A comparison of diffusion models and CycleGANs for virtual staining of slide-free microscopy images

Tanishq Mathew Abraham [1] **Richard Levenson** [2]

Abstract

Slide-free microscopy (SFM) methods can serve as a faster alternative to the standard histological examination of tissue specimens. However, SFM methods often provide images that differ from the hematoxylin- and eosin-stained (H&E) images commonly obtained in standard histology. Unpaired image-to-image translation has been explored for transforming SFM images into H&E images, a process known as virtual staining. Here, we compare a standard CycleGAN approach to a diffusion model-based approach for virtual staining of SFM images. We observe that the diffusion model approach, which relies on the inherent semantic preservation of the latent encodings, fails to outperform the standard CycleGAN approach, when tested on two different SFM datasets. This indicates that the semantic preservation of diffusion models is lacking for virtual staining tasks and additional regularization is needed.

1. Introduction

In histopathology, a tissue specimen is obtained from a patient, processed, prepared as a tissue slide, and viewed under a brightfield microscope (Bancroft & Gamble, 2008). These slides are stained with histological dyes that label important tissue structures: the extremely common hematoxylin and eosin (H&E) stain combination labels the nuclei and cytoplasm of cells. However, this typical workflow takes several hours, which limits the application of histopathology in more time-sensitive situations (like during a surgery) (Liu et al., 2022). Slide-free microscopy (SFM) systems have been developed to examine fresh tissue within a few minutes, but provide colors and contrasts that differ from standard brightfield image, which hinders interpretation by clinicians (Liu et al., 2022). This has motivated the use of deep learning, specifically generative adversarial networks, to transform SFM images into those that resemble H&E to enable easier interpretation by clinicians (Bai et al., 2023). This task is often referred to as virtual staining. However, diffusion models have not been previously examined in the context of virtual staining.

Here, diffusion models are applied for virtual staining of SFM images and compared to CycleGANs. As demonstrated in (Su et al., 2023), diffusion models inherently provide latent encodings that can be exploited for unpaired image-to-image translation. This approach is demonstrated here, and shown to provide satisfactory virtually stained SFM images, but are inferior to virtual staining with CycleGANs. This comparison is demonstrated for virtual staining of images taken with Microscopy with Ultraviolet Surface Excitation (MUSE) (Fereidouni et al., 2017) and Fluorescence Imitating Brightfield Imaging (Borowsky et al., 2023).

2. Methods

We define two image domains, one for SFM images (X), and one for H&E images (Y). We attempt to determine the transformation $G\colon X \to Y$.

2.1. Diffusion models

Diffusion models, or score-based generative models, are trained to reverse a diffusion process. It is also described by the following forward and reverse stochastic differential equations (Song et al., 2021):

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\mathrm{d}t + g(t)\mathrm{d}\mathbf{w},$$
$$\mathrm{d}\mathbf{x} = \left[\mathbf{f} - g^2\nabla_{\mathbf{x}} \log p_t(\mathbf{x})\right] \mathrm{d}t + g(t)\mathrm{d}\mathbf{w}$$

where $\mathbf{w}$ is the Wiener process, while $\mathbf{f}(\mathbf{x}, t)$ is termed the drift coefficient and $g(t)$ is the diffusion coefficient, both of which is defined by the type of diffusion model used. The $\nabla_{\mathbf{x}} \log p_t(\mathbf{x})$ "score" term is approximated by a neural network $\mathbf{s}_\theta(\mathbf{x}, t)$. The score term is learned via the following loss:

[1]Department of Biomedical Engineering, University of California, Davis [2]Department of Pathology, University of California, Davis Health. Correspondence to: Tanishq Mathew Abraham <tmabraham@ucdavis.edu>.

Workshop on Challenges in Deployable Generative AI at International Conference on Machine Learning (ICML), Honolulu, Hawaii, USA. 2023.

$$\mathcal{L}_{\text{diffusion}}(\theta) := \mathbb{E}_{t,\mathbf{x}_0,\boldsymbol{\epsilon}}\left[\left\|\boldsymbol{\epsilon} - \boldsymbol{s}_\theta\left(\sqrt{\alpha_t}\mathbf{x}_0 + \sqrt{1-\alpha_t}\boldsymbol{\epsilon}, t\right)\right\|_2\right]$$

where $\|\cdot\|_2$ is the L_2 norm, t is the timestep, α_t is some predefined variance parameterized by the timestep, $\boldsymbol{x}_0$ is an image from the training set, and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$. Images are sampled from the model through a discretization of the reverse stochastic differential equation.

