

Making HE Histopathological Images More Colorful by Conditional Flow Matching

Mohamed Salim Ben Omrane¹, Ségolène Martin², Mounir Kaaniche^{1,3}, Astrid Laurent-Bellue^{4,5,6}, Wided Miled Souid⁷, Catherine Guettier^{4,5,6}, and Jean-Christophe Pesquet¹

¹ Centre de Vision Numérique, Université Paris-Saclay, Inria, CentraleSupélec, Gif-sur-Yvette, France

² Inria, ENS de Lyon, CNRS, Université Claude Bernard Lyon 1, LIP, UMR 5668, France

³ Université Sorbonne Paris Nord, L2TI, UR 3043, Villetaneuse, F-93430, France

⁴ Department of Pathology, AP-HP, Hôpital Bicêtre, Le Kremlin-Bicêtre, France

⁵ Faculté de Médecine, Université Paris-Saclay, Le Kremlin-Bicêtre, France

⁶ INSERM U1193, Villejuif, France

⁷ Université de Carthage, Institut National des Sciences Appliquées et de Technologie (INSAT), Centre Urbain Nord 1080 Tunis, Tunisia

Abstract. Recently, virtual stain generation has emerged as a promising tool in histopathology, with the objective of enabling rapid and cost-effective staining procedures. While current staining protocols commonly rely on hematoxylin (H) and eosin (E), additional stains—such as saffron (S)—have demonstrated clinical value in pathological practice in some countries. In this paper, we investigate generative models to produce HES-stained images from standard HE-stained samples. The proposed method operates in the concentration domain, leveraging a stain separation algorithm to first extract concentration maps from the input HE images. A conditional Flow Matching framework is then developed to digitally generate the saffron stain, conditioned on the hematoxylin and eosin concentration maps. Finally, the RGB HES digital image is reconstructed according to the Beer–Lambert law. Extensive experiments on private and benchmark multi-tissue datasets demonstrate that the proposed approach produces high-fidelity HES images with strong visual realism. The source code as well as the subjective evaluation tool are available at: <https://github.com/MedSalimBenOmrane/CFM-HES>.

Keywords: Histopathology, stain translation, generative models, flow matching.

1 Introduction

Tissue examination performed in histopathology plays a crucial role in the diagnosis and prognosis of multiple cancers [26,12]. To enhance tissue contrast and highlight specific biological structures, different staining techniques are applied. The most common worldwide technique uses hematoxylin (H) and eosin

(E), producing HE-stained whole slide images (WSIs) [8]. In some countries, such as France and Belgium, saffron (S) is further included, resulting in HES images. While hematoxylin shows cell nuclei and eosin stains cytoplasm and extra-cellular structures, saffron aims to highlight collagen fibers of connective tissue in the extra-cellular structures. The presence of collagen is an important parameter for the diagnosis of different tumor and non-tumor diseases, such as the identification of fibrosis in certain liver diseases [6].

However, the HES staining protocol is more expensive than the HE staining protocol due to the high cost of saffron. Moreover, most existing histopathology image analysis tools are developed for a single staining protocol (HE or HES), which hinders their generalization performance and limits their clinical impact. To address these limitations, it is of high interest to design efficient methods mapping HE to HES images, and inversely. While HE images can be easily obtained from HES images by resorting to stain separation algorithms [24] and removing the saffron component, the inverse process, which constitutes the objective of this paper, is much more challenging.

Histopathological image generation has attracted considerable attention in recent years and an extensive review of existing methods can be found in [15]. This topic has been explored in three main contexts. The first context, known as stain translation, aims to produce a new stained image from a given stained one. For instance, in [5], the authors predict the saffron from HE images by training a UNet model, and then reconstruct the HES image based on stain deconvolution. Generative Adversarial Networks (GAN), particularly CycleGAN [32] and Pix2Pix [11], have been widely used to transform HE stains into other special stains, like periodic acid-Schiff (PAS) and Masson’s trichrome (MT) in [9] and immunohistochemistry (IHC) in [13,30,27,14,17]. The second context, known as virtual staining, consists in generating stained images from label-free microscopic images of unstained samples [4,31]. Finally, the third context aims to produce large-scale synthetic histopathological images based on latent diffusion models [28,29] and wavelet diffusion models [1].

