

Patch-to-Global: Random Patch Diffusion for Globally Consistent Megapixel Artifact Inpainting in Whole Slide Images

Hyeseong Lee¹, Eunsu Kim¹, D M BAPPY², Ho Heon Kim³, Youngsuk Lee¹, Se Young Chun⁴, Jang-Hwan Choi⁵, Sung Hak Lee⁶, and Sangjeong Ahn^{1,2}✉

¹ Department of Biomedical Informatics, Korea University College of Medicine, Seoul, Korea {gotjd709, kes311az, younnngsuk, vanitas80}@korea.ac.kr

² Department of Pathology, Korea University Anam Hospital, College of Medicine, Korea University, Seoul, Korea dewan8511@gmail.com

³ Seegene Medical Foundation, Seoul, Korea hoheon0509@gmail.com

⁴ Department of ECE, Seoul National University, Seoul, Korea sychun@snu.ac.kr

⁵ Department of Artificial Intelligence, Ewha Womans University, Seoul, Korea choij@ewha.ac.kr

⁶ Department of Hospital Pathology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea hakjjang@catholic.ac.kr

Abstract. Although deep learning has advanced Whole Slide Image (WSI) Analysis, tissue artifacts like bubbles and folds often cause silent failures by concealing essential morphology. Current pathology image restoration methods are mostly restricted to small patches, struggling to maintain global structural coherence at a megapixel scale. We introduce **RestorePath**, a framework for globally consistent megapixel-scale inpainting that reconstructs diagnostic structures in histological image to prevent incorrect high-confidence predictions and lower error rates. Our model utilizes a Latent Diffusion Model (LDM) conditioned on Pathology Foundation Model (PFM) embeddings, integrating Large Kernel Attention (LKA) to manage long-range dependencies during random patch diffusion. Enhanced by Distance-Weighted Interpolation (DWI) and an Adaptive Guidance Scale (AGS), RestorePath ensures structural consistency and fidelity by modulating information from surrounding patches. Evaluations across TCGA-BRCA, BACH, and Camelyon16 datasets for images ranging from 512 to 4608 pixels demonstrate state-of-the-art performance in maintaining histological consistency. RestorePath significantly improves downstream Computational Pathology (CP) tasks, outperforming both raw artifact images and the conventional Detect-and-Discard (D&D) approach. The code is available at <https://github.com/PathfinderLab/RestorePath>.

Keywords: Histological Artifact Restoration · Megapixel-Scale Inpainting · Latent Diffusion Model · Computational Pathology.

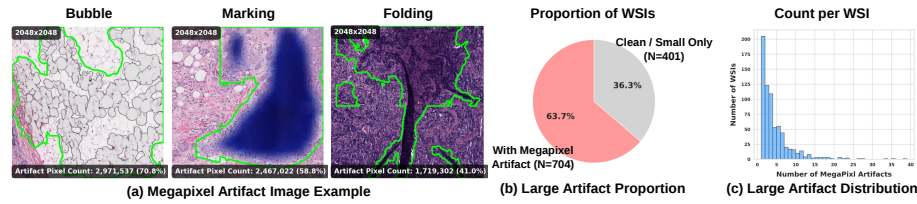


Fig. 1. Examples and distribution of megapixel artifacts in WSIs. (a) Examples of bubbles, markings, and tissue folding. (b) Prevalence of megapixel artifacts, with 63.7% of WSIs containing at least one megapixel-scale corruption. (c) Distribution of artifact counts per WSI, showing that many slides contain multiple large-scale artifacts.

1 Introduction

Pathological diagnosis remains the gold standard for clinical decision-making and research. Digital pathology has revolutionized this process by transforming glass slides into Whole Slide Images (WSIs) [4]. However, the immense scale of these images imposes a significant cognitive burden on pathologists, which has accelerated the development of deep learning algorithms and Pathology Foundation Models (PFMs) for automated analysis [16,15].