Song et al. introduced denoising diffusion implicit models (Song et al., 2021), a sampling procedure for diffusion models that can be done very efficiently in a deterministic manner. Essentially, the SDE can instead be expressed by a deterministic ordinary differential equation (ODE) that has the same marginal densities as the diffusion SDE.

$$\mathrm{d}\mathbf{x} = \left[\mathbf{f}(\mathbf{x}, t) - \frac{1}{2}g(t)^2\nabla_{\mathbf{x}}\log p_t(\mathbf{x})\right]\mathrm{d}t$$

This ODE is termed the probability flow (PF) ODE. As it is deterministic, the sampling of the PF ODE can be run in reverse to provide uniquely identifiable latent encodings $\mathbf{x}_0$. This process is sometimes referred to as DDIM inversion.

It has been observed that this latent encoding contains semantic information that is preserved when using the same latent encoding for sampling with two separate diffusion models. Exploiting the preservation of semantic information, Dual Diffusion Implicit Bridges (DDIBs) (Su et al., 2023) simply perform DDIM inversion with a diffusion model trained on datapoints from domain X, and DDIM sampling with the obtained latent encoding $\mathbf{x}_0$, using the diffusion model trained on domain Y. Here, a U-Net (Ronneberger et al., 2015) was trained with an ϵ-parameterization objective (Ho et al., 2020) and a cosine noise schedule with N=1000 timesteps. A single neural network was trained on both X and Y images, with a specified condition $\mathbf{c}$, resulting in a conditional diffusion model $\mathbf{s}_\theta(\mathbf{x}, t, \mathbf{c})$. The conditioning is implemented via sinusoid embeddings that are added to the timestep embedding. The neural network was trained from scratch for 1000 epochs at a learning rate of 1e-4 with a batch size of 4 using an AdamW optimizer. During inference, the DDIM sampler was used with 150 timesteps.

2.2. CycleGANs

In the CycleGAN framework, we have two tasks. One task is to learn a generator $G_X\colon X \to Y$ that maps $x \in X$ to $y \in Y$. The auxiliary task is to learn a generator $G_Y\colon Y \to X$. Additionally, we have the adversarial discriminators D_X and D_Y. D_X discriminates between the fake outputs of G_X and real images from domain Y. Conversely, D_Y discriminates between the fake outputs of G_Y and real images from domain X.

CycleGAN exploits the cycle-consistency property that $G_Y(G_X(x)) \approx x$ and $G_X(G_Y(y)) \approx y$. This constraint can be expressed as the following loss:

$$\begin{aligned}\mathcal{L}_{\text{cycle}}\left(G_X, G_Y\right) =& \mathbb{E}_{x\sim p_{\text{data}}(x)}\left[\left\|G_Y\left(G_X(x)\right) - x\right\|_1\right] \\ &+ \mathbb{E}_{y\sim p_{\text{data}}(y)}\left[\left\|G_X\left(G_Y(y)\right) - y\right\|_1\right]\end{aligned}$$

where $\|\cdot\|_1$ is the L_1 norm. Additionally, the GANs are trained with the traditional adversarial losses (Zhu et al., 2020). Finally, for regularization, we impose an "identity" constraint:

$$\begin{aligned}\mathcal{L}_{\text{identity}}\left(G_X, G_Y\right) =& \mathbb{E}_{y\sim p_{\text{data}}(y)}\left[\left\|G_X(y) - y\right\|_1\right] \\ &+ \mathbb{E}_{x\sim p_{\text{data}}(x)}\left[\left\|G_Y(x) - x\right\|_1\right]\end{aligned}$$

The generator architecture is a ResNet-based fully convolutional network described in (Zhu et al., 2020). A 70x70 PatchGAN (Isola et al., 2017) is used for the discriminator. The same loss function and optimizer as described in the original paper (Zhu et al., 2020) was used. The learning rate was fixed at 2e-4 the first 100 epochs and linearly decayed to zero in the next 100 epochs, like (Zhu et al., 2020).

