Translating Light-Sheet Microscopy Images to Virtual H&E Using CycleGAN

Jana

January 11, 2026

Introduction

Light sheet microscopy enables the rapid volumetric imaging of biological tissues using fluorescence labelling. Although this technique preserves fine structural detail, histopathological interpretation is traditionally performed on haematoxylin and eosin (H&E)-stained images. Overcoming this gap between the two modalities can significantly improve interpretability for pathologists.

The available fluorescence channels are:

- TO-PRO-3 (C01): Nuclear stain
- Eusion (C02): Cytoplasmic/extracellular contrast

This project aims to translate two-channel light-sheet fluorescence images into virtual H&E-stained images without paired ground-truth data.

Method Selection for Virtual Staining

Several approaches exist for virtual staining. Classical methods and supervised models are limited by their inability to handle complex textures or by the need for paired datasets. More recent techniques include Contrastive Unpaired Translation (CUT), which in many cases outperforms CycleGAN by reducing hallucinations, preserving tissue structure better, and enabling faster training, though it remains relatively complex and requires a large number of images. Advanced conditional diffusion models can produce very realistic textures and high visual quality, but they are data intensive, difficult to control in terms of morphology, and generally more complex than needed for this task. CycleGAN remains the most common choice for virtual staining because it does not require paired data and learns a mapping between two image domains, making it well suited for translating two-channel fluorescence images into realistic H&E images using unpaired training with datasets such as MHIST.

CycleGAN [1] was selected because it enables bidirectional domain translation without paired supervision and preserves structural consistency through cycle consistency loss. It has also been widely validated for histopathology and virtual staining tasks.

Data Preparation and Preprocessing

Source Domain (Two-channel light-sheet microscopy images (TO-PRO-3 + Eusion)): The source domain consists of fluorescence microscopy images with two channels: C01 (TO-PRO-3 nuclear staining) and C02 (Eosin cytoplasm channel). Both channels are stored as grayscale TIFF files. To create RGB inputs, we combine the channels using a semantic color mapping: C01 maps to the blue channel (Hematoxylin like), C02 maps to the green channel (Eosin like), and the red channel receives a 30% contribution from C02. Before combination, each channel is normalized to reduce the impact of outliers and intensity variations, then scaled to [0, 1] and converted to RGB format.

For Target Domain, the MHIST dataset is used (available at `https://bmirds.github.io/MHIST/`), which contains high quality H&E stained histopathology images suitable for learning realistic color and texture distributions.Jana The dataset contains 3,152 H&E images. Some examples are displayed in Figure 1.

All images are resized to 256×256 using bilinear interpolation. During training, data augmentation includes random horizontal and vertical flips (p=0.5) and color jitter (brightness and contrast ±0.1). Images are converted to tensors and normalized to [-1, 1] using mean=0.5 and std=0.5 to match the generator's tanh output range.

