Novelty Assessment Report

Paper: Multihead Mixture of Experts for Classification of Gigapixel Pathology Images

PDF URL: https://openreview.net/pdf?id=S5Io33pc78

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-01

Abstract

Multiple Instance Learning (MIL) is the predominant approach for classifying gigapixel whole-slide images in computational pathology. MIL follows a sequence of 1) extracting patch features, 2) applying a linear layer to obtain task-specific patch features, and 3) aggregating the patches into a slide feature for classification. While substantial efforts have been devoted to optimizing patch feature extraction and aggregation, none have yet addressed the second point, the critical layer which transforms general-purpose features into task-specific features. We hypothesize that this layer constitutes an overlooked performance bottleneck and that stronger representations can be achieved with a low-rank transformation tailored to each patch's phenotype, yielding synergistic effects with existing MIL approaches. To this end, we introduce MAMMOTH, a parameter-efficient, multi-head mixture of experts module designed to improve the performance of any MIL model with minimal alterations to the total number of parameters. Across 8 MIL methods and 19 different tasks, we find that this improvement to the task-specific transformation has a larger effect on performance than the choice of aggregation method. For instance, when equipped with MAMMOTH, even simple methods such as max or mean pooling attain higher average performance than any method with the standard linear layer. Finally, we identify Instance-Gradient Interference (IGI)—a limitation where heterogeneous instances produce conflicting gradients when processed by a single linear layer—and show that MAMMOTH effectively mitigates IGI by decoupling gradient flows between experts, yielding consistent performance gains in 130 of the 152 examined configurations.

Core Task Landscape

This paper addresses: **Classification of Gigapixel Whole-Slide Images in Computational Pathology**

A total of **50 papers** were analyzed and organized into a taxonomy with **23 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Multiple Instance Learning Architectures and Aggregation Methods**
- **Feature Extraction and Representation Learning**
- **Vision-Language Models and Multimodal Learning**
- **End-to-End and Full-Resolution Processing Methods**
- **Training Strategies and Optimization Techniques**
- **Generative Models and Synthetic Data Creation**
- **Preprocessing and Infrastructure for WSI Analysis**
- **Application-Specific and Clinical Deployment Studies**
- **Hybrid and Multi-Feature Fusion Approaches**

Complete Taxonomy Tree

- Classification of Gigapixel Whole-Slide Images in Computational Pathology Survey Taxonomy
- Multiple Instance Learning Architectures and Aggregation Methods
  - Attention-Based MIL and Transformer Architectures (5 papers)
  - [4] Attention-driven multi-scale analysis for accurate tumor classification in digital pathology. (Junze Huang, 2025) View paper
  - [11] Advances in multiple instance learning for whole slide image analysis: Techniques, challenges, and future directions (Wang Jun, 2024) View paper
  - [39] Multi-scale attention-based multiple instance learning for classification of multi-gigapixel histology images (Made Satria Wibawa, 2022) View paper
  - [40] Gigapixel histopathological image analysis using attention-based neural networks (Brancati Nadia, 2021) View paper
  - [47] Magnifying networks for histopathological images with billions of pixels (Neofytos Dimitriou, 2024) View paper
  - Prototype and Concept-Based MIL (3 papers)
  - [6] Label-free Concept Based Multiple Instance Learning for Gigapixel Histopathology (Sun, 2025) View paper
  - [9] Prototype-Based Multiple Instance Learning for Gigapixel Whole Slide Image Classification (Sun, 2025) View paper
  - [46] PAMIL: Prototype Attention-Based Multiple Instance Learning for Whole Slide Image Classification (Jiashuai Liu, 2024) View paper
  - Graph-Based MIL and Spatial Modeling (2 papers)
  - [22] From Pixels to Histopathology: A Graph-Based Framework for Interpretable Whole Slide Image Analysis (A Weers, 2025) View paper
  - [45] Dynamic Graph Representation for WSI Classification: A Knowledge-Aware Attention Mechanism for Enhanced Computational Pathology (T. Rajesh Kumar, 2025) View paper
  - Mixture of Experts and Task-Specific Transformations ★ (2 papers)
  - [0] Multihead Mixture of Experts for Classification of Gigapixel Pathology Images (Anon et al., 2026) View paper
  - [33] Learning Heterogeneous Tissues with Mixture of Experts for Gigapixel Whole Slide Images (Junxian Wu, 2025) View paper

