Correspondence

https://doi.org/10.1631/jzus.B2500012

Seeing the macro in the micro: a diffusion model-based approach for style transfer in cellular images

Jiayi CAI[1], Yong HE[1], Feng LIU[2], Byung-Ho KANG[3], Xuping FENG[1]✉

[1]*College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China*
[2]*College of Life Sciences, Nanjing Agricultural University, Nanjing 210095, China*
[3]*Cellular and Molecular Biology Program, State Key Laboratory of Agrobiotechnology, School of Life Sciences, The Chinese University of Hong Kong, New Territories, Hong Kong, China*

The internal structures of cells as the basic units of life are a major wonder of the microscopic world. Cellular images provide an intriguing window to help explore and understand the composition and function of these structures. Scientific imagery combined with artistic expression can further expand the potential of imaging in educational dissemination and interdisciplinary applications. This study presents an innovative diffusion model-based approach for style transfer in cellular images, combining scientific rigor with artistic expression. By leveraging training-free large-scale pre-trained diffusion models, the proposed method integrates the intricate morphological and textural features of cellular images with diverse artistic styles. Key techniques such as the inversion of Denoising Diffusion Implicit Models (DDIMs), Adaptive Instance Normalization (AdaIN), self-attention style injection, and attention temperature scaling ensure the preservation of cellular structures while enhancing visual expressiveness. The results showcase the potential of this strategy for interdisciplinary applications, enriching both the visualization and educational dissemination of cellular imagery through compelling storytelling and aesthetic appeal.

Unedited

While delving into the microcosm, it becomes clear that the small often mirrors the vast. Shen Fu, in his essay "Childhood Amusement," once wrote, "I recall as a child, I could stare directly at the sun, observe the minutest details, and examine the intricate patterns of small objects with curiosity, often finding joy beyond the object itself." This delight in observing fine details leads to a deeper understanding of the world (Shen, 2018). When we focus on the smallest elements of something grand, it may initially feel unfamiliar. However, when we relate them to something familiar, we can begin to uncover the “delight beyond the object.” The microcosm and the macrocosm, though seemingly distinct, are profoundly interconnected. As we marvel at the infinite vastness of the universe, gazing at distant galaxies and stars, have we ever considered that these unreachable cosmic entities might, in some sense, echo the narratives unfolding in the tiny dust particles or bacteria around us?

Cells, as fundamental building blocks of life, reveal the boundless wonders of the microscopic world. Within the microcosm of cells, countless fascinating stories may quietly unfold (Gao and Liu, 2019). This article seeks to infuse cellular images with a sense of storytelling and artistic beauty via the application of style transfer techniques. By doing so, it explores the aesthetic expression of cellular structures and their potential within the interdisciplinary realms of science and art. Guided by this vision, this project employs a training-free style transfer method based on large-scale pre-trained diffusion models (Chung et al., 2024), leveraging their

✉ Xuping FENG, fengxp@zju.edu.cn
Xuping FENG, https://orcid.org/0000-0001-9575-6916
Jiayi CAI, https://orcid.org/0009-0009-0532-0015

Received Jan. 8, 2025; Revision accepted Mar. 6, 2025;
Crosschecked xxx. xx, 20xx; Published online xxx. xx, 20xx

© Zhejiang University Press 2025

exceptional capabilities in generating high-quality images. To achieve an efficient fusion of content and style, it integrates these models with the morphological complexity and unique textural features of cellular images.

Specifically, the proposed method uses the inversion of Denoising Diffusion Implicit Models (DDIMs) to map cellular images and style images from the pixel space to the latent noise space, generating their corresponding latent representations (Fig. 1). On this basis, Adaptive Instance Normalization (AdaIN) is applied to adjust the channel statistics of the cellular images to match those of the style images, producing an initial noise representation that integrates cellular content and artistic style (Kim et al., 2024). This process ensures that during the reverse diffusion phase, the spatial structure of cellular images (e.g., nuclei, cell membranes, cytoplasm, etc.) is well-preserved while successfully incorporating the texture and color features of the style images.

