

Magnification-Scheduled Tile Routing for Whole-Slide Tumor Grading

Authors

Jason Wong^{1*}, Michelle Chan², Kevin Lau³

Affiliations

Department of Pathology, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Queen Mary Hospital, Pokfulam Road, Hong Kong SAR, China

*Corresponding author: wong@hku.hk

Abstract

Whole-slide pathological images contain gigapixel-level visual information, but only a small proportion of tissue regions are directly relevant to tumor grading. Processing all tiles at the same magnification creates heavy computational burden and may delay clinical screening. This study investigates efficient whole-slide tumor grading through magnification-scheduled tile routing. We propose HistoRoute, a tile-level routing model that assigns different magnification paths according to tissue foreground ratio, nuclear density, glandular boundary irregularity, and tumor-region uncertainty. Low-risk stromal and background regions are processed at low magnification, suspicious epithelial regions are examined at medium magnification, and high-uncertainty tumor margins are retained for high-magnification transformer encoding. A multiple-instance learning classifier then aggregates routed tile features into slide-level predictions. The study used 7,840 whole-slide images from breast, colorectal, and gastric cancer cohorts, including 3.2 million annotated tissue tiles and 18,600 pathologist-marked tumor regions. HistoRoute examined only 28.4% of all candidate high-resolution tiles while maintaining slide-level diagnostic performance. For tumor grading, the macro-AUC reached 0.914, compared with 0.919 for the full-tile transformer and 0.872 for a uniform low-resolution model. In external testing across two hospitals, grade-level agreement with pathologist labels improved from 0.781 to 0.836 Cohen’s kappa compared with random tile sampling. The average slide-processing time decreased from 96.5 seconds to 38.7 seconds, and disk read volume was reduced by 52.1%. Small invasive foci below 0.5 mm were still detected with 89.3% sensitivity. These findings suggest that magnification-aware tile routing can reduce computational cost while preserving clinically important pathological structures.

Keywords: Whole-slide imaging; tumor grading; tile routing; computational pathology; efficient transformer; multiple-instance learning; medical image analysis

1. Introduction

Whole-slide images (WSIs) preserve gigapixel-level tissue information and have become an important data source for computer-assisted diagnosis, tumor grading, biomarker assessment, and

large-scale pathological research. A single WSI, however, may contain tens of thousands of image tiles when processed at diagnostic magnification, resulting in substantial computational, memory, and storage requirements. This problem becomes more pronounced in slide-level learning because only a small proportion of regions may contain diagnostically decisive morphology. Multiple-instance learning (MIL) has therefore been widely adopted to learn slide-level representations without requiring exhaustive tile-level annotations. Representative approaches such as CLAM use attention-based clustering to identify discriminative instances [1], DSMIL combines instance-level and bag-level classifiers to improve slide representation [2], and TransMIL models correlations among tissue instances through transformer-based aggregation [3]. More broadly, adaptive computational systems have shown that allocating heterogeneous subtasks to specialized processing components can improve both efficiency and output quality [4]. This principle is particularly relevant to WSI analysis, where different tissue regions may not require the same computational resolution. Background areas, adipose tissue, and relatively homogeneous stroma can often be characterized at low magnification, whereas tumor margins, atypical glands, poorly differentiated regions, and small invasive foci may require substantially greater spatial detail. Advances in self-supervised learning and hierarchical modeling have further improved feature extraction from histopathology images. CTransPath learns transferable histological representations through transformer-based self-supervised pretraining [5], while HIPT captures tissue information through hierarchical representations spanning local cellular structures and larger tissue contexts [6]. DTFD-MIL improves slide-level classification through pseudo-bag construction and feature distillation [7], and H2-MIL incorporates hierarchical information to better model relationships among tissue regions at different spatial scales [8]. These methods demonstrate that integrating local morphology with broader tissue architecture can improve pathological classification. Nevertheless, many pipelines still begin by densely extracting high-resolution tiles from most or all foreground tissue. Feature selection, attention weighting, or instance pruning is then performed only after expensive image decoding and feature extraction have already occurred. As a result, substantial computational resources may be consumed by regions that contribute little to the final grading decision.