  - Efficient Patch Selection and Hierarchical MIL (4 papers)
  - [8] EvoPS: Evolutionary Patch Selection for Whole Slide Image Analysis in Computational Pathology (Saya Hashemian, 2025) View paper
  - [13] Fast and Accurate Gigapixel Pathological Image Classification with Hierarchical Distillation Multi-Instance Learning (Jiuyang Dong, 2025) View paper
  - [30] Deep learning-based sparse whole-slide image analysis for the diagnosis of gastric intestinal metaplasia (Braatz, 2022) View paper
  - [37] Efficient Classification of Very High Resolution Images (Nouyed, 2024) View paper
- Feature Extraction and Representation Learning
  - Self-Supervised and Foundation Models for Pathology (4 papers)
  - [1] A whole-slide foundation model for digital pathology from real-world data (Hanwen Xu, 2024) View paper
  - [12] Scaling vision transformers to gigapixel images via hierarchical self-supervised learning (Chen, 2022) View paper
  - [41] PLUTO: Pathology-Universal Transformer (Juyal, 2024) View paper
  - [48] A self-supervised framework for learning whole slide representations (Hou, 2024) View paper
  - Alternative Architectures for Feature Encoding (2 papers)
  - [24] Vim4Path: Self-Supervised Vision Mamba for Histopathology Images (Ali Nasiri-Sarvi, 2024) View paper
  - [25] Gigapixel whole-slide images classification using locally supervised learning (Zhang Jing-wei, 2022) View paper
  - Feature Extractor Selection and Fine-Tuning (3 papers)
  - [19] Evolutionary deep feature selection for compact representation of gigapixel images in digital pathology (Azam Asilian Bidgoli, 2022) View paper
  - [21] Task-Specific Fine-Tuning via Variational Information Bottleneck for Weakly-Supervised Pathology Whole Slide Image Classification (Honglin Li, 2023) View paper
  - [29] Rethinking pre-trained feature extractor selection in multiple instance learning for whole slide image classification (Bryan Wong, 2025) View paper
- Vision-Language Models and Multimodal Learning
  - Vision-Language Alignment and Prompt Learning (5 papers)
  - [5] Slide-Level Prompt Learning with Vision Language Models for Few-Shot Multiple Instance Learning in Histopathology (Tomar, 2025) View paper
  - [20] MGPATH: Vision-Language Model with Multi-Granular Prompt Learning for Few-Shot WSI Classification (Nguyen Anh Tien, 2025) View paper
  - [26] Few-Shot Learning from Gigapixel Images via Hierarchical Vision-Language Alignment and Modeling (Kim Jong Woo, 2025) View paper
  - [44] PathAlign: A vision-language model for whole slide images in histopathology (Ahmed, 2024) View paper
  - [49] PathFLIP: Fine-grained Language-Image Pretraining for Versatile Computational Pathology (Fengchun Liu, 2025) View paper
  - Report Generation and Multimodal Captioning (3 papers)
  - [7] Generating dermatopathology reports from gigapixel whole slide images with HistoGPT (M. Tran, 2025) View paper
  - [10] Generating highly accurate pathology reports from gigapixel whole slide images with HistoGPT (Tran, 2024) View paper
  - [38] PathM3: A Multimodal Multi-Task Multiple Instance Learning Framework for Whole Slide Image Classification and Captioning (Qifeng Zhou, 2024) View paper
  - Multimodal Conversational and Interactive Systems (2 papers)
  - [18] Slidechat: A large vision-language assistant for whole-slide pathology image understanding (Ying Chen, 2025) View paper
  - [23] PathBench: Advancing the Benchmark of Large Multimodal Models for Pathology Image Understanding at Patch and Whole Slide Level (Yuxuan Sun, 2025) View paper
- End-to-End and Full-Resolution Processing Methods
  - Distributed and Scalable Full-Resolution Training (2 papers)
  - [2] Scaling resolution of gigapixel whole slide images using spatial decomposition on convolutional neural networks (Aristeidis Tsaris, 2023) View paper
  - [3] When an image is worth 1,024 x 1,024 words: A case study in computational pathology (Wang Wen-hui, 2023) View paper
  - All-in-Memory End-to-End Slide Modeling (1 papers)
  - [27] Beyond multiple instance learning: Full resolution all-in-memory end-to-end pathology slide modeling (Campanella, 2024) View paper
- Training Strategies and Optimization Techniques
  - Imbalanced Learning and Contrastive Strategies (1 papers)
  - [35] SC-MIL: Supervised Contrastive Multiple Instance Learning for Imbalanced Classification in Pathology (Dinkar Juyal, 2023) View paper
  - Data Augmentation and Latent Space Manipulation (1 papers)
  - [28] Controllable Latent Space Augmentation for Digital Pathology (Scalbert, 2025) View paper
  - Bayesian and Collaborative Learning Frameworks (1 papers)
  - [32] Bayesian Collaborative Learning for Whole-Slide Image Classification (Jinâ□□Gang Yu, 2023) View paper
- Generative Models and Synthetic Data Creation (2 papers)
  - [31] High-Resolution Histopathology Whole Slide Image Generation Using Wavelet Diffusion Model (Abdullah, 2025) View paper
  - [34] Diffusion-based generation of histopathological whole slide images at a gigapixel scale (Harb, 2024) View paper
- Preprocessing and Infrastructure for WSI Analysis
  - Color Normalization and Stain Standardization (1 papers)
  - [15] Fast GPU-enabled color normalization for digital pathology (Goutham Ramakrishnan, 2019) View paper
  - Imaging Hardware and Scanning Technologies (1 papers)
  - [17] Rapid 3D imaging at cellular resolution for digital cytopathology with a multi-camera array scanner (MCAS) (KangHyun Kim, 2024) View paper
- Application-Specific and Clinical Deployment Studies
  - Cancer Subtyping and Prognostic Prediction (2 papers)
  - [16] A weakly supervised deep learning approach for guiding ovarian cancer treatment in prognosis: gigapixel histopathology analysis for personalized therapeutic â□¦ (I Rasool, 2024) View paper
  - [43] Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis (Xi Wang, 2020) View paper
  - Multi-Target Segmentation and Tissue Analysis (1 papers)

