



Lomonosov Moscow State University,  
Faculty of Computational Mathematics and Cybernetics,  
Faculty of Medicine,  
Laboratory of Mathematical Methods of Image Processing

 **CBSP**  
**2024**

9th International Conference on Biomedical Imaging, Signal  
Processing (ICBSP 2024), Hong Kong, October 18-20, 2024

# Tissue type classification for whole slide histological images with graph convolutional neural network

Zhongao Sun, A. Khvostikov, A.Krylov,



October 19, 2024

# About me

## Khvostikov Alexander

- ▶ Ph.D., researcher at Lomonosov Moscow State University, Laboratory of Mathematical Methods of Image Processing, conducted by Professor Andrey S. Krylov.
- ▶ Research interests: image processing and analysis, computer vision, medical images, machine learning, deep learning, hybrid methods.
- ▶ <https://istina.msu.ru/profile/xubiker/>



## Acknowledgments:

- ▶ This work is supported by RSCF grant 22-41-02002;

# Digital Pathology

Digital Pathology includes the collection, management, exchange and interpretation of pathological information, including slides and digital data.

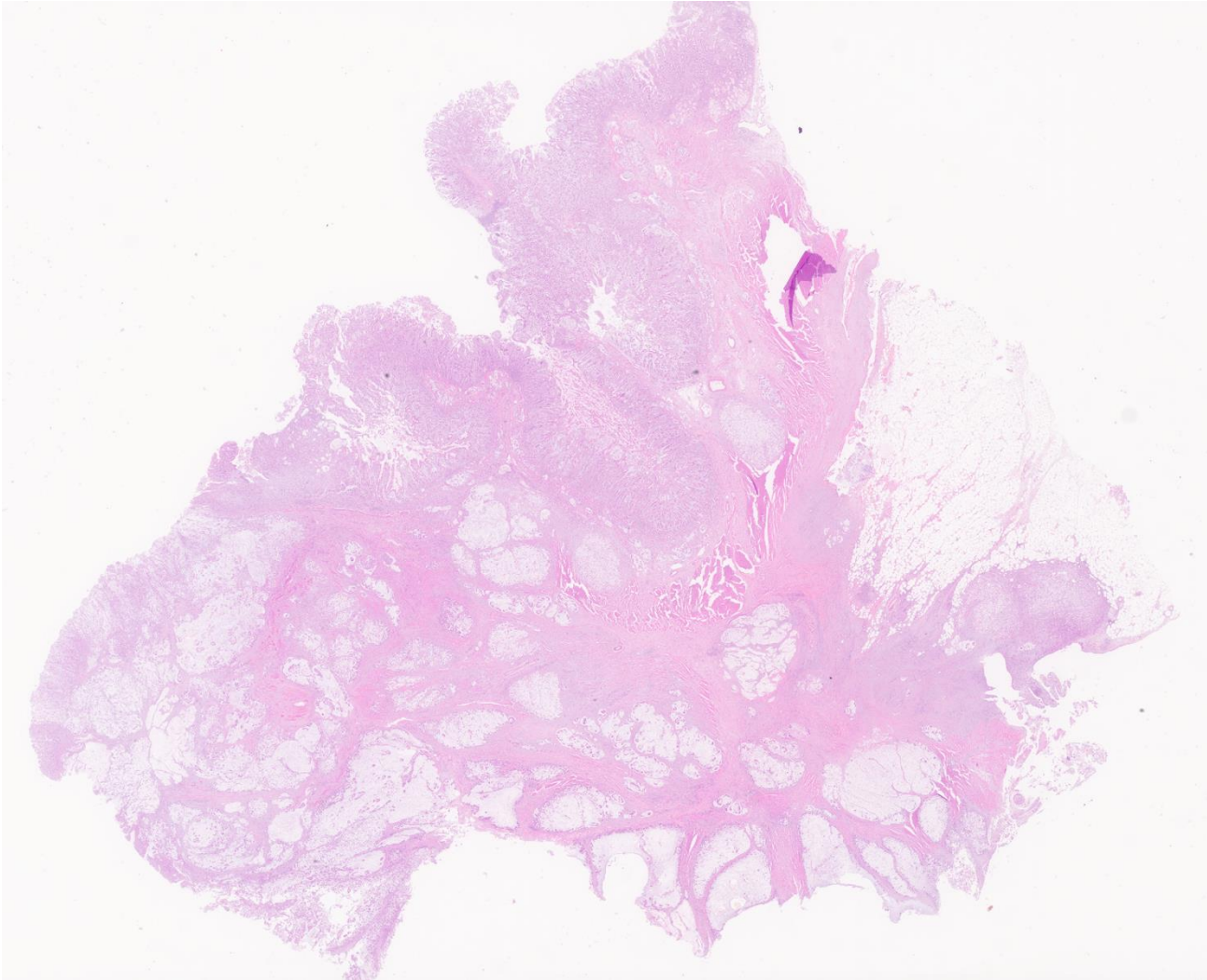
The main stages of the study:

- ▶ taking a biopsy;
- ▶ cutting, placing on glass;
- ▶ staining;
- ▶ glass scanning;
- ▶ image analysis by a histologist;



The images obtained by the scanner are used by histologists for morphological diagnostics (analysis of cellular structures and finding structural abnormalities).

