Whole-Slide Image Compression and On-Demand Viewing using Reference-Based Super-Resolution

Wenbo Yang, Seungho Shin, Richard Y. Zhu, Zhou Wang
University of Waterloo
200 University Avenue West Waterloo, ON, Canada N2L 3G1
{w243yang,s7shin,ry2zhu,zhou.wang}@uwaterloo.ca

Abstract

The transition from glass slide to digital pathology has been hindered by the vast storage and transmission requirements of archiving and remote-accessing whole-slide images (WSIs): a small regional hospital may generate hundreds to thousands of gigabytes of WSIs per day. Neural compression methods show good promise but suffers from potential hallucinations. Here we examine and benchmark modern statistical compression methods and develop a WSI archival pipeline that incorporates both traditional and machine learning techniques, achieving 93% reduction in space compared to conventional JPEG compression. The proposed pipeline is designed to be fallback-safe: a computer without machine-learning capabilities can still view archived WSIs without content alteration, while substantially more details at target regions on-demand may be recovered by our real-time reference-based super-resolution upsampling method.

1. Introduction

The field of digital pathology was initiated by the first automated high-resolution whole-slide image (WSI) system near the beginning of the millennium. Digital pathology and WSIs eliminate constant microscope observation and enable flexible image signal processing (ISP) and digital archiving methods [7, 8]. Combined with advances in computational and storage equipment, digital pathology's advantage have made it increasingly adopted by hospitals and health networks for cancer diagnosis [6, 9].

However, digital pathology's adoption is hindered by the storage and transmission requirements for archiving and remote-accessing WSIs; for example, a small regional hospital typically generates hundreds to thousands of gigabytes of WSIs per day. Current storage burdens stem from suboptimal JPEG compression and redundant information storage (typically only 10%-20% of patches in a WSI are critical to diagnosis). While artificial neural network (ANN)-based compression shows strong potential, the medical community has limited trust in it due to concerns about hallucinations that might generate non-existent features and mislead diagnosis.

To address these challenges while maintaining pathologists' trust, we propose a WSI archiving and viewing solution that is fallback-safe: computers without machine-learning (ML) capabilities can still decode and view WSIs without altered content (though at potentially reduced quality). To improve compression rate (CR), we benchmarked modern image compression algorithms including WebP [24], JPEG XL [2], and AVIF [4] on various settings to identify optimal methods and parameters for compressing WSI patches. Building on these results, we propose a WSI archiving pipeline that incorporates the optimal compression settings found, eliminates white spaces and stores regions not critical to diagnosis at reduced scale. The combined archiving solution achieves 92.7% space reduction compared to conventional JPEG compression.

A pathologist may need to view archived WSIs for reviewing, regulatory purposes, or retrospective studies. For regions stored at reduced scale, diagnostically non-critical details may be lost, creating annoyance to pathologists. However, the details can be reconstructed through super-resolution (SR) methods using high-resolution patches as references. To provide seamless viewing, we developed viewer software that performs "lazy" real-time SR: upsampling only the necessary on-screen regions.

Our contributions can be summarized as follows:

- A comprehensive benchmark of modern image compression algorithms on whole-slide images (WSIs) with practical recommendations for clinical deployment.
- A practical WSI archival pipeline that balances readability, clinical trust, and storage and computational efficiency.
- An efficient software implementation for on-demand viewing of archived WSIs with an intuitive user interface, seamlessly integrated into the proposed pipeline.

2. Technical Background

This section provides an overview of key technologies in the Whole-Slide Image (WSI) compression pipeline based on the literature and our experience.

2.1. Whole-Slide Image Compression

Whole-Slide Images (WSIs) are high-resolution scans of histological or cytological slides used in digital pathology. In typical clinical settings, histology slides are scanned at 40x magnification (around 0.2 microns per pixel [14]), resulting in tiled images of 1–30 gigapixels [19, 23], occupying tens of gigabytes when uncompressed [14], and 0.5–6 GB when compressed with JPEG (Quality=90). A hospital in a small city typically generates hundreds of WSIs per day, creating significant storage requirements.

Several recent efforts address this storage burden. Some neural network-based methods [5, 19] focus on facilitating downstream tasks such as classification or segmentation without focus on perceptual quality, which only allows low-quality reconstruction of the original WSI or no reconstruction at all. Recently, a lossless WSI compression framework [17] was proposed, but was impractical for archival use due to its compression ratio (CR) being lower than currently accepted lossy method.