    - [14] A Review of Advanced Deep Learning Methods of Multi-Target Segmentation for Breast Cancer WSIs (Qiaoyi Xu, 2025) View paper
    - Scalable Deployment and Validation Studies (2 papers)
    - [36] Scalable deep learning artificial intelligence histopathology slide analysis and validation (Colin Greeley, 2024) View paper
    - [50] A Decade of GigaScience: The Challenges of Gigapixel Pathology Images (G. Litjens, 2022) View paper
- Hybrid and Multi-Feature Fusion Approaches (1 papers)
    - [42] Coca-Mil: Attention-Based Handcrafted-Deep Feature Fusion in Computational Pathology (Paras Goel, 2024) View paper

Narrative

Core task: classification of gigapixel whole-slide images in computational pathology. The field has evolved into a rich ecosystem of approaches organized around several major branches. Multiple Instance Learning (MIL) architectures and aggregation methods form a central pillar, encompassing attention-based networks, mixture-of-experts designs, and hierarchical aggregation schemes that handle the enormous number of patches extracted from each slide. Feature extraction and representation learning focuses on self-supervised and foundation models that produce robust patch embeddings, while vision-language models and multimodal learning integrate textual reports or clinical metadata to enhance interpretability. End-to-end and full-resolution processing methods attempt to bypass the patch-based bottleneck by operating directly on high-resolution data, whereas training strategies and optimization techniques address sample efficiency, contrastive learning, and domain adaptation. Generative models and synthetic data creation explore augmentation and slide synthesis, preprocessing and infrastructure tackle color normalization and scalable pipelines, application-specific studies target particular cancer types or clinical workflows, and hybrid multi-feature fusion approaches combine complementary representations.

Within the MIL landscape, a particularly active line of work explores mixture-of-experts and task-specific transformations, where models dynamically route or weight different subnetworks to capture tissue heterogeneity. Multihead Mixture Experts[0] exemplifies this direction by employing multiple expert heads to handle diverse histological patterns, closely related to Mixture Experts Tissues[33], which similarly partitions the feature space according to tissue type. These methods contrast with simpler attention pooling schemes like those in Attention Based Networks[40] or Gigapixel Vision Transformers[12], which apply uniform aggregation across all patches. Meanwhile, vision-language approaches such as Slide Prompt Learning[5] and HistoGPT Reports[7] introduce textual guidance to steer classification, and full-resolution methods like Full Resolution Memory[27] and Scaling Gigapixel Resolution[2] bypass patch extraction altogether. The original paper sits squarely in the mixture-of-experts branch, emphasizing adaptive specialization over monolithic aggregation, and shares conceptual ground with Mixture Experts Tissues[33] while differing from prompt-based or end-to-end alternatives that do not explicitly model expert diversity.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Learning Heterogeneous Tissues with Mixture of Experts for Gigapixel Whole Slide Images