2.3. Slide-free microscopy datasets

In this work, two slide-free microscopy datasets were utilized:

1. Urothelial carcinoma (kidney) MUSE-to-H&E: The H&E data came from a region in a single whole-slide image of human kidney with urothelial cell carcinoma. The MUSE data came from a single surface image of similar tissue. We obtained 512x512 tiles from the images, resulting in 344 H&E tiles and 136 MUSE tiles for training, and 100 MUSE tiles for testing. The tiles were randomly cropped into 256x256 images when loaded into the model for training. During testing, the full 512x512 tiles were provided as inputs to the models. As observed in previous studies (Abraham et al., 2020; Bai et al., 2023), color- and intensity-inverted MUSE images greatly improves CycleGAN results, so all tiles were inverted.
2. Benign breast FIBI-to-H&E: The H&E data came from a region in a single whole-slide image of human benign breast. The FIBI data came from a single surface image of similar tissue. We obtained 512x512 tiles from the images, resulting in 832 H&E tiles and 900 FIBI tiles for training, and 567 FIBI tiles for testing.

2.4. Evaluation

A classifier (a ResNet18) was trained to quantitatively evaluate how "real" the outputs of the models look. We used

accuracy from the classifier to quantitatively compare the quality of the generative models. We trained a separate critic on the predictions for each model to keep results independent. Each critic were trained for 20 epochs with a 0.001 learning rate (one-cycle learning rate schedule). Each dataset consisted of "fake" H&E images generated from the test set and real H&E images from the train set. It was a balanced dataset with an 80/20 dataset split. The test set accuracy after 20 epochs is reported.

In addition to reporting the classifier accuracies, we report the Fréchet Inception Distance (FID) (Heusel et al., 2017).

2.5. Implementation details

Diffusion model code was implemented with PyTorch 1.12 (Paszke et al., 2019), and fastai (Howard & Gugger, 2020) libraries. The CycleGANs were trained with the UPIT library (Abraham, 2023). FID evaluation was performed with pytorch-fid (mseitzer, 2023). All experiments were run on an NVIDIA A100 80GB.

3. Results

Figure 1 shows results on the test dataset applying either the DDIB or CycleGAN approach for virtual staining of MUSE images. While DDIB produces H&E-like images, it is apparent that while some aspects of the original MUSE image is preserved by DDIB (*green arrow*), many of the nuclei that are present in the original MUSE image either are missing or have an incorrect shape (*red arrow*). Additionally, many nuclei and other details are hallucinated, and not originally present in the MUSE Image. Figure 2 shows results on the test dataset applying either DDIB or CycleGAN approach for virtual staining of FIBI images. Here, it is apparent that DDIB misses many of the features in the original image, especially in the lower half of the image. In contrast, CycleGAN is able to preserve these features and appropriately convert them to H&E-like contrasts. Overall, this seems to indicate that while some semantic features are preserved in the latent encodings, many finer-grained details are not being preserved in the latent encoding or reconstructed by the diffusion model.

To quantitatively evaluate the model, the external classifier accuracy and Frechet Inception Distance (FID) is reported in Table 1 for the MUSE-to-H&E dataset, and Table 2 for the FIBI-to-H&E dataset. For both metrics, a lower score indicates better model performance. Although a different classifier with more layers is used, the accuracy is similar to (Abraham et al., 2020) for urothelial carcinoma MUSE-to-H&E, which reported 72%. As demonstrated in both Table 1 and Table 2, the diffusion model approach fails to outperform CycleGANs for virtual H&E staining.

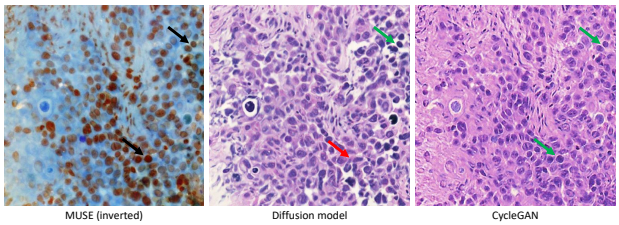

Figure 1. Comparison of a CycleGAN and a diffusion model for virtual staining of MUSE images (urothelial cell carcinoma). A red arrow indicates error, a green arrow indicates correct conversion.