Recent studies have attempted to reduce this redundancy through hard-instance mining, representative tile selection, adaptive sampling, and multi-scale analysis [9,10]. Such approaches are particularly important because diagnostically meaningful regions are often sparse and spatially heterogeneous. A small cluster of malignant cells or a narrow invasive boundary may occupy only a tiny fraction of the entire tissue area. Aggressive uniform sampling can therefore reduce computational cost but simultaneously increase the probability of excluding clinically important morphology [11,12]. SiiMIL addresses this problem by emphasizing small and informative lesions that may be underrepresented during conventional instance sampling [13,14]. MG-Trans uses multi-scale modeling to reduce redundant high-resolution processing while preserving contextual information. These developments indicate that computational efficiency should not be achieved

simply by decreasing the total number of tiles; instead, processing resources should be concentrated on regions in which additional spatial resolution is most likely to change the diagnostic interpretation [15,16]. The emergence of pathology foundation models has further strengthened WSI representation learning. Models such as UNI, CONCH, and Prov-GigaPath provide transferable visual or vision-language representations across multiple tissue types and pathological tasks [17,18]. Their pretrained features can reduce dependence on task-specific annotation and improve downstream classification performance. However, stronger feature encoders do not fully resolve the computational problem associated with slide-scale inference. When thousands or tens of thousands of tiles are retained, repeated image decoding, disk access, high-resolution feature extraction, and MIL aggregation remain costly [19,20]. Moreover, many existing workflows use a predefined magnification level for all tissue regions or perform multi-scale extraction according to fixed sampling rules. Such strategies do not explicitly account for the fact that pathological information density varies substantially across a slide [21,22]. Uniform high-magnification processing may allocate excessive resources to low-risk tissue, whereas overly aggressive low-resolution sampling may obscure subtle glandular atypia, infiltrative tumor borders, isolated malignant clusters, or other small structures that are critical for grading. Another limitation concerns the evidence used to evaluate computationally efficient pathology models. Many studies focus primarily on classification accuracy within a single dataset, cancer type, or institutional cohort [23,24]. Such evaluations may not fully reflect clinical deployment, where tissue preparation, staining protocols, scanner manufacturers, image compression, and disease morphology can vary substantially between hospitals. Computational efficiency is also frequently reported through theoretical measures such as the number of retained instances or floating-point operations, while practical quantities such as slide processing time, disk-read volume, and the ability to preserve very small invasive lesions receive less attention. An efficient WSI system should therefore be assessed not only by whether it reduces the number of processed tiles, but also by whether the reduction remains reliable across institutions and tissue types and whether diagnostically important small lesions are retained during adaptive sampling.

This study proposes HistoRoute, a magnification-scheduled tile-routing framework for computationally efficient whole-slide tumor grading. Instead of extracting all foreground regions at the same resolution and discarding uninformative tiles afterward, HistoRoute determines the required magnification before expensive high-resolution processing. Tissue foreground ratio, nuclear density, changes in gland boundaries, and model-estimated tumor uncertainty are jointly used to characterize the diagnostic complexity of individual regions. Areas with low morphological risk are routed to low-magnification processing, suspicious epithelial or structurally heterogeneous regions are analyzed at medium magnification, and uncertain tumor margins or regions with a high probability of containing diagnostically important fine structures are selectively escalated to high magnification. The resulting multi-resolution features are subsequently integrated through an MIL classifier to generate slide-level tumor grades. This

design aims to align computational expenditure with regional diagnostic demand rather than treating all tissue as equally informative. The proposed framework is evaluated using 7,840 WSIs from breast, colorectal, and gastric cancer cohorts, allowing the routing strategy to be examined across substantially different glandular structures, tumor growth patterns, and tissue environments. The evaluation considers slide-level grading performance, agreement on external-hospital data, end-to-end processing time, disk-read volume, and the detection of invasive foci smaller than 0.5 mm. Through these experiments, the study investigates whether adaptive magnification scheduling can reduce redundant high-resolution processing without sacrificing clinically relevant morphological information. The broader objective is to establish a more resource-aware paradigm for computational pathology in which image resolution is treated as a dynamically allocated resource rather than a fixed preprocessing parameter. Such a strategy may improve the scalability of WSI analysis for large retrospective cohorts and routine clinical workloads while preserving sensitivity to small but diagnostically important tumor regions.

2. Materials and Methods

2.1. Study Samples and Whole-Slide Images

The study included 7,840 whole-slide images from breast, colorectal, and gastric cancer cohorts. All tissue samples were fixed in formalin, embedded in paraffin, and stained with hematoxylin and eosin. The dataset contained about 3.2 million tissue tiles and 18,600 tumor regions marked by pathologists. The slides covered different tumor grades, tissue sizes, stromal levels, and tumor proportions. Slides with severe scanning errors, large blurred areas, incomplete tissue regions, or missing grade labels were removed. Slides from the same patient were assigned to only one data subset. Two hospital cohorts were kept for external testing to examine model performance under different tissue preparation and scanner settings.

