic and higher quality. The FID score is calculated using the following formulation:

$$I_{FID} = ||\vec{m}1 - \vec{m}2||^2_2 + Tr(C1 + C2 - 2 \times \sqrt{(C1 \times C2)}) \quad (10)$$

here $\vec{m}1$ and $\vec{m}2$ represent the mean values of the real and virtually stained patches, and C1 and C2 are the covariance matrices. Tr denotes the trace function, which sums the diagonal elements of the matrix. In our comparison of Xstainer's performance with two established style transfer models, CUT and FastCUT, the FID score indicated that Xstainer achieved a high degree of similarity in the virtual staining tasks. Detailed comparisons of these results are presented in Section 2. All FID scores were computed on the whole test set. Since each score was calculated once per stain-tissue-method combination, without repeated runs or subsampling, they represent deterministic values with no statistical variability.

3.2.8 | Patch-Level VTT

A patch-level VTT was conducted using kidney tissue samples to evaluate the proficiency of pathologists in distinguishing between real and virtually stained patches. The study involved 150 patches, consisting of 25 real and 25 virtual patches for each of the three stains (PAS, MT, and JMS). These patches were randomly presented to a panel of 12 experienced pathologists.

The primary objective of the VTT was to assess the accuracy of pathologists in correctly identifying real and virtual patches. The results of this survey aimed to provide insight into the challenges associated with virtual staining techniques and their potential implications for diagnostic workflows. The inability to distinguish between real and virtual stains could indicate how closely the AI-generated stains resemble traditional methods, particularly in terms of diagnostic reliability.

3.2.9 | Slide-Level Staining Quality Assessment

In addition to patch-level evaluations, pathologists reviewed WSIs (30 slides per stain type) and assessed overall staining quality, including nuclear detail, cytoplasmic contrast, and extracellular matrix visualization. To comprehensively assess WSIs from a pathologist's perspective, we surveyed 90 kidney tissue WSIs. The dataset comprised 30 WSIs for each stain type (MT, PAS, JMS) that underwent successful virtual slide-level staining. Ten pathologists from various countries, including Türkiye and the United Kingdom, actively participated in the survey. We aimed to evaluate the overall staining performance of the AI and its ability to visualize key diagnostic features, including nuclear and cytoplasmic detail and extracellular fibrosis. A scoring system was employed, with scores of 1, 2, 3, and 4 representing "unacceptable", "acceptable", "very good", and "perfect" staining, respectively. This systematic evaluation provided a comprehensive understanding of the strengths and potential limitations of the virtual staining approach.

3.2.10 | Reader Study

To evaluate the impact of virtual staining on the slide-level diagnostic process for renal transplant rejection, a reader study was conducted with five experienced nephropathologists. The study consisted of two phases, enabling comparison between virtual and traditional staining methods. In Phase 1, the pathologists were presented with a set of 30 patients' 4 WSIs, including real stained H&E images and virtual stained JMS, MT, and PAS images. Their task was to provide a diagnosis based on these images. In this study, C4d scores were used as an external reference to support the evaluation of Xstainer's virtual stains in the context of antibody-mediated rejection diagnosis. Board-certified nephropathologists reviewed the virtual stains generated by Xstainer alongside known C4d scores, using these scores to inform their overall diagnostic assessments. The C4d scoring, based on traditional histochemical staining, ranged from 0 to 3: "0" indicating no C4d staining, "1" minimal focal staining, "2" moderate focal staining, and "3" strong diffuse staining. While Xstainer did not create virtual C4d stains, the existing C4d scores allowed pathologists to cross-reference and verify the alignment of the virtual stains with established diagnostic indicators.

Following a 3-week washout period to minimize potential bias, Phase 2 was initiated. The same set of patients' WSIs was again presented to the pathologists in this phase. However, only the real stained images were included this time. The pathologists were asked to provide diagnoses using the real-stained images. This two-phase approach enabled a direct comparison of the pathologists' diagnoses between virtual staining and traditional staining methods. The pathologists' diagnostic process involved determining the presence of rejection as the first question, followed by identifying the type of rejection in two additional questions: cellular, humoral, or a combination of both and acute, chronic, or chronic active.

