OPEN ACCESS

Artificial Intelligence LETTER

The Innovation Informatics

Unlocking efficient gigapixel whole slide pathological image compression via diagnostic-aware attention guidance

Xiang Liu,[1,2] Fuqiang Chen,[1,2] Jiaye He,[4] Hairong Zheng,[1,3] and Wenjian Qin[1,3,*]

[1]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

[2]University of Chinese Academy of Science, Beijing 101400, China

[3]State Key Laboratory of Biomedical Imaging Science and System, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

[4]National Innovation Center for Advanced Medical Devices, Shenzhen 518055, China

*Correspondence: wj.qin@siat.ac.cn

Received: December 29, 2025; Accepted: July 6, 2026; Published Online: July 9, 2026; https://doi.org/10.59717/j.xinn-inform.2026.100053

Citation: Liu X., Chen F., He J., et al. (2026). Unlocking efficient gigapixel whole slide pathological image compression via diagnostic-aware attention guidance. *The Innovation Informatics* **2**:100053.

Dear Editor,

Gigapixel-level whole slide imaging (WSI) is fundamental for quantitative cancer analysis and biomarker discovery, but its massive data volume creates critical storage and transmission bottlenecks. Conventional monomodal compression techniques struggle to balance compression efficiency and image quality, particularly failing to account for WSI's unique characteristics—including ultra-high resolution, irregular blank backgrounds, and variable diagnostic relevance of tissue regions. To address this, we propose an attention-guided hierarchical compression strategy (AttHSCS) that achieves high-fidelity and efficient compression of large-size WSIs by prioritizing diagnostically critical regions. Experiments on liver, breast, and colon pathology datasets demonstrate that AttHSCS significantly improves compression ratios while maintaining high segmentation performance, providing a high-efficiency and practical compression tool for digital pathology applications in common solid tumor scenarios.

INTRODUCTION

Whole slide imaging enables simultaneous multi-pathologist review, remote diagnosis, and computer-aided diagnosis (CAD), but its cellular-level precision leads to terabyte-scale data volumes per patient-especially with emerging 3D pathology technologies. Biomedical images exhibit semantic redundancy, a trait that can be leveraged to enhance compression efficiency, yet existing deep learning-based compression methods (autoencoders, recurrent neural networks, generative adversarial networks) focus on universal image restoration rather than WSI-specific features. Most WSI compression studies also neglect the varying diagnostic value of tissue regions, wasting resources on non-critical areas like blank backgrounds. To fill this gap, we designed AttHSCS, a flexible framework that segments WSIs into high- and low-diagnostic-value regions via attention maps and applies differentiated compression parameters to each region.[1-3]

MATERIALS AND METHODS

Directly processing gigapixel WSIs causes memory exhaustion and degraded recovery quality. We adopted a chunked processing strategy, dividing WSIs into 1024 × 1024 patches for independent compression and storage. A metadata file records each WSI's resolution, shape, and patch positions, ensuring accurate reconstruction during decompression. End-to-end processing of a typical gigapixel WSI averages ~18 minutes on a single NVIDIA RTX 3090 GPU, with attention map generation contributing < 5 % of total computation.

AttHSCS generates an attention map to score each patch's diagnostic value. Patches with zero attention scores are classified as low-value; the remaining patches are sorted by score, with the top 50 % designated as high-value. High-value patches use high-quality compression parameters (to preserve diagnostic integrity), while low-value patches use low-quality parameters (to maximize compression efficiency). The framework supports flexible selection of attention generation strategies (deep learning, information entropy, pathologist annotations) and compression methods (VAE-based deep learning,[4] JPEG, WebP).

We developed an entropy-based attention strategy for superior universality (avoiding the need for labeled training data). First, WSIs are split into patches, and blank background patches (≥ 80 % white pixels) are filtered out (attention score = 0). For remaining patches, we convert RGB images to grayscale and calculate information entropy using pixel intensity histograms:

$$entropy = -\sum_{i=0}^{255} histogram_i * log_2(histogram_i + e^{-12}) \quad (1)$$

Entropy values are normalized via Z-Score and logistic function to generate final attention scores (0–1), which are mapped back to patch positions to form the WSI attention map. Tumor regions (high variability) exhibit high entropy, while stroma and background regions (low variability) show low entropy-enabling precise identification of critical diagnostic areas.