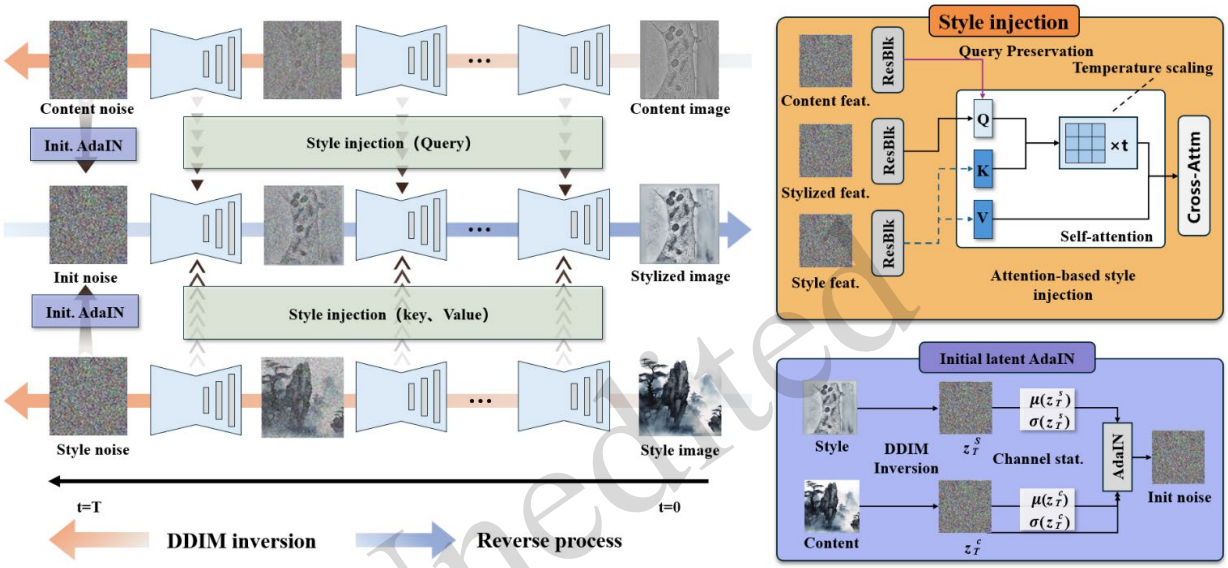

Fig. 1 Diagram of the proposed style transfer method. First, the content image and the style image are inverted into the latent noise space, resulting in content noise and style noise, respectively. Then, the latent AdaIN is utilized to combine the content noise and style noise, forming the stylized image noise. During the reverse diffusion process using the initial noise, the information of content and style is injected through attention-based style injection and Attention Temperature Scaling.

During the reverse diffusion process, this method further employs a self-attention style injection mechanism (Zhang et al., 2021). By replacing the key and value features of the cellular images with those of the style images, the style textures are injected into the cellular images. At the same time, a query preservation mechanism is used to protect the unique morphological features of the cellular images (Tumanyan et al., 2023), ensuring that the generated images achieve stylized expression while retaining cellular structural features. Additionally, key-value feature replacement may cause the attention map to become overly smooth, which can reduce detail sharpness. To address this issue, we introduce Attention Temperature Scaling to enhance the clarity and visual expressiveness of the generated images. This technical design not only improves the style consistency and content fidelity of the results but also provides greater flexibility and controllability for the artistic expression of cellular images.

This approach offers three key innovations, providing robust technical support for the style transfer and artistic expression of cellular images:

1. Training-free implementation, significantly lowering the barrier to adoption: Traditional style transfer methods typically rely on large datasets and extensive model training, which not only demand substantial computational resources but also limit the flexibility and applicability of these methods in real-world scenarios. In contrast, the proposed approach leverages a large-scale pre-trained diffusion model framework, directly utilizing the generative capabilities learned during pre-training, without requiring additional optimization or

Unedited

fine-tuning. This characteristic simplifies the implementation process of style transfer, reducing both time and resource costs. Moreover, the training-free nature of this method enables seamless application across diverse research contexts, greatly expanding its practical utility.