Authors: Junxian Wu, Minheng Chen, Xinyi Ke, Tianwang Xun, Xiaoming Jiang, et al. (8 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Analyzing gigapixel Whole Slide Images (WSIs) is challenging due to the complex pathological tissue environment and the absence of target-driven domain knowledge. Previous methods incorporated pathological priors to mitigate this issue but relied on additional inference steps and specialized workflows, restricting scalability and the modelâs capacity to identify novel outcome-related factors. To address these challenges, we propose a plug-and-play Pathology-Aware Mixture-of-Experts (PAMoE) mod...

Relationship Analysis

Both papers belong to the Mixture of Experts and Task-Specific Transformations category, using MoE architectures to handle tissue heterogeneity in WSI classification. While the original paper (MAMMOTH) focuses on replacing the task-specific linear layer with a multihead soft MoE using slot-based pooling and low-rank experts for parameter efficiency, the candidate paper (PAMoE) introduces pathology-aware expert choice routing that supervises experts using pre-extracted tissue prototypes and allows experts to select patches rather than assigning experts to patches. The key distinction is that MAMMOTH emphasizes parameter-efficient soft routing with matrix factorization, whereas PAMoE uses supervised expert preferences based on pathological priors with expert choice routing.

Contributions Analysis

Overall novelty summary. The paper introduces MAMMOTH, a parameter-efficient multihead mixture-of-experts module that transforms general-purpose patch features into task-specific representations before aggregation in MIL pipelines. It resides in the 'Mixture of Experts and Task-Specific Transformations' leaf, which contains only two papers total (including this one). This is a notably sparse research direction within the broader MIL landscape, suggesting that explicit modeling of task-specific transformations via mixture-of-experts architectures remains underexplored compared to attention-based aggregation or prototype methods, which occupy more densely populated leaves.

The taxonomy reveals that neighboring leaves focus on attention mechanisms (five papers), prototype learning (three papers), graph-based spatial modeling (two papers), and hierarchical patch selection (four papers). These directions emphasize aggregation strategies or interpretability rather than the intermediate transformation layer. The paper's focus on the linear projection between feature extraction and aggregation diverges from these neighboring approaches, which largely treat this step as a fixed operation. The scope note for this leaf explicitly excludes standard attention aggregation and prototype methods, positioning MAMMOTH as addressing a distinct bottleneck in the MIL pipeline that other branches do not target.

Among 26 candidates examined across three contributions, no refutable prior work was identified. The MAMMOTH module itself was assessed against nine candidates with zero refutations, the Instance-Gradient Interference analysis against seven candidates with zero refutations, and the task-specific transformation bottleneck hypothesis against ten candidates with zero refutations. This suggests that within the limited search scope—primarily top-K semantic matches and citation expansion—no prior work directly anticipates the combination of multihead mixture-of-experts applied specifically to the task-specific transformation layer in MIL. The single sibling paper in the same leaf (Mixture Experts Tissues) likely addresses tissue-type partitioning rather than per-patch adaptive transformations, though detailed comparison would require full-text review.

Given the sparse taxonomy leaf, absence of refutable candidates among 26 examined papers, and the paper's focus on an intermediate MIL component that neighboring methods treat as fixed, the work appears to occupy a relatively unexplored niche. However, the limited search scope means this assessment reflects top-K semantic proximity rather than exhaustive coverage of the computational pathology literature. The novelty claim rests on the specific architectural choice—multihead mixture-of-experts for task-specific transformations—rather than on introducing mixture-of-experts or MIL concepts themselves, which are established in the field.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: MAMMOTH: Multihead Mixture of Experts module for MIL

Description: The authors propose MAMMOTH, a plug-and-play mixture of experts architecture that replaces the standard linear layer in Multiple Instance Learning frameworks. It uses multihead processing, soft expert assignment, and low-rank decomposition to transform general-purpose patch features into task-specific features for gigapixel pathology image classification while maintaining parameter efficiency.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Uni-med: a unified medical generalist foundation model for multi-task learning via connector-MoE

URL: View paper

Brief Assessment

Uni Med[69] focuses on multi-modal medical image-text learning with a connector-MoE between visual encoder and LLM, not on multiple instance learning for gigapixel pathology images. The architectural contexts and problem domains are fundamentally different.