However, such computational models often faces significant obstacles from tissue artifacts such as bubbles, folding, and markings (Fig. 1(a)). These artifacts arise inevitably during tissue acquisition and scanning, often leading to increased error rates [27] or silent failures where models generate incorrect predictions with high confidence [23]. By concealing cellular details or misleading models into misidentifying morphology as different cell types, artifacts severely undermine diagnostic reliability. Furthermore, The process of tissue recutting and its physical rescanning is frequently impractical due to high costs, time constraints, and the risk of tissue depletion [7,22].

To mitigate these issues, automated Quality Control (QC) tools like HistoQC [11] and GrandQC [25] have been developed. These tools typically adopt a detect-and-discard (D&D) strategy to identify and exclude defective regions. However, this strategy is detrimental in needle-in-a-haystack scenarios where discarding sparse lesions risks losing indispensable diagnostic evidence. While diffusion-based restoration methods such as Artifusion [7] and ArtiDiffuser [22] have emerged to address this, they remain restricted to patch-level reconstruction. Our analysis of the TCGA-BRCA [24] dataset reveals that 63.7% of WSIs contain megapixel-scale artifacts (Fig. 1(b)), which underscores the necessity for high-resolution restoration that extends far beyond the limitations of individual patches. Although DiffInfinite [1] proposed a Random Diffusion Inpainting approach for high-resolution images, its primary focus is on synthesizing new tissue rather than achieving globally-aligned restoration. Furthermore, while LRDM [5] and ZoomLDM [26] successfully generate high-resolution images using PFM-based feature embeddings, they have not yet attempted globally consistent inpainting that maintains strict alignment with the surrounding healthy tissue.

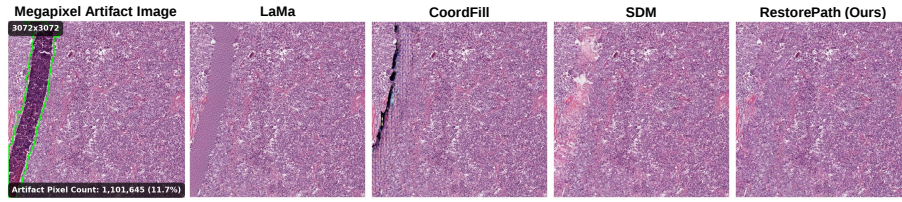


Fig. 2. Qualitative comparison of megapixel artifact restoration performance. Unlike existing methods such as LaMa, CoordFill, and SDM, RestorePath successfully reconstructs artifact regions while preserving the histological context and histological patterns of surrounding tissues.

In this study, we redefine the restoration of megapixel-scale artifacts as a globally consistent inpainting task to ensure robust performance in downstream Computational Pathology (CP) analyses. We propose RestorePath, a novel framework that leverages an inpainting Latent Diffusion Model (LDM) specifically optimized for the complex histological patterns found in large-scale WSIs. Our framework introduces two primary innovations for high-fidelity restoration. First, we integrate Large Kernel Attention (LKA) into a PFM-conditioned LDM to effectively handle long-range correlations in high-resolution images. Second, we implement a random patch diffusion strategy utilizing Distance Weighted Interpolation (DWI) and an Adaptive Guidance Scale (AGS). DWI ensures structural consistency by interpolating conditioning signals from nearby healthy regions, even when patches are entirely obscured. Meanwhile, AGS optimizes the process by dynamically adjusting guidance scales based on local content and image complexity. Together, these mechanisms enable RestorePath to achieve seamless global restoration that aligns with the surrounding histological tissue.

Using CONCH [13]-based FID [8] (FID_C) and embedding similarity (Emb Sim), we demonstrate that RestorePath preserves histological context and pathological information across TCGA BRCA [24], BACH [17], and Camelyon16 [3] more effectively than existing high-resolution inpainting methods. To further illustrate the practical utility of our framework in CP research, we conducted evaluations on downstream tasks, specifically breast cancer classification and lymph node metastasis prediction using Multiple Instance Learning (MIL) [10,14,20]. Our results show that RestorePath significantly reduces error rates in the BACH dataset and mitigates silent failures in the needle-in-a-haystack scenarios of Camelyon16. By consistently outperforming artifact images and the D&D strategy, our framework establishes itself as a robust image QC tool for CP.

