

Virtual 3D H&E Staining from Phase-contrast Back-illumination Interference Tomography

Anthony A. Song¹, Boyan Zhou¹, Mayank Golhar¹, Marisa Morakis¹,
Alexander Baras², and Nicholas J. Durr¹

¹ Department of Biomedical Engineering, Johns Hopkins University,
Baltimore, MD, USA

asong18@jhu.edu, ndurr@jhu.edu

² Department of Pathology, Johns Hopkins Hospital,
Baltimore, MD, USA

Abstract. Three-dimensional (3D) histopathology of unprocessed tissues has the potential to transform disease management by enabling volumetric characterization of tissue microarchitecture and in-vivo assessment. Back-illumination Interference Tomography (BIT) is a new phase microscopy technology that provides rapid, non-destructive volumetric imaging of unprocessed tissues. However, translating BIT volumes into clinically interpretable H&E images remains challenging, particularly due to shift-variant contrast and the absence of quantitative validation benchmarks. We introduce **HistoBIT3D**, the first voxel-wise paired BIT and fluorescence-labeled nuclei dataset, enabling quantitative evaluation of structural preservation in unsupervised virtual staining against ground-truth nuclear distributions. Using this dataset, we present a novel virtual staining framework that translates BIT volumes with shift-variant contrast into realistic H&E volumes by leveraging bidirectional multiscale content consistency and cross-domain style reuse to enhance structural fidelity and perceptual realism. Our method achieves state-of-the-art realism metrics while significantly improving 3D nuclei segmentation accuracy and boundary preservation under zero-shot Cellpose evaluation. Together, these contributions establish a quantitatively validated, structurally faithful, and scalable pipeline for 3D virtual H&E staining, advancing the paradigm of slide-free, volumetric computational histopathology. Our data and code are available at: https://github.com/aasong113/HistoBIT3D_VirtualStaining.

Keywords: Virtual staining · Computational Pathology · Phase microscopy · 3D Dataset · Multimodal dataset · Generative models.

1 Introduction

Histopathology is the gold standard for identifying structural and cellular changes that differentiate normal from diseased tissue. The current workflow uses hematoxylin and eosin (H&E) staining of sliced two-dimensional (2D) tissue samples and analysis under bright-field microscopy. However, 2D tissue samples only give

a small glimpse of the complex three-dimensional (3D) morphological structures. A shift towards 3D pathology may allow for better characterization of tissue microstructures for downstream pathological understanding. 3D pathology has advanced the study of pancreatic tumorigenesis and cardiovascular microanatomy using serially sectioned H&E reconstruction (CODA)[12, 11], enabled clinical risk assessment in prostate cancer and Barrett esophagus via open-top light-sheet microscopy[18, 16], and supported skin-biopsy analysis using two-photon fluorescence microscopy[4]. However, current 3D histopathological techniques require complex instrumentation and labor-intensive sample preparation, including sectioning, staining, or optical clearing, making them difficult to deploy for real-time intraoperative analysis and incompatible with in-vivo imaging. To address these limitations, approaches based on intrinsic phase contrast, such as optical coherence tomography (OCT) [23], oblique back-illumination microscopy (OBM) [7], and back-illumination interference tomography (BIT) [14], have emerged as label-free 3D imaging modalities.

A major challenge with pathology applications of phase microscopy is that the contrast is fundamentally different from that of H&E. Phase microscopy maps refractive index gradients, while H&E captures the distribution of color stains mediated by chemical and molecular interactions. Generative adversarial networks (GANs) have been developed to generate "virtually stained" H&E-like images from other microscopy techniques[1, 23, 2, 19]. Initially, virtual staining employed supervised learning strategies and was shown to successfully reconstruct histological features across the input and target domains[15, 17]. However, this requires precisely aligned supervised image pairs from both imaging modalities, which is often impractical due to tissue degradation, staining quality, and hardware limitations. To overcome this limitation, CycleGAN [27] has become a popular unsupervised framework that allows virtual staining of thick, unstained tissue using corresponding unpaired H&E images [25, 1, 19, 23]. Cycle consistency encourages preservation of spatial structure but often fails to achieve precise pixel-wise alignment during translation. Extensions to CycleGAN have addressed this limitation by incorporating saliency consistency constraints [13] and feature map information consistency losses [24], improving structural fidelity in virtual staining tasks. However, these approaches have primarily been validated on modalities exhibiting shift-invariant contrast, such as fluorescence and differential interference contrast. BIT images, however, have shift-variant contrast that changes with the specimen's position relative to the microscope. Furthermore, there are currently no methods for quantitatively benchmarking the results of unsupervised 3D virtual H&E staining. Qualitative metrics like Fréchet Inception Distance (FID) and Kernel Inception Distance (KID) provide little insight into content preservation during image translation.