2. Spatially-Aware Mixture of Experts with Log-Logistic Survival Modeling for Whole-Slide Images

URL: View paper

Brief Assessment

Spatially Aware Experts[73] focuses on survival prediction from whole-slide images using spatial clustering and log-logistic survival modeling, not on replacing linear layers in MIL frameworks with mixture of experts for general classification tasks.

3. Deep multi-instance learning using multi-modal data for diagnosis of lymphocytosis

URL: View paper

Brief Assessment

Multi Modal Lymphocytosis[72] focuses on lymphocytosis diagnosis using blood cell images with a mixture-of-experts for combining image and clinical data modalities, not on general-purpose pathology image classification with task-specific feature transformation for gigapixel WSIs.

4. A Mixture-of-Experts Decision Support System for Digital Pathology

URL: View paper

Brief Assessment

Mixture Experts Decision[70] focuses on decision support systems for digital pathology classification tasks but does not provide sufficient technical detail in the available context to assess whether it addresses the specific architectural innovations of MAMMOTH (multihead processing, soft expert assignment, low-rank decomposition for task-specific patch feature transformation).

5. Mome: Mixture of multimodal experts for cancer survival prediction

URL: View paper

Brief Assessment

Mome[67] focuses on multimodal survival prediction combining WSIs and genomic data, not on replacing linear layers in MIL frameworks for patch feature transformation in classification tasks.

6. Multimodal Gated Mixture of Experts Using Whole Slide Image and Flow Cytometry for Multiple Instance Learning Classification of Lymphoma

URL: View paper

Brief Assessment

Multimodal Gated Lymphoma[71] focuses on multimodal learning combining whole slide images and flow cytometry for lymphoma classification, not on general-purpose MIL architectures with mixture of experts for pathology.

7. Mamba-HMIL: Hierarchical Multiple Instance Learning via State Space Model for Whole Slide Image Diagnosis

URL: View paper

Brief Assessment

Mamba HMIL[74] focuses on state space models (Mamba) with hierarchical feature extraction for WSI classification, not on mixture of experts architectures for task-specific feature transformation in MIL frameworks.

8. Learning Heterogeneous Tissues with Mixture of Experts for Gigapixel Whole Slide Images

URL: View paper

Brief Assessment

Mixture Experts Tissues[33] focuses on pathology-aware routing with tissue-specific priors for survival prediction, while MAMMOTH addresses task-specific feature transformation in MIL without requiring additional priors during inference. The architectural designs and objectives differ substantially.

9. M4: Multi-Proxy Multi-Gate Mixture of Experts Network for Multiple Instance Learning in Histopathology Image Analysis

URL: View paper

Brief Assessment

Multi Proxy Experts[68] focuses on multi-task learning for simultaneous prediction of multiple genetic mutations using a multi-gate mixture-of-experts architecture, while MAMMOTH targets single-task performance improvement by replacing the task-specific linear layer with a multihead soft mixture of experts for morphological specialization.

Contribution 2: Identification of Instance-Gradient Interference (IGI)

Description: The authors identify a previously unrecognized limitation called Instance-Gradient Interference, where heterogeneous patch instances create conflicting gradient updates in standard linear layers. They demonstrate that MAMMOTH addresses this issue by routing different instances to separate experts, enabling decoupled gradient flows.

This contribution was assessed against **7 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Occlusion-aware tracking for drones using neural methods

URL: View paper

Brief Assessment

Occlusion Aware Tracking[63] focuses on visual object tracking for drones under occlusion conditions, not on gradient interference in multiple instance learning for pathology images. The technical domains are entirely different.

---

2. Disentangled Multi-modal Learning of Histology and Transcriptomics for Cancer Characterization

URL: View paper

Brief Assessment

Disentangled Multi Modal[65] addresses multi-modal heterogeneity in histology-transcriptomics integration, not gradient interference from heterogeneous instances in multiple instance learning frameworks for pathology images.

---

3. SLAM-AGS: Slide-Label Aware Multi-Task Pretraining Using Adaptive Gradient Surgery in Computational Cytology

URL: View paper

Brief Assessment

SLAM AGS[62] addresses gradient conflicts in multi-task learning between self-supervised and weakly supervised objectives, not gradient interference from heterogeneous instances within a single linear layer as described in the original paper's IGI concept.