2 Method

RestorePath is designed to restore megapixel artifact regions in high-resolution images based on a random patch diffusion (Fig 3). We introduce the model architecture and training process in Section 2.1. Subsequently, we detail the methodology for applying this model to high-resolution images in Section 2.2.

2.1 Training

LDM with Large Kernel Attention. To effectively reconstruct histological patterns, we employ an LDM for inpainting that takes a masked latent z_m , a binary mask m , and a feature embedding c as inputs. The embedding c is extracted from the original image using a pre-trained PFM to provide a delicate morphological prior. While the cross-attention blocks in the LDM U-Net backbone facilitate precise inpainting by incorporating the condition c . However, they are insufficient for maintaining long-range correlations necessary to inpaint megapixel artifacts in high-resolution images. To address this, we replace the cross-attention blocks in the middle stage and the first block of the decoder with **Large Kernel Attention (LKA)**. These specific locations represent the transition where the model begins synthesizing global spatial coherence from compressed representations. LKA factorizes large-scale convolutions into spatial local, spatial long-range, and channel-wise element to expand the receptive field without prohibitive computational costs. The training objective is defined as:

$$\mathcal{L}_{\text{RestorePath}} = \mathbb{E}_{z,m,c,t,\epsilon} [\|\epsilon - \epsilon_\theta(z_t, z_m, m, t, c)\|_2^2] \quad (1)$$

where z_t is the noisy latent at timestep t and ϵ is the Gaussian noise.

2.2 Inference

Random Patch Diffusion with Distance Weighted Interpolation. To restore artifacts reaching megapixel scales while maintaining global histological coherence, we employ a **Random Patch Diffusion** strategy. Let $J \in \mathbb{R}^{P_H \times P_W \times 3}$ represent a high-resolution image divided into a grid of uniform size $H \times H$. This division yields a grid of dimensions $N_H \times N_W$, where $N_H = P_H/H$ and $N_W = P_W/H$. We define (k, l) as the discrete grid coordinates such that $0 \leq k < N_H$ and $0 \leq l < N_W$. A feature embedding $y_{k,l}$ is extracted from each grid cell using a PFM. A valid set of grid positions $\mathcal{S}$ is defined as $\{(k, l) \mid r_{k,l} \leq \tau\}$, where $r_{k,l}$ represents the artifact ratio in the corresponding grid cell and τ is the threshold to determine artifact-free regions. During the inference process, a patch of size $H \times H$ is randomly sampled at a continuous spatial coordinate $p = (y, x)$ within J . To guarantee the model obtains adequate morphological context despite the complete obstruction of the sampled patch by artifacts, we calculate a localized conditioning vector c_p employing **Distance Weighted Interpolation (DWI)**. This vector is derived as a weighted summation of the valid embeddings in $\mathcal{S}$:

$$c_p = \frac{\sum_{(k,l) \in \mathcal{S}} w((k, l), p) \cdot y_{k,l}}{\sum_{(k,l) \in \mathcal{S}} w((k, l), p)} \quad (2)$$

The spatial weight $w((k, l), p)$ reflects the proximity of the sampled patch to surrounding healthy tissue and is defined as:

$$w((k, l), p) = \frac{1}{\|(k \cdot H + \frac{H}{2}, l \cdot H + \frac{H}{2}) - p\|_2^\alpha + \delta} \quad (3)$$

In this formulation, $(k \cdot H + \frac{H}{2}, l \cdot H + \frac{H}{2})$ represents the center coordinates of grid cell (k, l) , α is a distance attenuation factor, and δ is a small constant used to avoid division by zero. By propagating feature embeddings from valid neighboring regions into artifact-heavy zones, RestorePath ensures that the inpainted histological patterns remain contextually and biologically consistent with the global histological environment.