This work falls within the first category of methods mentioned above. More precisely, we propose to generate HES stained images from HE stained samples leveraging recent advances on generative models. In this respect, a recent stain separation algorithm is applied to the input HE images to extract their associated stain concentration maps. Then, a conditional Flow Matching is developed to produce the saffron stain component. Finally, the input H&E components are merged with the generated saffron through the Beer-Lambert law, yielding the final HES stained images. Extensive experiments on two datasets, with different types of tissue, show the superiority of the proposed approach. Validation at a high-resolution level as well as pathologist subjective quality assessment are also conducted, showing the ability of the proposed method to operate on whole slide images and demonstrating its applicability in digital pathology practice.

2 Methodology

2.1 Motivation

The reconstruction of saffron-stained images from hematoxylin–eosin (HE) data can be naturally formulated as a *generation problem*, where the goal is to synthesize plausible HES information consistent with a given HE observation. This setting motivates the use of modern generative models, which have recently shown strong performance in high-dimensional image synthesis and inverse problems.

In recent years, flow-based generative models, including diffusion models [10,25] and Flow Matching (FM) [16,2,18], have emerged as powerful approaches. These methods rely on ordinary or stochastic differential equation formulations and learn either score functions or time-dependent vector fields that define probability flows between a source distribution (e.g., white Gaussian) and the data distribution. As these approaches have been found to be closely related and even equivalent under certain assumptions, we focus here on the FM model. This choice is mainly driven by practical considerations. FM learns a constant-velocity field along the straight-line path, leading to more stable training and much faster inference compared to diffusion models. For instance, while a standard DDPM [10] requires about 1000 time steps, FM produces a high-quality generation in 50 ODE steps, which is a clear advantage for whole-slide deployment.

The choice of representation space is critical. Rather than operating in the RGB domain, we adopt a *concentration-based representation* derived from color deconvolution, in which the image formation process becomes approximately linear with respect to stain concentrations. This transformation provides a physically meaningful space, simplifying the representation of hematoxylin, eosin, and saffron components and facilitating learning.

In addition, we propose a *Conditional Flow Matching (CFM)* based approach for the HES reconstruction task. The pipeline of the proposed approach is shown in Fig. 1. Thus, our method relies on three main steps, named stain separation, conditional flow matching, and stain reconstruction. The proposed methodology will be described in what follows.

2.2 Concentration representation via stain separation

Let $I \in \mathbb{R}^{N \times 3}$ be a given vectorized stained image in the RGB color space, where N represents the number of pixels. The image intensity at each pixel can be expressed, through the Beer-Lambert law [23], as follows:

$$I = I_0 \exp(-HW), \quad (1)$$

where $I_0 = 255$ is the incident light, $W \in \mathbb{R}^{r \times 3}$ and $H \in \mathbb{R}^{N \times r}$ are referred to as stain column-vector matrix and stain concentration matrix, respectively. Here, r denotes the number of stains. For instance, for a HE (resp. HES) stained image, r is equal to 2 (resp. 3). Note that the r rows of W can be interpreted

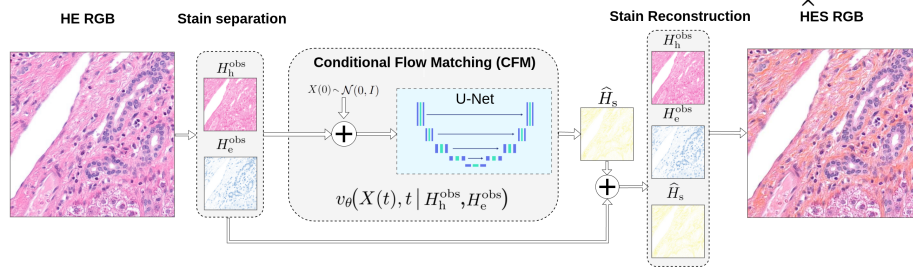


Fig. 1. Overview of the proposed framework composed of (1) stain separation for extracting the stain concentration maps, (2) conditional flow matching for saffron generation, and (3) stain reconstruction to produce the final HES stained image.

as the RGB values of the different stains, and the r columns of H correspond to the concentration maps of these stains. To linearize the relationship between the observed image and the matrices W and H , the input image I is transformed in the Optical Density (OD) space, yielding

$$\begin{aligned} V &= -\log \left(\frac{I}{I_0} \right) \\ &= HW. \end{aligned} \quad (2)$$