# Digital Pathology. WSI



*An example of WSI.*

A histological whole slide image (WSI) sample:

- ▶ optical magnification 40x;
- ▶ resolution  $111552 \times 90473$  ( $\sim 10^{10}$  px);
- ▶ tiff file: ~3GB;
- ▶ even after 16x downsampling the resolution is  $6972 \times 5654$  px;

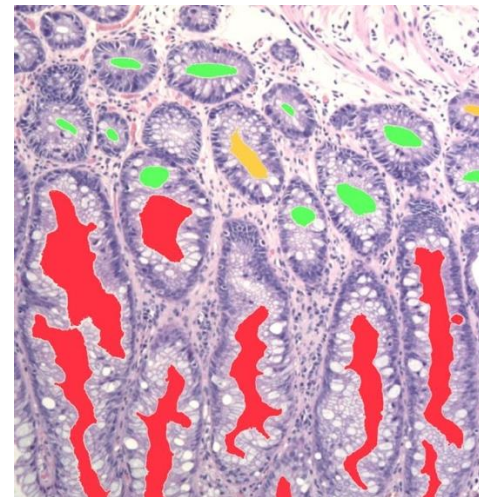
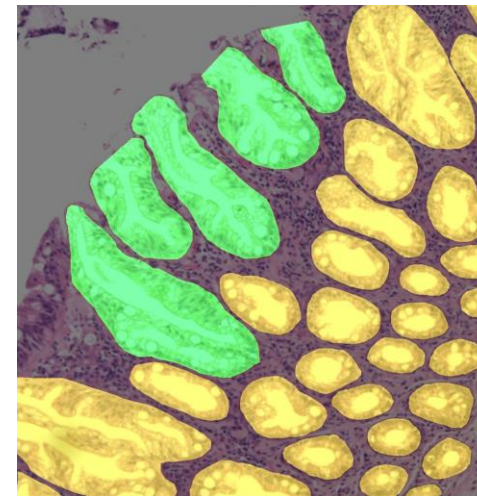
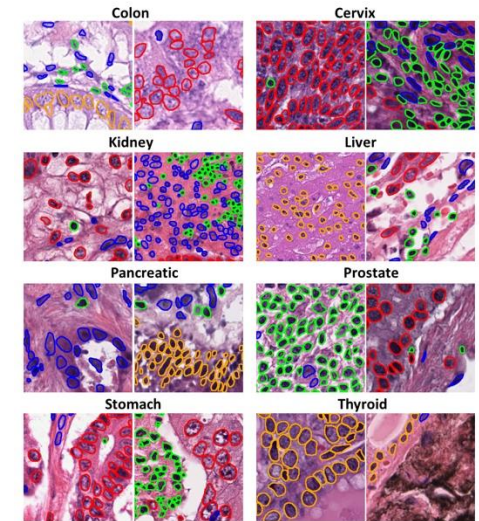
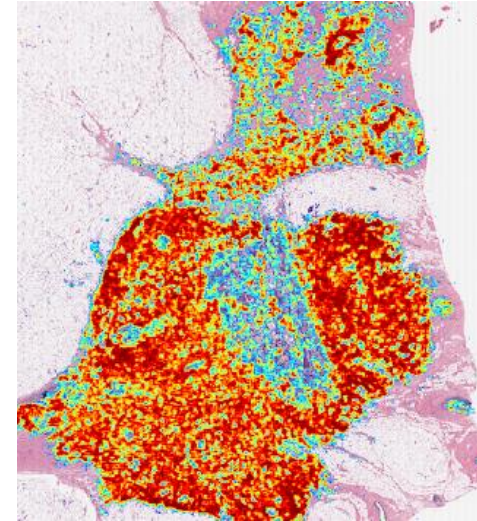
# Problems of tissue type classification for WSIs

## ► Low accuracy of traditional CNN models

- ❖ Due to the memory limitation, the WSI cannot be loaded into neural network, so we must split WSI into small fragments and sent them into the network.
- ❖ But in this case, we lose information about spatial relationships so the performance of CNNs decrease.

## ► GNNs are not yet used for fragment classification

- ❖ Deep neural networks based on graphs (graph neural networks, GNNs) are more effective for describing spatially-connected data.
- ❖ But existing methods yet have attempted to apply GNNs only to whole slide histological images on classifying the label of a set of instances rather than individual instances.