We evaluated AttHSCS on three public datasets: PAIP2020 colorectal cancer (47 training, 31 validation WSIs, 40× magnification),[6] PAIP2019 liver cancer (50 training, 10 validation WSIs, 20× magnification),[5] and Tiger breast cancer (93 labeled WSIs, 40× magnification).[7] The bmshj2018-factorized VAE-based compression algorithm was integrated into the framework, trained with four lambda values (0.0001, 0.0005, 0.0025, 0.0125) to balance compression rate and image quality. Experiments assessed two core metrics: (1) compression performance (PSNR, SSIM, MS-SSIM, BPP, compression ratio); (2) impact on downstream segmentation tasks (Dice coefficient, IOU, Accuracy) using a U-Net model.

RESULTS

The AttHSCS achieves higher compression ratios than the hyperprior baseline across all anatomical sites and parameter combinations. For example, on colon WSIs, AttHSCS (λ = 0.0125 + 0.0025) reaches a compression ratio of 154, compared to 110 for the hyperprior method (λ = 0.0125). While AttHSCS shows slightly lower average PSNR than the baseline (due to aggressive compression of low-value background patches), MS-SSIM and SSIM metrics are comparable. Visualization of decompressed images confirms that AttHSCS preserves the quality of high-value tumor regions even at high compression ratios (up to 332× for colon WSIs).

AttHSCS outperforms the hyperprior baseline on downstream tumor segmentation tasks across all three datasets, with particularly strong gains at high compression rates. For colon WSIs, AttHSCS (λ = 0.0125 + 0.0025) achieves a compression ratio of 154 while maintaining segmentation metrics (Dice, IOU, Accuracy) equivalent to or better than the hyperprior method at a lower ratio (110×). At extreme compression (λ = 0.0005 + 0.0001), AttHSCS segmentation results significantly surpass the hyperprior baseline, with minimal false-positive regions. Statistical analysis shows that AttHSCS preserves 65 %–79.7 % of tumor regions while only retaining 17.2 %–32.6 % of total WSI area—validating its ability to prioritize diagnostically critical regions.

Tumor segmentation is the fundamental WSI quantitative task, requiring intact fine pathological features (cell boundaries, nuclear morphology, tissue structure). AttHSCS's superior performance at ultra-high compression ratios validates its effective retention of core diagnostic features, the basis for all downstream WSI analyses, and it generalizes well to other clinical tasks: (1) Tumor classification/grading: high-quality tumor region compression preserves key morphological features (glandular distortion, cellular atypia) for accurate grading; (2) Biomarker prediction (MSI, Ki-67, IDH): stable MS-SSIM/SSIM/PSNR in tumor regions ensures pixel-level feature integrity for prediction. Its modular design enables direct downstream model replace-