Therefore, to generate the color image associated to each stain, we need to estimate W and H from the given observation V . This problem is known as stain separation, or color deconvolution, in the literature [20,24]. In this respect, the stain color-vector matrix W can be experimentally obtained from pure stains as investigated in [23,24]. Then, the stain concentration matrix H can be estimated using a stain separation technique. In this work, we resort to a recent approach, based on an unrolled projected gradient algorithm incorporating a smooth total variation regularization to ensure the smoothness of separated images [24]. This method has shown state-of-the-art performance compared to other separation algorithms.

Once the stain matrix H is estimated, each stain image is produced using the exponential model. For example, for a HES stained image, using $W = [W_h \ W_e \ W_s]^\top \in \mathbb{R}^{3 \times 3}$ and $H = [H_h \ H_e \ H_s] \in \mathbb{R}^{N \times 3}$, the reconstructed RGB images, associated to hematoxylin I_h , eosin I_e and saffron I_s , are given by

$$I_h = I_0 \exp \left(-H_h W_h^\top \right) \quad (3)$$

$$I_e = I_0 \exp \left(-H_e W_e^\top \right) \quad (4)$$

$$I_s = I_0 \exp \left(-H_s W_s^\top \right). \quad (5)$$

2.3 Conditional flow matching based reconstruction approach

Given the HE WSIs, we extract RGB patches and map them to stain concentrations through color deconvolution as described in Section 2.2. Let

$$Y = [H_h^{\text{obs}} H_e^{\text{obs}}] \quad (6)$$

denote the *conditioning observation* in the concentration space, i.e., the hematoxylin and eosin concentration maps obtained from the HE-stained patch. For training purposes, we consider paired data that requires their associated ground-truth HES concentrations denoted by

$$H^{\text{GT}} = [H_h^{\text{GT}} H_e^{\text{GT}} H_s^{\text{GT}}]. \quad (7)$$

Thus, our main goal is to estimate the missing saffron concentration map through *Conditional Flow Matching* (CFM). In this work, we use the term Conditional Flow Matching [7] to explicitly denote the learning of a *conditional* velocity field and the estimation of a conditional data distribution. We note that in the literature, CFM is also a generic term for Flow Matching.

In our setting, p_0 is a standard Gaussian over saffron maps, i.e., $X(0) \sim p_0 = \mathcal{N}(0, I)$, while the target corresponds to the conditional distribution of saffron concentrations given the observation Y , i.e., $X(1) \sim p(H_s | Y)$. Sampling is performed by solving the conditional probability-flow ODE:

$$\frac{dX(t)}{dt} = v_\theta(X(t), t | Y), \quad t \in [0, 1], \quad (8)$$

where v_θ is parameterized by a neural network.

During training, each pair (Y, H_s^{GT}) defines a conditional transport problem. We draw $X(0)$ from a zero mean i.i.d. Gaussian distribution and sample $t \sim \mathcal{U}[0, 1]$. Following the straight-line probability path used in FM [16], we set

$$X(t) = (1 - t)X(0) + tH_s^{\text{GT}}, \quad (9)$$

which induces the target velocity $V(t) = H_s^{\text{GT}} - X(0)$. The conditional velocity field is learned by minimizing

$$\mathcal{L}_{\text{CFM}}(\theta) = \mathbb{E}_{t, X(0), Y} \|v_\theta(X(t), t | Y) - (H_s^{\text{GT}} - X(0))\|_2^2. \quad (10)$$

In practice, conditioning is implemented by concatenating the observation Y with the current state $X(t)$ at the network input (channel-wise), together with a time embedding.

At inference time, we sample $X(0) \sim \mathcal{N}(0, I)$ and integrate (8) from $t = 0$ to $t = 1$, yielding the estimated saffron stain $\hat{H}_s$. While the input stains are kept fixed, the HES concentration matrix can be estimated as follows

$$\hat{H} = [H_h^{\text{obs}} H_e^{\text{obs}} \hat{H}_s]. \quad (11)$$

Finally, the RGB reconstructed HES patches are obtained from $\hat{H}$ using (1), while enforcing non-negativity on $\hat{H}_s$ (e.g., by clipping).