A Wilcoxon signed-rank test was conducted to assess the accuracy achieved using solely virtually stained WSIs compared to real WSIs. The Wilcoxon signed-rank test is a nonparametric test that compares paired samples to determine whether their population means differ. The test statistic is calculated as follows:

$$W = \sum_{i=1}^{n} \mathrm{sgn}(x_i - y_i).R_i \quad (11)$$

where x_i and y_i are paired observations, sgn $(x_i - y_i)$ is the sign function of the difference, and R_i represents the rank of the absolute differences $|x_i - y_i|$ in ascending order, excluding zeros. W is the sum of the signed ranks. As presented in Section 2, the Wilcoxon test analyses underscore the diagnostic reliability of virtually stained WSIs, further demonstrating the potential of Xstainer to augment traditional pathological workflows.

3.3 | Computational Hardware and Software

WSI processing (patching and stitching) was performed on Xeon multi-core CPUs and Nvidia V100 GPUs in Python (version 3.6.13). Deep virtual staining networks were trained and tested on an Nvidia V100 GPU for each tissue–stain combination using PyTorch (version 1.9.0).

4 | Discussion

In this study, we introduced the Xstainer tool, which enables virtual staining of H&E-stained kidney, liver, and stomach biopsies. We leveraged the unique OmniST dataset, comprising WSIs from various staining techniques and different centers, to demonstrate the clinical effectiveness of the Xstainer tool.

Although H&E staining remains the foundation of routine histopathological evaluation, it has well-documented limitations in visualizing certain diagnostically relevant structures. For example, H&E is insufficient for accurately detecting basement membrane thickening, subtle mesangial matrix expansion, and early-stage fibrosis, which are often critical in evaluating renal pathology, especially in transplant rejection and glomerular diseases [35]. In clinical practice, special stains such as PAS and JMS are routinely employed to enhance the visualization of basement membranes and glycoprotein-rich structures, features that remain indistinct on H&E slides alone. Xstainer addresses this limitation by enabling the virtual transformation of H&E-stained images into multiple special stain equivalents, such as PAS and JMS, without requiring additional tissue sections or physical staining. This not only preserves the integrity of the original tissue block but also allows rapid, on-demand access to enhanced contrast and diagnostic visibility. In our reader study, nephropathologists could recognize histologically relevant features—particularly those involving basement membrane pathology—more clearly when aided by virtual PAS and JMS slides, supporting the potential of Xstainer to augment conventional H&E analysis in diagnostically complex cases. In summary, while H&E remains indispensable, virtual staining provides an efficient, scalable, and resource-conscious solution to overcome its interpretive limitations, especially when physical restaining is not feasible due to tissue or time limitations.

We employed a GAN with a SAB to achieve virtual staining. GANs consist of a generator and a discriminator that compete to generate realistic images and distinguish between real and generated images. We incorporated a spatial attention block into the generator, enabling the network to focus on relevant image regions. This enhances the model's ability to reproduce realistic and diverse output virtual images by emphasizing key histological features and spatial relationships within the tissue [36, 37]. By doing so, the model ensures accurate staining while maintaining the structural and diagnostic integrity of the virtual slides.

We conducted four distinct evaluations to assess the performance of the Xstainer model in virtual staining tasks. The FID, a quantitative metric, measured the similarity between feature distributions of real and generated images. Our model achieved the lowest FID scores, indicating high resemblance between real and virtual images [38, 39]. Patch-level VTT analysis involved 12 pathologists distinguishing real and virtual patches, revealing comparable or higher rates of assigning virtual images as real than real images. SSQA showed that pathologists rated the virtual staining of MT, PAS, and JMS as "very good" for overall staining performance and for successful visualization of nuclear detail, cytoplasmic detail, and extracellular fibrosis. The clinical effectiveness of this tool was rigorously evaluated through a comprehensive reader study conducted with five experienced nephropathologists. C4d has been recognized as a hallmark of antibody-mediated rejection, and integrating C4d staining into routine clinical practice has greatly enhanced understanding of antibody-mediated mechanisms in allograft rejection. In this study, C4d scores served as an external benchmark to support the diagnostic evaluation of Xstainer's virtual stains. While Xstainer does not directly produce virtual C4d stains, the availability of C4d scores as a reference helped nephropathologists assess antibody-mediated rejection. By reviewing Xstainer-generated virtual stains alongside known C4d scores, pathologists could cross-reference tissue morphology with established markers of immune response, thereby reinforcing the reliability of the virtual staining approach. This combination highlights the potential for Xstainer to assist in the diagnostic process by providing supplementary digital staining insights that can be used alongside established clinical markers, such as C4d [40].