Lossy neural compression methods optimizing for perceptual quality have been explored through autoencoders (AEs). VQ-VAE2 [18] and diffusion models' AEs shown reasonable performance on WSI compression, but still underperformed JPEG in PSNR and SSIM [22], while having other practical drawbacks [14, 23].

Aside from limited perceptual quality in reconstruction, neural compression methods receive limited trust from pathologists due to their black-box nature. In response, some non-neural methods have been proposed. Faghani et al. [7] proposed densely packing non-empty regions of a WSI into a JPEG image with smaller dimensions. While this method can be incorporated into our framework, its significant performance improvement is only observed on WSIs with large blank regions. Additionally, AVIF [4] can reduce the method's space conservation by up to 25x. Helin et al. [10] measured WSI compression performance of JPEG 2000, an already-superseded codec, in a setting resembling practical use cases, except having a target quality (JPEG Quality=80) unacceptable to some pathologists. A comprehensive benchmark of modern codecs is warranted.

2.2. Super-Resolution for Histology

Several studies [1, 13] found super-resolution (SR) architectures designed for natural images (e.g., ESRGAN [21] and EDSR [16]) perform well on WSIs. Pathologists confirmed that SRed images are acceptable for diagnosis [1, 13]. However, in clinical settings, relevant high-resolution (HR) patches are often available. These single-image SR methods do not utilize such information.

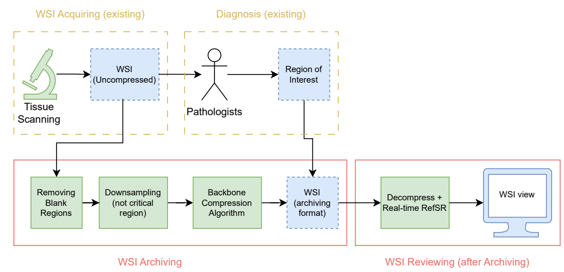


Figure 1. The proposed pipeline for whole-slide image (WSI) acquisition, compression and on-demand viewing. Our contributions are focused on the latter two stages.

2.3. Modern Statistical Image Compression

The JPEG compression standard was proposed in 1992 [20] and has been widely used since and remains the most compression algorithm for WSIs. WebP is an open format proposed by Google in 2010 [24], which provides around 30% increase in compression ratio (CR) compared to JPEG and has since gained wide support on all mainstream browsers [3]. In this decade, some new image compression standards have been proposed. JPEG XL features higher CR than JPEG, while providing some backward compatibility [2]. However, Google dropped Chrome's support in 2023 [3], despite being one of its main proponents. HEIC is the image format based on HEVC [11], which has yet to be supported by major web browsers except Safari [3]. In contrast, AVIF [4], backed by AOMedia, is widely supported in modern browsers [3]. It reportedly has better compression performance than JPEG XL when the target quality is lower than MS-SSIM 0.990 [2].

2.4. Reference-Based Super-Resolution

Conventional single-image super-resolution (SISR) models receive a low-resolution (LR) image as input and predict a high-resolution (HR) image as output. However, in real-world applications, HR images of related contents are often available, which provide additional details for reconstruction, leading to the development of reference-based super-resolution (RefSR) models. While many RefSR models for natural images have been proposed in recent years, most use C^2-matching [12] as their baseline.

3. Methodology

3.1. Overview

Under existing settings, pathology slides are first scanned through high-resolution scanners to obtain whole-slide images (WSIs) that are uncompressed or compressed with a quick but sub-optimal algorithm. The results are then sent to a pathologist for analysis, during which they label diagnosis-relevant regions (regions of interest, RoI). According to microscopy technologists we interviewed, the region of interest (RoI) typically occupies approximately 10-20% of the total tissue area.

Figure 1 depicts our proposed WSI archiving and on-demand viewing procedure after WSI acquisition and diagnosis. To preserve fallback viewability on systems without hardware for deep-learning (DL) models and maintain trust among pathologists, we use statistical image compression methods and bicubic downsampling for storage reduction. When a WSI is ready for archival storage, our method removes blank tiles and downsamples regions outside the RoI to a lower resolution (by 4x). It then compresses the WSI using a backbone statistical image compression algorithm selected through a systematic benchmark. When a pathologist needs to review an archived WSI, the compressed WSI is first decompressed and displayed in our proposed user interface. When a region is zoomed in beyond its stored resolution, a reference-based super-resolution (RefSR) method upsamples the region in real time.

The rest of this section details the selection of compression algorithms (Sec. 3.2) and the implementation of real-time viewing for archived WSIs (Sec. 3.3), including the user interface and RefSR method usage.