2.2. Experimental Design and Control Models

HistoRoute was compared with a full-tile Transformer, a uniform low-magnification model, and a random tile-sampling model. The full-tile Transformer processed all candidate tissue tiles at high magnification and was used as the accuracy control. The uniform low-magnification model processed all tissue regions at one low magnification and was used to measure the effect of limited image detail. The random sampling model selected the same number of tiles as HistoRoute without using tissue features. Ablation tests were also performed by removing nuclear density, gland boundary irregularity, tumor uncertainty, and tissue foreground ratio separately. All models used the same training cases, test cases, optimizer, feature size, and slide-level labels. This setting allowed direct comparison of the tile-routing methods under the same training conditions.

2.3. Image Processing and Quality Control

Tissue regions were separated from the white slide background using color and intensity thresholds. The selected tissue areas were divided into non-overlapping tiles at low, medium, and high magnification. Color normalization was used to reduce staining differences among slides. Random rotation, horizontal flipping, vertical flipping, and small color changes were applied only during training. Tiles with severe blur, pen marks, tissue folds, or very low tissue content were removed before feature extraction. Pathologist-marked tumor regions were used to check whether important regions remained after tile routing. Two pathologists reviewed unclear tumor boundaries and small invasive foci used in the test set. Slide-level tumor grades followed the final pathology reports. Processing time and disk-read volume were measured under the same hardware and storage settings for all models.

2.4. Tile Routing and Model Formulation

For each candidate tile, HistoRoute calculated a routing score using tissue foreground ratio, nuclear density, gland boundary irregularity, and tumor uncertainty. The routing score was defined as:

$$R_i = \alpha F_i + \beta N_i + \gamma G_i + \delta U_i$$

where R_i is the routing score of tile i , F_i is the tissue foreground ratio, N_i is nuclear density, G_i is gland boundary irregularity, and U_i is tumor uncertainty. The terms α , β , γ , and δ are weight parameters. Tiles with low R_i were processed at low magnification. Tiles with medium R_i were sent to the medium-magnification path. Tiles with high R_i were kept for high-magnification Transformer encoding. The routed tile features were then combined by a multiple-instance learning classifier. The slide-level output was calculated as:

$$\hat{y} = \sum_{i=1}^N a_i f_i$$

where f_i is the feature vector of tile i , a_i is its attention weight, and N is the number of retained tiles. The routing thresholds and weight parameters were selected using the validation set and kept unchanged during final testing.

2.5. Model Evaluation and Statistical Analysis

Tumor-grading performance was measured by macro-AUC, sensitivity, specificity, and Cohen's kappa. Small invasive foci below 0.5 mm were tested separately because they are more likely to be lost during tile reduction. Model cost was measured by the proportion of high-magnification tiles, mean processing time per slide, GPU memory use, and disk-read volume. External testing used slides from two hospitals that were not included in model training. Results from repeated experiments were reported as mean $\pm$ standard deviation. Paired comparisons were performed using the same test slides. The paired t-test was used when the differences were close to a normal distribution. The Wilcoxon signed-rank test was used for other cases. Statistical

significance was set at $p < 0.05$. Tumor-grading accuracy and processing cost were considered together when comparing the models.

3. Results and Discussion

3.1. Tumor-Grading Performance

HistoRoute reached a macro-AUC of 0.914 for tumor grading, compared with 0.919 for the full-tile Transformer and 0.872 for the uniform low-magnification model. The AUC difference between HistoRoute and the full-tile model was only 0.005. In contrast, the low-magnification model showed a larger loss in grading accuracy. This indicates that low-resolution tissue structure alone cannot provide enough detail for tumor grading [25,26]. Nuclear shape, gland structure, and tumor margins still require higher magnification in selected areas. Fig.1 compares grading accuracy with processing cost. HistoRoute kept nearly the same grading performance as the full-tile Transformer while using fewer high-resolution tiles. Liang et al. (2026) also reported that WSI magnification affects both cancer recognition and inference time, and that the highest magnification is not always necessary for all tissue regions.