The reader study plays a crucial role in assessing the clinical utility and validity of the generated virtual images, providing valuable insights into the performance and diagnostic accuracy of Xstainer. The results of this study demonstrate that Xstainer can produce virtually stained images that closely resemble their traditionally stained counterparts without compromising pathologists' diagnostic accuracy. These findings highlight the significance of the reader study in validating the clinical effectiveness of virtual staining and affirming its potential as a valuable tool in pathology practice. One significant consideration is the statistical power of the analysis. While the test did not show a statistically significant difference in the accuracy of rejection diagnosis between virtual and real staining, this result should be interpreted with caution, as it may reflect low statistical power due to the small patient sample. Future studies with larger sample sizes are necessary to achieve more robust conclusions and reduce the possibility of type II errors. The method's potential, supported by evidence from transplantation pathologists, could influence end users' perspectives on Xstainer and contribute to its broader adoption.

Our study significantly contributes to the rapidly expanding field of virtual staining, offering a faster, cost-effective, and expressive alternative to conventional staining methods. Virtual staining brings several advantages, including reduced biopsy requirements, preserved tissue integrity, time and resource savings, improved image contrast and resolution, and the ability to visualize multiple biomarkers simultaneously [41]. It also addresses the limitations and challenges associated with traditional staining, such as variability, reproducibility, standardization, quality control, and interpretation [42, 43].

Xstainer offers distinct financial and operational advantages through its unique capability to directly transform digitized routine H&E-stained images into virtual histochemical stains. According to current real-world procurement data from pathology laboratories in Türkiye, conventional histochemical staining typically costs $12–14 per slide, whereas in higher-cost regions such as the United States, the average cost ranges from $20 to 40 per slide. A comparative overview of estimated costs for traditional histochemical staining versus Xstainer-based virtual staining in the USA and Türkiye is provided in Table 2.

Xstainer significantly reduces or eliminates associated costs, such as reagents, technical labor, and additional tissue sectioning, by

TABLE 2 | Estimated cost comparison of traditional vs. virtual staining in the USA and Türkiye.

Cost item	Traditional staining (USA)	Xstainer (USA)	Traditional staining (Türkiye)	Xstainer (Türkiye)
Cost per histochemical slide	$30 (range $20–40)	~$3 (digitization only)	$13 (range $12–14)	~$3 (digitization only)
Annual histochemical staining volume (example)	50,000 slides/year	50,000 slides/year	10,000 slides/year	10,000 slides/year
Annual staining costs	$1,500,000	$150,000	$130,000	$30,000
Hardware + Software maintenance (annual)	—	$2,000	—	$1,000
License + Software updates (annual)	—	$1,000	—	$500
Total annual costs	**$1,500,000**	**$153,000**	**$130,000**	**$31,500**
Estimated total savings per year	**—**	**$1,347,000**	**—**	**$98,500**

digitally converting H&E images into histochemical staining visualizations. This results in estimated savings of approximately $10–11 per slide in Türkiye, with even greater potential savings of $20–35 per slide in regions with higher baseline costs, such as the US. Based on current laboratory expenditures, preliminary estimations suggest potential cost savings of approximately 70–90% per histochemical slide when virtual staining is utilized (Table 2). These savings primarily result from eliminating additional reagents, labor, and physical slide preparation, translating into considerable economic advantages, particularly in high-volume pathology laboratories.