# Without data, algorithms mean nothing → PATH-DT-MSU

PATH-DT-MSU consists of several subsets:

- ▶ S1, S2
- ▶ WSS1, WSS2
- ▶ WSR1

In this work we used PATH-DT-MSU WSS2-v2:


- ▶ 10 whole slide images at x40 magnification (resolution ~ 110,000 x 90,000 pixels)
- ▶ 5 classes (AT, BG, LP, MM, TUM) annotated with polygons
- ▶ Annotated train set: 1560, 7098, 533, 895, 1303 million pixels
- ▶ Annotated test set: 1086, 8032, 318, 743, 1199 million pixels



MMIP  
CMC MSU

**Laboratory of Mathematical Methods of Image Processing**

Chair of Mathematical Physics  
Faculty of Computational Mathematics and Cybernetics  
Lomonosov Moscow State University



[Home](#) | [People](#) | [Research](#) | [Publications](#) | [Activities](#) | [Software](#)

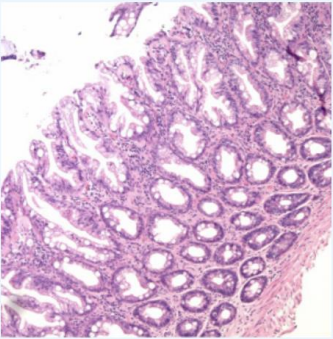
**Laboratory**

- Home
- ▶ People
- ▼ Research
  - ▼ Biomedical imaging
    - Blinking fluorescence
    - Cell tracking
    - ▼ Histology
      - **PATH-DT-MSU**
    - ▶ MRI
  - ▼ Image enhancement
    - Resampling & Superresolution
    - Grid warping
    - Rank image filtering
  - ▼ Image analysis
    - Basic edges
    - Edge width
    - ▶ Ringing effect
    - Keypoints & descriptors
  - ▼ Geology
    - LumenStone
  - ▼ Medical imaging
    - Ophthalmology
    - Tomography
  - ▼ Audio

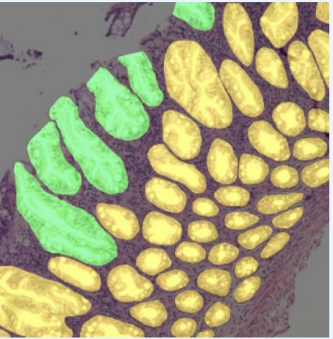
## PATH-DT-MSU Dataset

### Description

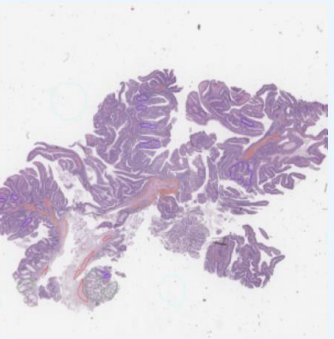
Real biopsy and surgical material from various parts of the human digestive tract was used for paraffin blocks preparation. Microscopic examination was performed using microscope Leica DM2500 (Leica Microsystems, Germany). Microscope Leica DM4000B/DFC495 and scanner Leica SCN400 (Leica Microsystems, Germany) were used for high resolution histological images acquisition.