3.2. Selection of Compression Algorithm

We conducted experiments on 512×512 tissue patches from six WSIs (that contains around 88k patches) provided by Huron Digital Pathology. JPEG compression (Quality = 90), the most commonly used method in clinical settings, is first applied to the dataset to obtain a reference baseline. We then use exhaustive search to obtain the optimal parameter group for each image patch for each tested compression algorithm. The optimal parameter group produces the smallest file size while maintaining at least the same quality as JPEG compression (SSIM ≈ 0.938). Benchmark results and optimal parameter statistics are reported in Sec. 4.1.

After the benchmark and optimal parameter group search, we selected AVIF as the underlying compression algorithm in our pipeline, due to its superior performance and wide adoption. We then measure the space reduction of resulting WSIs after adding each proposed improvement to our pipeline. Results are shown in Sec. 4.2.

3.3. Real-Time Reference-Based Super-Resolution for On-Demand Viewing

Since patches in our archived WSIs are stored at varying scales, existing viewing systems cannot be used. We implement a custom viewer (Fig. 2) that supports zooming, panning, and RoI selection and viewing for use by pathologists and technicians to view archived WSIs. Notably, WSI

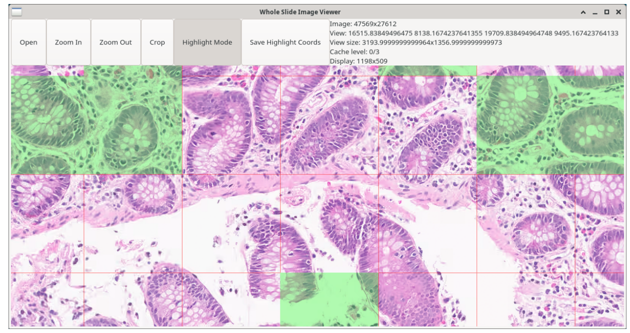


Figure 2. The UI of the proposed WSI viewer software. Regions highlighted in green are simulated RoI stored at full resolution, while the rest are stored downsampled.

patches outside the RoI are stored at lower resolution to preserve storage space, which reduces perceptual quality for details that are not critical to diagnosis. To restore perceptual quality, we retrain the C^2 matching model [12] on the SegPath dataset [15] and use it as the SR backbone. The backbone is distributed with the viewer software and is to be deployed on the client side. When a user zooms into a downsampled region beyond its stored resolution, the backbone RefSR model is called to upsample the region using the nearest high-resolution patch (in the RoI) as reference.

4. Experiments and Results

4.1. Benchmark of Compression Algorithms

We obtained six WSIs (that contains around 88k patches) from Huron Digital Pathology for benchmarking modern image compression algorithms. Since space overhead created by white spaces is handled separately in our pipeline, we exclude them from the search of optimal parameters for each compression algorithm. We first filter out empty (white) patches and patches that contain significant portions of background, leaving 25,849 patches of 512×512 pixels. To establish a baseline, we compress the patches using standard JPEG compression (Quality = 90), which yields an average SSIM of 0.938 and a compression rate (CR) of 12.14.

Due to the storage constraints for exhaustive search, we randomly selected 500 out of the $25,849$ patches for benchmarking WebP, JPEG XL, and AVIF. Each algorithm is tested on all combinations of frequently adjusted parameters, and the optimal parameter group for each patch is selected and recorded. The "optimal" parameter group is defined as the group that yields the smallest file size while maintaining at least the same quality as the JPEG baseline (SSIM $= 0.938$). Table 1 shows that AVIF outperforms the other two algorithms in terms of space reduction and has better parameter consistency between the two better-performing algorithms. In addition, we found the following group of

Table 1. Comparison of compression performance of different algorithms on WSI patches. The SSIM score for JPEG is used as the target quality for other algorithms. The space saving is calculated against JPEG. The largest group indicates the percentage of patches whose optimal parameter group is the most popular one among all patches. The first-3 groups indicate the percentage of patches whose best three parameter groups are among most popular one. Size of the groups indicates parameter consistency.

Comp. Algo.	Mean SSIM	Space saving	Largest group	First-3 groups	Time/ patch
JPEG	0.938	-	-	-	15.7 ms
WebP	0.940	35.2%	**26.2%**	**71.0%**	62.7 ms
AVIF	0.941	**71.5%**	*20.8%*	*27.0%*	140.5 ms
JPEG XL	0.940	*62.8%*	14.6%	24.0%	129.8 ms

Table 2. Comparison of space reduction provided by each step in the proposed archiving pipeline.