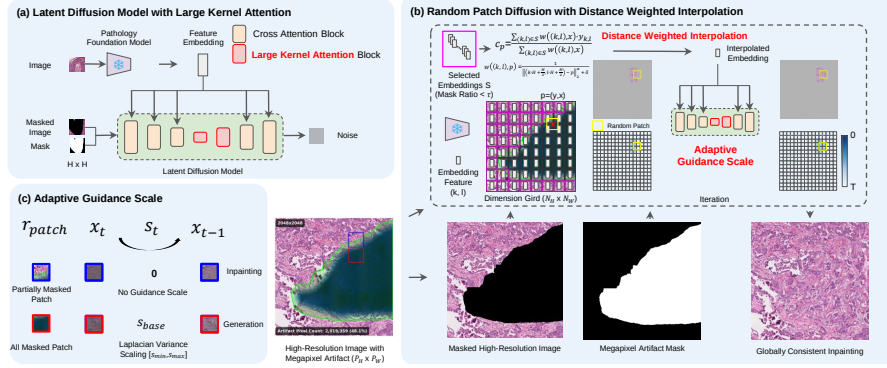


Fig. 3. Workflow of RestorePath. (a) Training phase of the Latent Diffusion Model with Large Kernel Attention. (b, c) Megapixel artifact inpainting process utilizing Random Patch Diffusion with Distance Weighted Interpolation (DWI) and Adaptive Guidance Scale (AGS).

Adaptive Guidance Scale. Restoration quality is highly dependent on the guidance scale. We propose an **Adaptive Guidance Scale (AGS)** to dynamically modulate influence of the condition c based on the patch content. We quantify image complexity by calculating the Laplacian Variance (V_{lap}) based on the artifact-free regions of the high-resolution image J with respect to its artifact mask M . This complexity is then mapped to a base scale $s_{base} = f(V_{lap}(J, M))$ ranging from s_{min} to s_{max} to reflect intrinsic histological details. The applied scale s_t for each patch is dynamically determined by its local artifact ratio r_{patch} . We define k as the artifact occupancy threshold that distinguishes between the inpainting and generation modes.

$$s_t = \begin{cases} s_{base} & \text{if } r_{patch} > k \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The final noise prediction follows the classifier-free guidance formulation:

$$\tilde{\epsilon}_\theta = \epsilon_\theta(z_t, z_m, m, t, \emptyset) + s_t \cdot (\epsilon_\theta(z_t, z_m, m, t, c) - \epsilon_\theta(z_t, z_m, m, t, \emptyset)) \quad (5)$$

This mechanism prioritizes original pixel preservation in partially masked patches while enforcing strong contextual priors in heavily damaged regions.

Table 1. Quantitative comparison of inpainting performance across different datasets.

Methods	TCGA BRCA			Camelyon16			BACH		
	LPIPS _M	FID _C	Emb	LPIPS _M	FID _C	Emb	LPIPS _M	FID _C	Emb
	(↓)	(↓)	Sim(↑)	(↓)	(↓)	Sim(↑)	(↓)	(↓)	Sim(↑)
LaMa	0.4492	117.54	0.7995	0.3610	83.28	0.8858	0.4199	135.61	0.8000
CoordFill	0.6829	235.22	0.6723	—	—	—	—	—	—
SDM	0.4555	39.69	0.9103	0.3780	54.24	0.9155	<u>0.4390</u>	53.21	0.9114
RP (w/o LKA)	<u>0.4474</u>	<u>32.18</u>	<u>0.9258</u>	<u>0.3732</u>	45.91	0.9289	0.4414	47.33	0.9135
RP (w/o DWI)	0.5920	110.48	0.8225	0.3757	<u>44.72</u>	<u>0.9304</u>	0.4422	44.03	0.9232
RestorePath	0.4467	29.66	0.9317	0.3754	44.57	0.9308	0.4418	<u>44.37</u>	<u>0.9201</u>

3 Experiments

3.1 Datasets and implementation details

We used three breast pathology datasets: TCGA BRCA, Camelyon16, and BACH. For all datasets, we extracted training patches at a resolution of 256×256 pixels ($20\times$ magnification) and utilized the UNI [2] PFM for feature embedding. Training masks were equally sampled from GrandQC [25] and *failure modes* [23], including artifacts such as markings, tissue folding, bubbles, and tissue tears.

