MicroMIL: Graph-Based Multiple Instance Learning for Context-Aware Diagnosis with Microscopic Images

Jongwoo Kim[1*], Bryan Wong[1*], Huazhu Fu[2], Willmer Rafell Quiñones Robles[1], Youngsin Ko[3], Mun Yong Yi[1†]
[1] KAIST, South Korea, [2] IHPC, A*STAR, Singapore, [3] Seegene Medical Foundation, South Korea
* Equal contribution, † Corresponding author

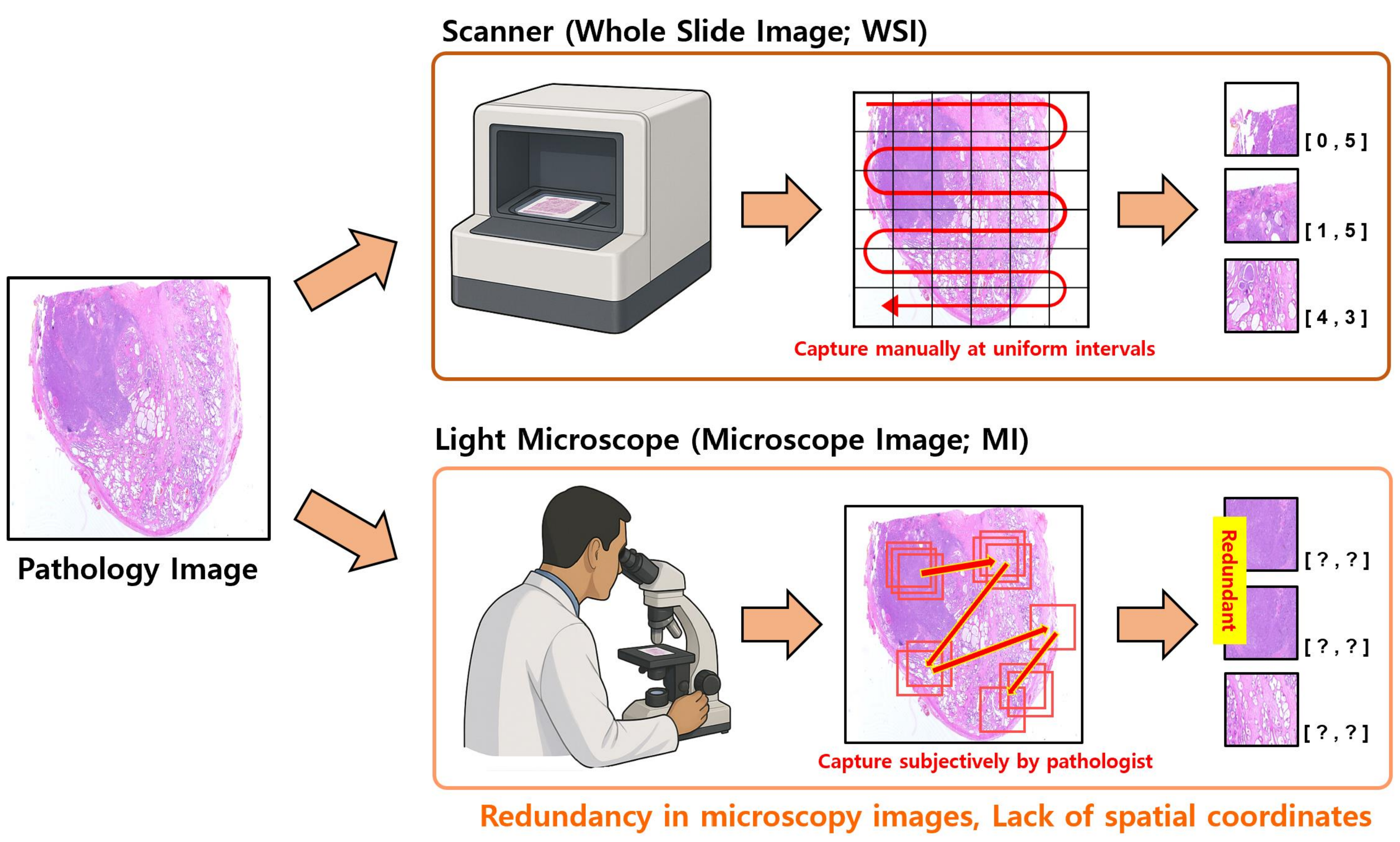
Introduction

Whole Slide Image (WSI)