Backbone algorithm	Removing blank	Downsampling non-RoI	Space saving
JPEG	-	-	-
JPEG	✓	-	25.80%
AVIF	-	-	79.07%
AVIF	✓	-	80.09%
AVIF	✓	✓	92.75%

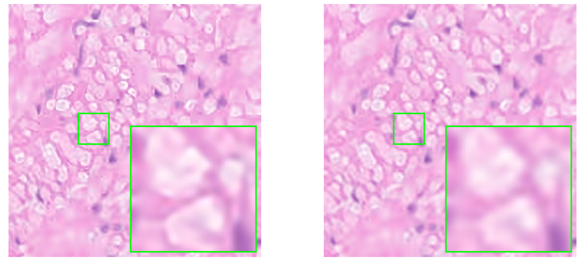
AVIF parameters works well on most patches in terms of CR and SSIM: `'aq_mode': 3, 'end_usage': q, 'cq_level': 30, 'sharpness': 0`. Further considering the wide adaptation of AVIF (Sec. 2.3), we select AVIF as the underlying compression algorithm in our proposed pipeline.

4.2. Performance of the Final Archiving Settings

Based on interviewed domain experts' assumptions, we randomly selected 20% of tissue-containing patches as the RoI. We then sequentially add each step of our proposed archiving pipeline to the WSIs and measure the overall space reduction compared to JPEG. Table 2 shows that downsampling non-RoI patches and AVIF significantly reduce storage size. On the other hand, removing white spaces contributes significantly to reducing JPEG's storage size but has limited effect on AVIF due to its optimised codec design. However, white space removal is still included in our pipeline to speed up the compression process.

4.3. Reference-Based Super-Resolution

The C^2-matching implementation in Zhang et al. [25]'s code is used for training. We train the model on the SegPath [15] for 75,000 iterations. Figure 3 shows a comparison between the upsampling results of the trained RefSR model and bicubic upsampling on a downsampled and AVIF-compressed WSI patch. The RefSR model reconstructs sharper textures and more details than bicubic upsampling, providing less visual distraction to pathologists. In addition, RefSR on each 512x512 patch only requires 0.12 second on a RTX 3080 GPU, which is fast enough for real-time zooming and viewing on a FHD screen. In addition, in our field test, the model achieves real-time zooming on a 4K screen on a typical workstation supplied by microscopy scanner vendors (equipped with an NVIDIA A6000 GPU).



(a) RefSR (b) Bicubic Upsampling

Figure 3. Upsampling results of RefSR and bicubic upsampling on the same downsampled AVIF-compressed WSI patch. RefSR provides sharper results with more details.

5. Conclusion and Future Works

In this paper, we conducted a comprehensive benchmark of statistics-based modern image compression algorithms and proposed a practical WSI archival and viewing pipeline based on the benchmark results. The archival pipeline consists of non-machine-learning methodologies that reduced the storage size by 92.75% compared to conventional JPEG compression while maintaining perceptual quality, fallback-safety, and clinical trust. The viewer software utilizes real-time reference-based super-resolution to reconstruct diagnostically non-critical details on-demand, enabling seamless viewing of archived WSIs. The drastic space saving and the efficiency of the proposed pipeline warrant its practical adoption in medical institutions.

Despite the pipeline's advantages in practice, some future improvements can be made: The RefSR model requires the diagnostically critical reference patches to be labelled. Currently, these regions of interest (RoIs) are manually selected by pathologists, which can sometimes be time-consuming. Automating the selection process using machine learning models could further streamline the workflow. Alternatively, a low-birate vector can be generated from the original WSI patch prior to downsampling, creating a reduced reference to guide super-resolution. This method has potentials to eliminate the need for RoI labelling. As the medical community gains greater confidence in neural compression methods, techniques such as implicit neural representations (INRs) could be explored, as they can effectively exploit their WSIs' redundancies.

Acknowledgments

This work is supported in part by the Ontario Research Fund - Research Excellence (ORF-RE) program, and the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (DG) and Canada Research Chair (CRC) Programs.