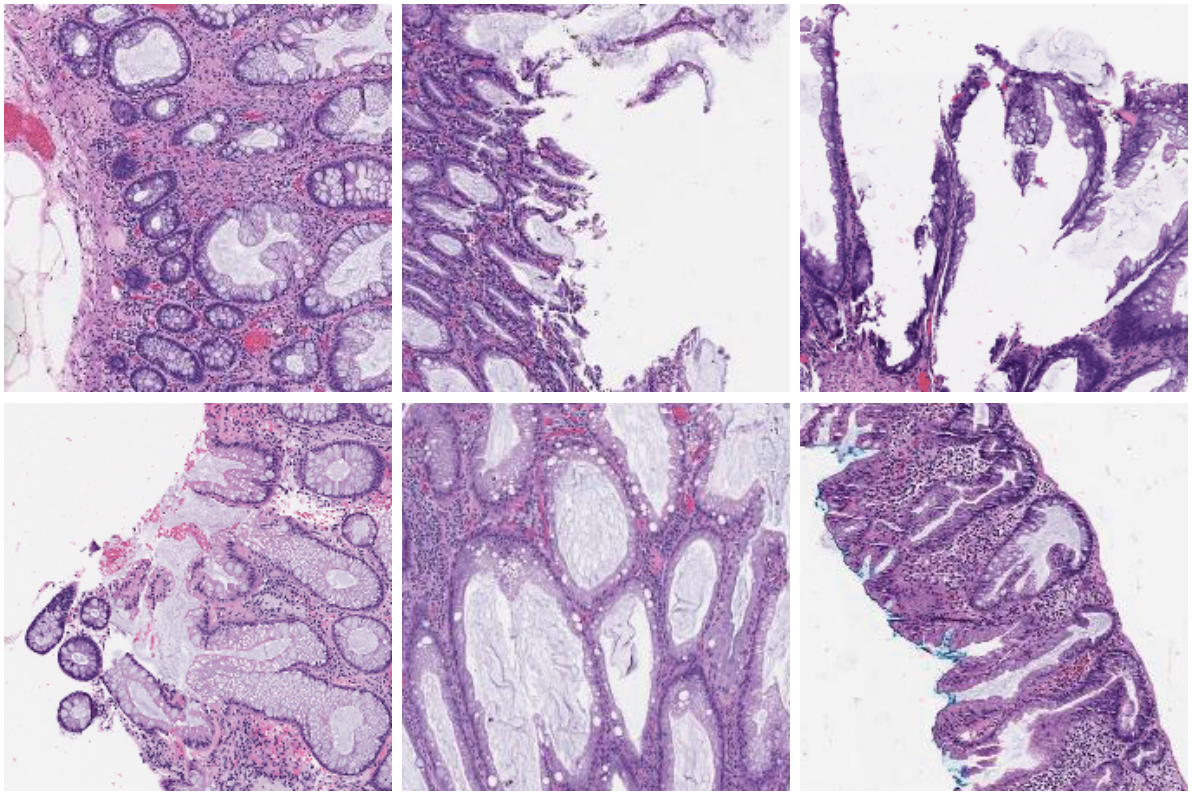

Figure 1. Example H&E images from the MHIST dataset. Each image shows different tissue regions used as the reference H&E domain in CycleGAN training.

Model Architecture

The model includes two generators and two discriminators trained adversarially, as shown in Figure 2.

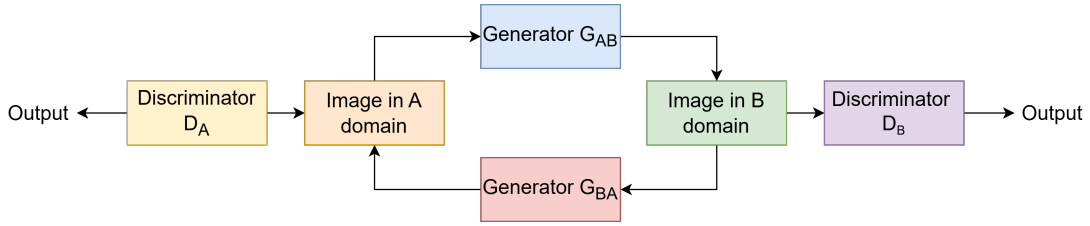

Figure 2. Overview of the proposed virtual H&E generation pipeline. A is Source Domain, B is Target Domain.

Generator Architecture

Each generator is a ResNet based encoder-decoder. The encoder: initial 7×7 convolution with reflection padding (64 filters), two downsampling blocks (stride 2) increasing filters

to 256, and 9 residual blocks with instance normalization and ReLU. The decoder: two transposed convolution upsampling blocks (stride 2) reducing filters back to 64, and a final 7×7 convolution with tanh activation producing RGB outputs in [-1, 1]. Reflection padding is used throughout to reduce boundary artifacts.

Discriminator Architecture

Each discriminator has: an initial 4×4 convolution (stride 2, 64 filters) with LeakyReLU (0.2), two additional downsampling blocks (stride 2) increasing filters to 256 with instance normalization, a final 4×4 convolution (stride 1) producing a single-channel output map, and LeakyReLU (0.2) throughout. This design focuses on local patch-level realism.

