Self-Attention Guided CycleGAN for Enhanced Reconstruction of Fibrosis Features in Pathological Images

Cheng Xianda[a], **Nicholas Syn**[a], **Cheng Han Ng**[a], **Aileen Wee**[c], **Gwyneth Shook Ting Soon**[a], **Hirokazu Takahashi**[b], **Ho Khek-Yu**[a], **Mark Muthiah**[a]

[a] *Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore* cheng.xianda@outlook.sg

[b] *Saga University Hospital, Faculty of Medicine, Saga University, Saga-shi, Japan*

[c] *Department of Pathology, National University Hospital, Singapore*

* Presenting author

1. Introduction

Non-alcoholic fatty liver disease (NAFLD) is rapidly becoming the most prevalent cause of liver disease worldwide.[?] Pathological examination remains the diagnostic gold standard, using Hematoxylin and Eosin (H&E) for morphology and Masson's Trichrome (MT) for fibrosis assessment.[1] However, traditional MT staining poses multiple obstacles—time, labor, sampling error risks, and inconsistent quality.[2]

Digital pathology and machine learning promise new solutions for virtual staining Fig.1. Pix2Pix relies on paired datasets to map H&E to MT, while CycleGAN[3] operates on unpaired data but struggles with fibrosis detail. We propose extending CycleGAN through Region of Interest (ROI)[4] extraction and self-attention[5], improving the model's detection and representation of fibrotic regions, thereby offering enhanced clinical utility in virtual staining.

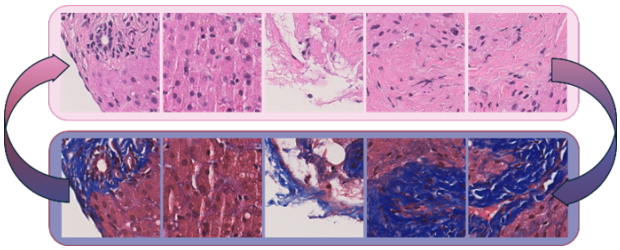

Fig. 1: Virtual staining schematic diagram of H&E and MT staining pathological images

2. Related Work

2.1 GAN in Pathology Image Generation

The Pix2Pix model, proposed by Isola et al[6], employs a conditional Generative Adversarial Network (GAN) architecture, combining adversarial loss with L1 reconstruction loss. However, it requires strictly paired data. In contrast, CycleGAN[7], achieves unpaired data transformation through cycle consistency loss, significantly reducing the difficulty of data collection. Despite this advantage, CycleGAN has limited performance in depicting detailed features of specific lesion areas.

2.2 Attention Mechanisms in Virtual Staining

Zhang[8] introduced the Self-Attention GAN, which incorporates self-attention layers to enable the model to learn long-range dependencies within images. Methods like AttentionGAN[9] integrate spatial and channel attention modules, effectively increasing the focus on Regions of Interest (ROI) and enhancing the quality of virtual staining. However, there remains room for improvement in accurately expressing fibrosis features.

3. Methodology

3.1 Dataset

In this study, we used paraffin-embedded liver biopsy sections from 68 anonymized patients (fibrosis stages F1–F4). Each was stained with H&E and MT, then scanned at 40× to produce whole-slide images. We aligned the two stain modalities using VALIS[10] registration method. From both H&E and MT slides, a 512×512 sliding window extracted tiles and resized to 256×256. In each MT tile, a 48×48 bounding box (bbox) marked fibrotic regions (purple areas labeled positive /1, otherwise negative /0), with the same bbox applied to H&E.

3.2 Model Architecture

Drawing inspiration from Region-guided CycleGAN [11], we fuse self-attention and ROI discrimination into a multi-level framework consisting of an enhanced generator and two discriminators. The generator, based on a modified ResNet backbone, inserts self-attention layers to adaptively focus on fibrotic lesions. A global PatchGAN discriminator ensures overall style consistency, while an ROI discriminator uses RoIAlign to judge synthetic quality specifically in fibrotic areas. This design preserving both global and local fidelity.

4. Experiments and Results

4.1 Experimental Design and Baseline Method

This study aims to validate the effectiveness of our proposed ROI-SAGAN and compare it with the native CycleGAN. We use the IoU (Intersection over Union)[12] metric to measure the accuracy of key areas and combine visual comparisons to evaluate the

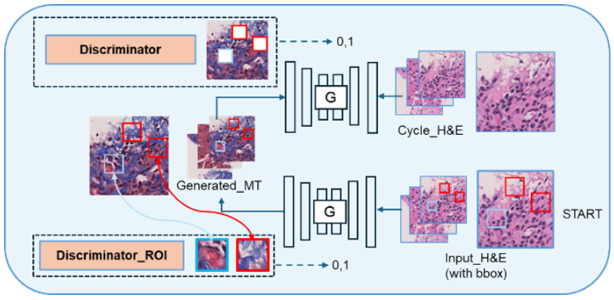

Fig. 2: H&E tiles are processed through a generator to generate MT-stained tiles, with discriminators trained on labeled fibrosis/non-fibrosis regions refining the output; cycle consistency loss ensures fidelity between original H&E and regenerated tiles.

fidelity of generated images in fibrosis regions. To ensure a fair comparison, we trained and tested both methods using the same dataset and data split strategy. During the testing phase, we quantitatively analyze the IoU of fibrosis regions between the virtual staining results and real MT images, combined with visual comparisons of lesion areas, to verify the effectiveness of ROI-SAGAN in preserving pathological features.

4.2 Main Result

This study compares fibrosis region generation quality between ROI-SAGAN and the native CycleGAN, using IoU to measure accuracy at key lesion sites. As shown in Table 1, the native CycleGAN achieves an IoU of 0.254, whereas ROI-SAGAN reaches 0.510, indicating that explicit ROI extraction and self-attention improve detail reproduction and overall staining fidelity in fibrosis areas.