sample image from S1 set



visualization of corresponding gland and "open" gland annotations from S1 set



preview of whole slide image from WSS1 with visualization of annotation

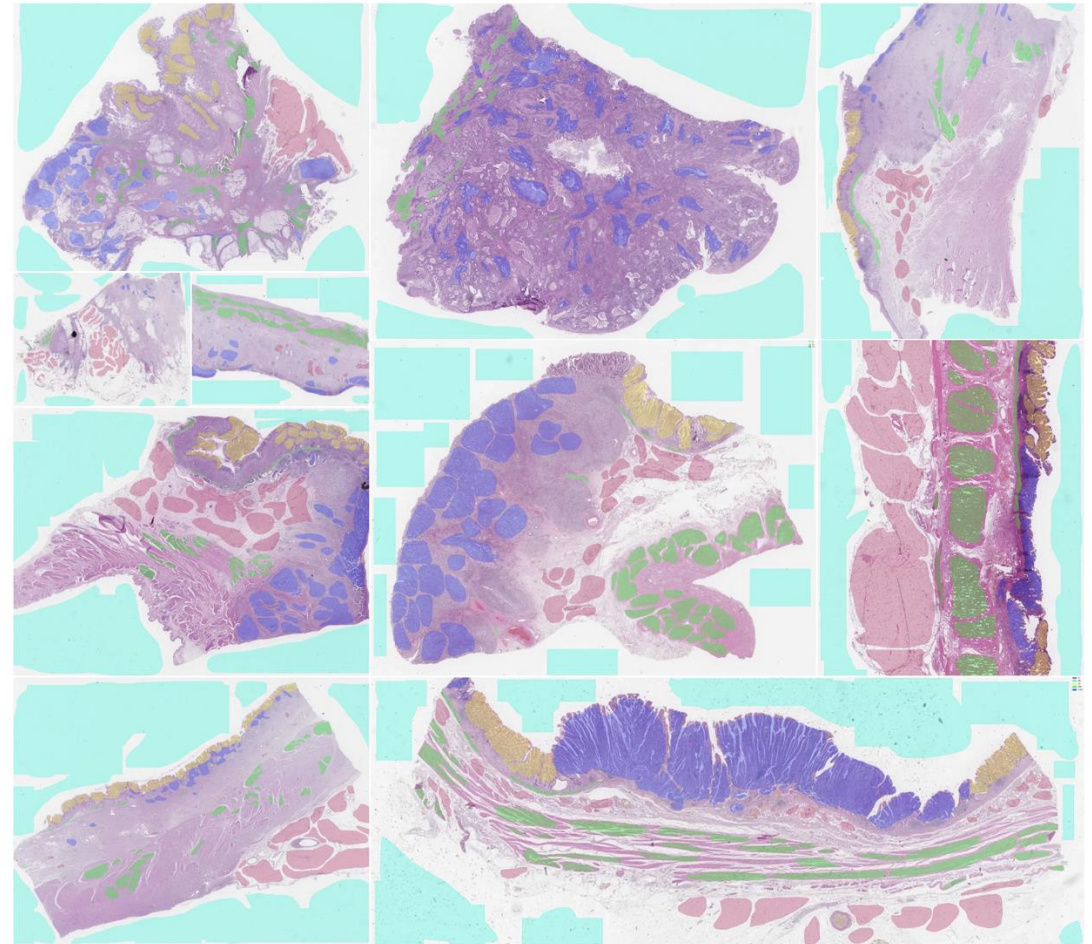
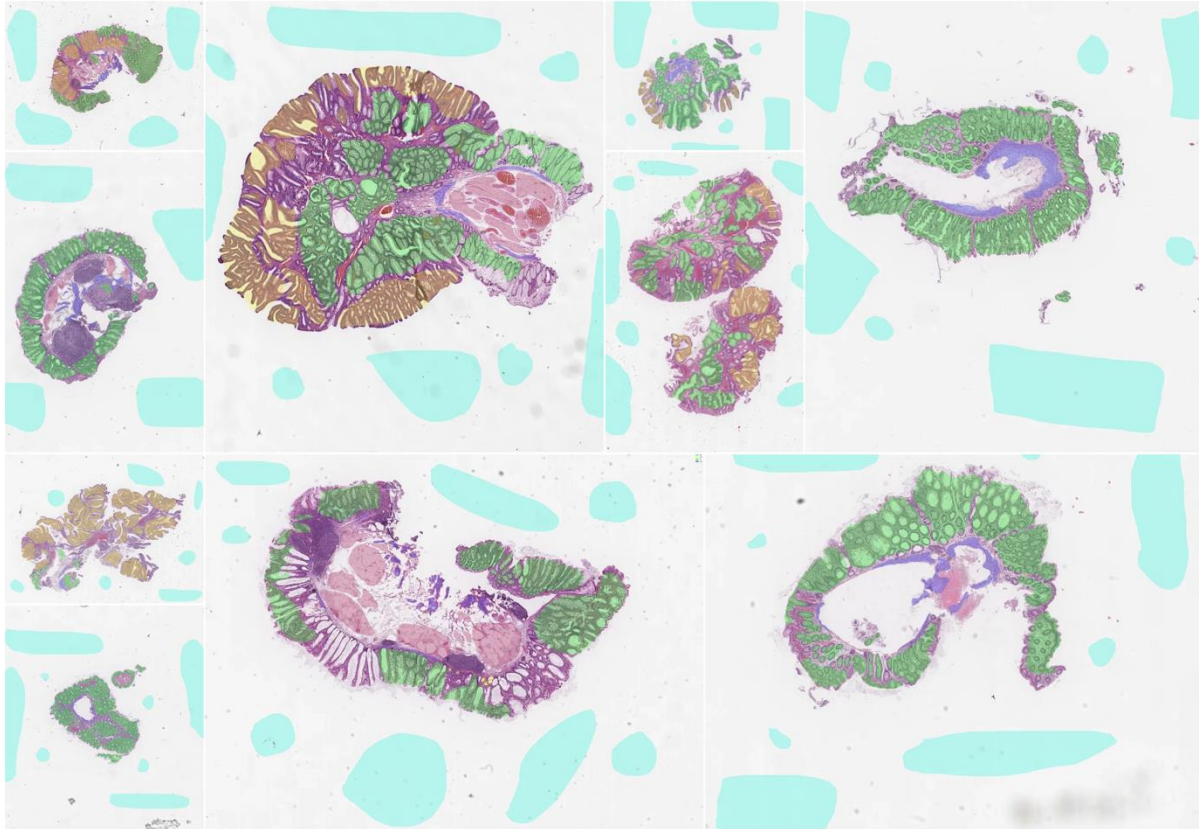
PATH-DT-MSU dataset was created to unite high-quality histological images of different parts of human gastrointestinal tract. It consists of several sets, containing different image types and annotations for different target tasks:

- **S1**: histological images of colon with annotations for segmentation;
- **S2**: histological images of stomach with annotations for segmentation (planned);
- **S3**: histological images of esophagus with annotations for segmentation (planned);
- **WSS1**: whole slide images of colon polyps with annotations for segmentation;
- **WSS2**: whole slide images of gastric cancer with annotation for segmentation;
- **WSR1**: whole slide images with annotations for registration (planned);

All images are annotated, the annotation type for different sets is different according to the image type and target task. All images in sets for convenience are already split into train and test samples.