- Gold standard in computational pathology
- Provides high-resolution tissue insights
- **Limitations**: costly scanners, large storage demands, heavy computation

Light Microscope

- Low-cost and widely used worldwide
- **Potential**: enable AI-powered diagnostics in low-resource settings



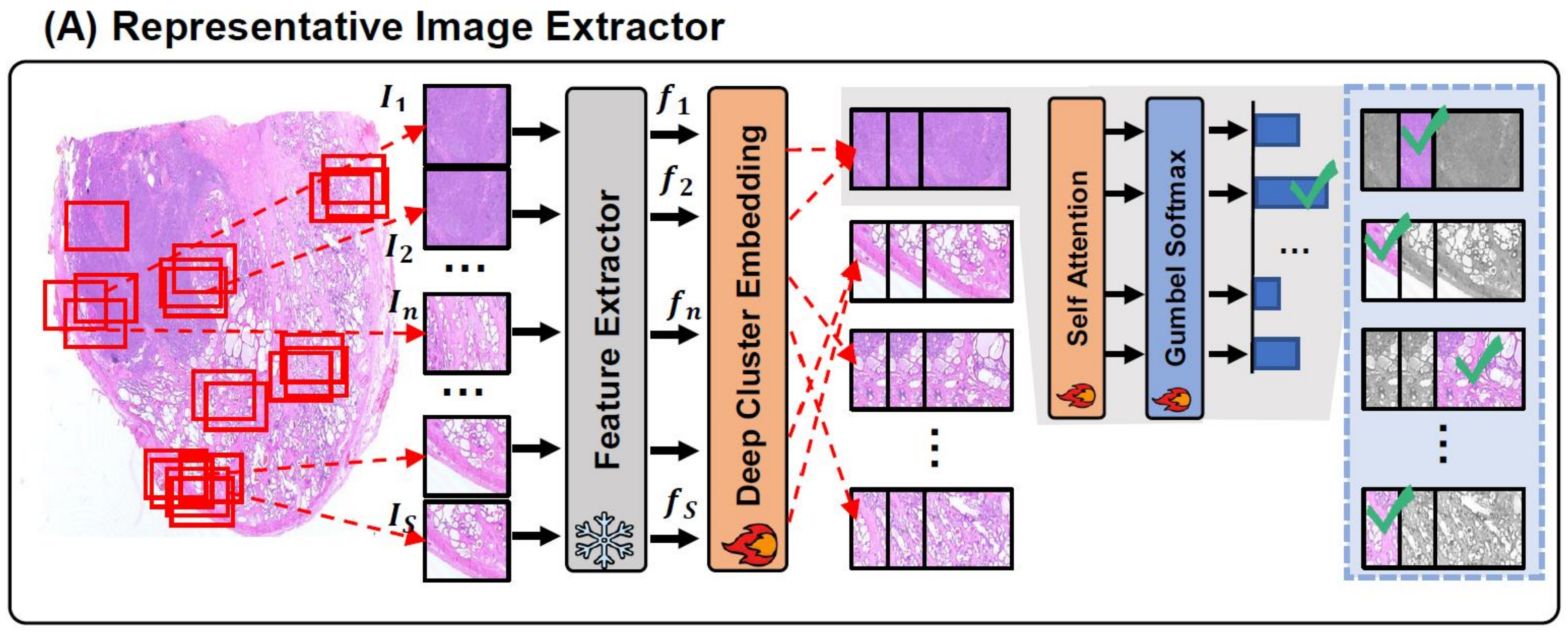
Challenge 1. High Redundancy
Subjective captures by pathologists create redundant instances, reinforcing connections among similar images, and narrowing focus to local details

Challenge 2. Lack of Spatial Coordinates
Microscopy images lack positional metadata, making it infeasible to consider contextual information in pathology image analysis