---

4. Local Self-attention-based Hybrid Multiple Instance Learning for Partial Spoof Speech Detection

URL: View paper

Brief Assessment

Local Self Attention[61] addresses gradient conflicts in partially spoofed speech detection using MIL, but focuses on segment-level gradient issues in audio processing rather than the heterogeneous patch instance problem in gigapixel pathology images that IGI describes.

---

5. MGIML: Cancer Grading With Incomplete Radiology-Pathology Data via Memory Learning and Gradient Homogenization

URL: View paper

Brief Assessment

MGIML[60] addresses gradient conflicts between missing-modality situations in multi-modal learning, not gradient interference from heterogeneous patch instances in multiple instance learning for pathology images.

---

6. Facial Analysis of Dyadic Interactions Using Multiple Instance Learning

URL: View paper

Brief Assessment

Facial Dyadic Interactions[66] focuses on multiple instance learning for facial expression analysis in romantic couple interactions, not on gradient interference phenomena in pathology image classification. The candidate addresses a completely different domain and does not discuss gradient conflicts or heterogeneous patch instances.

---

7. MoMIL: Mixture of Multi-instance Learners for Modeling Multiple Compound Activities in High Content Imaging

URL: View paper

Brief Assessment

MoMIL[64] focuses on mixture of multi-instance learners for high content imaging in drug discovery, not on gradient interference phenomena in pathology image classification. The technical contexts are fundamentally different.

Contribution 3: Task-specific transformation as performance bottleneck

Description: The authors identify the task-specific linear transformation layer in MIL pipelines as a critical but previously unexplored performance bottleneck. They demonstrate that improving this transformation yields larger performance gains than changing aggregation methods, showing consistent improvements in 130 of 152 examined configurations.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Shared-specific feature learning with bottleneck fusion transformer for multi-modal whole slide image analysis

URL: View paper

Brief Assessment

Bottleneck Fusion Transformer[54] focuses on multi-modal fusion between WSIs and tabular clinical data using bottleneck tokens for cross-modal knowledge transfer, not on identifying task-specific linear transformations as performance bottlenecks in single-modal MIL pipelines.

---

2. Flow-MIL: Constructing Highly-expressive Latent Feature Space For Whole Slide Image Classification Using Normalizing Flow

URL: View paper

Brief Assessment

Flow MIL[56] focuses on normalizing flow-based latent space transformations for feature aggregation in MIL, not on analyzing task-specific linear transformation layers as performance bottlenecks compared to aggregation methods.

---

3. Bayesian Collaborative Learning for Whole-Slide Image Classification

URL: View paper

Brief Assessment

Bayesian Collaborative Learning[32] addresses memory bottleneck issues in MIL through collaborative learning between patch and MIL classifiers, not the task-specific linear transformation layer that the original paper identifies as a bottleneck.

---

4. Attention-challenging multiple instance learning for whole slide image classification

URL: View paper

Brief Assessment

Attention Challenging MIL[51] focuses on attention mechanism concentration issues in MIL aggregation, not on the task-specific linear transformation layer that the original paper identifies as a bottleneck.

5. Mitigating Representation Bottlenecks in Multiple Instance Learning

URL: View paper

Brief Assessment

Mitigating Representation Bottlenecks[58] addresses aggregation bottlenecks in MIL (compressing diverse patches into single slide embeddings), not the task-specific linear transformation layer that the original paper identifies as a bottleneck.

6. Nciemil: Rethinking decoupled multiple instance learning framework for histopathological slide classification

URL: View paper

Brief Assessment

Nciemil[53] focuses on information redundancy from gradient operations, instance concatenation, and channel dimensions in MIL frameworks, not on comparing task-specific transformation layers versus aggregation methods as performance bottlenecks.

7. A Study of Temporal Contextual Semantic Enhanced Fusion Module for AnomalousVideo Detection

URL: View paper

Brief Assessment

Temporal Contextual Semantic[57] focuses on video anomaly detection using temporal context aggregation and dynamic batch normalization, not on task-specific transformation layers in multiple instance learning pipelines for pathology images.