Cycle-Consistent Architecture

The model includes two generators: G_A2B (fluorescence → H&E) and G_B2A (H&E → fluorescence), and two discriminators: D_A (fluorescence domain) and D_B (H&E domain). The forward cycle (A→B→A) and backward cycle (B→A→B) enforce cycle consistency, ensuring that translating an image to the target domain and back reconstructs the original image. Identity mappings are computed for both domains to preserve domain specific characteristics when the input is already in the target domain.

Results

The Figure 3 shows virtual H&E generation from light-sheet microscopy image. Each column represents a different type of image. Each row corresponds to a different slice. From these images, it can be observed that the CycleGAN has learned to map the nuclear and cytoplasmic signals to the characteristic H&E colors. Overall tissue morphology and structure are largely preserved, demonstrating that the model captures the spatial relationships of nuclei and cytoplasm from the fluorescence channels.

The Figure 4 illustrates how CycleGAN outputs evolve across training epochs. Each column shows the virtual H&E image generated at a specific epoch, while each row corresponds to the same slices. Early in training, the images may exhibit incomplete or uneven staining and less clear tissue structure.

As training progresses, the representation of H&E changes slightly and becomes visually more consistent. However, through comparison across epochs, it can also be observed that the CycleGAN generated results exhibit certain drawbacks, including variability in staining intensity, occasional loss of fine structural details, and the introduction of artificial texture patterns that are not present in the original fluorescence images. Due to the lack of ground truth H&E images, the quality of the generated virtual H&E cannot be quantitatively validated in this study and is assessed primarily through visual inspection. If paired ground truth H&E data were available, quantitative performance metrics such as cell count agreement or dice score and intersection of union (IOU) of the stained cells (positive cells) in the virtually generated image with respect to the real image [2].

Limitations and Possible Improvements

A primary limitation of this work is the relatively small size of the fluorescence source domain, which consists of only 256 paired fluorescence images. This limited diversity