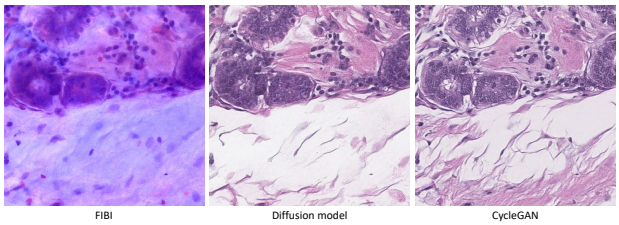

Figure 2. Comparison of a CycleGAN and a diffusion model for virtual staining of FIBI images (benign breast).

4. Discussion

In this study, virtual staining of SFM images with H&E using diffusion models was explored and compared to GANs. It was demonstrated both qualitatively and quantitatively that the performance of a diffusion model approach (DDIB) was inferior to a standard GAN approach (CycleGAN). To the author's best knowledge, this serves as the first exploration of diffusion models for virtual staining in the literature.

CycleGANs have been explored for a variety of virtual staining use-cases (Abraham et al., 2020; Kang et al., 2022; Combalia et al., 2019). Although CycleGANs are a fairly basic approach without many additional constraints or specialized loss functions, studies have observed CycleGANs outperform other, often more sophisticated GAN-based approaches for histopathology tasks (Zingman et al., 2023; Altini et al., 2023; Salido et al., 2023). Here, we observe a similar phenomenon even with a diffusion model-based approach.

The results here suggest that the inherent latent encodings provided by diffusion models and DDIM inversion does not appropriately maintain the semantic information of the original image. Future work will therefore further explore the incorporation of semantic constraints, likely in a form

Table 1. Urothelial carcinoma MUSE-to-H&E

Model	External critic accuracy ↓	FID ↓
CycleGAN	75%	97.2
DDIB	100%	127.0

Table 2. Benign breast FIBI-to-H&E

Model	External critic accuracy ↓	FID ↓
CycleGAN	77%	61.7
DDIB	100%	151.0

similar to classifier guidance (Dhariwal & Nichol, 2021). Regularization using self-supervised vision transformer features during the reverse diffusion process (Kwon & Ye, 2023) was attempted with no improvement in performance (data not shown), but other constraints, especially those designed specifically for histology images, may achieve more success.

Note that even if diffusion models manage to outperform existing GAN-based approaches, high latency may hinder the practical applications of diffusion models for virtual staining, where real-time feedback to clinicians is desired. CycleGANs require four neural function evaluations (NFE) per optimization step during training, but only one NFE for inference. In contrast, diffusion models require only one NFE per optimization step during training but numerous NFEs (N=150 used here) for inference. With DDIB specifically, 150 timesteps are performed for DDIM inversion to obtain the latent encoding, followed by an additional 150 timesteps for output image generation. For this reason, inference takes about 100 seconds for DDIB compared to 1 second with a CycleGAN for a 512x512 tile on an NVIDIA A100. Distillation approaches (Meng et al., 2022; Song et al., 2023) may help improve inference times and enable practical application of diffusion models for virtual staining.

A major challenge with training unpaired image-to-image translation models is the lack of adequate quantitative metrics. Most approaches rely on crowdsourcing approaches like Amazon Mechanical Turk to rate the quality of the generated images. However, this is infeasible when attempting to generate difficult-to-interpret domain-specific images. Some virtual staining tasks utilize paired data (Rivenson et al., 2019a;b), allowing for direct comparison to ground truth with metrics like PSNR and SSIM (Wang et al., 2004). Unfortunately, SFM images pixel-paired to H&E images are not available. The use of FID and an external classifier provides some relevant feedback, but doesn't fully capture the presence of lack of semantic preservation. Manual inspection of generated samples compared to the input images still remains one of the best forms of model evaluation.

While the results demonstrated here demonstrate that DDIB fails to outperform CycleGANs, we believe alternative diffusion model-based approaches may eventually surpass GANs. Additionally, it may be possible that combining GANs and diffusion models (i.e. incorporating some form of adversarial training) may result in better results. Therefore, we believe applying diffusion models to virtual staining is still a fruitful direction and requires further research.

References

Abraham, T., Shaw, A., O'Connor, D., Todd, A., and Levenson, R. Slide-free MUSE Microscopy to H&E Histology Modality Conversion via Unpaired Image-to-Image Translation GAN Models. *ArXiv*, 2020.