2. Integration of scientific rigor and artistic expression: The presented approach achieves a harmonious balance between preserving the scientific accuracy of cellular images and imparting them with the uniqueness of artistic expression. During the style transfer process, the key morphological features (e.g., nuclei, cell membranes, cytoplasm) and spatial structure information of the cellular images are retained, ensuring that the generated results faithfully reflect the scientific content of the original cells. At the same time, the artistic representation of the cellular images is enhanced by incorporating the texture, color tone and visual elements of the target style. This integration of scientific and artistic elements provides an innovative solution for an enhanced visualization, dissemination and educational use of cellular images.

3. Efficient and adjustable style modulation capabilities: The proposed method introduces multiple adjustable parameters, granting users flexible control over the balance between style and content in the generated images. For instance, the style-content trade-off parameter determines the degree of style application while preserving the original structure, and attention temperature scaling enhances detail sharpness by adjusting the self-attention map intensity. Additionally, the feature blending ratio in query preservation fine-tunes the integration of content and style. These parameters function like a "sliding rheostat" for style modulation, allowing users to tailor the strength of the style transfer and the fidelity to the original content according to specific requirements. This feature not only enhances the diversity of the generated results but also meets the varying demands of different application scenarios.

We have dressed a number of cellular images in styles such as blue-and-white porcelain, Monet's brushstrokes, and traditional Chinese ink painting. The following images illustrate the results of style transfer, showcasing various artistic transformations.

For example, when the visual style of a durian fruit is transferred onto a chloroplast image, we obtain a "durian-style" cellular image (Fig. 2).

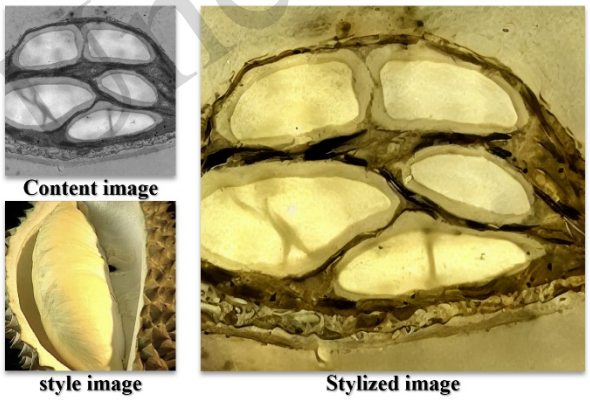

Fig. 2 Durian-style cellular image of chloroplast.

Similarly, applying the artistic style of Beauties with Hairpin Flowers to tobacco cell images produces a mural-like image, where the black dots resemble explorers venturing deeper into a tunnel (Fig. 3).

Unedited

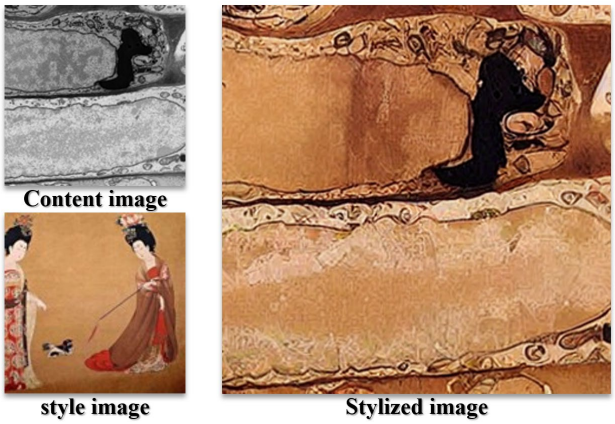

Fig. 3 Tobacco cell image in mural style "The Explorer."

The strong sense of spatial extension in the chloroplast transmission electron microscopy image in Fig. 4 is particularly striking. After processing, the resulting image on the right exhibits the texture of traditional Chinese ink painting, which we named “Fishing Alone on a Cold River in the Snow.” The thylakoids in the chloroplast resemble layers of overlapping mountain ridges, while the ribosomes within the chloroplast appear like pine trees standing firm in the cold wind. The oval-shaped starch granule in the lower-right corner resembles a small boat drifting on the cold river. Looking deeper into the image feels as though an infinite series of peaks extend far into the distance.