To address these concerns, we introduce HistoBIT3D, the first voxel-wise registered 3D dataset of BIT and fluorescence-labeled nuclei, and an innovative GAN framework for realistic, content-preserving virtual H&E staining of BIT volumes. By introducing a 3D quantitative validation dataset with a novel architecture focused on content preservation and realism, we establish a foundation

for 3D, unsupervised, fast, simple, stain-free, high-fidelity pathology imaging. The three main contributions are summarized as follows:

(1) **The first voxel-wise paired 3D Back-illumination Interference Tomography (BIT) and fluorescent nuclei dataset.** This dataset enables quantitative validation of structural preservation in 3D virtual staining against ground-truth nuclear distributions.

(2) **A novel virtual staining framework for translating BIT volumes with shift-variant contrast into virtual H&E volumes.** Our model incorporates a bidirectional multiscale content consistency loss that aligns hierarchical features across translation cycles to enforce structural fidelity.

(3) **Enhanced perceptual realism through efficient cross-domain style reuse.** We leverage AdaIN-based injection of learned H&E style features into the BIT-to-H&E generator to improve visual fidelity.

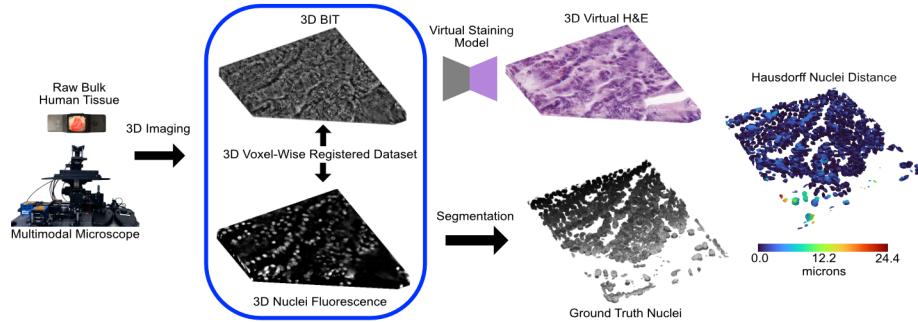


Fig. 1. Overview of the 3D virtual staining pipeline for Back-illumination Interference Tomography (BIT). From left to right: optical setup for volumetric acquisition of voxel-wise paired BIT and fluorescence nuclei data from bulk human tissue; virtual H&E generation using our GAN-based framework (top row); 3D fluorescence nuclei segmentation results (bottom row); and quantitative evaluation via Hausdorff distance comparing segmented nuclei from virtual H&E with ground-truth fluorescence volumes.

2 Methodology

2.1 Multi-modal BIT and Fluorescence 3D Pathology Dataset

We introduce HistoBIT3D, a multimodal slide-free tissue dataset combining Back-illumination Interference Tomography (BIT) with voxel-wise paired 3D fluorescence imaging of nuclei and cytoplasmic structures, alongside unpaired 2D FFPE H&E images for virtual staining. To our knowledge, this is the first dataset providing volumetric fluorescence ground truth registered to label-free imaging, enabling quantitative benchmarking of structural preservation, particularly nuclear morphology, in unsupervised virtual staining. HistoBIT3D includes paired

BIT–nuclear fluorescence duodenum volumes with corresponding unpaired 2D H&E, additional 3D BIT-only duodenum datasets (crypts and muscularis) with matched 2D H&E, and a 2D normal and cancer kidney dataset for cross-tissue generalization, with each subset comprising approximately 5,000 images of size 512×512 . This study was approved by the Johns Hopkins Medicine Institutional Review Board (IRB) and all participants provided informed consent.