3 Experiments and Results

3.1 Datasets

Our experiments are conducted on two datasets. The first is a private dataset constructed from liver tissue at Kremlin-Bicêtre hospital, France. It consists of 25 paired HE/HES WSIs corresponding to different patients. To mitigate leakage and overfitting caused by strong intra-slide correlations, we adopt a *slide-level* split by using 21 WSIs pairs for training, 2 for validation, and 2 for testing. From these WSIs, we extracted 30,000 paired patches of size 256×256 pixels, divided into 24,285 for training, 2,547 for validation, and 3,146 for testing. The second is the multi-tissue WSSB dataset, which is publicly available [3]. More precisely, we retained the breast and lung subsets, composed of 6 and 4 images of size 2000×2000 , respectively. From each subset, 14 patches of size 500×500 have been extracted. It is worth pointing out that the WSSB dataset is only used for test purposes due to the lack of HES ground truth, and the obtained results are evaluated using subjective tests.

3.2 Comparison methods and implementation details

Our CFM model is implemented using a U-Net backbone architecture [22] and trained for 150 epochs using Adam optimizer, while enabling automatic mixed precision (AMP) and using a batch size of 16 and a learning rate of 5×10^{-5} . At inference time, the sampling is performed by setting the time steps to 50. The proposed model will be compared to different state-of-the-art methods, including CycleGAN [32], Pix2Pix [11], PnP-FM [19], Pyramid pix2pix [17] and PredS [5]. While CycleGAN and Pix2Pix are considered among the most representative image translation methods, PnP-FM is a recent algorithm which combines Plug-and-Play (PnP) framework with Flow Matching for solving inverse problems. Furthermore, the pyramid pix2pix approach has recently been developed for HE to IHC staining, while PredS has been specifically developed for HE to HES staining. In particular, PredS shares the same pipeline as our CFM, including stain separation, concentration domain, and Beer-Lambert reconstruction. It differs by its reliance on a deterministic U-Net regression rather than a conditional generative model. For fair comparison, all these models have been also trained on the liver HE/HES dataset.

3.3 Results

Quantitative results: The different methods have been first evaluated using the liver test dataset. The comparison between the generated HES image and the ground truth is performed using traditional and perceptual metrics, including PSNR, SSIM, LPIPS, and PieAPP [21]. The results are shown in Table 1. Thus, among baseline image translation methods, CycleGAN seems to yield lower LPIPS and PieAPP values, while Pix2Pix and PnP-FM appear better

Table 1. Performance of the different HES generation methods on the liver test dataset.

Model	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PieAPP $\downarrow$
CycleGAN [32]	27.65	0.957	0.048	0.82
Pix2Pix [11]	28.68	0.959	0.052	1.31
PnP-FM [19]	28.05	0.964	0.066	1.60
Pyramid pix2pix [17]	29.15	0.952	0.050	0.88
PredS [5]	31.34	0.978	0.047	1.06
CFM (Ours)	32.06	0.968	0.030	0.66

in terms of PSNR and SSIM metrics. Regarding specific stain translation techniques, Pyramid pix2pix and PredS lead to better reconstruction performance. While PredS yields better results in terms of PSNR, SSIM and LPIPS compared to Pyramid pix2pix, the latter shows lower average PieAPP value. Most importantly, the proposed CFM model outperforms all the remaining methods in terms of PSNR, LPIPS, and PieAPP metrics. For instance, compared to the second best method PredS, our method achieves an average improvement of 0.7 dB in terms of PSNR and 0.4 in terms of PieAPP. These results confirm that the proposed CFM model generates HES images, which are both accurate and perceptually close to the ground truth data.

Qualitative results: Figure 2 illustrates the input HE image and its corresponding HES ground truth as well as the different generated HES images for a given sample taken from the liver dataset. Beyond patch-level comparison, we also consider high-resolution evaluation by applying the different methods on a larger tile of size 2048×2048 . The obtained high-resolution images revealed some mosaic artifacts (i.e., visible seams and color/intensity discontinuities at patch boundaries) with image translation methods CycleGAN, Pix2Pix, and PnP-FM. Moreover, these baselines as well as the stain translation methods may exhibit local texture inconsistencies and stain-intensity variations. In contrast, our CFM model produces more spatially consistent tiles, with reduced seam visibility and more stable saffron-related intensity across neighboring patches.