Abraham, T. M. UPIT - A fastai/PyTorch package for unpaired image-to-image translation., April 2023. URL https://github.com/tmabraham/UPIT.

Altini, N., Marvulli, T. M., Zito, F. A., Caputo, M., Tommasi, S., Azzariti, A., Brunetti, A., Prencipe, B., Mattioli, E., De Summa, S., and Bevilacqua, V. The role of unpaired image-to-image translation for stain color normalization in colorectal cancer histology classification. *Computer Methods and Programs in Biomedicine*, 234:107511, June 2023. ISSN 0169-2607. doi: 10.1016/j.cmpb.2023.107511. URL https://www.sciencedirect.com/science/article/pii/S0169260723001761.

Bai, B., Yang, X., Li, Y., Zhang, Y., Pillar, N., and Ozcan, A. Deep learning-enabled virtual histological staining of biological samples. *Light: Science & Applications*, 12(1):57, March 2023. ISSN 2047-7538. doi: 10.1038/s41377-023-01104-7. URL https://www.nature.com/articles/s41377-023-01104-7. Number: 1 Publisher: Nature Publishing Group.

Bancroft, J. D. and Gamble, M. *Theory and Practice of Histological Techniques*. Churchill Livingstone, Edinburgh, sixth edition edition, 2008. ISBN 978-0-443-10279-0. doi: https://doi.org/10.1016/B978-0-443-10279-0.50001-4. URL https://www.sciencedirect.com/science/article/pii/B9780443102790500014.

Borowsky, A. D., Levenson, R. M., Gown, A. M., Morningstar, T., Fleury, T. A., Henderson, G., Schaberg, K., Sybenga, A. B., Glassy, E. F., Taylor, S. L., and Fereidouni, F. A Pilot Validation Study Comparing Fluorescence-Imitating Brightfield Imaging, A Slide-Free Imaging Method, With Standard Formalin-Fixed, Paraffin-Embedded Hematoxylin-Eosin–Stained

Tissue Section Histology for Primary Surgical Pathology Diagnosis. *Archives of Pathology & Laboratory Medicine*, May 2023. ISSN 0003-9985. doi: 10.5858/arpa.2022-0432-OA. URL https://doi.org/10.5858/arpa.2022-0432-OA.

Combalia, M., Pérez-Anker, J., García-Herrera, A., Alos, L., Vilaplana, V., Marqués, F., Puig, S., and Malvehy, J. Digitally Stained Confocal Microscopy through Deep Learning. In *International Conference on Medical Imaging with Deep Learning*, pp. 121–129. PMLR, May 2019. URL http://proceedings.mlr.press/v102/combalia19a.html. ISSN: 2640-3498.

Dhariwal, P. and Nichol, A. Diffusion Models Beat GANs on Image Synthesis. Technical Report arXiv:2105.05233, arXiv, June 2021. URL http://arxiv.org/abs/2105.05233. arXiv:2105.05233 [cs, stat] type: article.

Fereidouni, F., Harmany, Z. T., Tian, M., Todd, A., Kintner, J. A., McPherson, J. D., Borowsky, A. D., Bishop, J., Lechpammer, M., Demos, S. G., and Levenson, R. Microscopy with ultraviolet surface excitation for rapid slide-free histology. *Nature Biomedical Engineering*, 1(12):957–966, December 2017. ISSN 2157-846X. doi: 10.1038/s41551-017-0165-y. URL https://www.nature.com/articles/s41551-017-0165-y. Number: 12 Publisher: Nature Publishing Group.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://papers.nips.cc/paper/2017/hash/8a1d694707eb0fefe65871369074926d-Abstract.html.

Ho, J., Jain, A., and Abbeel, P. Denoising Diffusion Probabilistic Models, December 2020. URL http://arxiv.org/abs/2006.11239. arXiv:2006.11239 [cs, stat].

Howard, J. and Gugger, S. Fastai: A Layered API for Deep Learning. *Information*, 11(2):108, February 2020. doi: 10.3390/info11020108. URL https://www.mdpi.com/2078-2489/11/2/108. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. Image-to-Image Translation with Conditional Adversarial Networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5967–5976, July 2017. doi: 10.1109/CVPR.2017.632. ISSN: 1063-6919.