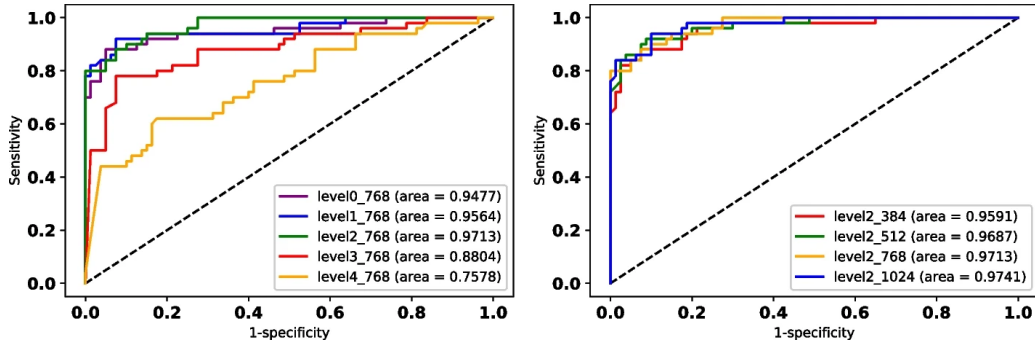


Fig. 1. Tumor-grading accuracy and processing cost of HistoRoute and the control models.

3.2. High-Magnification Tile Use and Processing Cost

HistoRoute processed only 28.4% of candidate tiles at high magnification. Other regions remained at lower magnification when their grading risk was low. Mean slide-processing time decreased from 96.5 s to 38.7 s, a reduction of 59.9%. Disk-read volume decreased by 52.1%. These results indicate that WSI processing cost depends on both model computation and the amount of image data read from storage. This is important for gigapixel slides because loading high-resolution tiles can take a large part of the total processing time. Gensbigler et al. (2025) also found clear differences in prediction time across WSI magnification levels. HistoRoute avoids using one fixed magnification for the whole slide and keeps high-resolution processing only for regions that need more tissue detail [27,28].

3.3. External Hospital Testing and Tile Selection

In the two external hospital cohorts, HistoRoute reached a Cohen's kappa of 0.836 with pathologist grades, compared with 0.781 for random tile sampling. The increase of 0.055 indicates

better grade agreement after tissue-based tile selection. Random sampling can miss small tumor areas because all candidate tiles have the same selection chance. HistoRoute gives higher priority to regions with dense nuclei, irregular gland boundaries, and uncertain tumor features. As shown in Fig.2, low-risk regions remain at low magnification, while suspicious regions receive medium- or high-magnification processing. Previous MIL studies have also found that tissue-guided attention can locate tumor-related regions more reliably than unguided sampling. The external results suggest that HistoRoute is less affected by differences in slide appearance between hospitals [29,30].