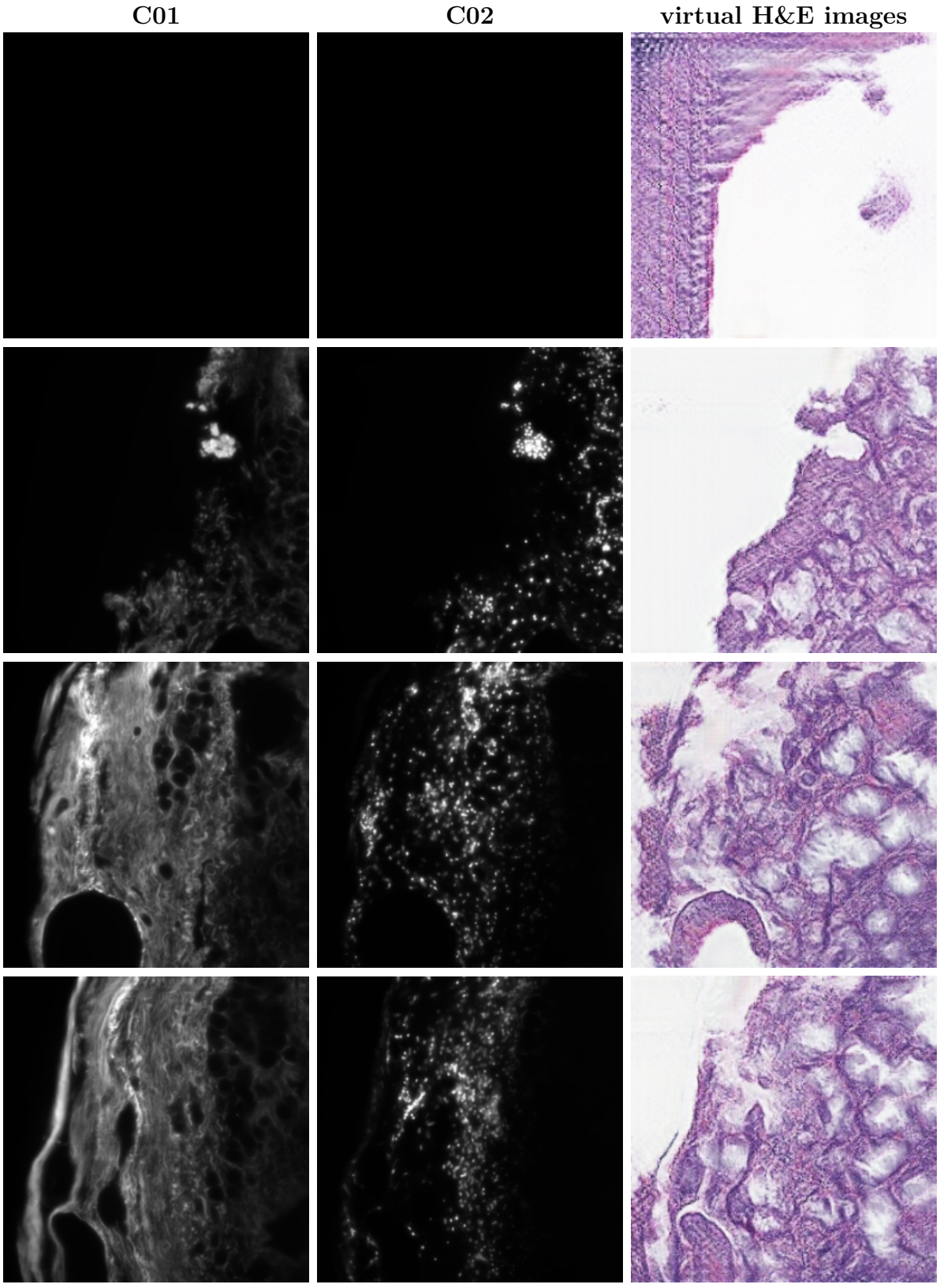

Figure 3. Virtual H&E generation from light-sheet microscopy. Columns show TO-PRO-3 (nuclear), Eusion, and CycleGAN generated virtual H&E images. Each row corresponds to a different slice.Results are shown from the CycleGAN model at epoch .