TCGA BRCA. We utilized 1,105 WSIs for which GrandQC masks are available, partitioning them into 884 for training and 221 for testing. Training patches were extracted exclusively from artifact-free regions as defined by GrandQC. For validation, we selected 620 regions (2048×2048 pixels) from the test set, with synthetic artifacts generated via *failure modes* to match the training distribution.

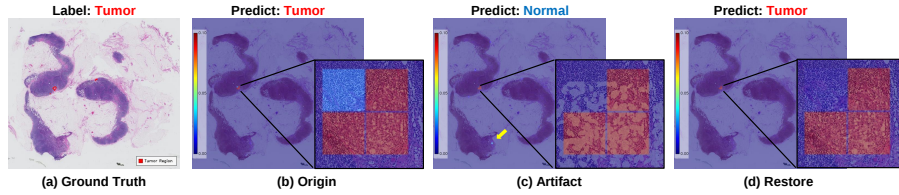
Camelyon16. This training dataset consists of 110 tumor and 160 normal cases. To address the Needle-in-a-haystack nature, tumor patches were extracted with a 32-pixel overlap within lesions, while normal regions were 20% subsampled. The test set includes 38 tumor regions (ranging from 512 to 4608 pixels) selected from 20 small-lesion cases (out of 50 tumor WSIs) and 20 normal regions (2048×2048 pixels) selected from 20 cases (out of 80 normal WSIs), all masked with synthetic bubbles to ensure uniform occlusion of the pathological context.

BACH. This dataset comprises 400 images across four classes: Normal, Benign, In Situ, and Invasive. We assigned 80 images per class for training and 20 for testing. Training patches were extracted with a 32-pixel overlap. Test artifacts were restricted to bubbles to evaluate the restoration of pathological context.

Implementation Details. To ensure dataset-specific optimization, we trained separate models for each dataset using hyperparameters according to the LRDM [5] framework. In addition, we set $H = 256$, $\tau = 0.3$, $\alpha = 2$, and $\delta = 10^{-6}$. For the AGS, s_{min} , s_{max} , and k are assigned values of 2, 7, and 0.99, respectively.

Table 2. Performance of MIL models on the Camelyon16 needle-in-a-haystack cases.

Methods	ABMIL		TransMIL		CLAM _{MB}	
	Acc.	AUC	Acc.	AUC	Acc.	AUC
Origin	85.0 $\pm$ 1.6	89.6 $\pm$ 2.6	81.5 $\pm$ 5.6	86.2 $\pm$ 2.9	84.5 $\pm$ 2.4	88.5 $\pm$ 3.1
Artifact	73.5 $\pm$ 4.6	82.3 $\pm$ 2.7	67.5 $\pm$ 6.1	77.2 $\pm$ 4.6	72.5 $\pm$ 5.7	83.3 $\pm$ 3.4
D&D	55.0 $\pm$ 2.2	70.8 $\pm$ 5.0	53.5 $\pm$ 2.5	59.4 $\pm$ 6.2	53.0 $\pm$ 1.9	71.6 $\pm$ 4.2
Restored	84.0 $\pm$ 1.2	90.3 $\pm$ 1.9	78.5 $\pm$ 6.0	85.0 $\pm$ 3.1	83.0 $\pm$ 3.3	90.3 $\pm$ 2.1

**Fig. 4.** MIL attention heatmaps on Camelyon16. (a) Ground Truth, (b) original, (c) artifact, and (d) restored images. Artifacts (c) cause attention dispersion (yellow arrow) and a silent failure (misclassified as Normal). RestorePath successfully refocuses attention on the lesion, ensuring accurate tumor prediction.