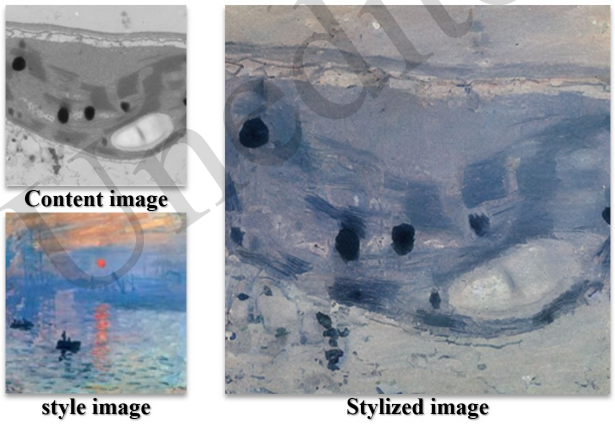

Fig. 4 Thylakoids in chloroplast: "Fishing Alone on a Cold River in the Snow."

Figure 5 presents an image that, after undergoing style transfer, exudes the delicate aesthetic of ink painting diffusion. In this image, the nuclear membrane and cell membrane form a tree trunk-like structure; the mitochondria resemble ladybugs perched on the trunk; and the Golgi bodies seem like lenticels on the bark. One cannot help but imagine this as a weathered yet resilient tree, standing tall against the elements.

Unedited

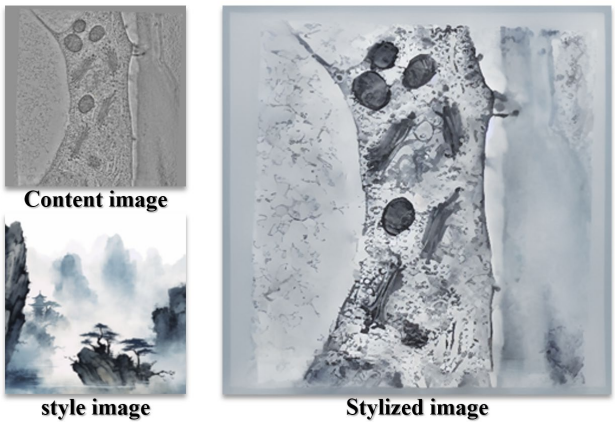

Fig. 5 "Withered Tree Awaits Spring."

Data availability statement

Data will be made available on request.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (226-2024-00038), China.

Author contributions

Xuping FENG and Jiayi CAI designed and conceptualized the study. Jiayi CAI implemented the algorithm and analysed the image data. Feng LIU and Byung-Ho KANG provided cell data. Xuping FENG and Jiayi CAI wrote the paper. Yong HE and Xuping FENG reviewed the paper.

Compliance with ethics guidelines

Jiayi CAI, Yong HE, Feng LIU, Byung-Ho KANG and Xuping FENG declare that they have no conflict of interest.
This article does not contain any studies with human or animal subjects performed by any of the authors.

Unedited

References

Chung J, Hyun S, Heo J-P, et al., 2024. Style injection in diffusion: A training-free approach for adapting large-scale diffusion models for style transfer. 2024 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, CVPR 2024, p.8795-8805. https://doi.org/10.1109/CVPR52733.2024.00840

Gao SW, Liu F, 2019. Novel insights into cell cycle regulation of cell fate determination. J Zhejiang Univ Sci B, 20(6):467-475. https://doi.org/10.1631/jzus.B1900197

Kim B, Shin W, Jung Y, et al., 2024. Explicitly color-inspired neural style transfer using patchified adain. CMES - Computer Modeling in Engineering and Sciences, 141(3):2143-2164. https://doi.org/https://doi.org/10.32604/cmes.2024.056079

Shen F, 2018. Six Chapters of A Floating Life. China Publishing House, Beijing, China, p.53-54 (in Chinese).

Tumanyan N, Geyer M, Bagon S, et al., 2023. Plug-and-play diffusion features for text-driven image-to-image translation. 2023 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, CVPR, p.1921-1930. https://doi.org/10.1109/CVPR52729.2023.00191

Zhang T, Gao Y, Gao F, et al., 2021. Arbitrary style transfer with parallel self-attention. 2020 25TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR), p.1406-1413. https://doi.org/10.1109/ICPR48806.2021.9412049