<http://imaging.cs.msu.ru/en/research/histology/path-dt-msu>

# PATH-DT-MSU. WSS2 v2

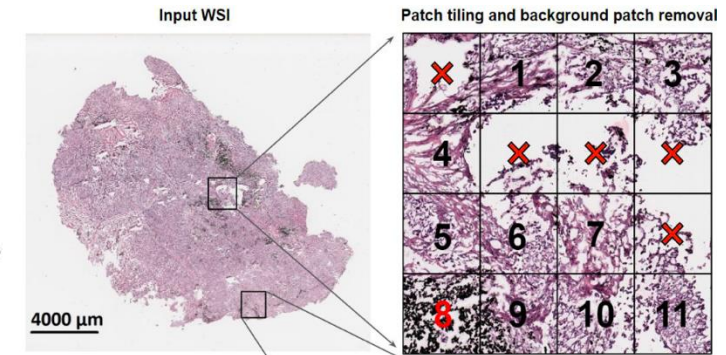
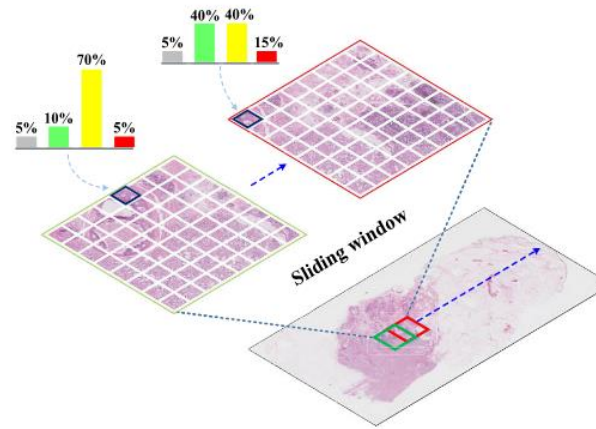


<http://imaging.cs.msu.ru/en/research/histology/path-dt-msu>

# Sub-graph sampling strategy

## ► Existing sub-graph sampling strategies

Most related works apply GNN on WSIs sample the sub-graph with a large sliding windows, which is sub-optimal. This is because most ground truth patches from this window are the same, and unannotated areas are also included. This prevent the model from fully utilizing location information and learning valuable features.

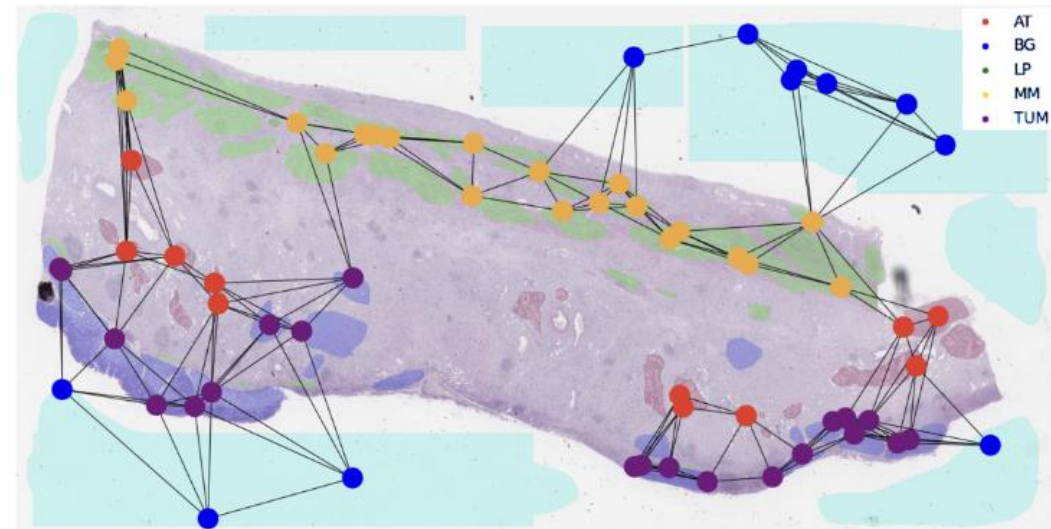


## ► Proposed sampling strategy

We select patches only from annotated regions and the probability of a region being selected is related to its area:

$$\hat{p}_i = \begin{cases} \frac{1}{N} + \left( \frac{S_i^{-1}}{\sum_{j=1}^N S_j^{-1}} - \frac{1}{N} \right) \cdot c, & \text{if } c \in [-1, 0], \\ \frac{1}{N} + \left( \frac{S_i}{\sum_{j=1}^N S_j} - \frac{1}{N} \right) \cdot c, & \text{if } c \in [0, 1], \end{cases}$$