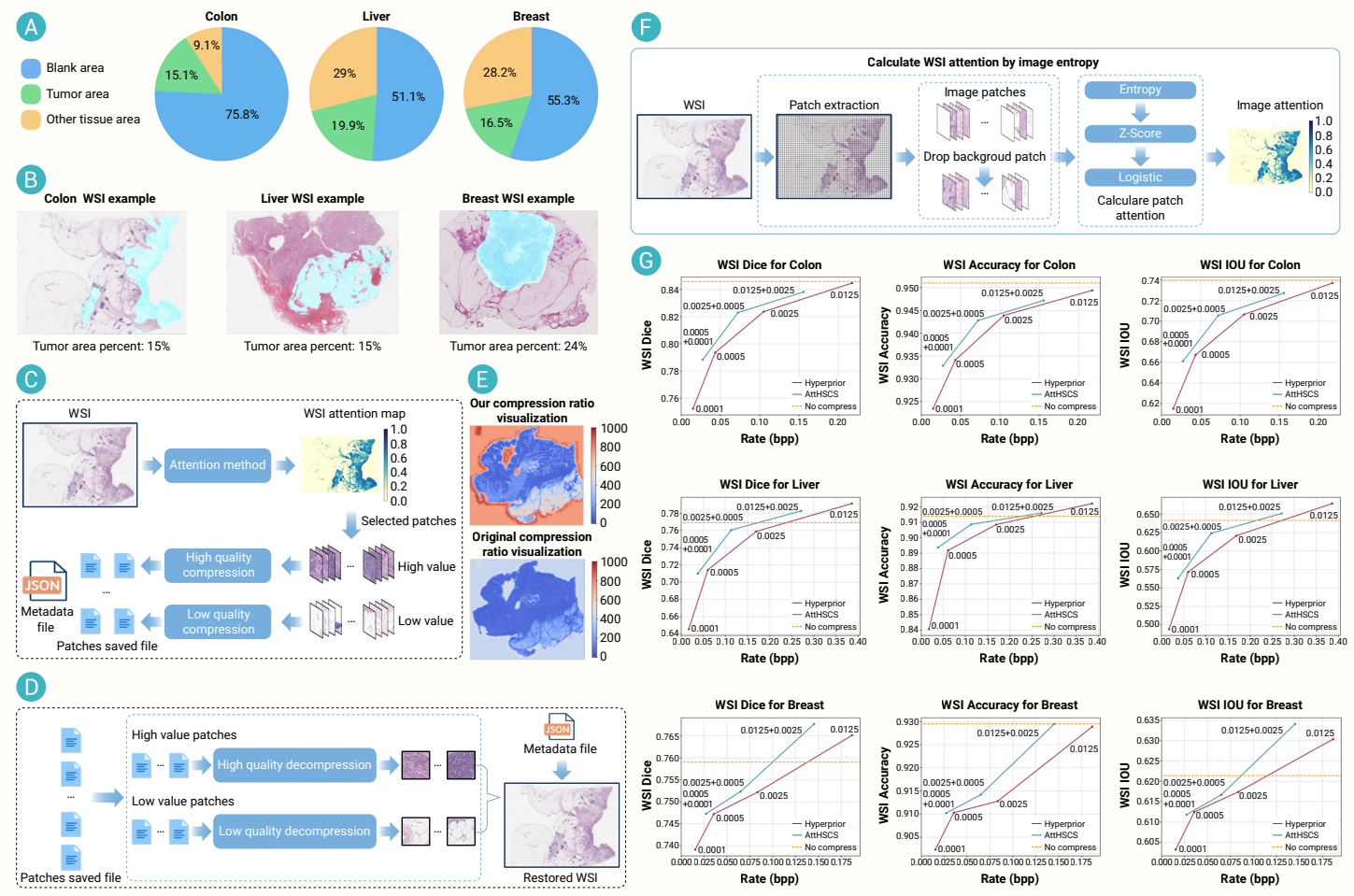

Figure 1. Overview of the AttHSCS model and experimental results (A) Distribution of blank background, tumor, and other tissue regions across three anatomical sites. (B) WSI example from the datasets with tumor areas highlighted in blue. (C) Workflow of AttHSCS hierarchical compression: attention map generation → high/low-value region segmentation → differentiated compression parameter application. (D) Decompression framework (inverse process of compression). (E) Compression ratio visualization of different regions under different methods (top: AttHSCS; bottom: baseline method), showing AttHSCS achieves higher compression ratios for background and tissue regions. (F) WSI attention map calculation strategy based on image information entropy. (G) Performance comparison of AttHSCS (Dice, IOU, Accuracy) across three anatomical sites. X-axis: BPP (compression rate, left = higher compression); point values: compression parameter λ.

ment without modifying the core compression framework, adapting to diverse clinical needs.

We further compared AttHSCS with JPEG 2000 to highlight its clinical value: (1) Compression efficiency: JPEG 2000's global uniform compression yields only 67x for WSIs, while AttHSCS's hierarchical strategy achieves 154x for colorectal cancer WSIs (up to 332x in extreme cases); (2) Diagnostic feature preservation: JPEG 2000 causes block effects, edge blurring and texture loss in tumor regions at high compression, while AttHSCS maintains baseline MS-SSIM/SSIM and excellent segmentation performance; (3) WSI applicability: JPEG 2000 lacks targeted chunked processing and metadata recording, leading to memory exhaustion and incomplete decompression for gigapixel WSIs, while our chunked strategy is compatible with mainstream digital pathology systems and ensures precise reconstruction. JPEG 2000 is suitable for low-ratio WSI preview, while AttHSCS addresses core needs of large-scale WSI storage, remote diagnosis and multi-center data sharing. Notably, AttHSCS's hierarchical strategy is orthogonal to base codec advances; replacing bmshj2018 with newer neural codecs can further boost performance within the same framework.