BIT imaging was performed using 660 nm LED back-illumination through the objective back aperture to generate transmission-like phase contrast from interference between diffracted and backscattered light near the focal plane. Registered fluorescence images of bulk tissue stained with Hoechst (nuclei) and eosin (cytoplasm) were acquired through the same optical path and camera under oblique UV excitation [6]. Volumetric stacks spanning 30–50 μm were collected for both modalities. BIT images were background-subtracted ($\sigma = 30$) and scaled to 8-bit for training. To reduce shift-variant contrast due to focal-plane dependence, the network input consisted of a three-channel stack (original, inverted, original) to emphasize structural over intensity information. Fluorescence volumes were processed with 3D PSF deconvolution to suppress bulk scattering.

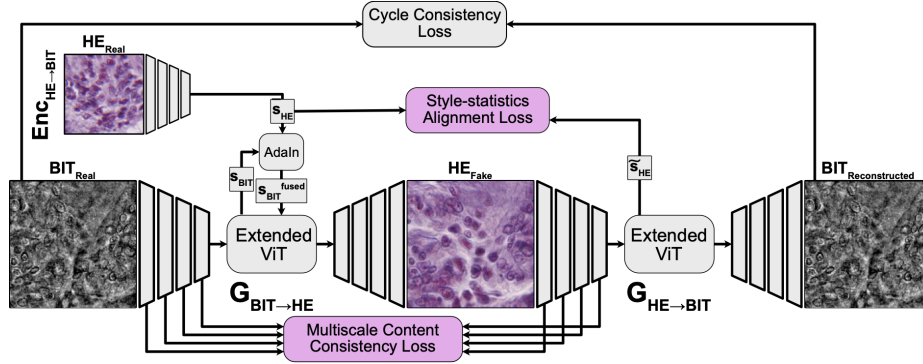


Fig. 2. Architecture showing bidirectional multiscale content consistency for structurally faithful BIT-to-H&E translation and AdaIN-based cross-domain style injection for enhanced realism.

2.2 Vision Transformer CycleGAN

Prior work has demonstrated that CycleGAN-based architectures [1], including variants with spatial consistency constraints [13, 24], are effective for virtual H&E staining of label-free phase-contrast data due to their ability to preserve feature correspondence through cycle consistency while translating domain-specific appearance. However, biomedical imaging demands preservation of global tissue architecture beyond local texture realism, motivating the incorporation of non-local pattern learning mechanisms [5]. Vision transformers (ViTs) have shown

improved modeling of long-range structural dependencies, such as relative nuclear arrangement and glandular organization, in virtual staining tasks [9]. To capture global context, we build off UVCGANv2 [21], which extends the classical CycleGAN framework with a ViT-based generator architecture. We enforce structural preservation using multiscale feature maps from the U-Net encoder together with spatial tokens from the ViT bottleneck as a content-consistency objective. Furthermore, UVCGANv2’s extended transformer design disentangles spatial and style representations, enabling source-driven style modulation through the reuse of target-domain style tokens to enhance staining realism while maintaining geometric fidelity. In this training paradigm, our model performs 2D-to-2D virtual staining of an axial stack, enabling 3D histopathology.

2.3 Multiscale Content Consistency Loss

To preserve content fidelity between virtually stained outputs and input BIT images, the learned feature representations must retain precise spatial information from the input domain [24]. In virtual staining, this is particularly important because the desired transformation is primarily a domain/style change, while underlying tissue morphology should remain unchanged. Motivated by hierarchical representation learning in modern vision backbones [3], we introduce a bidirectional multiscale content consistency loss that aligns intermediate generator features across both translation cycles. An additional motivation is shift-variant contrast in BIT imaging: nuclei can appear dark or light depending on focal-plane position, even when their underlying morphology is unchanged.