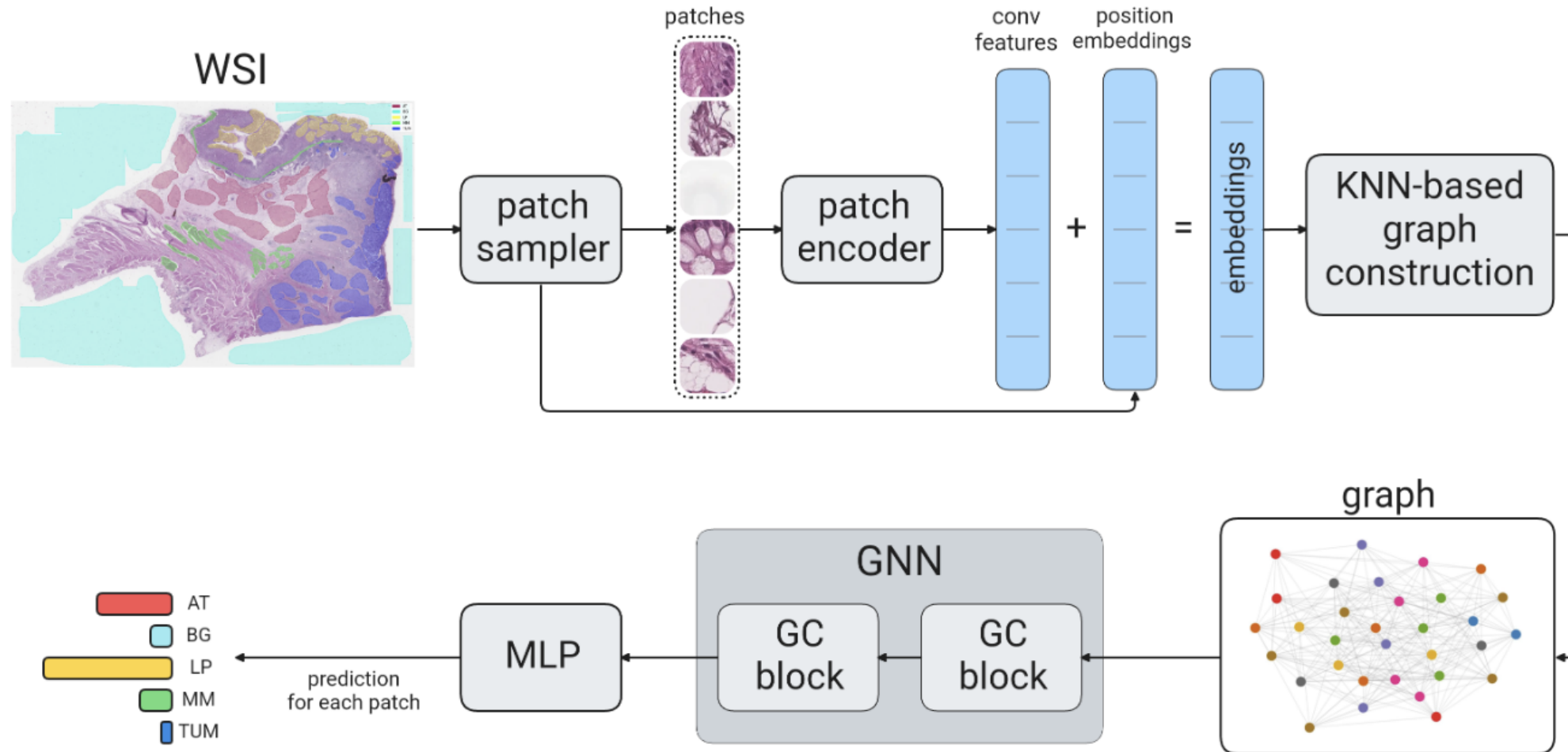
$$p_i = \hat{p}_i / \sum_{j=1}^N \hat{p}_j,$$





# Proposed model architecture

The proposed pipeline consists of patch sampling step, feature embedding step, graph constructing step, 2 GC (attention) blocks and MLP head:



## Proposed model architecture. Graph convolution block (GC block)

### ► Feature embeddings:

The input of the graph neural network can be defined as follow:

$$X_{in} = X + p_x,$$

where  $X$  is the feature embedding formed with the patch-based encoder and positional embedding  $p_x$  is the absolute position encoding (generated with one linear layer from coords).

### ► Graph attention:

For each edge  $e_{uv}$  that connects nodes  $u$  and  $v$  from the  $X_{in}$  we calculate edge attention:

$$\alpha_{uv} = \frac{\exp(\sigma(L[\mathbf{W}f_u || \mathbf{W}f_v]))}{\sum_{k \in N_i} \exp(\sigma(L[\mathbf{W}f_u || \mathbf{W}f_k]))},$$

Where  $f_u, f_v, f_k$  are the features of nodes  $u, v, k$ .  $N_i$  is a set of neighbor nodes,  $||$  is the concatenation operator.  $L()$  is a linear transformation and  $\sigma()$  is a RELU function.

The final output of GC block is obtained as:

$$f'_u = \sigma \left( \sum_{v \in N_i} \alpha_{uv} f_v \right) + f_u.$$

# Main Results

## 1. Effectiveness of graph-based approach:

Graph-based models, including ours, demonstrate a significant improvement over patch-based models in Macro-F1 score, highlighting the importance of incorporating spatial and neighborhood information for tissue classification.

## 2. Improved graph sampling:

great performance gain compared to previous graph-based methods indicates the effectiveness of the proposed graph sampling strategy in capturing relevant location and neighborhood context;

## 3. Success in challenging categories:

best performance in classifying MM and TUM tissues (have been historically challenging);

method \ F1-score	AT	BG	LP	MM	TUM	type	Macro F1
ResNet50[15]	0.83	0.98	0.58	0.82	0.70	patch-based	0.78
DenseNet121[23]	0.89	0.99	0.81	0.84	0.80	patch-based	0.83
Efficient Net[24]	0.84	0.98	0.74	0.80	0.81	patch-based	0.82
Mobile Net[25]	0.82	0.99	0.60	0.84	0.67	patch-based	0.78
SR+CLS [5]	0.86	0.99	0.86	0.75	0.71	patch-based	0.83
Graph V-net [11]	0.86	1.00	0.90	0.89	0.87	graph-based	0.91
GraphSAGE [26]	0.87	1.00	0.87	0.89	0.86	graph-based	0.90
EdgeConv2d [27]	0.88	1.00	0.81	0.91	0.85	graph-based	0.89
MRConv2d [28]	0.91	1.00	0.78	0.91	0.84	graph-based	0.90
SGFormer[29]	0.92	1.00	0.87	0.91	0.87	graph-based	0.91
Ours	<b>0.96</b>	<b>1.00</b>	0.85	<b>0.94</b>	<b>0.88</b>	graph-based	<b>0.92</b>