DISCUSSION

In this study, diagnostic value is defined by tumor tissue regions (high entropy, high texture variability) as core high-diagnostic-value regions, the focus of clinical cancer pathology diagnosis and quantitative analysis. Stromal regions and tumor microenvironment (TME) are low-entropy non-tumor regions with critical prognostic significance, and aggressive compression may limit tumor-stroma interaction studies. Our framework only applies aggressive compression to blank background patches (≥ 80 % white pixels); stromal/TME regions are low-value but non-blank patches with non-zero entropy scores, retaining basic compression quality (PSNR > 30 dB, SSIM > 0.90), preserving sufficient integrity for nuclear-cytoplasmic ratio calculations and stromal assessments. AttHSCS's modular design allows fine-tuning of low-value region compression parameters to enhance stromal/TME feature preservation for specific clinical research needs.

The 50 % high-value patch threshold was identified as the Pareto optimal solution via pre-experiments (10 %, 30 %, 50 %, 70 %, 90 %) evaluating compression ratio and segmentation Dice coefficient: increasing the threshold to 50 % significantly improves segmentation performance, while further increases cause a sharp drop in compression ratio with marginal gains. This threshold adapts well to heterogeneous tumor burdens: for high-burden samples, it covers most core tumor regions (low-value region dominated by blank background) to maintain efficiency; for low-burden samples, it includes auxiliary diagnostic non-tumor tissues (glands, morphologically relevant stroma) into high-value regions, avoiding rare tumor patch misclassification and ensuring feature preservation, while achieving a far superior compression ratio to the baseline. Flexible threshold tuning (60 %–70 % for ultra-low-burden rare tumors) is supported without modifying the core framework, verifying parameter robustness. Sensitivity analysis shows ultra-low-burden samples (< 10 % tumor area) maintain Dice within 2 % of baselines, while ultra-high-burden samples (> 50 %) sustain ratios above 130×.

AttHSCS's entropy-based attention strategy is based on the correlation between high diagnostic importance (core tumor regions) and high texture variability (high entropy), valid for the tested colorectal, liver and breast

cancers. However, this proxy has limitations in specific scenarios: diagnostically critical regions (mucinous adenocarcinoma mucus regions, homogeneous neuroendocrine tumor regions) may have low entropy, while non-diagnostic regions (tissue folding, staining unevenness, slide impurities) may show high entropy variability, leading to attention map misclassification. Three framework-compatible optimization strategies address such cases: (1) Fuse entropy with pathology-specific morphological features (cell density, nuclear-cytoplasmic ratio, glandular structure) via lightweight pre-trained models for correction; (2) Support manual attention map fine-tuning by pathologists for rare tumors and special samples; (3) Fine-tune entropy normalization parameters with small labeled datasets for low-entropy tumor types. Preliminary testing on mucinous samples shows morphological feature fusion recovers ~85 % of misclassified low-entropy patches, maintaining ratios above 120× with Dice within 3 % of standard cases.

A data-driven adaptive threshold strategy based on tumor burden estimation is proposed as a future optimization direction to enhance generalizability. This algorithm calculates the proportion of high-entropy tumor-related patches in WSIs from attention maps, then dynamically adjusts the threshold: increase to 60 %–70 % for ultra-low-burden samples (< 10 % high-entropy patches) to include all rare tumor patches, and decrease to 40 %–50 % for ultra-high-burden samples (> 50 % high-entropy patches) to maintain compression efficiency. Only a simple statistical module needs to be added after attention map generation, without changing the core chunked processing and hierarchical compression structure, enabling easy clinical implementation. The fixed 50 % threshold is robust for common solid tumors, and the adaptive strategy will extend AttHSCS's applicability to rare tumors and extreme tumor burden samples.