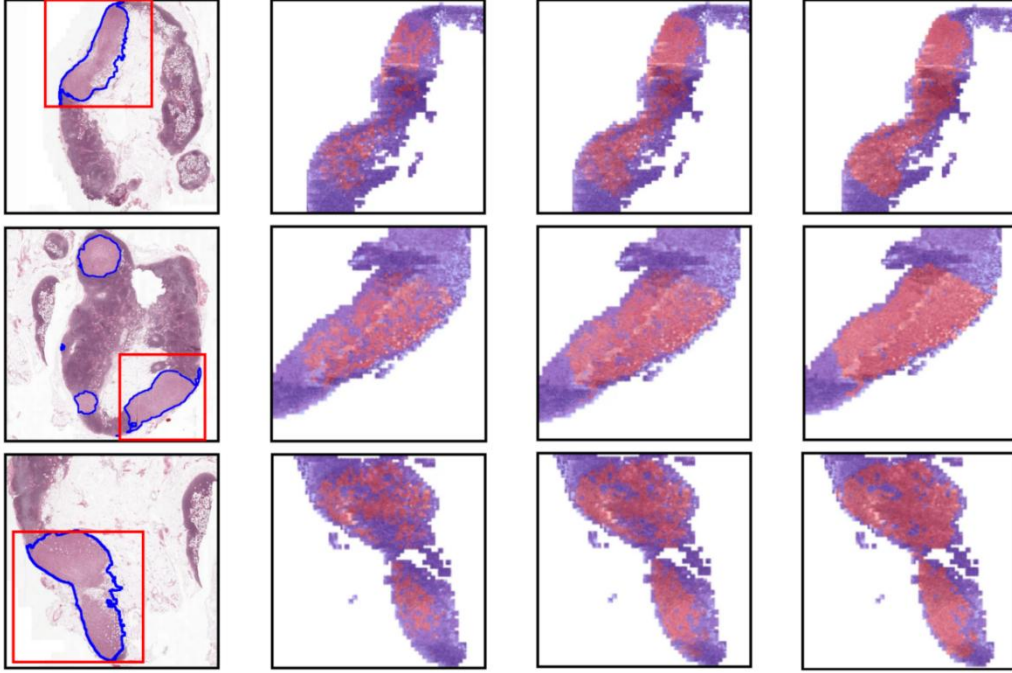


Fig. 2. Tile-routing results at three magnification levels in representative whole-slide images.

3.4. Detection of Small Invasive Foci

HistoRoute achieved 89.3% sensitivity for invasive foci below 0.5 mm, although less than one-third of candidate tiles were processed at high magnification. Small invasive regions may occupy only a few tiles and can be missed by random or fixed-ratio sampling. Previous WSI studies have also reported that tile selection affects the detection of small tumor regions. HistoRoute reduces this problem by keeping uncertain tumor margins and dense epithelial regions in the high-magnification path. However, 10.7% of small foci were still missed. This remains an important limit of the current routing method. Further studies should include more tumor types, scanners, staining methods, and external hospitals to test the stability of the routing rules.

4. Conclusion

This study developed HistoRoute for whole-slide tumor grading with magnification-based tile routing. The model selects low, medium, or high magnification according to tissue foreground ratio, nuclear density, gland boundary changes, and tumor uncertainty. HistoRoute achieved a

macro-AUC of 0.914, compared with 0.919 for the full-tile Transformer, while only 28.4% of candidate tiles were processed at high magnification. Mean slide-processing time decreased from 96.5 s to 38.7 s, and disk-read volume decreased by 52.1%. In external hospital testing, Cohen's kappa increased from 0.781 with random sampling to 0.836. Sensitivity for invasive foci below 0.5 mm reached 89.3%. These results show that tissue regions can be processed at different magnification levels without a large loss in grading accuracy. HistoRoute can reduce image-reading and computing costs in digital pathology. However, some small invasive foci were still missed, and the study included a limited number of hospitals and scanner types. Future studies should test the method on more cancer types, scanners, staining settings, and external clinical datasets.

References

1. Butt, N. A., Sarwat, D., Noya, I. D., Tutusaus, K., Samee, N. A., & Ashraf, I. (2026). Benchmarking multiple instance learning architectures from patches to pathology for prostate cancer detection and grading using attention-based weak supervision. *Scientific Reports*, 16(1), 11535.
2. Jiao, Y., Shi, T., Zhao, B., & Wang, A. (2026). Retrieval-Guided Structured Reasoning and Interpretable Representation Learning for Large-Scale Video–Language Models.
3. Liang, S., Du, Y., Chen, W., & Liu, Z. (2026). Order Allocation and Emergency Transportation Decisions Amid Multi-Source Procurement Disruptions.
4. Zhang, Z., Wang, J., Li, Z., Wang, Y., & Zheng, J. (2025). Annocoder: A mti-agent-based code generation and optimization model. *Symmetry*, 17(7), 1087.
5. Kavitha, A. (2026, June). Cross-Domain Transfer Learning and Self-Supervised Representation Learning for Automated Histopathological Image Classification in Colorectal Cancer Grading System. In 2026 7th International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 2044-2050). IEEE.
6. Bao, Y., Qiu, Y., & Wang, H. (2026). FLARE: A Real-Time Framework for Detecting Malicious Account Activities at Internet Scale. Available at SSRN 7185659.
7. Wang, S., Feng, Y., & Fang, X. (2026, May). A Large Language Model-Enabled Multi-Agent Collaboration Method for Complex Task Solving. In 2026 6th International Symposium on Computer Technology and Information Science (ISCTIS) (pp. 253-256). IEEE.
8. Brunet, E. (2026). Contributions to Resource-Efficient HPC: A Data-Centric Approach (Doctoral dissertation, Institut polytechnique de Paris).
9. Yang, Q., & Du, Y. (2026). Predicting Rollback Risks and Controlling Gradual Rollout Traffic for Online Ranking Feature Deployments.
10. Zhao, Z., & Welsch, R. E. (2026). Point-in-Time Financial RAG with Frozen LLMs and Market-Feedback Adaptive Retrieval. arXiv preprint arXiv:2605.31201.