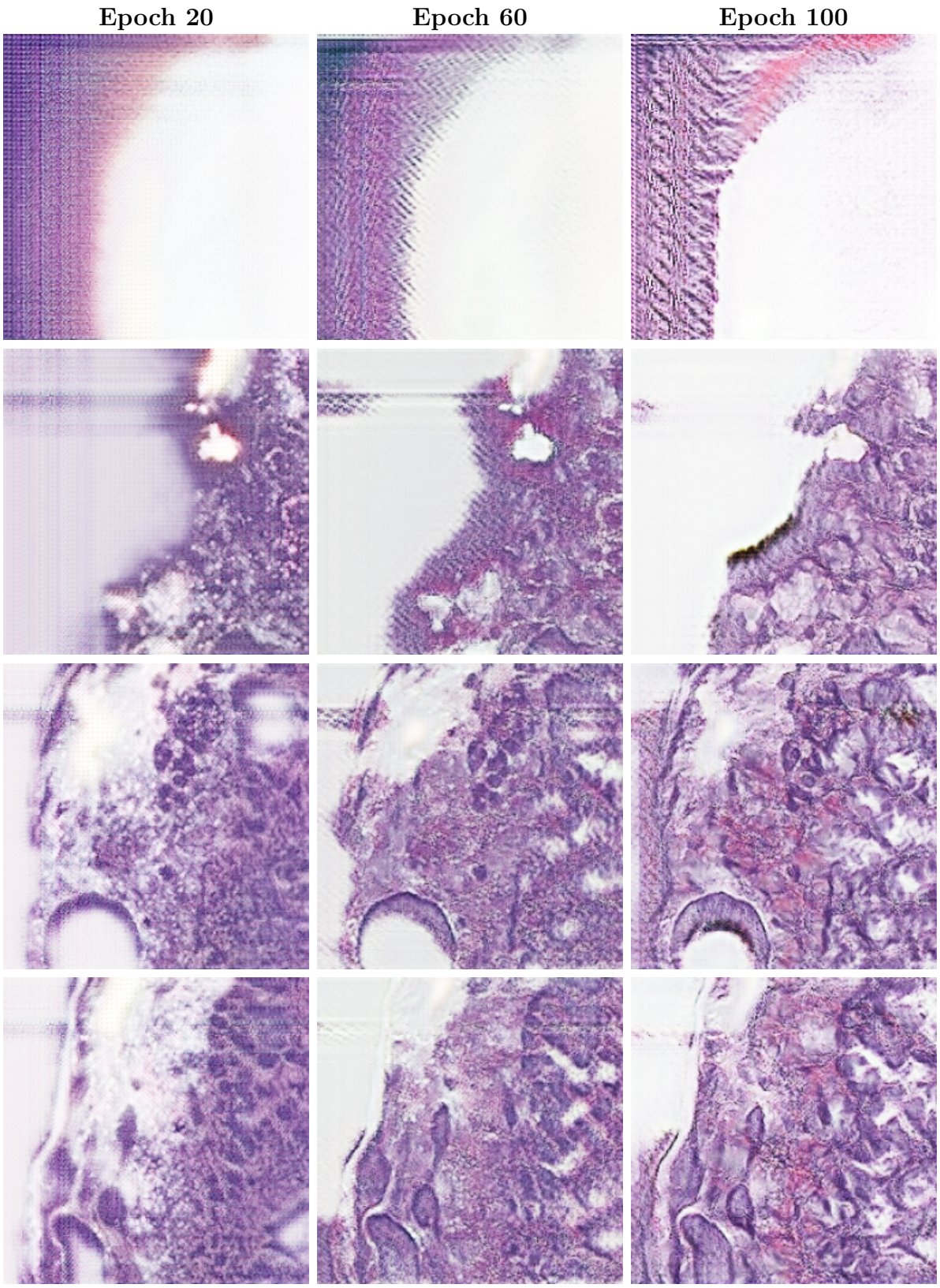

Figure 4. Comparison of CycleGAN generated virtual H&E images across training epochs. Columns show outputs generated at different training epochs, while each row corresponds to the same tissue slice.

may restrict the model's ability to generalize across different tissue regions and imaging conditions, increasing the risk of overfitting. Although the H&E target domain contains 3,152 images, the quality and diversity of these reference images strongly influence the realism of the generated virtual stains. Biases or limited variability in the H&E dataset can lead to suboptimal color mapping and texture representation.

Another limitation is the lack of explicit confidence estimates or quality control mechanisms for the generated virtual H&E images. The model produces outputs deterministically without providing uncertainty measures, making it difficult to automatically identify low-quality results. As a consequence, some generated images may exhibit artifacts, unrealistic textures, or color inconsistencies, particularly during early training stages or when processing challenging tissue structures.

Several improvements can be explored in future work. Incorporating quantitative evaluation metrics, such as structural similarity measures and task-specific metrics tailored to virtual H&E image quality, would enable more objective assessment of model performance. Expanding the training dataset with additional fluorescence samples and more diverse H&E images would likely improve robustness and generalization. Finally, alternative unpaired translation models, such as Contrastive Unpaired Translation (CUT), could be investigated to reduce hallucination effects and further improve structural preservation in virtual staining results.

Bibliography

[1] J. Vasiljević, Z. Nisar, F. Feuerhake, C. Wemmert, and T. Lampert, "CycleGAN for Virtual Stain Transfer: Is Seeing Really Believing?", *Artificial Intelligence in Medicine*, vol. 133, p. 102420, 2022.

[2] Dubey, Shikha, et al. "Structural cycle gan for virtual immunohistochemistry staining of gland markers in the colon." International Workshop on Machine Learning in Medical Imaging. Cham: Springer Nature Switzerland, 2023.