Our model consists of two generators $G_{\text{BIT} \rightarrow \text{H\&E}}$ and $G_{\text{H\&E} \rightarrow \text{BIT}}$ operating on inputs $x_{\text{BIT}} \in R^{H \times W \times 3}$ and $y_{\text{H\&E}} \in R^{H \times W \times 3}$. Let $\phi_k^{\text{BIT} \rightarrow \text{H\&E}}(\cdot)$ and $\phi_k^{\text{H\&E} \rightarrow \text{BIT}}(\cdot)$ denote the feature maps extracted from the k -th encoder stage of the respective generators, corresponding to downsampling scales $k \in \{1, 2, 4, 16\}$. The full-resolution map ($k = 1$) preserves fine nuclear boundaries and microtexture, intermediate scales encode mid-level morphology, and the ViT bottleneck $k = 16$ captures global tissue organization, stabilizing histologic context despite focal-plane-dependent intensity shifts. To reduce memory overhead, we sample a subset of channels $\mathcal{C}_s$ at later scales. We then enforce bidirectional feature consistency:

$$\mathcal{L}_{\text{BIT}} = \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} \left\| \phi_{k, \mathcal{C}_s}^{\text{BIT, H\&E}}(x_{\text{BIT}}) - \text{sg} \left(\phi_{k, \mathcal{C}_s}^{\text{H\&E, BIT}}(G_{\text{BIT} \rightarrow \text{H\&E}}(x_{\text{BIT}})) \right) \right\|_1, \quad (1)$$

$$\mathcal{L}_{\text{H\&E}} = \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} \left\| \phi_{k, \mathcal{C}_s}^{\text{H\&E, BIT}}(y_{\text{H\&E}}) - \text{sg} \left(\phi_{k, \mathcal{C}_s}^{\text{BIT, H\&E}}(G_{\text{H\&E} \rightarrow \text{BIT}}(y_{\text{H\&E}})) \right) \right\|_1. \quad (2)$$

The multiscale consistency term is $\mathcal{L}_{\text{msc}} = \mathcal{L}_{\text{BIT}} + \mathcal{L}_{\text{H\&E}}$, $\mathcal{L} = \lambda_{\text{msc}} \mathcal{L}_{\text{msc}}$. Here, $\text{sg}(\cdot)$ denotes the stop-gradient operator, which treats one branch as a fixed target while updating the opposite branch to improve optimization stability.

2.4 Cross-Domain Style Fusion

UVCGANv2 promotes explicit style-content disentanglement at the transformer bottleneck by separating global appearance information from spatially structured content features, enabling controlled style modulation via Adaptive Instance Normalization (AdaIN)[8]. In this formulation, the injected bottleneck token is interpreted as a global style token rather than a spatial content token, allowing appearance statistics to be modulated independently of tissue geometry. During reverse passes $G_{\text{H\&E} \rightarrow \text{BIT}}$ on real H&E samples $y_{\text{H\&E}}$, we extract the final extra token $s_{\text{H\&E}}$, representing the current H&E style embedding at the ViT bottleneck. We update a running-mean prototype $\bar{s}_{\text{H\&E}}$ via exponential moving average with factor α . This prototype provides a stable estimate of the global H&E style distribution, is checkpointed during training, and reused at inference to ensure consistent cross-dataset staining.

For forward passes $G_{\text{BIT} \rightarrow \text{H\&E}}$ with input x_{BIT} , we fuse the BIT style token s_{BIT} with $\bar{s}_{\text{H\&E}}$ via Adaptive Instance Normalization (AdaIN) [8] in token space:

$$\tilde{s}_{\text{BIT}} = \text{AdaIN}(s_{\text{BIT}}, \bar{s}_{\text{H\&E}}) = \sigma(\bar{s}_{\text{H\&E}}) \frac{s_{\text{BIT}} - \mu(s_{\text{BIT}})}{\sigma(s_{\text{BIT}})} + \mu(\bar{s}_{\text{H\&E}}), \quad (3)$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ compute statistics across the embedding dimension. To gradually reduce style injection during training, the fusion weight follows a cosine decay schedule. This mechanism enables $\text{BIT} \rightarrow \text{H\&E}$ synthesis to retain a stable H&E-domain style signature without requiring paired H&E images at inference time [10].