To demonstrate framework flexibility, we integrated the CLAM unsupervised deep learning method for attention map generation, which also outperformed the hyperprior baseline—confirming that AttHSCS's hierarchical design, not a specific attention strategy, is the key to its superiority. This modularity supports diverse clinical needs with interchangeable attention strategies.

CONCLUSION

AttHSCS is a flexible, efficient hierarchical compression framework that effectively alleviates the storage and transmission bottleneck of WSIs in digital pathology and provides a feasible technical solution. By prioritizing diagnostically critical regions via attention maps, it achieves significantly higher compression ratios than conventional methods in the tested colorectal, liver and breast cancer WSI scenarios while preserving downstream segmentation performance. The entropy-based attention strategy offers good universal applicability for common solid tumors without labeled training data, and the modular design supports diverse compression and attention methods. Limitations include the need for broader clinical validation and attention optimization for rare tumor types. Future work will integrate advanced codecs, adaptive thresholds, and real-time processing for enhanced clinical applicability.

REFERENCES

1. Ghaznavi F., Evans A., Madabhushi A., et al. (2013). Digital imaging in pathology: Whole-slide imaging and beyond. *Annu. Rev. Pathol. Mech. Dis.* **8**:331–359. DOI:10.1146/annurev-pathol-011811-120902
2. Nakagawa K., Moukheiber L., Celi L.A., et al. (2023). AI in pathology: What could possibly go wrong. *Semin. Diagn. Pathol.* **40**:100–108. DOI:10.1053/j.semdp.2023.02.006
3. Artuğer F. and Özkaynak F. (2023). Chaotic quantization based JPEG for effective compression of whole slide images. *Vis. Comput.* **39**:5609–5623. DOI:10.1007/s00371-022-02684-y
4. Zhu X., Song J., Gao L., et al. (2022). Unified multivariate gaussian mixture for efficient neural image compression. *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* **2022**:17612–17621. DOI:10.48550/arXiv.2203.10897
5. Kim Y.J., Jang H., Lee K., et al. (2021). PAIP 2019: Liver cancer segmentation challenge. *Med. Image Anal.* **67**:101854. DOI:10.1016/j.media.2020.101854
6. Kim K., Lee K., Cho S., et al. (2023). PAIP 2020: Microsatellite instability prediction in colorectal cancer. *Med. Image Anal.* **89**:102886. DOI:10.1016/j.media.2023.102886
7. TIGER: Grand Challenge (2022). https://tiger.grandchallenge.org (Accessed Nov-2022)

FUNDING AND ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Science and Technology's key research and development program (2023YFF0723400), Shenzhen-Hong Kong Joint Lab on Intelligence Computational Analysis for Tumor Imaging (E3G111), Guangdong Youth Talent program (2024TQ08A386) and the Youth Innovation Promotion Association CAS (2022365). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

X.L and WJ.Q conceived and designed the study. X.L and FQ.C handled the data preprocessing. X.L developed the experimental code, conducted the experiments, and obtained the results. Additionally, X.L and WJ.Q wrote the manuscript and generated the figures. FQ.C, JY.He HR.Zheng and WJ.Q made some optimizations and comments on the picture. WJ.Q supervised the work and obtained the funding. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

Hairong Zheng is the Editor-in-Chief of The Innovation Informatics and was blinded from reviewing or making final decisions on the manuscript. Peer review was handled independently of this member and their research group. The other authors declare no conflicts of interest.

ETHICAL STATEMENT AND PATIENT CONSENT

This study utilized publicly available, fully anonymized whole-slide image datasets from three grand challenge competitions: TIGER , PAIP 2020, and PAIP 2019. All datasets were originally collected with appropriate ethical approvals or waivers granted by the respective institutional review boards of the contributing institutions, and all patient data were fully de-identified prior to public release. As this study constitutes a secondary analysis of pre-existing, publicly available, anonymized data involving no direct interaction with human participants, no additional ethical approval was required.

DATA AND CODE AVAILABILITY

We used three publicly available datasets, which can be obtained from the following websites: https://tiger.grand-challenge.org/, https://paip2020.grand-challenge.org/, and https://paip2019.grand-challenge.org/. No internal patient data were specifically collected for this study. This study relies on retrospective analysis of anonymized whole-slide images.