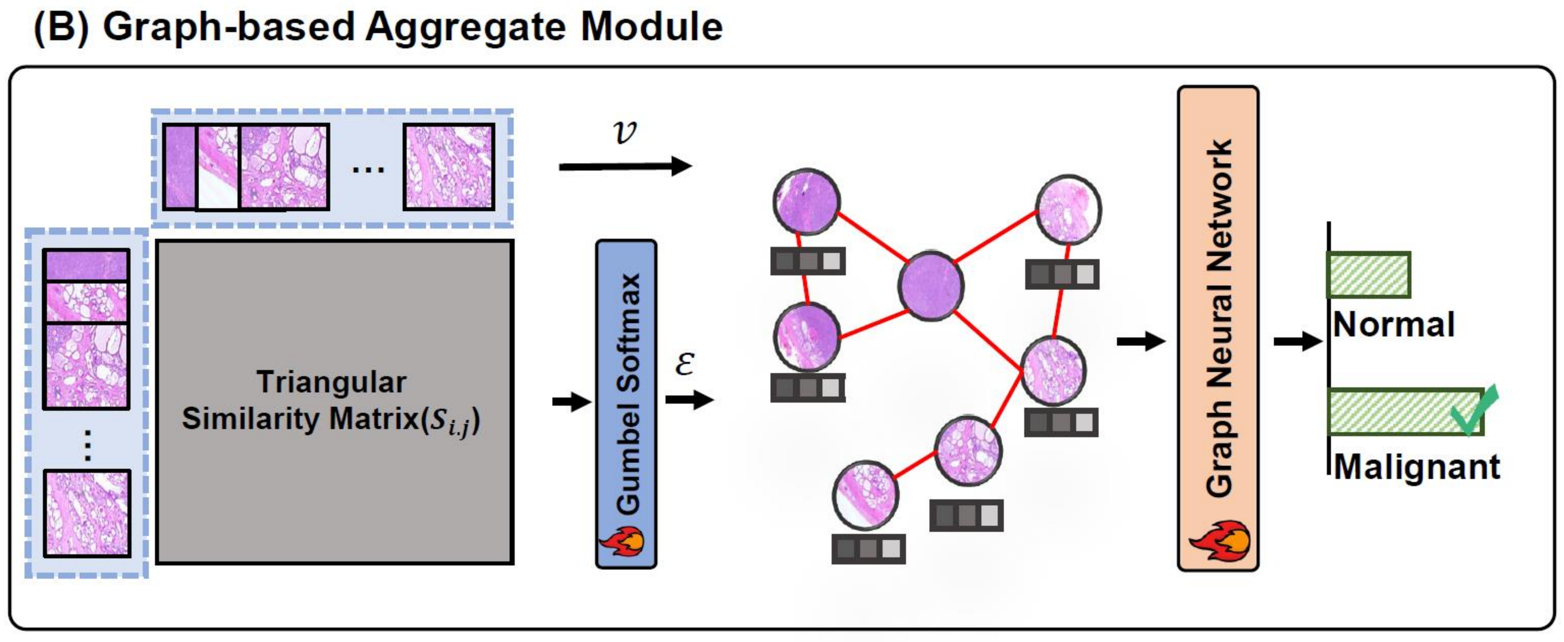
Existing MIL models are not designed for addressing these challenges

Methodology

(A) Representative Image Extractor



(B) Graph-based Aggregate Module



Module 1. Representative Image Extractor (RIE)
Employs DCE and hard Gumbel-Softmax to dynamically reduce redundancy and select representative images in an end-to-end manner

Module 2. Graph-based Aggregation
Builds a graph with representative images as nodes and edges computed via cosine similarity, and leverages GNNs to capture contextual information