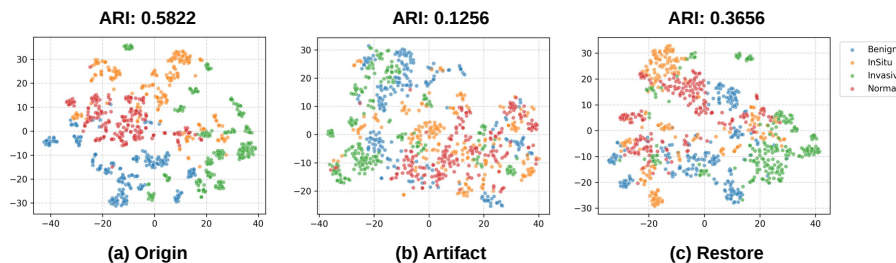
3.2 Performance evaluation

We evaluate quality via $LPIPS_M$, FID_C , and Emb Sim. $LPIPS_M$ is the LPIPS metric applied specifically to masked images. FID_C replaces standard Inception-based FID with a CONCH PFM to capture pathological distributions more effectively. Emb Sim measures the cosine similarity between CONCH features to verify diagnostic preservation. RestorePath is compared against high-resolution image inpainting models including LaMa [21], CoordFill [12], and an SDM [18] inpainting model applied with a random patch diffusion. Qualitative evaluations also demonstrate that RestorePath provides the most globally consistent results (Fig. 2). Ablation studies evaluating LKA and DWI are summarized in Table 1.

Discussion. Our framework achieves state-of-the-art performance in FID_C and Emb Sim, confirming superior preservation of pathological features. While RestorePath leads in LPIPS for TCGA-BRCA, LaMa excels on Camelyon16 and BACH. This discrepancy stems from artifact shapes. TCGA-BRCA features opaque artifacts like markings and folds that entirely obscure underlying structures, making the context-aware interpolation of DWI essential. Conversely, air bubbles in other datasets allow localized textures to remain visible through gaps between bubbles, favoring texture-based models like LaMa. Consequently, RestorePath is most impactful when surrounding context is the primary source of information.

Table 3. Performance of classification model on the BACH (2048×1536) datasets.

Methods	ResNet50		DenseNet121		MobileNetv2	
	Acc.	AUC	Acc.	AUC	Acc.	AUC
Origin	86.2 ± 1.8	96.0 ± 0.3	86.8 ± 2.6	96.4 ± 0.6	87.0 ± 1.7	96.5 ± 0.1
Artifact	61.0 ± 3.7	88.3 ± 1.1	68.0 ± 7.1	90.6 ± 2.6	73.2 ± 4.7	91.2 ± 0.6
D&D	64.2 ± 4.1	89.3 ± 1.1	69.8 ± 7.8	91.5 ± 2.6	73.0 ± 3.9	92.1 ± 0.7
Restore	75.5 ± 1.3	94.4 ± 0.7	78.0 ± 2.8	94.0 ± 0.7	79.8 ± 2.2	93.9 ± 0.7

**Fig. 5.** t-SNE and ARI analysis on the BACH. (a) Origin: Baseline clusters with $ARI = 0.5822$. (b) Artifact: Reduced class separability with $ARI = 0.1256$. (c) Restore: RestorePath improves the ARI to 0.3656 by mitigating artifact-induced distortions.

3.3 Downstream tasks

Multiple Instance Learning (Camelyon16). We evaluated lymph node metastasis prediction using ABMIL [10], TransMIL [20], and CLAM [14] with feature embeddings extracted from 256×256 pixel patches at $10\times$ magnification. After 5-fold cross-validation on the training WSIs, models were tested on artifact-affected and restored WSIs. For D&D strategy affected patches were zero-filled (black). For ABMIL, attention heatmaps were generated to visualize the restoration effect (Fig 4). While Artifact and D&D cases showed significant drops from the Origin, RestorePath achieved performance within the standard deviation of the Origin slides, with ABMIL and CLAM even showing slight improvements in AUC (Table 2). This validates restoration as superior to conventional strategies.

