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

PathScribe: new software to work with whole slide histological images for education and research

A. Khvostikov, V. Ippolitov, A.Krylov,

I.Mikhailov, P.Malkov

October 21, 2023

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/

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


An example of WSI.

A histological whole slide image (WSI) sample:

▶ optical magnification 40x;

▶ resolution 111552×90473 (~10^{10} px);

▶ tiff file: ~3GB;

▶ even after 16x downsampling the resolution is 6972×5654 px;

Tasks and problems of histological image analysis

▶ Image segmentation

    ★ semantic / object segmentation ★ tissue type recognition
    ★ identifying individual objects

▶ Image registration and processing

    ★ registration of adjacent slices ★ «restaining»

▶ Calculation of morphometric characteristics

    ★ determination the shape of lumen glands ★ counting the number
    of nucleoli ★ identification of tumour and invasion depth

▶ Analysis and search tasks

    ★ finding artefacts ★ content-based retrieval

...

▶ The problem of obtaining labelled data

▶ WSI data processing problems
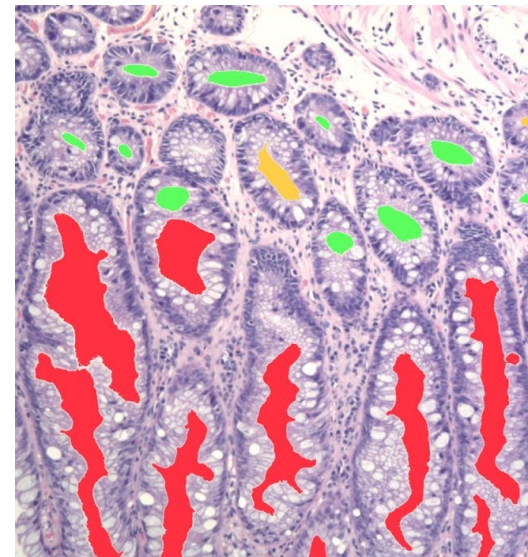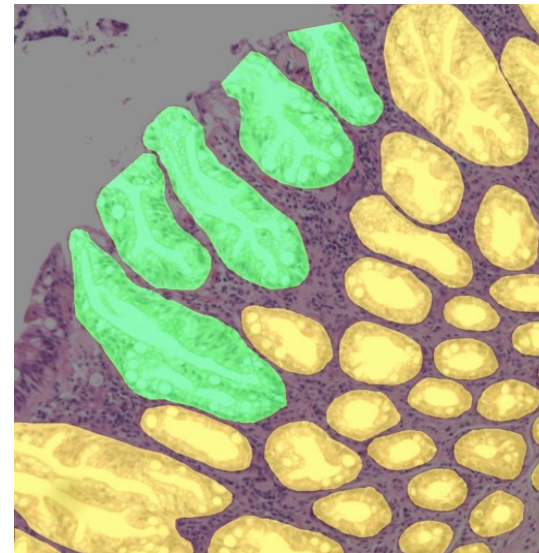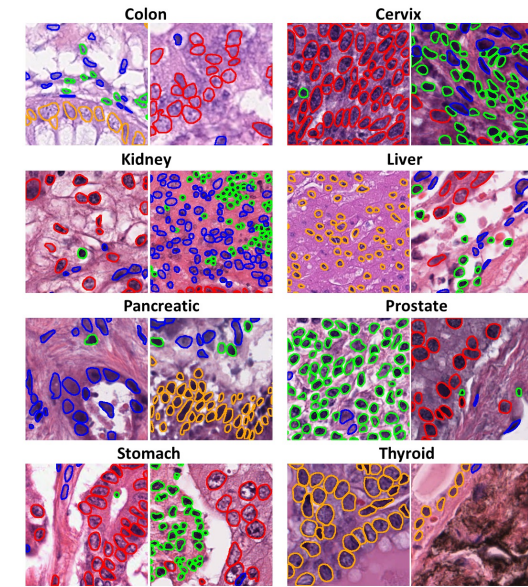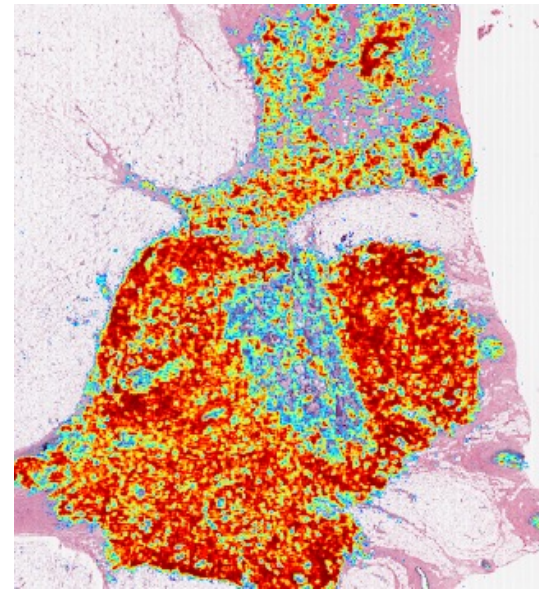
Without data, algorithms mean nothing → PATH-DT-MSU