Expert validation: The two best methods CFM and PredS have been subjectively evaluated on liver and WSSB datasets. Five pathologists have been invited to assess the subjective quality of 14 reconstructed HES images, for each organ and each model, based on the input HE stained image and its corresponding HES ground truth when available (in case of liver samples). This results in a subjective dataset of 84 images. Pathologists were asked to provide one quality score among four levels that correspond to poor, medium, high, and very high quality. The score is given based on the overall HES reconstruction, while taking into account the intensity of the generated saffron stain and its distribution throughout the generated image.

Table 2 shows the average rate (across pathologists) of each quality level for each tissue organ. It can be observed that the proposed model outperforms PredS in the liver and breast organs, producing a very high quality level on all

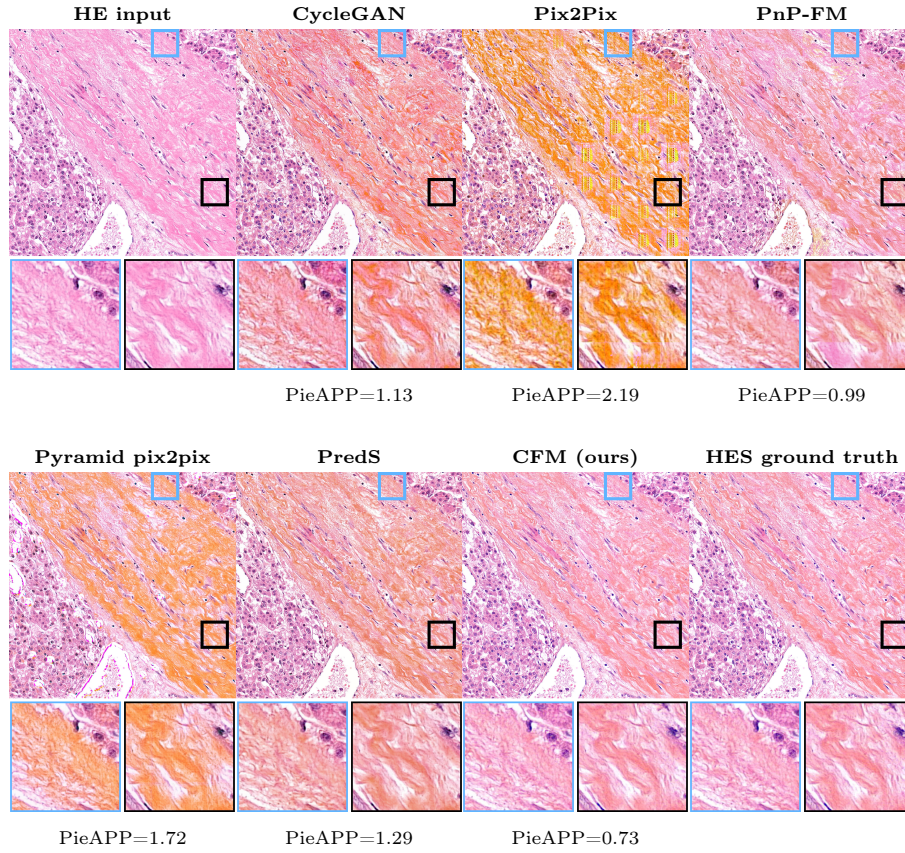


Fig. 2. Reconstructed HES patches and high-resolution images for the different generation methods.

Table 2. Expert validation results on liver and WSSB datasets.

Organ	CFM (Ours)				PredS			
	Very high	High	Medium	Poor	Very high	High	Medium	Poor
Liver	100%	0%	0%	0%	38.6%	45.7%	15.7%	0%
Breast	30%	47.1%	22.9%	0%	30%	45.7%	20%	4.3%
Lung	2.9%	20%	17.1%	60%	0%	24.3%	34.3%	41.4 %
ALL	44.3%	22.4%	13.3%	20%	22.9%	38.6%	23.3%	15.2%

liver samples and a high or very high quality (77.1% for both levels) on the breast organ. However, the proposed model seems to be less efficient on the lung organ. This may be explained by differences in tissue architecture and cytology in these organs. Note that the subjective evaluation tool is available at⁸. The

⁸ <https://subjective-hes-image-evaluation.vercel.app/>

latter illustrates the set of reconstructed HES using the two best models. A validation at a large-scale level (of size $20,480 \times 24,576$ pixels) is also provided through the subjective tool.