To further align H&E-domain style distributions, we introduce a bottleneck style-statistics loss. Let $s_{\text{H\&E}}^r$ denote the style token from $G_{\text{H\&E} \rightarrow \text{BIT}}(y_{\text{H\&E}})$, cached with stop-gradient, and $s_{\text{H\&E}}^f$ the token from the synthesized pathway $G_{\text{H\&E} \rightarrow \text{BIT}}(G_{\text{BIT} \rightarrow \text{H\&E}}(x_{\text{BIT}}))$. We match first- and second-order statistics between real and generated style tokens via

$$\mathcal{L}_{\text{style}} = \|\mu(s_{\text{H\&E}}^f) - \text{sg}(\mu(s_{\text{H\&E}}^r))\|_2^2 + \|\sigma(s_{\text{H\&E}}^f) - \text{sg}(\sigma(s_{\text{H\&E}}^r))\|_2^2, \quad (4)$$

This encourages the generated branch to match the real H&E style distribution while restricting gradient flow to the synthesized pathway [8].

3 Experiments and Results

3.1 Implementation Details

We pretrained the generator backbone as a joint masked autoencoder on unpaired BIT and FFPE H&E images for 50 epochs, with a batch size of 4. The full model was trained on BIT input and the FFPE-HE target style domain, using 60% of the 512×512 patches from our HistoBIT3D dataset with a batch size of 2 on a single NVIDIA GeForce RTX 3090 GPU for 170 epochs. Loss weights are $\lambda_{\text{cycle}} = 10$, $\lambda_{\text{idt}} = 0.5$, $\lambda_{\text{msc}} = 1.0$, and $\lambda_{\text{style}} = 1.0$.

3.2 Realism and 3D content preservation

To validate realistic and content-preserving stain transfer, we compared our model with state-of-the-art virtual staining methods, including CycleGAN [27], STABLE [24], CycleDiffusion [22], and baseline UVCGANv2 (Base)[21]. We selected widely used unsupervised image translation metrics, FID and KID, to benchmark image quality and similarity. To assess the accuracy of virtual staining in 3D, we used zero-shot Cellpose [20] to segment nuclei in our voxel-wise-registered fluorescent and virtual H&E volumes ($130\mu m \times 83\mu m \times 20\mu m$). The fluorescent labeled nuclei volume acts as our ground truth nuclei locations, as Hoechst[6] is a well-known nucleic acid stain. We characterized the fidelity of our virtually stained 3D volume with conventional 3D segmentation metrics-foreground voxel-level 3D DICE, population-level 95th percentile Hausdorff Distance (HD95) on global foreground boundaries, and mean per-instance nuclei volume as a pathology-based morphological metric. We calculate 3D metrics by first translating our 2D masks into a 3D volume using u-Segment3D[26].