Table 1: ROI-SAGAN and CycleGAN mean IoU on Test dataset

Patient-ID	CycleGAN-Fibrosis-IoU	ROI-SAGAN-Fibrosis-IoU
NEMB078	0.294	0.716
NEMB079	0.313	0.710
NEMB018	0.253	0.370
NEMB085	0.510	0.680
NEMB020	0.112	0.457
NEMB074	0.155	0.340
NEMB050	0.143	0.305
NEMB059	0.314	0.464
NEMB076	0.194	0.526
Average	0.254	**0.510**

Figure 3 shows the same pathological sample under both methods, with red boxes highlighting fibrosis discrepancies for the native CycleGAN vs. real staining, and green boxes comparing ROI-SAGAN to the real staining. ROI-SAGAN offers more precise detail in fibrosis edges and internal structures, while retaining a style closer to real MT images. Hence, explicit ROI focus plus self-attention enhances lesion feature recovery and consistency, providing a stronger foundation for virtual staining applications and subsequent diagnostics.

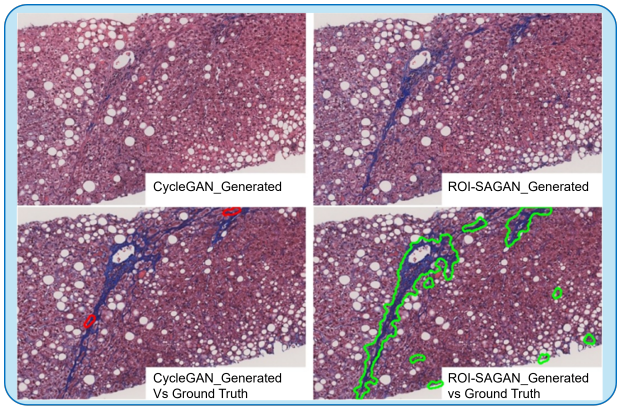

Fig. 3: Cyclegan and ROI-SAGAN staining ranges compared to real MT stained image

4.3 Image Generated Evaluation

Figure 4 illustrates the generate results of ROI-SAGAN. In regions with fibrosis, ROI-SAGAN presents more refined color gradients and edge handling, closely matching real stained images. These findings confirm the approach's viability for pathological staining conversion, providing a higher-quality synthetic basis for diagnostic analysis. This framework not only improves the fidelity of virtual staining but also increases the reliability of synthesized images for clinical use.

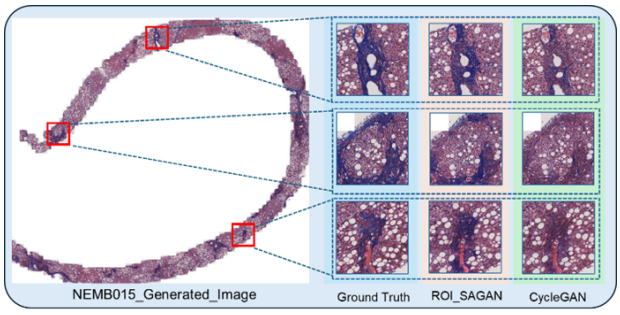

Fig. 4: MT virtual staining WSI image and fibrosis part images

5. Conclusion

We proposed ROI-SAGAN, which augments CycleGAN with fibrosis ROI extraction and self-attention, significantly improving virtual staining in fibrotic regions. Experiments show ROI-SAGAN achieves an IoU of **0.510**, outperforming native CycleGAN's 0.254, especially in color transitions and edge detail. This enhanced fidelity highlights the method's potential for automated pathological imaging and advanced diagnostic support in liver fibrosis analysis.

References

[1] Bing Qu and Zheng Li. Exploring non-invasive diagnostics for metabolic dysfunction-associated fatty liver disease. *World Journal of Gastroenterology*, 30(28):3447–3451, 2024.

[2] Junxian Wu, Xinyi Ke, XIAOMING JIANG, Huanwen Wu, Youyong Kong, and Lizhi Shao. Leveraging tumor heterogeneity: Heterogeneous graph representation learning for cancer survival prediction in whole slide images. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 64312–64337. Curran Associates, Inc., 2024.

[3] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017.

[4] Shreyas Malakarjun Patil, Li Tong, and May D. Wang. Generating region of interests for invasive breast cancer in histopathological whole-slide-image. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 723–728, 2020.

[5] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Andy Gomez, Lukasz Kaiser, and Ilya Polosukhin. Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 2017.

[6] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5967–5976, 2017.

[7] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2242–2251, 2017.

[8] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 7354–7363. PMLR, 09–15 Jun 2019.

[9] Youssef Alami Mejjati, Christian Richardt, James Tompkin, Darren Cosker, and Kwang In Kim. Unsupervised attention-guided image-to-image translation. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

[10] Chandler Gatenbee and Alexander R.A. Anderson. Valis: Virtual alignment of pathology image series for multi-gigapixel whole slide images, June 2024.

[11] Joseph Boyd, Irène Villa, Marie-Christine Mathieu, Eric Deutsch, Nikos Paragios, Maria Vakalopoulou, and Stergios Christodoulidis. Region-guided cyclegans for stain transfer in whole slide images. In Linwei Wang, Qi Dou, P. Thomas Fletcher, Stefanie Speidel, and Shuo Li, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, pages 356–365, Cham, 2022. Springer Nature Switzerland.

[12] Chih-Ying Wu, Wei-Chang Yeh, Shiaw-Meng Chang, Che-Wei Hsu, and Zi-Jie Lin. The application of deep learning to accurately identify the dimensions of spinal canal and intervertebral foramen as evaluated by the iou index. *Bioengineering*, 11(10), 2024.