4 Conclusion

We introduced a concentration-domain framework for HE to HES stain translation that explicitly targets the missing saffron map. Our approach combines stain separation to obtain hematoxylin/eosin concentration maps with a Conditional Flow Matching model that generates the saffron concentration conditioned on the input stains. The final RGB HES image is reconstructed via the Beer-Lambert model. Various experiments on two datasets, with different types of tissue, demonstrate that the proposed CFM approach achieves good reconstruction performance. Future work aims at further improving the generalization to unseen tissues and adapting the proposed approach to other staining protocols.

Acknowledgments. We would like to thank the surgical team of the Centre Hépatobiliaire of the Hôpital Paul Brousse, and the technicians of the Pathology department of the Hôpital Bicêtre.

Disclosure of Interest. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Abdullah, M., Huang, T., Lee, I., Ahn, E.: High-resolution histopathology whole slide image generation using wavelet diffusion model. In: International Symposium on Biomedical Imaging (ISBI). Houston, TX, USA (April 2025) [2](#)
2. Albergo, M.S., Vanden-Eijnden, E.: Building normalizing flows with stochastic interpolants. In: Int. Conference on Learning Representations (ICLR) (2023) [3](#)
3. Alsubaie, N., Trahearn, N., Raza, S.E.A., Snead, D., Rajpoot, N.M.: Stain deconvolution using statistical analysis of multi-resolution stain colour representation. PLOS One **17**(1), e0169875 (2017) [6](#)
4. Bai, B., Yang, X., Li, Y., et al.: Deep learning-enabled virtual histological staining of biological samples. Light: Science and Applications **12** (2023) [2](#)
5. Balezo, G., Bertram, C.A., Tilmant, C., Petit, S., Hadj, S.B., Fick, R.H.J.: Enabling collagen quantification on he-stained slides through stain deconvolution and restained he-hes. In: International Symposium on Biomedical Imaging (ISBI). Cartagena, Colombia (September 2023) [2](#), [6](#), [7](#)
6. Bedossa, P.: Utility and appropriateness of the fatty liver inhibition of progression (flip) algorithm and steatosis, activity, and fibrosis (saf) score in the evaluation of biopsies of nonalcoholic fatty liver disease. Hepatology **60**, 565–575 (2014) [2](#)
7. Chemseddine, J., Hagemann, P., Steidl, G., Wald, C.: Conditional Wasserstein distances with applications in Bayesian OT flow matching. Journal of Machine Learning Research **26**(141), 1–47 (2025) [5](#)
8. Feldman, A.T., Wolfe, D.: Tissue processing and hematoxylin and eosin staining. Methods in Molecular Biology **1180**, 31–43 (2014) [2](#)