8. Reducing Cross-Sensor Domain Gaps in Tactile Sensing via Few-Sample-Driven Style-to-Content Unsupervised Domain Adaptation

URL: View paper

Brief Assessment

Cross Sensor Tactile[52] addresses cross-sensor domain adaptation in tactile sensing using style-to-content transformation methods, not task-specific transformation layers in multiple instance learning pipelines for pathology images.

9. An investigation into the effectiveness of bottleneck based input control compared to aggregate input control

URL: View paper

Brief Assessment

Bottleneck Input Control[59] investigates bottleneck release strategies in manufacturing/production systems, not task-specific transformation layers in multiple instance learning pipelines for computational pathology.

10. From Segments to Concepts: Interpretable Image Classification via Concept-Guided Segmentation

URL: View paper

Brief Assessment

Segments to Concepts[55] focuses on concept-guided segmentation and multiple instance learning for interpretable image classification, not on analyzing task-specific transformation layers versus aggregation methods in MIL pipelines for pathology images.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Multihead Mixture of Experts for Classification of Gigapixel Pathology Images View paper
- [1] A whole-slide foundation model for digital pathology from real-world data View paper
- [2] Scaling resolution of gigapixel whole slide images using spatial decomposition on convolutional neural networks View paper
- [3] When an image is worth 1,024 x 1,024 words: A case study in computational pathology View paper
- [4] Attention-driven multi-scale analysis for accurate tumor classification in digital pathology. View paper
- [5] Slide-Level Prompt Learning with Vision Language Models for Few-Shot Multiple Instance Learning in Histopathology View paper
- [6] Label-free Concept Based Multiple Instance Learning for Gigapixel Histopathology View paper
- [7] Generating dermatopathology reports from gigapixel whole slide images with HistoGPT View paper
- [8] EvoPS: Evolutionary Patch Selection for Whole Slide Image Analysis in Computational Pathology View paper
- [9] Prototype-Based Multiple Instance Learning for Gigapixel Whole Slide Image Classification View paper
- [10] Generating highly accurate pathology reports from gigapixel whole slide images with HistoGPT View paper
- [11] Advances in multiple instance learning for whole slide image analysis: Techniques, challenges, and future directions View paper
- [12] Scaling vision transformers to gigapixel images via hierarchical self-supervised learning View paper
- [13] Fast and Accurate Gigapixel Pathological Image Classification with Hierarchical Distillation Multi-Instance Learning View paper
- [14] A Review of Advanced Deep Learning Methods of Multi-Target Segmentation for Breast Cancer WSIs View paper
- [15] Fast GPU-enabled color normalization for digital pathology View paper
- [16] A weakly supervised deep learning approach for guiding ovarian cancer treatment in prognosis: gigapixel histopathology analysis for personalized therapeutic â¦ View paper
- [17] Rapid 3D imaging at cellular resolution for digital cytopathology with a multi-camera array scanner (MCAS) View paper
- [18] Slidechat: A large vision-language assistant for whole-slide pathology image understanding View paper
- [19] Evolutionary deep feature selection for compact representation of gigapixel images in digital pathology View paper
- [20] MGPATH: Vision-Language Model with Multi-Granular Prompt Learning for Few-Shot WSI Classification View paper
- [21] Task-Specific Fine-Tuning via Variational Information Bottleneck for Weakly-Supervised Pathology Whole Slide Image Classification View paper
- [22] From Pixels to Histopathology: A Graph-Based Framework for Interpretable Whole Slide Image Analysis View paper
- [23] PathBench: Advancing the Benchmark of Large Multimodal Models for Pathology Image Understanding at Patch and Whole Slide Level View paper
- [24] Vim4Path: Self-Supervised Vision Mamba for Histopathology Images View paper