**Table 1: Comparison of the results from different convolution-based and graph-based models for tissue type classification using the PATH-DT-MSU WSS2v2 dataset.**

# Ablation studies. Encoder

## ► Ablations on the patch-based encoder:

1. Pre-training and fine-tuning are crucial;
2. Encoder architecture has limited impact;

Encoder \ F1-score	AT	BG	LP	MM	TUM	Macro F1
Resnet50 (from scratch)	0.90	0.99	0.87	0.82	0.87	0.89
Resnet50 *	0.91	1.00	0.87	0.87	0.88	0.91
Resnet50 **	0.93	1.00	0.89	0.93	0.88	<b>0.93</b>
DenseNet121 **	0.92	1.00	0.88	0.93	0.89	0.92
SR+CLS **	0.91	1.00	0.90	0.91	0.88	<b>0.93</b>

\* - pretrained on ImageNet

\*\* - pretrained on ImageNet + finetuned on PATH-DT-MSU WSS2v2

# Ablation studies. Graph construction

## ► Ablations on the graph construction strategy:

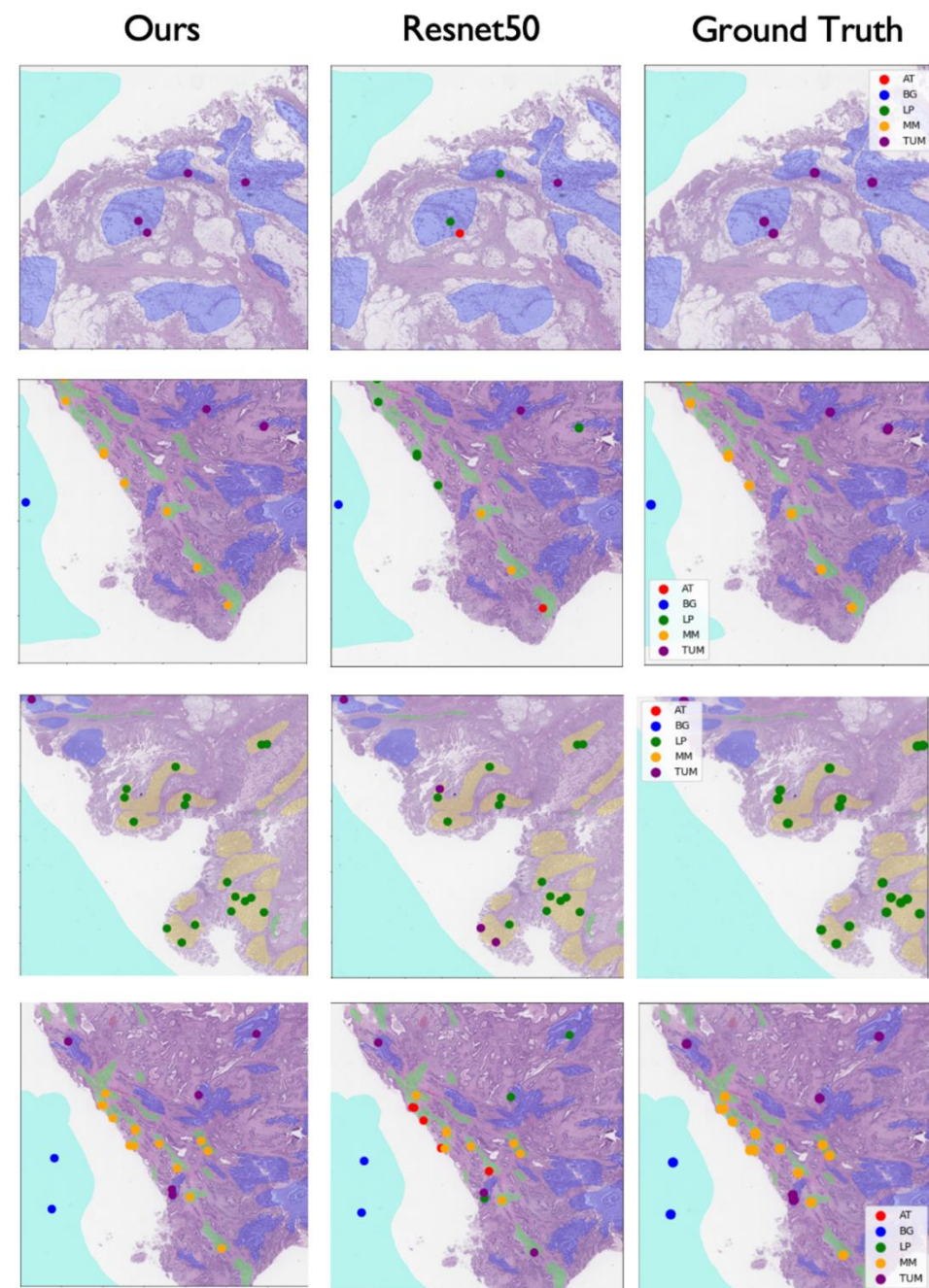
1. Spatial distance outperforms feature similarity.
2. KNN by distance with  $k = 5$  yields best results