9. Haan, K.D., Zhang, Y., Zuckerman, J.E., et al.: Deep learning-based transformation of h&e stained tissues into special stains. *Nature Com.* **12** (2021) [2](#)
10. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: *Advances in Neural Information Processing Systems (NeurIPS)*. vol. 33, pp. 6840–6851. Vancouver, Canada (December 2020) [3](#)
11. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. p. 1125–1134. Honolulu, HI, USA (2017) [2](#), [6](#), [7](#)
12. Laurent-Bellue, A., Sadraoui, A., Claude, L., Calderaro, J., Posseme, K., Vibert, E., Cherqui, D., Rosmorduc, O., Lewin, M., Pesquet, J.C., Guettier, C.: Deep learning classification and quantification of pejorative and nonpejorative architectures in resected hepatocellular carcinoma from digital histopathologic images. *The American Journal of Pathology* **194**(9), 1684–1700 (2024) [1](#)
13. Li, F., Hu, Z., Chen, W., Kak, A.: Adaptive supervised patchnce loss for learning h&e-to-ihc stain translation with inconsistent groundtruth image pairs. In: Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., Taylor, R. (eds.) *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. pp. 632–641. Springer Nature, Cham (2023) [2](#)
14. Li, J., Dong, J., Huang, S., Li, X., Jiang, J., Fan, X., Zhang, Y.: Virtual immunohistochemistry staining for histological images assisted by weakly-supervised learning. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA (June 2024) [2](#)
15. Lin, W., Hu, Y., Zhu, R., Wang, B., Wang, L.: Virtual staining for pathology: Challenges, limitations and perspectives. *Intelligent Oncology* **1**, 105–119 (2025) [2](#)
16. Lipman, Y., Chen, R.T.Q., Ben-Hamu, H., Nickel, M., Le, M.: Flow matching for generative modeling. In: *International Conference on Learning Representations (ICLR)*. Kigali, Rwanda (May 2023) [3](#), [5](#)
17. Liu, S., Zhu, C., Xu, F., Jia, X., Shi, Z., Jin, M.: Bci: Breast cancer immunohistochemical image generation through pyramid pix2pix. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. USA (June 2022) [2](#), [6](#), [7](#)
18. Liu, X., Gong, C., Liu, Q.: Flow straight and fast: Learning to generate and transfer data with rectified flow. In: *Int. Conf. on Learning Representations (2023)* [3](#)
19. Martin, S., Gagneux, A., Hagemann, P., Steidl, G.: PnP-Flow: Plug-and-play image restoration with flow matching. In: *International Conference on Learning Representations (ICLR)* (2025) [6](#), [7](#)
20. Moyes, A., Zhang, K., Ji, M., Zhou, H., Crookes, D.: Unsupervised deep learning for stain separation and artifact detection in histopathology images. In: *Medical Image Understanding and Analysis*. pp. 221–234 (April 2020) [4](#)
21. Prashnani, E., Cai, H., Mostofi, Y., Sen, P.: PieAPP: Perceptual Image-Error Assessment Through Pairwise Preference. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1808–1817. UT, USA (June 2018) [6](#)
22. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. pp. 234–241. Springer, Cham (2015) [6](#)
23. Ruifrok, A., Johnston, D.: Quantification of histochemical staining by color deconvolution. *Analytical and Quantitative Cytology and Histology* **23**(4), 291–299 (2001) [3](#), [4](#)
24. Sadraoui, A., Laurent-Bellue, A., Kaaniche, M., Benazza-Benyahia, A., Guettier, C., Pesquet, J.C.: Unrolled projected gradient algorithm for stain separation in digital histopathological images. In: *IEEE International Conference on Image Processing (ICIP)*. pp. 2814–2819. United Arab Emirates (October 2024) [2](#), [4](#)

25. Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B.: Score-based generative modeling through stochastic differential equations. In: International Conference on Learning Representations (ICLR) (2021) [3](#)
26. Wang, X., Zhao, J., Marostica, E., et al.: A pathology foundation model for cancer diagnosis and prognosis prediction. *Nature* **634**, 970–978 (2024) [1](#)
27. Wei, L., Hua, S., Zhang, S., Zhang, X.: Derestainer: H&e to ihc pathological image translation via decoupled staining channels. In: MICCAI Workshop on Deep Generative Models. Marrakesh, Morocco (October 2024) [2](#)
28. Yellapragada, S., Graikos, A., Li, Z., Triaridis, K., et al.: Pixcell: A generative foundation model for digital histopathology images. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). USA (June 2025) [2](#)
29. Yellapragada, S., Graikos, A., Triaridis, K., Prasanna, P., Gupta, R.R., Saltz, J., Samaras, D.: Zoomldm: Latent diffusion model for multi-scale image generation. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA (June 2025) [2](#)
30. Zhang, W., Hui, T.H., Tse, P.Y., Hill, F., Lau, C., Li, X.: High-resolution medical image translation via patch alignment-based bidirectional contrastive learning. In: Linguraru, M.G., Dou, Q., Feragen, A., Giannarou, S., Glocker, B., Lekadir, K., Schnabel, J.A. (eds.) *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. pp. 178–188. Springer Nature Switzerland, Cham (2024) [2](#)
31. Zhang, Y., Huang, L., Liu, T., Cheng, K., de Haan, K., Li, Y., Bai, B., Ozcan, A.: Virtual staining of defocused autofluorescence images of unlabeled tissue using deep neural networks. *Intelligent Computing* (2022) [2](#)
32. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). p. 2223–2232. Honolulu, HI, USA (July 2017) [2](#), [6](#), [7](#)