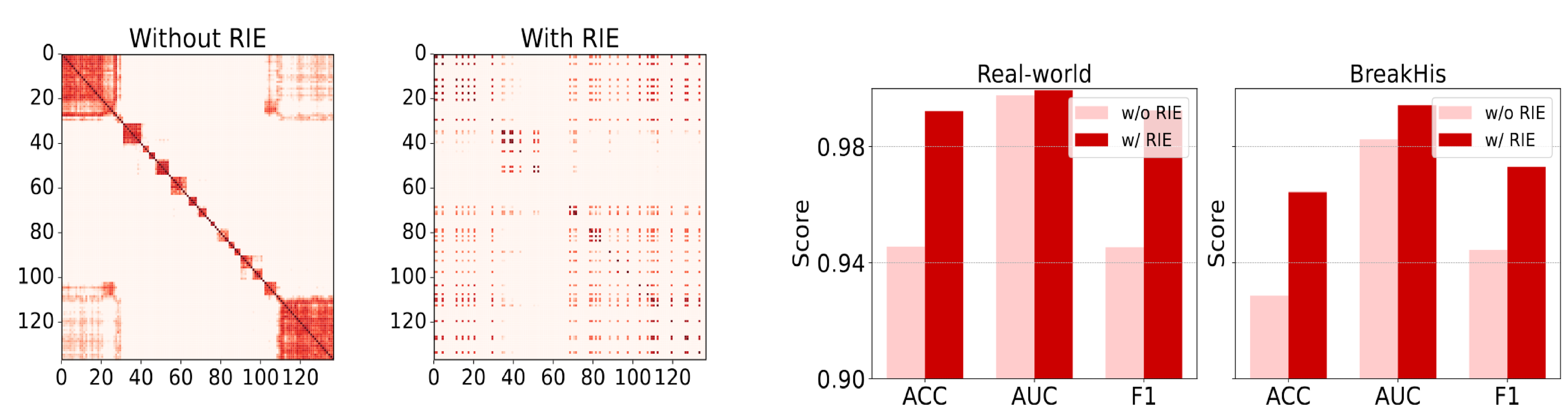
Results

State-of-the-Art Performance

	Real-world (Seegene)			BreakHis		
Model	ACC	AUC	F1	ACC	AUC	F1
ABMIL [ICML'18] [12]	0.9444	0.9764	0.9433	0.8929	0.8947	0.8805
MS-DA-MIL [CVPR'20] [10]	0.9556	0.9829	0.9514	0.8929	0.9591	0.9268
DSMIL [CVPR'21] [16]	0.9444	0.9829	0.9440	0.8214	0.8947	0.8155
CLAM [Nat BioMed'21] [18]	0.9556	0.9873	0.9552	0.9286	0.9298	0.9181
TransMIL [NeurIPS'21] [22]	0.9778	0.9873	0.9776	0.8929	0.9825	0.9268
DTFD-MIL [CVPR'22] [27]	0.9611	0.9901	0.9607	0.9286	0.9766	0.9222
IBMIL [CVPR'23] [17]	0.9611	0.9894	0.9606	0.9286	0.9532	0.9181
ACMIL [ECCV'24] [28]	0.9611	0.9893	0.9606	0.8929	0.9474	0.8857
MicroMIL	**0.9922**	**0.9994**	**0.9925**	**0.9643**	**0.9942**	**0.9730**

- Existing MIL models are designed for scanner-based WSIs, without accounting for high redundancy or missing spatial coordinates
- By addressing the unique challenges of light microscopy, MicroMIL is well-suited for patient diagnosis

Effectiveness of Representative Image Extractor (RIE)



- Without RIE: Limited region interactions, as shown in heatmaps
- With RIE: Enhances diverse region interactions, improving context understanding and performance

Robustness to Image Redundancy

	(1) T10 → B10			(2) B10 → T10			(3) T10 → T10		
Model	ACC	AUC	F1	ACC	AUC	F1	ACC	AUC	F1
ABMIL	0.8090	0.8592	0.7966	0.9213	0.9592	0.9210	0.9091	0.9229	0.9089
MS-DA-MIL	0.9101	0.9526	0.9248	0.9213	0.9642	0.9248	0.9091	0.9348	0.9209
DSMIL	0.9101	0.9474	0.9092	0.9326	0.9755	0.9319	0.9091	0.9438	0.9089
CLAM	0.9326	0.9796	0.9315	0.9213	0.9770	0.9194	0.9318	0.9521	0.9318
TransMIL	0.9213	0.9776	0.9212	0.8989	0.8526	0.8963	0.9318	0.8854	0.9309
DTFD-MIL	0.9438	0.9658	0.9428	0.9213	0.9750	0.9203	0.9318	0.9583	0.9318
IBMIL	0.9326	0.9709	0.9319	0.9326	0.9719	0.9319	0.9318	0.9458	0.9315
ACMIL	0.9434	0.9704	0.9431	0.9213	0.9689	0.9210	0.9318	0.9521	0.9315
MicroMIL	**0.9663**	**0.9842**	**0.9630**	**0.9551**	**0.9801**	**0.9542**	**0.9545**	**0.9958**	**0.9524**

- Evaluates MicroMIL's robustness using a 0.995 similarity threshold across redundancy shifts
- Demonstrates robust performance in both high-and low-redundancy scenarios compared to baselines

Conclusion

- MicroMIL is the first weakly-supervised MIL framework designed for conventional light microscopy images
- It combines DCE and hard Gumbel-Softmax to dynamically reduce redundancy, select representative instances, and construct context-aware graph representations without spatial coordinates

Paper Jongwoo Kim Bryan Wong