Graph construction strategy	Macro F1	Recall	Accuracy
KNN by features (k=3)	0.88	0.89	0.89
KNN by features (k=5)	0.89	0.89	0.89
KNN by distance (k=3)	0.90	0.91	0.91
KNN by distance (k=5) *	<b>0.92</b>	<b>0.92</b>	<b>0.92</b>
Fully Connected [9, 11, 12]	0.85	0.86	0.86

\* Fully connected means constructing graph from large sliding window

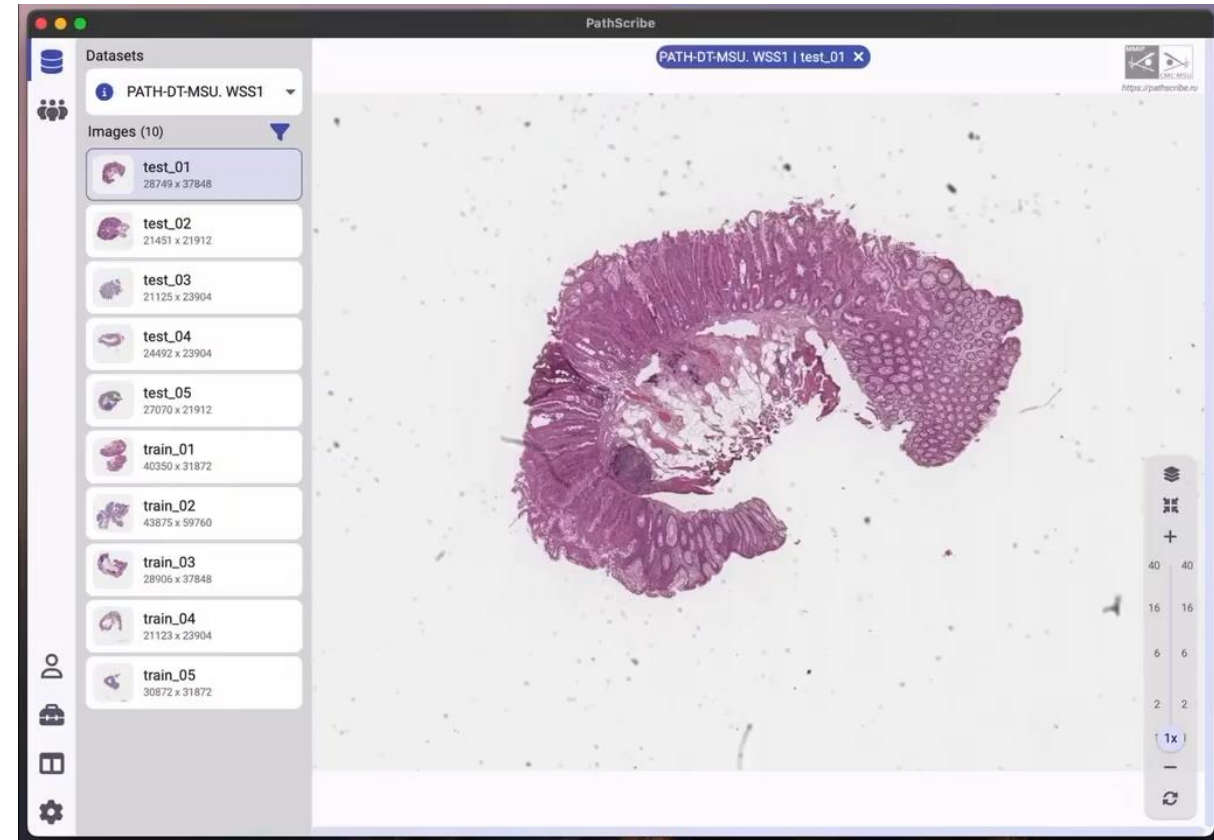
# Conclusion

1. We presented a novel graph-based method for tissue classification in histological WSIs that leverage GCNNs and a new sub-graph sampling strategy;
2. The proposed method demonstrates superior performance compared to state-of-the-art patch-based and graph-based methods (**macro F1 = 0.92** for PATH-DT-MSU WSS2-v2 test set);



# Future work

1. Inference on whole slide images:
  - ▶ extract all patches and construct multiple sub-graphs;
  - ▶ due to class imbalance use the same node in multiple sub-graphs, with its prediction being averaged;
  - ▶ uniting all nodes' prediction, obtain a coarse segmentation of the entire WSI
2. More detailed testing on various datasets;
3. Integrating the proposed method of WSI segmentation into PathScribe software (<https://pathscribe.ru/>);



*Demo video of main desktop view of PathScribe for macOS.*



Thank you for  
your attention!