Therefore, Xstainer streamlines operational workflows and substantially reduces the financial burden of histochemical staining. Its ability to digitally transform routine H&E images into specialized histochemical stains highlights its potential as an innovative, cost-effective tool applicable in pathology laboratories worldwide.

Unlike previous studies, we demonstrated the adaptability of the Xstainer tool using WSIs from six different centers (Figure 3) and a wide range of tissue-stain combinations (Figure 2). Additionally, we incorporated a final stitching step in the Xstainer workflow to generate virtually stained WSIs in minutes, facilitating integration into traditional pathology laboratory workflows. Moreover, we curated the patient-based OmniST dataset, comprising 1,646 WSIs stained with H&E, MT, PAS, JMS, or TB, which will expedite data preparation and foster the development of new methods in digital pathology, particularly virtual staining. While our study demonstrates the high visual realism of virtual stains generated by Xstainer, as evidenced by low FID scores and pathologist evaluations, realism alone does not ensure diagnostic utility. Diagnostic accuracy was specifically evaluated in the reader study. However, we acknowledge that future studies should directly compare diagnostic outcomes using H&E alone versus H&E combined with virtual stains to more clearly quantify the added value of virtual staining.

While our study provides compelling evidence, it is essential to acknowledge some limitations.

The pathologists' opinions in the evaluation may not represent the entire pathology community. Therefore, conducting a larger and more diverse reader study involving pathologists from different backgrounds and experiences would be valuable in gathering additional feedback and ensuring broader applicability. While stain translation, such as from H&E to other histochemical stains, is a powerful technique, it may not capture every structural detail visible in the original H&E stain. The translation's efficacy depends on the complexity and uniqueness of the features each stain presents. This limitation arises because specific chemical and molecular interactions, detectable with traditional staining methods, cannot be fully inferred from visual information alone in H&E-stained images. While Xstainer demonstrates the ability to generate visually realistic and diagnostically practical virtual stains, some localized morphological differences were observed between the virtual and real stains, including minor variations in lumen size and glomerular structure, and occasional artifacts in certain regions (e.g., in virtual JMS and MT stains). These discrepancies arise from the generative network's attempt to synthesize the staining characteristics of the target stain while adapting features from the H&E input. Critically, these variations were found to be diagnostically insignificant for high-level tasks, such as identifying fibrosis or assessing glomerular abnormalities, and did not affect pathologists' diagnostic accuracy in our evaluations. Overall, while these minor differences do not currently affect the diagnostic utility of the virtual stains, our findings demonstrate the robustness and promise of Xstainer as a transformative tool in histopathology. Further studies, including comparisons of diagnostic accuracy between H&E alone and H&E combined with virtual stains, will help quantify the added value of virtual staining in enhancing diagnostic workflows.

This study presents Xstainer, a novel virtual staining tool that distinguishes itself from existing methods in several key aspects. Its ability to produce multiple stain types efficiently places it ahead of many current methods, which are often limited to specific stains or tissue types. Moreover, it offers versatile capabilities, enabling the generation of a wide range of histochemical stains. Its clinical relevance is evidenced by a comprehensive two-phased reader study involving experienced nephropathologists, which showed no statistically significant difference compared to traditional staining, suggesting comparable performance within the sample size limits. Xstainer also shows promising quantitative performance, as indicated by lower FID scores than those of generic style transfer models. While the quantitative results are promising, it is essential to recognize that comparisons with other methods should be contextualized within the model's

design constraints. Efficiency is another highlight, potentially saving valuable time and resources. Considering these advantages, Xstainer stands as a promising and practical solution for virtual staining in the field of pathology.

In conclusion, our Xstainer tool demonstrates the ability to generate high-quality, realistic, and clinically valid virtual-stained images from H&E-stained biopsies. The performance evaluation, encompassing quantitative measures, qualitative assessments, and diagnostic accuracy, highlights the potential of virtual staining to overcome limitations associated with conventional staining methods. Despite the promising results, further refinement and broader clinical validation are necessary for the widespread adoption of virtual staining. Virtual staining offers a promising avenue to enhance the efficiency and accuracy of histopathological diagnosis, benefiting researchers and pathologists across various clinical settings. Future research will focus on refining the tool, addressing feedback from a larger cohort of pathologists, and exploring its applicability in broader clinical contexts.