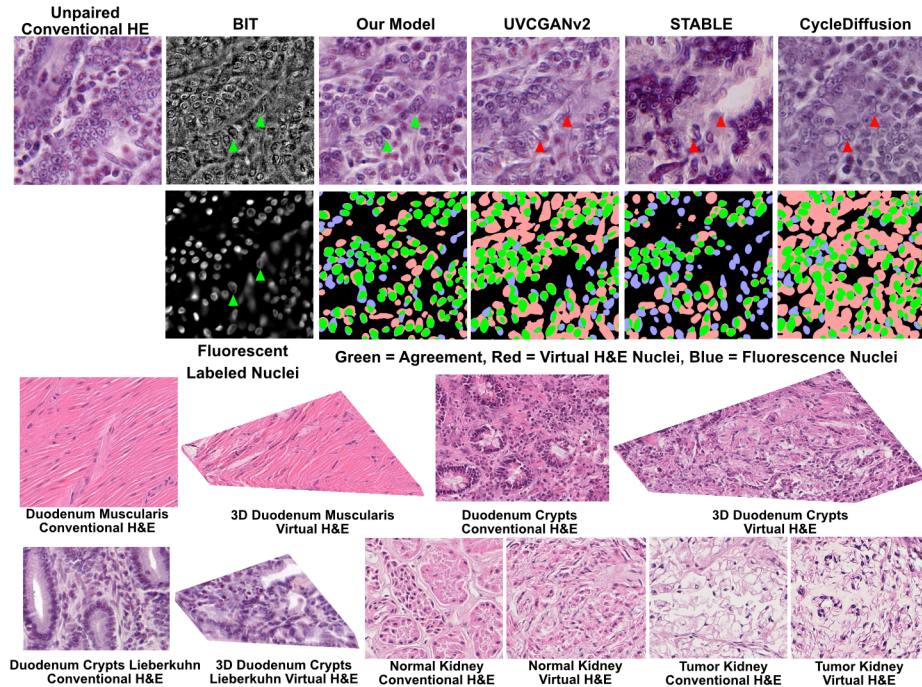


Fig. 3. Row 1: Comparison of virtual H&E results from our method and baseline models against unpaired conventional H&E. **Row 2:** Zero-shot Cellpose [20] segmentation of virtual H&E images compared with ground-truth nuclei from fluorescence imaging. Green arrows highlight nuclei whose structural content is preserved. **Rows 3–4:** Diverse tissue samples from the HistoBIT3D dataset, virtually stained from BIT into H&E.

Table 1. Comparison with SOTA and ablation study for realism and fidelity.

	FID↓	KID↓	3D DICE ↑	Nuc. Volume (μm^3)	HD95 (μm) ↓
CycleGAN	136.01	0.1556	0.360	377.2	5.59
STABLE[24]	86.54	0.0684	0.418	331.9	5.38
CycleDiffusion[22]	106.83	0.1138	0.359	233.2	4.77
UVCANv2[21]	95.80	0.0821	0.515	386.7	4.67
Base+Style	<u>62.45</u>	<u>0.0434</u>	<u>0.583</u>	<u>416.6</u>	<u>4.22</u>
Our Model	60.69	0.0417	0.594	408.3	4.04

***Bold** and underlined indicate best and second best results, respectively.

Ground Truth Nuclei Volume is **405.6** μm^3 .

The performance of our framework is demonstrated qualitatively in Fig.3 and quantitatively summarized in Table1. Row 1 of Fig.3 qualitatively compares our model with baseline approaches and conventional unpaired H&E, which can be quantitatively validated by our model’s FID (60.69) and KID (0.0417) outperforming baseline methods. Our model produces faithful translations of the structural features observed in the input BIT image, as confirmed by alignment with the corresponding fluorescent nuclei ground truth. Row 2 of Fig.3 demonstrates improved zero-shot Cellpose segmentation for our model, reflected by higher 3D Dice (0.594), nuclei volume consistent with ground truth fluorescence nuclei volume (408.3 μm^3 vs. 405.6 μm^3), and lower HD95 (4.04 μm), indicating superior boundary preservation. The success of zero-shot segmentation further suggests that the generated H&E images are sufficiently realistic to be reliably recognized by Cellpose. Rows 3–4 of Fig. 3 present multiple virtually stained samples from the HistoBIT3D dataset, including duodenum tissue, normal kidney, and tumor kidney.

To evaluate the impact of cross-domain style fusion and multiscale content consistency, we conducted an ablation study starting from the UVCANv2 base model and incrementally adding components: style injection and, finally, the full multiscale consistency loss (Our Model). We observe a substantial improvement in realism and 3D segmentation metrics with the addition of cross-domain style fusion, indicating higher-fidelity translation compared to the baseline, with further gains when combined with the multiscale content consistency loss.

4 Conclusions

In this work, we advance 3D virtual histopathology by introducing HistoBIT3D, the first voxel-wise paired 3D BIT–fluorescence dataset, enabling quantitative validation of structural preservation against ground-truth nuclear distributions. We further propose a virtual staining framework that enforces structural integrity through bidirectional multiscale feature alignment and enhances perceptual realism via cross-domain style fusion. Future work will address contrast discrepancies between fluorescence and BIT to improve feature correspondence, incorporate fluorescence-guided semi-supervision, and refine downstream segmentation with task-specific tuning. To foster continued progress, HistoBIT3D

will be made publicly available. Together, these contributions establish a validated and structurally faithful pipeline for 3D virtual H&E staining, promoting the adoption of slide-free, volumetric computational histopathology.