PATH-DT-MSU 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);

Now the dataset contains 80 "usual" images (3263x2442, 2176x1632) with full annotation into 2 classes of glands and 20 whole slide images with polygonal annotations of 5 tissue classes.

The **total size** of the dataset is **22 GB**.

Laboratory of Mathematical Methods of Image Processing
Chair of Mathematical Physics
Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University

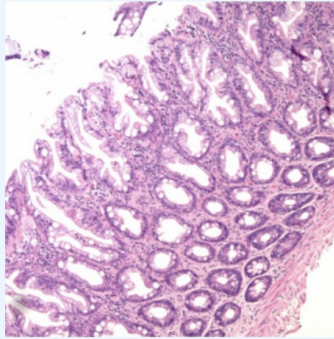Home | People | Research | Publications | Activities | Software

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
    - Keypoins & descriptors
  - Geology
    - LumenStone
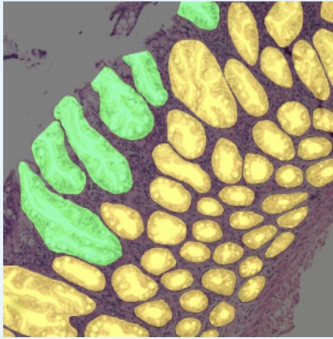 - Medical imaging
    - Ophtalmology
    - Tomography
  - Audio
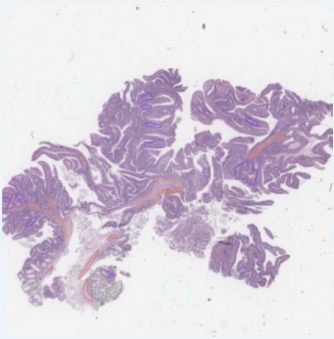
PATH-DT-MSU Dataset

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

Real world problems of WSI analysis algorithms

Main steps of applied research in digital pathology:

▶ collecting data,

▶ annotating and labelling data,

▶ applying of the developed algorithms,

▶ visualizing the results,

▶ share the results.