Kang, L., Li, X., Zhang, Y., and Wong, T. T. W. Deep learning enables ultraviolet photoacoustic microscopy based histological imaging with near real-time virtual staining. *Photoacoustics*, 25:100308, March 2022. ISSN 2213-5979. doi: 10.1016/j.pacs.2021.100308. URL https://www.sciencedirect.com/science/article/pii/S2213597921000689.

Kwon, G. and Ye, J. C. Diffusion-based image translation using disentangled style and content representation, 2023.

Liu, Y., Levenson, R. M., and Jenkins, M. W. Slide Over: Advances in Slide-Free Optical Microscopy as Drivers of Diagnostic Pathology. *The American Journal of Pathology*, 192(2):180–194, February 2022. ISSN 0002-9440. doi: 10.1016/j.ajpath.2021.10.010. URL https://www.sciencedirect.com/science/article/pii/S0002944021004715.

Meng, C., Gao, R., Kingma, D. P., Ermon, S., Ho, J., and Salimans, T. On Distillation of Guided Diffusion Models, October 2022. URL http://arxiv.org/abs/2210.03142. arXiv:2210.03142 [cs].

mseitzer. FID score for PyTorch, May 2023. URL https://github.com/mseitzer/pytorch-fid. original-date: 2018-02-10T12:33:06Z.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F. d., Fox, E., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems 32*, pp. 8026–8037. Curran Associates, Inc., 2019. URL http://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance.pdf.

Rivenson, Y., Liu, T., Wei, Z., Zhang, Y., de Haan, K., and Ozcan, A. PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light: Science & Applications*, 8(1):1–11, February 2019a. ISSN 2047-7538. doi: 10.1038/s41377-019-0129-y. URL https://www.nature.com/articles/s41377-019-0129-y. Number: 1 Publisher: Nature Publishing Group.

Rivenson, Y., Wang, H., Wei, Z., de Haan, K., Zhang, Y., Wu, Y., Günaydın, H., Zuckerman, J. E., Chong, T., Sisk, A. E., Westbrook, L. M., Wallace, W. D., and Ozcan, A. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nature Biomedical Engineering*, 3(6):466–477, June 2019b. ISSN 2157-846X. doi: 10.1038/s41551-019-0362-y. URL https://www.nature.

com/articles/s41551-019-0362-y. Number: 6 Publisher: Nature Publishing Group.

Ronneberger, O., Fischer, P., and Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F. (eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lecture Notes in Computer Science, pp. 234–241, Cham, 2015. Springer International Publishing. ISBN 978-3-319-24574-4. doi: 10.1007/978-3-319-24574-4_28.

Salido, J., Vallez, N., González-López, L., Deniz, O., and Bueno, G. Comparison of deep learning models for digital H&E staining from unpaired label-free multispectral microscopy images. *Computer Methods and Programs in Biomedicine*, 235:107528, June 2023. ISSN 0169-2607. doi: 10.1016/j.cmpb.2023.107528. URL https://www.sciencedirect.com/science/article/pii/S0169260723001931.

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. Score-Based Generative Modeling through Stochastic Differential Equations, February 2021. URL http://arxiv.org/abs/2011.13456. arXiv:2011.13456 [cs, stat].

Song, Y., Dhariwal, P., Chen, M., and Sutskever, I. Consistency Models. March 2023. URL https://arxiv.org/abs/2303.01469v1.

Su, X., Song, J., Meng, C., and Ermon, S. Dual Diffusion Implicit Bridges for Image-to-Image Translation, March 2023. URL http://arxiv.org/abs/2203.08382. arXiv:2203.08382 [cs].

Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13 (4):600–612, April 2004. ISSN 1941-0042. doi: 10.1109/TIP.2003.819861. Conference Name: IEEE Transactions on Image Processing.

Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. *arXiv:1703.10593 [cs]*, August 2020. URL http://arxiv.org/abs/1703.10593. arXiv: 1703.10593.

Zingman, I., Frayle, S., Tankoyeu, I., Sukhanov, S., and Heinemann, F. A comparative evaluation of image-to-image translation methods for stain transfer in histopathology, April 2023. URL http://arxiv.org/abs/2303.17009. arXiv:2303.17009 [cs, eess].