Author Contributions

M.T. and F.N.K. led the planning, execution, and coordination of all research steps. M.T., F.N.K., D.D., and K.B. designed the experiments. K.B.O. and F.N.K. performed the statistical analysis of the results. D.D., F.N.K., G.S., B.S., C.U., U.P.H., F.O., Y.N.E., İ.T., and K.B. curated the OmniST dataset. F.N.K., M.T., D.D., and K.B. designed the figures. F.N.K., K.B.O., D.D., Y.A., G.E.K., E.C.K., D.F.K.W., and K.B. prepared the manuscript. M.T., K.B., Y.N.E., and D.D. supervised the research.

Acknowledgments

Mehmet Turan was supported by the Turkish Directorate of Strategy and Budget, Türkiye, under TAM project number 2007K12-873. Mehmet Turan and Fatma Nur Kinali gratefully acknowledge support from the Scientific and Technological Research Council of Turkey (TÜBİTAK) through the International Fellowship for Outstanding Researchers program. We also thank TÜBİTAK ULAKBİM for providing access to the Turkish National e-Science e-Infrastructure (TRUBA) cluster and data storage services, which facilitated the work of Mehmet Turan and Fatma Nur Kinali. Yavuz Nuri Ertas acknowledges support from the Turkish Academy of Sciences Distinguished Young Scientist Award (TÜBA-GEBİP) and Scientific and Technological Research Council of Türkiye (TÜBİTAK) Incentive Award. We appreciate the pathologists listed in the "Survey participants" section for their valuable time, insights, and feedback. Additionally, we thank Yeşim Yaşan, Talha Akbaş, and Daye Tangül for their assistance in collecting the dataset.

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All code is available at https://github.com/DeepMIALab/Xstainer. The OmniST dataset will be made available on request.

Survey Participants

Yasemin Yuyucu Karabulut, Havva Serap Toru, Ezgi Ayhan Cinar, Nihan Haberal, Enes Emre Asman, Saniye Sevim Tuncer, Ezgi Hacihasanoglu, Yelda Dere.

Affiliations: Faculty of Medicine, Mersin University, Mersin, Türkiye; Department of Pathology, Akdeniz University School of Medicine, Antalya, Türkiye; Department of Pathology, Van Research and Training Hospital, Van, Türkiye; Department of Pathology, Başkent University Faculty of Medicine, Ankara, Türkiye; Department of Pathology, Süleyman Demirel University School of Medicine, Isparta, Türkiye; İzmir Provincial Health Directorate, Bakırçay University Çiğli Regional Education Hospital; Yeditepe University School of Medicine, Department of Pathology, İstanbul, Türkiye; Muğla Sıtkı Koçman University, Faculty of Medicine, Department of Pathology, Muğla, Türkiye.

References

1. M. Titford, "The Long History of Hematoxylin," *Biotechnic & Histochemistry* 80 (2005): 73–78.

2. J. D. Bancroft and M. Gamble, *Theory and Practice of Histological Techniques* (Elsevier health sciences, 2008).

3. A. T. Feldman and D. Wolfe, "Tissue Processing and Hematoxylin and Eosin Staining," in *Histopathology*, ed. C. E. Day, (Springer New York, 2014), Vol. 1180, 31–43.

4. M. Golberg, J. Kobos, E. Clarke, et al., "Application of Histochemical Stains in Anatomical Research: A Brief Overview of the Methods," *Translational Research in Anatomy* 35 (2024): 100294.

5. Z. Y. Ma, X. F. Zhang, Y. Z. Hu, M. D. Zhu, J. Jin, and P. Qian, "Comparison of Staining Quality between Rapid and Routine Hematoxylin and Eosin Staining of Frozen Breast Tissue Sections: An Observational Study," *Journal of International Medical Research* 52 (2024): 03000605241259682.