All this actions use <u>different</u> software solutions with a <u>low level of integration</u>.
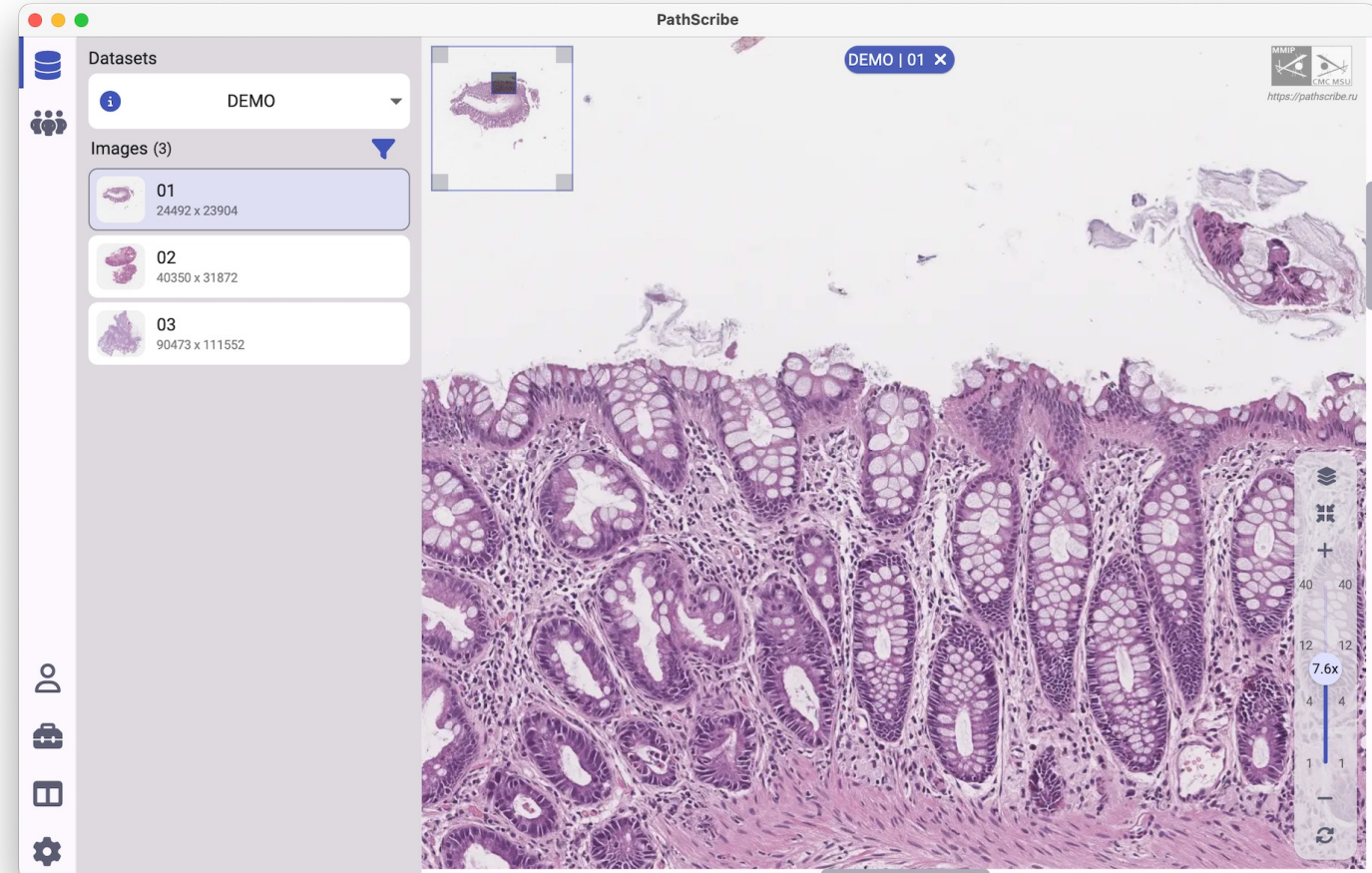
PathScribe

🔬 **PathScribe** is a cross-platform software for viewing, annotating, labelling and automatic analysis of histological images, including whole slide images.

Pathscribe is conceived and developed as a universal tool and can be used both for science and education purposes.

PathScribe implements an adaptive interface that allows to comfortably work with WSIs, both on a desktop computer or laptop, and on mobile devices (tablets, smartphones) with touch screens on different platforms:

▶ Windows,

▶ Android,

▶ MacOS,

▶ Linux.

Screenshot of the main desktop view of PathScribe for macOS. The displayed whole slide image is from DEMO dataset.

Comparison of different software for digital pathology

	Fiji [2]	QuPath [3]	ICY [4]	Orbit [5]	Cytomine [6]	Pathomation [7]	PathScribe (ours)
WSI support	yes	yes	limited	yes	yes	yes	yes
local / cloud working	local	local	local	local	cloud	local, cloud	cloud
desktop clients	Windows, Mac, Linux	Windows, Mac, Linux	Windows, Mac, Linux	Windows, Mac, Linux	web	web	Windows, Mac, Linux
mobile clients	no	no	no	no	no	no	yes (Android)
app runtime	VM (Java)	VM (Java)	VM (Java)	VM (Java)	web	web	native
annotate images	yes	yes	yes	yes	yes	yes	yes
image analysis methods	yes	yes	yes	yes	yes	yes	yes (in development)
education functionality	no	limited (via plugin)	no	no	yes	yes	yes
plugins	yes	yes	yes	no	yes	yes	no
collaborative work	no	no	no	no	yes	yes	yes (in development)

Data storage and transmission architecture

As part of PathScribe, we developed our own universal format (.psi) for storing histological images, as well as a converter.