Image Classification (BACH). We trained ResNet50 [6], DenseNet121 [9], and MobileNet-v2 [19] via 5-fold cross-validation on 256×256 patches. Hard-voting inference was performed across Origin, Artifact, D&D, and Restoration sets. The t-SNE analysis revealed that restoration significantly recovered class separability. Adjusted Rand Index (ARI) scores improved from 0.1256 in the artifact set to 0.3656 after restoration, which is a significant recovery toward the origin score of 0.5822 as illustrated in Fig 5. RestorePath recovers nearly half of the performance loss observed in Artifact and D&D cases. This maintains superior diagnostic integrity compared to traditional QC procedures (Table 3).

4 Conclusion

We propose RestorePath, a diffusion-based patch-to-global framework for globally consistent megapixel-scale artifact restoration in WSIs. Our model based on inpainting LDM with LKA blocks conditioned on PFM embeddings. For high-resolution processing, we employ random patch diffusion with DWI and AGS for context-aware restoration leveraging surrounding artifact-free tissue. Benchmarks and downstream experiments demonstrate state-of-the-art performance, outperforming both the direct use of artifact images and conventional D&D strategies. RestorePath serves as an advanced quality control tool to ensure diagnostic integrity in computational pathology.

Acknowledgements. This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korean government (MSIT) (Nos. RS-2026-25475860, RS-2025-02215813, and RS2024-00441029 for the Digital-Bio AI + X Global Innovative Talent Nurturing Project with Hands-on Experience). This work was also supported by the Commercialization Promotion Agency for R&D Outcomes (COMPA) grant funded by the MSIT (No. RS-2026-25539491). Additional support was provided by the High-Performance Computing Support Project, funded by the Government of the Republic of Korea (Ministry of Science and ICT) (Grant No. RQT-25-090105; S.A.).

Disclosure of Interests. The authors have no competing interests to disclose that are relevant to the content of this article.

References

1. Aversa, M., Nobis, G., Hägele, M., Standvoss, K., Chirica, M., Murray-Smith, R., Alaa, A.M., Ruff, L., Ivanova, D., Samek, W., et al.: Diffinfinite: Large mask-image synthesis via parallel random patch diffusion in histopathology. *Advances in Neural Information Processing Systems* **36**, 78126–78141 (2023)
2. Chen, R.J., Ding, T., Lu, M.Y., Williamson, D.F., Jaume, G., Song, A.H., Chen, B., Zhang, A., Shao, D., Shaban, M., et al.: Towards a general-purpose foundation model for computational pathology. *Nature medicine* **30**(3), 850–862 (2024)
3. Ehteshami Bejnordi, B., Veta, M., Johannes van Diest, P., Van Ginneken, B., Karssemeijer, N., Litjens, G., Van Der Laak, J.A., consortium, C., Hermsen, M., Manson, Q.F., et al.: Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *Jama* **318**(22), 2199–2210 (2017)
4. Farahani, N., Parwani, A.V., Pantanowitz, L.: Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathology and Laboratory Medicine International* pp. 23–33 (2015)
5. Graikos, A., Yellapragada, S., Le, M.Q., Kapse, S., Prasanna, P., Saltz, J., Samaras, D.: Learned representation-guided diffusion models for large-image generation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 8532–8542 (2024)
6. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 770–778 (2016)