6. K. Larson, H. H. Ho, P. L. Anumolu, and M. T. Chen, "Hematoxylin and Eosin Tissue Stain in Mohs Micrographic Surgery: A Review," *Dermatologic Surgery* 37 (2011): 1089–1099.

7. H. P. Cathro, S. S. Shen, and L. D. Truong, "Diagnostic Histochemistry in Medical Diseases of the Kidney," in *Seminars in Diagnostic Pathology* (Elsevier, 2018), Vol. 35, 360–369.

8. J. C. Iezzoni, "Diagnostic Histochemistry in Hepatic Pathology," in *Seminars in Diagnostic Pathology* (Elsevier, 2018), Vol. 35, 381–389.

9. A. B. Chowdhury and K. J. Mehta, "Liver Biopsy for Assessment of Chronic Liver Diseases: A Synopsis," *Clinical and Experimental Medicine* 23 (2022): 273–285.

10. T. V. Veuthey, M. G. Herrera, and V. I. Dodero, "Dyes and Stains: From Molecular Structure to Histological Application," *Frontiers in Bioscience-Landmark* 19 (2014): 91–112.

11. K. Rydell-Törmänen, K. Andréasson, R. Hesselstrand, et al., "Extracellular Matrix Alterations and Acute Inflammation; Developing in Parallel during Early Induction of Pulmonary Fibrosis," *Laboratory Investigation* 92 (2012): 917–925.

12. P. Dey, "Special Stains for the Carbohydrate, Protein, Lipid, Nucleic Acid and Pigments," in *Basic and Advanced Laboratory Techniques in Histopathology and Cytology* (Springer Nature Singapore, 2022), 83–99, https://doi.org/10.1007/978-981-19-6616-3_9.

13. K. S. Suvarna, C. Layton, and J. D. Bancroft, *Bancroft's Theory and Practice of Histological Techniques* (Elsevier health sciences, 2018).

14. S. Guo, N. D. Pottanat, J. L. Herrmann, and M. S. Schamberger, "Bartonella Endocarditis and Diffuse Crescentic Proliferative Glomerulonephritis with a Full-House Pattern of Immune Complex Deposition," *BMC Nephrology* 23 (2022): 181.

15. P. D. Walker, T. Cavallo, and S. M. Bonsib, "Practice Guidelines for the Renal Biopsy," *Modern Pathology* 17 (2004): 1555–1563.

16. G. Sridharan and A. A. Shankar, "Toluidine Blue: A Review of Its Chemistry and Clinical Utility," *Journal of Oral and Maxillofacial Pathology* 16 (2012): 251–255.

17. B. de Campos Vidal and M. L. S. Mello, "Toluidine Blue Staining for Cell and Tissue Biology Applications," *Acta Histochemica* 121 (2019): 101–112.

18. D. H. Kim, E. A. Song, S. W. Kim, and S. H. Hwang, "Efficacy of Toluidine Blue in the Diagnosis and Screening of Oral Cancer and Pre-Cancer: A Systematic Review and Meta-Analysis," *Clinical Otolaryngology* 46 (2021): 23–30.

19. K. de Haan, Y. Zhang, J. E. Zuckerman, et al., "Deep Learning-Based Transformation of H&E Stained Tissues into Special Stains," *Nature Communications* 12 (2021): 1–13.

20. Y. Rivenson, K. De Haan, W. D. Wallace, and A. Ozcan, "Emerging Advances to Transform Histopathology Using Virtual Staining," *BME Frontiers* 2020 (2020): 9647163.

21. C. Mercan, G. Reijnen-Mooij, D. T. Martin, et al., "Virtual Staining for Mitosis Detection in Breast Histopathology," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (IEEE, 2020), 1770–1774.

22. A. Lahiani, I. Klaman, N. Navab, S. Albarqouni, and E. Klaiman, "Seamless Virtual Whole Slide Image Synthesis and Validation Using Perceptual Embedding Consistency," *IEEE Journal of Biomedical and Health Informatics* 25 (2020): 403–411.