Main features of the developed format:

▶ store image in tiles,

▶ use pyramid of scales,

▶ effectively compress tiles,

▶ analyse tile content to remove the background.

All this reduces the size of the WSIs by **2.5-5 times** compared to the conventional formats.

We also store images in 2 levels of quality and allow users to switch between them (this gives **+20% speedup**).

Image formats currently supported by PathScribe converter:

❖ universal formats: `jpg, png, bmp, tif;`

❖ formats used in histological scanners: `svs, mrxs, scn, tiff;`
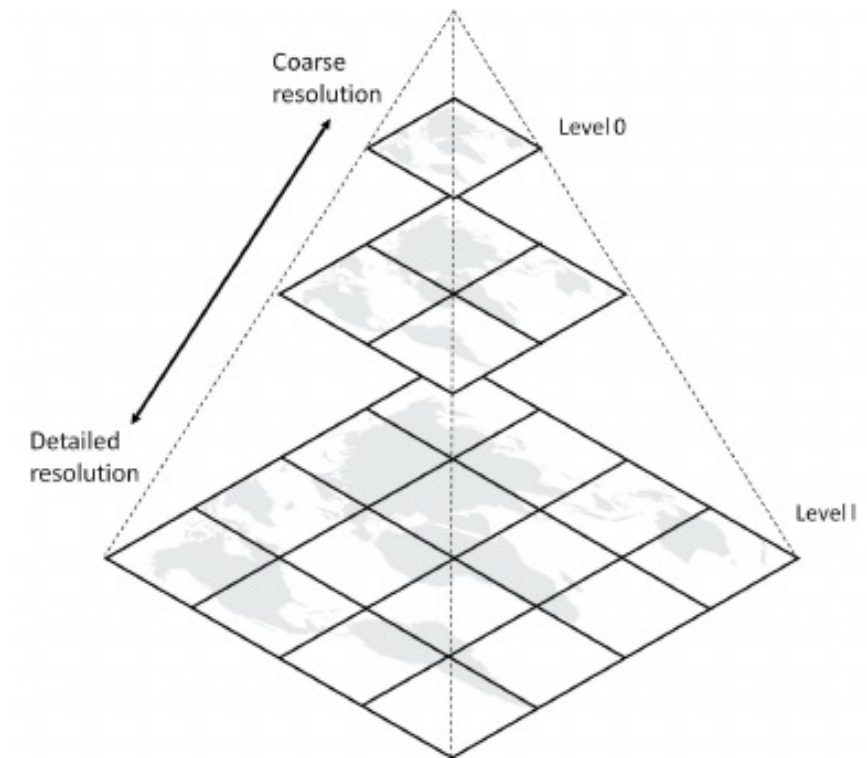
Representation of a pyramide of tiles for WSI.

Data storage and transmission architecture

PathScribe is designed for quick and easy working with WSIs. The obtained image loading speed of PathScribe is achieved through a few tricks:

▶ only the tiles that are currently needed are loaded,

▶ loading tiles and all meta information is performed asynchronously,

▶ PathScribe makes assumptions about where the user will move next and preloads corresponding tiles in background,

▶ tiles are loaded with a few parallel processes and are also cached on the device.