References

[1] Mehdi Afshari, Saba Yasir, Gary L Keeney, Rafael E Jimenez, Joaquin J Garcia, and Hamid R Tizhoosh. Single patch super-resolution of histopathology whole slide images: a comparative study. *Journal of Medical Imaging*, 10(1): 017501–017501, 2023. 2

[2] Jyrki Alakuijala, Jon Sneyers, Luca Versari, and Jan Wassenberg. JPEG white paper: JPEG XL image coding system, 2023. 1, 2

[3] Alexis Deveria, et al. Can i use... support tables for HTML5, CSS3, etc. 2

[4] AOMedia. AV1 image file format (AVIF) v1.2.0, 2025. 1, 2

[5] Agnes Barsi, Suvendu Chandan Nayak, Sasmita Parida, and Raj Mani Shukla. A deep learning-based compression and classification technique for whole slide histopathology images. *International Journal of Information Technology*, 16 (7):4517–4526, 2024. 2

[6] British Columbia Government News. New digital tools being tested to improve people's health, 2024. 1

[7] Shahriar Faghani, D Chamil Codipilly, Mana Moassefi, Prasad G Iyer, and Bradley J Erickson. Optimizing storage and computational efficiency: an efficient algorithm for whole slide image size reduction. *Mayo Clinic Proceedings: Digital Health*, 1(3):419–424, 2023. 1, 2

[8] Farzad Ghaznavi, Andrew Evans, Anant Madabhushi, and Michael Feldman. Digital imaging in pathology: whole-slide imaging and beyond. *Annual Review of Pathology: Mechanisms of Disease*, 8(1):331–359, 2013. 1

[9] Grand River Hospital. Grand river hospital celebrates $1m gift to launch digital pathology solution powered by local innovation, 2025. 1

[10] Henrik Helin, Teemu Tolonen, Onni Ylinen, Petteri Tolonen, Juha Napankangas, and Jorma Isola. Optimized jpeg 2000 compression for efficient storage of histopathological whole-slide images. *Journal of pathology informatics*, 9(1): 20, 2018. 2

[11] ISO/IEC. Information technology — high efficiency coding and media delivery in heterogeneous environments — part 12: Image file format, 2017. 2

[12] Yuming Jiang, Kelvin CK Chan, Xintao Wang, Chen Change Loy, and Ziwei Liu. Robust reference-based super-resolution via c2-matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2103–2112, 2021. 2, 3

[13] Aniwat Juhong, Bo Li, Cheng-You Yao, Chia-Wei Yang, Dalen W Agnew, Yu Leo Lei, Xuefei Huang, Wibool Piyawattanametha, and Zhen Qiu. Super-resolution and segmentation deep learning for breast cancer histopathology image analysis. *Biomedical Optics Express*, 14(1):18–36, 2022. 2

[14] Jason Keighley, Marc de Kamps, Alexander Wright, and Darren Treanor. Digital pathology whole slide image compression with vector quantized variational autoencoders. In *Medical Imaging 2023: Digital and Computational Pathology*, pages 344–353. SPIE, 2023. 2

[15] Daisuke Komura, Takumi Onoyama, Koki Shinbo, Hiroto Odaka, Minako Hayakawa, Mieko Ochi, Ranny Rahaningrum Herdiantoputri, Haruya Endo, Hiroto Katoh, Tohru Ikeda, Tetsuo Ushiku, and Shumpei Ishikawa. Restaining-based annotation for cancer histology segmentation to overcome annotation-related limitations among pathologists. *Patterns*, 4(2):100688, 2023. 3, 4

[16] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 136–144, 2017. 2

[17] Yu Mao, Jun Wang, Nan Guan, and Chun Jason Xue. Wise: A framework for gigapixel whole-slide-image lossless compression. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 29342–29351, 2025. 2

[18] Ali Razavi, Aaron van den Oord, and Oriol Vinyals. Generating diverse high-fidelity images with vq-vae-2. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2019. 2

[19] David Tellez, Jeroen van der Laak, and Francesco Ciompi. Gigapixel whole-slide image classification using unsupervised image compression and contrastive training. In *Medical Imaging with Deep Learning*, 2018. 2

[20] Gregory K Wallace. The jpeg still picture compression standard. *IEEE transactions on consumer electronics*, 38(1): xviii–xxxiv, 2002. 2

[21] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *The European Conference on Computer Vision Workshops (ECCVW)*, 2018. 2

[22] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 2

[23] Srikar Yellapragada, Alexandros Graikos, Kostas Triaridis, Zilinghan Li, Tarak Nath Nandi, Ravi K Madduri, Prateek Prasanna, Joel Saltz, and Dimitris Samaras. Pathology image compression with pre-trained autoencoders. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 442–452. Springer, 2025. 2

[24] James Zern, Pascal Massimino, and Jyrki Alakuijala. WebP Image Format. RFC 9649, 2024. 1, 2

[25] Lin Zhang, Xin Li, Dongliang He, Fu Li, Errui Ding, and Zhaoxiang Zhang. Lmr: A large-scale multi-reference dataset for reference-based super-resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 13118–13127, 2023. 4