11. Yuan, Y., Huang, J., Yin, J., & Xu, T. (2026). Confidence Calibration and Semantic Error Analysis for Ambiguous Coreference Resolution in User-generated Social Media Text. Available at SSRN 7393238.
12. Ferreira, F. R. T., & do Couto, L. M. (2025). Explainable machine learning framework integrating PPPM for automated cytological triage and personalized breast cancer risk stratification. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 14(1), 167.
13. Liang, S., Hsu, K. H., & Liu, Z. (2026). Rolling Scheduling and Capacity Balancing for Server Assembly Under Engineering Change Disturbances.
14. Bao, Y., & Wang, H. (2026). Large-scale Metrics Migration: A Systematic Approach to Rebuilding Production Measurement Infrastructure. Available at SSRN 7185660.
15. Yuan, Y., Yin, J., Huang, J., & Xu, T. (2026). Weakly Supervised Anomaly Detection and Privacy Risk Scoring in Community Message Threads.
16. Büchel, J., Vasilopoulos, A., Simon, W. A., Boybat, I., Tsai, H., Burr, G. W., ... & Sebastian, A. (2025). Efficient scaling of large language models with mixture of experts and 3D analog in-memory computing. *Nature Computational Science*, 5(1), 13-26.
17. Xu, D., Gui, H., & Chen, H. (2026, March). Research on Layered Control and Fault Recovery Mechanisms for Fast Charging Safety Diagnosis of High-Voltage Battery Systems Under Charging Network Interoperability Conditions. In *2026 3rd International Conference on Electrical Technology and Automation Engineering (ETAE)* (pp. 669-672). IEEE.
18. Zhang, Z. (2026). A Study on the Impact of Cost Presentation on Decision-making Bias and Cancellation Behavior in Online Subscriptions. Available at SSRN 7349839.
19. Qi, C., & Qiao, X. (2026). Building and Operating a Large Scale Multi-Agent System: A Case Study from Industry. Available at SSRN 6795198.
20. Casper, M. J. (2025). Human Brain Optimized Light Sheet Microscopy (HOLiS) for High-Throughput Cell Type Atlasing of Whole Human Brains (Doctoral dissertation, Columbia University).
21. Xiong, W., Guo, Y., & Zeng, Y. (2026). A Study on the Energy Efficiency of MEP Systems and Coordinated Dispatch Mechanisms for Multi-energy Systems in High-density Urban Buildings. Available at SSRN 7120518.
22. Hong, Z., Liang, S., Xu, T., & Chen, H. (2026). A Study on Failover Verification and Recovery Objective Prediction for Cross-Region Cloud Services.
23. Hamza, A., & Damaševičius, R. (2026). Deep learning for brain tumor segmentation and classification: a systematic review of methods and trends. *Computers, materials and continua.*, 86(1), 1-41.
24. Huang, J., Yin, J., Xu, T., & Yang, J. (2026). Quality Assessment and User Adoption Prediction for Enterprise-Level Financial Reporting Data Products. Available at SSRN 7179118.
25. Zheng, J., & Makar, M. (2022). Causally motivated multi-shortcut identification and removal. *Advances in Neural Information Processing Systems*, 35, 12800-12812.

26. Yao, G., Zheng, J., Wang, Z., Zhang, W., Han, R., Zhao, C., ... & Liu, R. (2026, March). V-pruner: A fast and globally-informed token pruning framework for vision transformer. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 40, No. 40, pp. 34396-34404).
27. Gensbigler, P. A., Foster, W., Kiemen, A. L., & Bever, G. S. (2025). An optimized, high-throughput workflow for the collection, processing, and visualization of histology data in comparative vertebrate morphogenesis. *Developmental Dynamics*.
28. Zhao, J., Fan, J., & Li, L. (2026). A Study on an Explainable Causal-Enhanced LLM Agent for Predicting the Forming Quality of Automotive Component Materials.
29. You, S. (2026). Verifiable Audit Mechanisms in AI Compliance Automation: Scalability. Available at SSRN 6547458.
30. Jimenez, M., Flandes, J., van der Heijden, E. H., Ng, C. S., Iding, J. S., Garcia-Hierro, J. F., ... & Krinsky, W. S. (2025). Safety and Feasibility of Pulsed Electric Field Ablation for Early-Stage Non-Small Cell Lung Cancer Prior to Surgical Resection. *Journal of Surgical Oncology*, 131(8), 1529-1542.