**P.S. And you can work with several WSIs simultaneously!**
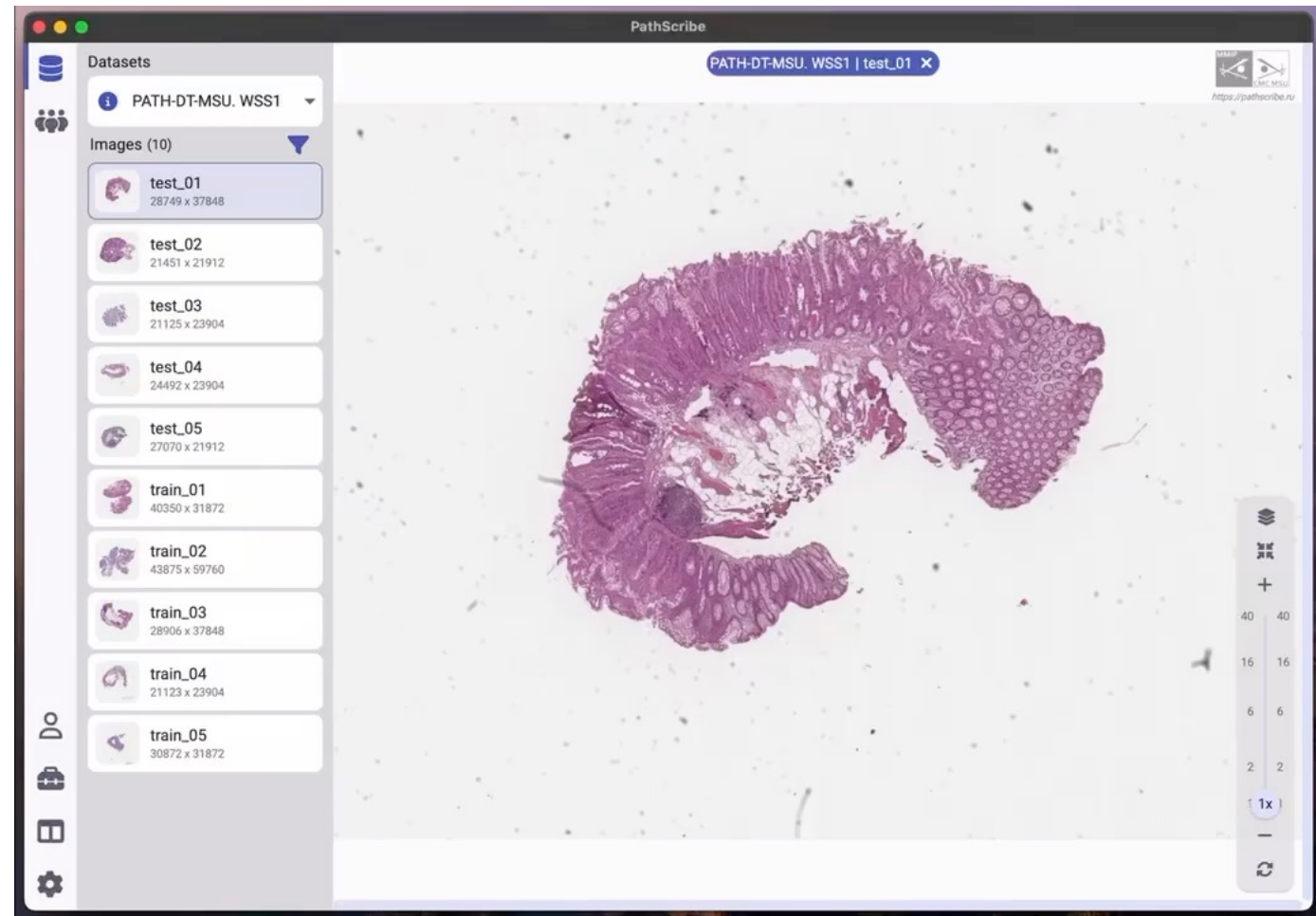
Demo video of main desktop view of PathScribe for macOS.

Educational features

PathScribe has been designed to simplify the work of histologists and to be a handy and versatile tool that can be used in education. It is useful for teaching and learning pathology since it:

▶ works on all kind of devices, **including smartphones and tablets**,

▶ has a flexible configurable system of fields and tags that allows to formalize and store the descriptions of images accordingly to all terminology that is used in histology,