- [25] Gigapixel whole-slide images classification using locally supervised learning View paper
- [26] Few-Shot Learning from Gigapixel Images via Hierarchical Vision-Language Alignment and Modeling View paper
- [27] Beyond multiple instance learning: Full resolution all-in-memory end-to-end pathology slide modeling View paper
- [28] Controllable Latent Space Augmentation for Digital Pathology View paper
- [29] Rethinking pre-trained feature extractor selection in multiple instance learning for whole slide image classification View paper
- [30] Deep learning-based sparse whole-slide image analysis for the diagnosis of gastric intestinal metaplasia View paper
- [31] High-Resolution Histopathology Whole Slide Image Generation Using Wavelet Diffusion Model View paper
- [32] Bayesian Collaborative Learning for Whole-Slide Image Classification View paper
- [33] Learning Heterogeneous Tissues with Mixture of Experts for Gigapixel Whole Slide Images View paper
- [34] Diffusion-based generation of histopathological whole slide images at a gigapixel scale View paper
- [35] SC-MIL: Supervised Contrastive Multiple Instance Learning for Imbalanced Classification in Pathology View paper
- [36] Scalable deep learning artificial intelligence histopathology slide analysis and validation View paper
- [37] Efficient Classification of Very High Resolution Images View paper
- [38] PathM3: A Multimodal Multi-Task Multiple Instance Learning Framework for Whole Slide Image Classification and Captioning View paper
- [39] Multi-scale attention-based multiple instance learning for classification of multi-gigapixel histology images View paper
- [40] Gigapixel histopathological image analysis using attention-based neural networks View paper
- [41] PLUTO: Pathology-Universal Transformer View paper
- [42] Coca-Mil: Attention-Based Handcrafted-Deep Feature Fusion in Computational Pathology View paper
- [43] Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis View paper
- [44] PathAlign: A vision-language model for whole slide images in histopathology View paper
- [45] Dynamic Graph Representation for WSI Classification: A Knowledge-Aware Attention Mechanism for Enhanced Computational Pathology View paper
- [46] PAMIL: Prototype Attention-Based Multiple Instance Learning for Whole Slide Image Classification View paper
- [47] Magnifying networks for histopathological images with billions of pixels View paper
- [48] A self-supervised framework for learning whole slide representations View paper
- [49] PathFLIP: Fine-grained Language-Image Pretraining for Versatile Computational Pathology View paper
- [50] A Decade of GigaScience: The Challenges of Gigapixel Pathology Images View paper
- [51] Attention-challenging multiple instance learning for whole slide image classification View paper
- [52] Reducing Cross-Sensor Domain Gaps in Tactile Sensing via Few-Sample-Driven Style-to-Content Unsupervised Domain Adaptation View paper
- [53] Nciemil: Rethinking decoupled multiple instance learning framework for histopathological slide classification View paper
- [54] Shared-specific feature learning with bottleneck fusion transformer for multi-modal whole slide image analysis View paper
- [55] From Segments to Concepts: Interpretable Image Classification via Concept-Guided Segmentation View paper
- [56] Flow-MIL: Constructing Highly-expressive Latent Feature Space For Whole Slide Image Classification Using Normalizing Flow View paper
- [57] A Study of Temporal Contextual Semantic Enhanced Fusion Module for AnomalousVideo Detection View paper
- [58] Mitigating Representation Bottlenecks in Multiple Instance Learning View paper
- [59] An investigation into the effectiveness of bottleneck based input control compared to aggregate input control View paper
- [60] MGIML: Cancer Grading With Incomplete Radiology-Pathology Data via Memory Learning and Gradient Homogenization View paper
- [61] Local Self-attention-based Hybrid Multiple Instance Learning for Partial Spoof Speech Detection View paper
- [62] SLAM-AGS: Slide-Label Aware Multi-Task Pretraining Using Adaptive Gradient Surgery in Computational Cytology View paper
- [63] Occlusion-aware tracking for drones using neural methods View paper
- [64] MoMIL: Mixture of Multi-instance Learners for Modeling Multiple Compound Activities in High Content Imaging View paper
- [65] Disentangled Multi-modal Learning of Histology and Transcriptomics for Cancer Characterization View paper
- [66] Facial Analysis of Dyadic Interactions Using Multiple Instance Learning View paper
- [67] Mome: Mixture of multimodal experts for cancer survival prediction View paper
- [68] M4: Multi-Proxy Multi-Gate Mixture of Experts Network for Multiple Instance Learning in Histopathology Image Analysis View paper
- [69] Uni-med: a unified medical generalist foundation model for multi-task learning via connector-MoE View paper
- [70] A Mixture-of-Experts Decision Support System for Digital Pathology View paper
- [71] Multimodal Gated Mixture of Experts Using Whole Slide Image and Flow Cytometry for Multiple Instance Learning Classification of Lymphoma View paper
- [72] Deep multi-instance learning using multi-modal data for diagnosis of lymphocytosis View paper
- [73] Spatially-Aware Mixture of Experts with Log-Logistic Survival Modeling for Whole-Slide Images View paper
- [74] Mamba-HMIL: Hierarchical Multiple Instance Learning via State Space Model for Whole Slide Image Diagnosis View paper