23. J.-Y. Zhu, 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 (ICCV)* (2017), 2223–2232.

24. Y. Zhao, R. Wu, and H. Dong, "Unpaired Image-to-Image Translation Using Adversarial Consistency Loss," in *Computer Vision – ECCV. 2020*, ed. A. Vedaldi, H. Bischof, T. Brox and J.-M. Frahm (Springer International Publishing, 2020), Vol. 12354, 800–815.

25. C. Yang, 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* 28 (2019): 4845–4856.

26. A. Rana, A. Lowe, M. Lithgow, et al., "Use of Deep Learning to Develop and Analyze Computational Hematoxylin and Eosin Staining of Prostate Core Biopsy Images for Tumor Diagnosis," *JAMA Network Open* 3 (2020): e205111–e205111.

27. S. Liu, C. Zhu, F. Xu, X. Jia, Z. Shi, and M. Jin, "Bci: Breast cancer immunohistochemical image generation through pyramid pix2pix," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2022), 1815–1824.

28. T. Park, A. A. Efros, R. Zhang, and J.-Y. Zhu, "Contrastive Learning for Unpaired Image-to-Image Translation," in *Computer Vision – ECCV. 2020*, ed. A. Vedaldi, H. Bischof, T. Brox and J.-M. Frahm (Springer International Publishing, 2020), Vol. 319–345.12354,

29. M. Y. Lu, 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," *Nature Biomedical Engineering* 5 (2021): 555–570.

30. K. B. Ozyoruk, S. Can, B. Darbaz, et al., "A Deep-Learning Model for Transforming the Style of Tissue Images From Cryosectioned to Formalin-Fixed and Paraffin-Embedded," *Nature Biomedical Engineering* 6 (2022): 1407–1419.

31. T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A Simple Framework for Contrastive Learning of Visual Representations," in *International Conference on Machine Learning*, (PMLR, 2020), 1597–1607.

32. X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (JMLR Workshop and Conference Proceedings, 2010), 249–256.

33. D. Ulyanov, "Instance Normalization: The Missing Ingredient for Fast Stylization," arXiv preprint arXiv:1607.08022, 2016.

34. M. Seitzer, pytorch-fid: FID Score for PyTorch (2020).

35. A. Katsuma, T. Yamakawa, Y. Nakada, I. Yamamoto, and T. Yokoo, "Histopathological Findings in Transplanted Kidneys," *Renal Replacement Therapy* 3 (2017): 6.

36. H. Emami, M. M. Aliabadi, M. Dong, and R. B. Chinnam, "SPA-GAN: Spatial Attention GAN for Image-to-Image Translation," *IEEE Transactions on Multimedia* 23 (2020): 391–401.

37. H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, Self-Attention Generative Adversarial Networks," in *International Conference on Machine Learning* (PMLR, 2019), 7354–7363.

38. A. Obukhov and M. Krasnyanskiy, "Quality Assessment Method for GAN Based on Modified Metrics Inception Score and Fréchet Inception Distance," in *Software Engineering Perspectives in Intelligent Systems*, ed. R. Silhavy, P. Silhavy and Z. Prokopovar (Springer International Publishing, 2020), Vol. 1294, 102–114.

39. A. Borji, "Pros and Cons of GAN Evaluation Measures," *Computer Vision and Image Understanding* 179 (2019): 41–65.

40. R. R. M. Corrêa, J. R. Machado, M. V. da Silva, et al., "The Importance of C4d in Biopsies of Kidney Transplant Recipients," *Clinical and Developmental Immunology* 2013 (2013): 1–8.

41. L. Jose, S. Liu, C. Russo, A. Nadort, and A. Di Ieva, "Generative Adversarial Networks in Digital Pathology and Histopathological Image Processing: A Review," *Journal of Pathology Informatics* 12 (2021): 43.

42. H. R. Tizhoosh and L. Pantanowitz, "Artificial Intelligence and Digital Pathology: Challenges and Opportunities," *Journal of Pathology Informatics* 9 (2018): 38.

43. A. Anghel, M. Stanisavljevic, S. Andani, et al., "A High-Performance System for Robust Stain Normalization of Whole-Slide Images in Histopathology," *Frontiers in Medicine* 6 (2019): 193.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.