Acknowledgments. Support by the Leon Troper Professorship in Computational Pathology.

Disclosure of Interests. NJD is a co-inventor on a provisional patent application assigned to Johns Hopkins University and Boston University. He may be entitled to future royalties from intellectual property related to the technologies described in this article.

References

1. Abraham, T.M., Casteleiro Costa, P., Serafini, C.E., Guang, Z., Zhang, Z., Neill, S., Olson, J.J., Levenson, R., Robles, F.E.: Label- and slide-free tissue histology using 3d epi-mode quantitative phase imaging and virtual h&e staining. *Optica* **10**(12), 1605–1616 (2023). <https://doi.org/10.1364/OPTICA.485994>
2. Almagro-Pérez, C., Peruzzi, N., Galambos, C., Song, A.H., Brunnström, H., Gawlik, K.I., Stampanoni, M., Tran-Lundmark, K., Lovric, G.: Histology-guided 3d virtual staining of microct-imaged lung tissue via deep learning. *bioRxiv preprint* (2025)
3. Chen, R.J., Chen, C., Li, Y., Chen, T.Y., Trister, A.D., Krishnan, R.G., Mahmood, F.: Scaling vision transformers to gigapixel images via hierarchical self-supervised learning. *arXiv preprint arXiv:2206.02647* (2022)
4. Ching-Roa, V.D., Huang, C.Z., Ibrahim, S.F., Smoller, B.R., Giacomelli, M.G.: Real-time analysis of skin biopsy specimens with 2-photon fluorescence microscopy. *JAMA Dermatology* **158**(10), 1175–1182 (2022). <https://doi.org/10.1001/jamadermatol.2022.3628>
5. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* (2020)
6. 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., Levenson, R.: Microscopy with ultraviolet surface excitation for rapid slide-free histology. *Nature Biomedical Engineering* **1**, 957–966 (2017). <https://doi.org/10.1038/s41551-017-0165-y>
7. Ford, T.N., Chu, K.K., Mertz, J.: Phase-gradient microscopy in thick tissue with oblique back-illumination. *Nature Methods* **9**(12), 1195–1197 (2012). <https://doi.org/10.1038/nmeth.2219>
8. Huang, X., Belongie, S.: Arbitrary style transfer in real-time with adaptive instance normalization. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. pp. 1501–1510 (2017)
9. Hussain, M.A., et al.: Vit-stain: Vision transformer-driven virtual staining for skin histopathology via global contextual learning. *PLOS ONE* **21**(2), e0341311 (2026). <https://doi.org/10.1371/journal.pone.0341311>

10. Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4401–4410 (2019)
11. Kathiriya, I.S., et al.: A disrupted compartment boundary underlies abnormal cardiac patterning and congenital heart defects. *Nature Cardiovascular Research* **5**, 67–83 (2026). <https://doi.org/10.1038/s44161-025-00755-6>
12. Kiemen, A.L., Braxton, A.M., Grahm, M.P., Han, K.S., Babu, J.M., Reichel, R., Jiang, A.C., Kim, B., Hsu, J., Amoa, F., Reddy, S., Hong, S.M., Cornish, T.C., Thompson, E.D., Huang, P., Wood, L.D., Hruban, R.H., Wirtz, D., Wu, P.H.: Coda: quantitative three-dimensional reconstruction of large tissues at cellular resolution. *Nature Methods* **19**(11), 1490–1499 (2022). <https://doi.org/10.1038/s41592-022-01650-9>
13. Li, X., Zhang, G., Qiao, H., Bao, F., Deng, Y., Wu, J., He, Y., Yun, J., Lin, X., Xie, H., Wang, H., Dai, Q.: Unsupervised content-preserving transformation for optical microscopy. *Light: Science & Applications* **10**, 44 (2021). <https://doi.org/10.1038/s41377-021-00484-y>
14. McKay, G.N., Mertz, J., Durr, N.J.: Back-illumination interference tomography for imaging weak scattering in thick tissues. *Optica* **12**(12), 1865–1872 (2025). <https://doi.org/10.1364/OPTICA.567543>
15. Park, J., et al.: Revealing 3d microanatomical structures of unlabeled thick cancer tissues using holotomography and virtual h&e staining. *Nature Communications* **16**, 4781 (2025). <https://doi.org/10.1038/s41467-025-19440-1>
16. Reddi, D.M., Barner, L.A., Burke, W., Gao, G., Grady, W.M., Liu, J.T.C.: Nondestructive 3d pathology image atlas of barrett esophagus with open-top light-sheet microscopy. *Archives of Pathology & Laboratory Medicine* **147**(10), 1164–1171 (2023). <https://doi.org/10.5858/arpa.2022-0133-OA>
17. Rivenson, Y., Wang, H., Wei, Z., Zhang, Y., Wu, T., Gunaydin, H., Zuckerman, J.E., Chong, T., Zhou, Z., Ozcan, A.: Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nature Biomedical Engineering* **3**, 466–477 (2019). <https://doi.org/10.1038/s41551-019-0362-y>
18. Serafin, R., Koyuncu, C., Xie, W., Huang, H., Glaser, A.K., Reder, N.P., Janowczyk, A., True, L.D., Madabhushi, A., Liu, J.T.C.: Nondestructive 3d pathology with analysis of nuclear features for prostate cancer risk assessment. *The Journal of Pathology* **260**(4), 390–401 (2023). <https://doi.org/10.1002/path.6090>
19. Song, A.A., Golhar, M., Morakis, M., Abdulrahim, M., Jayawardane, S., McKay, G.N., Xu, R., Baras, A., Durr, N.J.: Slide-free tissue histology using back-illumination interference tomography and virtual h&e staining. In: Optica Biophotonics Congress 2025, Technical Digest Series. p. paper NTh1C.3. Optica Publishing Group (2025)
20. Stringer, C., Wang, T., Michaelos, M., Pachitariu, M.: Cellpose: A generalist algorithm for cellular segmentation. *Nature Methods* **18**(1), 100–106 (2021). <https://doi.org/10.1038/s41592-020-01018-x>
21. Torbunov, D., Huang, Y., Tseng, H.H., Yu, H., Huang, J., Yoo, S., Lin, M., Viren, B., Ren, Y.: UVCAN v2: An improved cycle-consistent GAN for unpaired image-to-image translation. *arXiv preprint arXiv:2303.16280* (2023)
22. Wu, C.H., De la Torre, F.: Unifying diffusion models’ latent space, with applications to cyclediffusion and guidance (2022), *arXiv:2210.05559*
23. Yin, Z., et al.: Fast and label-free 3d virtual h&e histology via active phase modulation-assisted dynamic full-field optical coherence tomography. *npj Imaging* **3**, 12 (2025). <https://doi.org/10.1038/s44303-025-00068-0>

24. You, J.Y., Eom, M., Choi, T.I., Cho, E.S., Choi, J., Lee, M., Shin, C., Moon, J., Kim, E., Kim, P., Kim, C.H., Yoon, Y.G.: Preserving spatial and quantitative information in unpaired biomedical image-to-image translation. *Cell Reports Methods* **5**(6), 101074 (2025). <https://doi.org/10.1016/j.crmeth.2025.101074>
25. Zhang, Y., Kang, L., Wong, I.H.M., Dai, W., Li, X., Chan, R.C.K., Hsin, M.K.Y., Wong, T.T.W.: High-throughput, label-free and slide-free histological imaging by computational microscopy and unsupervised learning. *Advanced Science* **9**(2), e2102358 (2022). <https://doi.org/10.1002/advs.202102358>
26. Zhou, F.Y., Marin, Z., Yapp, C., et al.: Universal consensus 3d segmentation of cells from 2d segmented stacks. *Nature Methods* **22**, 2386–2399 (2025). <https://doi.org/10.1038/s41592-025-02887-w>
27. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. pp. 2242–2251 (2017)