7. He, Z., He, J., Ye, J., Shen, Y.: Artifact restoration in histology images with diffusion probabilistic models. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 518–527. Springer (2023)
8. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems* **30** (2017)
9. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708 (2017)
10. Ilse, M., Tomczak, J., Welling, M.: Attention-based deep multiple instance learning. In: International conference on machine learning. pp. 2127–2136. PMLR (2018)
11. Janowczyk, A., Zuo, R., Gilmore, H., Feldman, M., Madabhushi, A.: Histoqc: an open-source quality control tool for digital pathology slides. *JCO clinical cancer informatics* **3**, 1–7 (2019)
12. Liu, W., Cun, X., Pun, C.M., Xia, M., Zhang, Y., Wang, J.: Coordfill: Efficient high-resolution image inpainting via parameterized coordinate querying. In: Proceedings of the AAAI conference on artificial intelligence. vol. 37, pp. 1746–1754 (2023)
13. Lu, M.Y., Chen, B., Williamson, D.F., Chen, R.J., Liang, I., Ding, T., Jaume, G., Odintsov, I., Le, L.P., Gerber, G., et al.: A visual-language foundation model for computational pathology. *Nature medicine* **30**(3), 863–874 (2024)
14. Lu, M.Y., Williamson, D.F., Chen, T.Y., Chen, R.J., Barbieri, M., Mahmood, F.: Data-efficient and weakly supervised computational pathology on whole-slide images. *Nature biomedical engineering* **5**(6), 555–570 (2021)
15. McGenity, C., Clarke, E.L., Jennings, C., Matthews, G., Cartlidge, C., Freduah-Agyemang, H., Stocken, D.D., Treanor, D.: Artificial intelligence in digital pathology: a systematic review and meta-analysis of diagnostic test accuracy. *NPJ digital medicine* **7**(1), 114 (2024)
16. Niazi, M.K.K., Parwani, A.V., Gurcan, M.N.: Digital pathology and artificial intelligence. *The lancet oncology* **20**(5), e253–e261 (2019)
17. Polónia, A., Eloy, C., Aguiar, P.: Bach dataset: Grand challenge on breast cancer histology images. *Med. Image Anal* **2019**, 563 (2019)
18. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)
19. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4510–4520 (2018)
20. Shao, Z., Bian, H., Chen, Y., Wang, Y., Zhang, J., Ji, X., et al.: Transmil: Transformer based correlated multiple instance learning for whole slide image classification. *Advances in neural information processing systems* **34**, 2136–2147 (2021)
21. Suvorov, R., Logacheva, E., Mashikhin, A., Remizova, A., Ashukha, A., Silvestrov, A., Kong, N., Goka, H., Park, K., Lempitsky, V.: Resolution-robust large mask inpainting with fourier convolutions. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. pp. 2149–2159 (2022)
22. Wang, C., Qu, K., Li, S., Yu, Y., He, J., Zhang, C., Shen, Y.: Artidiffuser: A unified framework for artifact restoration and synthesis for histology images via counterfactual diffusion model. *Medical Image Analysis* **102**, 103567 (2025)
23. Wang, N.C., Kaplan, J., Lee, J., Hodgins, J., Udager, A., Rao, A.: Stress testing pathology models with generated artifacts. *Journal of Pathology Informatics* **12**(1), 54 (2021)

24. Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., Stuart, J.M.: The cancer genome atlas pan-cancer analysis project. *Nature genetics* **45**(10), 1113–1120 (2013)
25. Weng, Z., Seper, A., Pryalukhin, A., Mairinger, F., Wickenhauser, C., Bauer, M., Glamann, L., Bläker, H., Lingscheidt, T., Hulla, W., et al.: Grandqc: A comprehensive solution to quality control problem in digital pathology. *Nature Communications* **15**(1), 10685 (2024)
26. Yellapragada, S., Graikos, A., Triaridis, K., Prasanna, P., Gupta, R., Saltz, J., Samaras, D.: Zoomldm: Latent diffusion model for multi-scale image generation. In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. pp. 23453–23463 (2025)
27. Zhang, Y., Sun, Y., Li, H., Zheng, S., Zhu, C., Yang, L.: Benchmarking the robustness of deep neural networks to common corruptions in digital pathology. In: *International conference on medical image computing and computer-assisted intervention*. pp. 242–252. Springer (2022)