▶ can filter images by tags,

▶ has collab mode for performing presentations and collaborative work with WSIs *(in development)*.
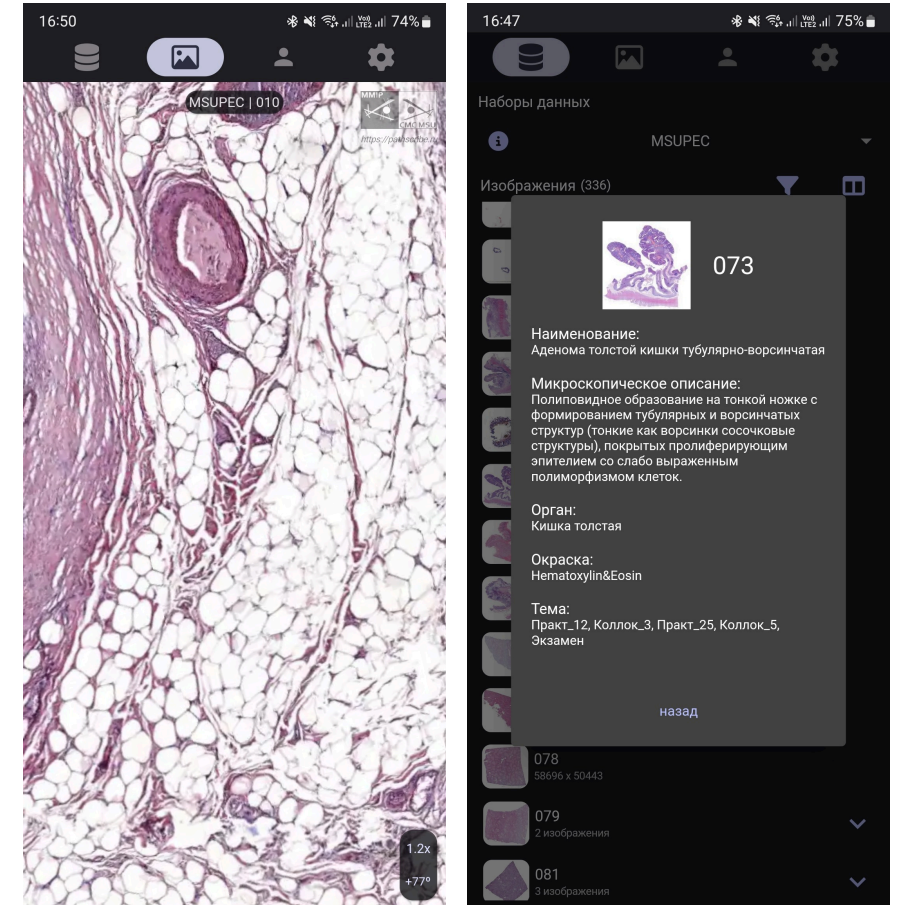
*Screenshots of PathScribe (educational collection) Android client, mobile view.*

Educational features

PathScribe is used in teaching the pathology course at faculty of Fundamental Medicine, Lomonosov Moscow State University **since fall 2022.**

The course includes:

➢ 34 lectures, 28 practical classes (10 to 30 images each),

➢ 6 staged colloquia (sum of all images of several practice lessons),

➢ exam (100 images),

➢ 336 WSIs in total.

More than 100 students of Lomonosov Moscow State University use PathScribe to study pathology **annually.**
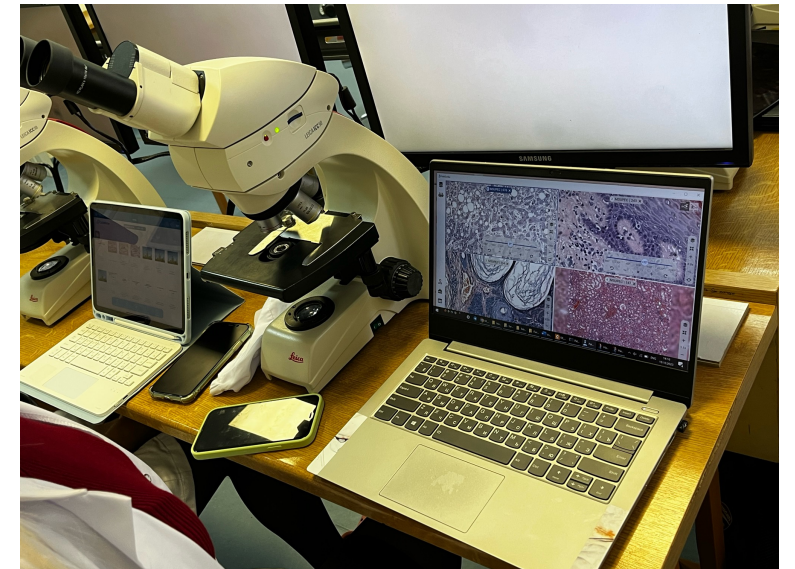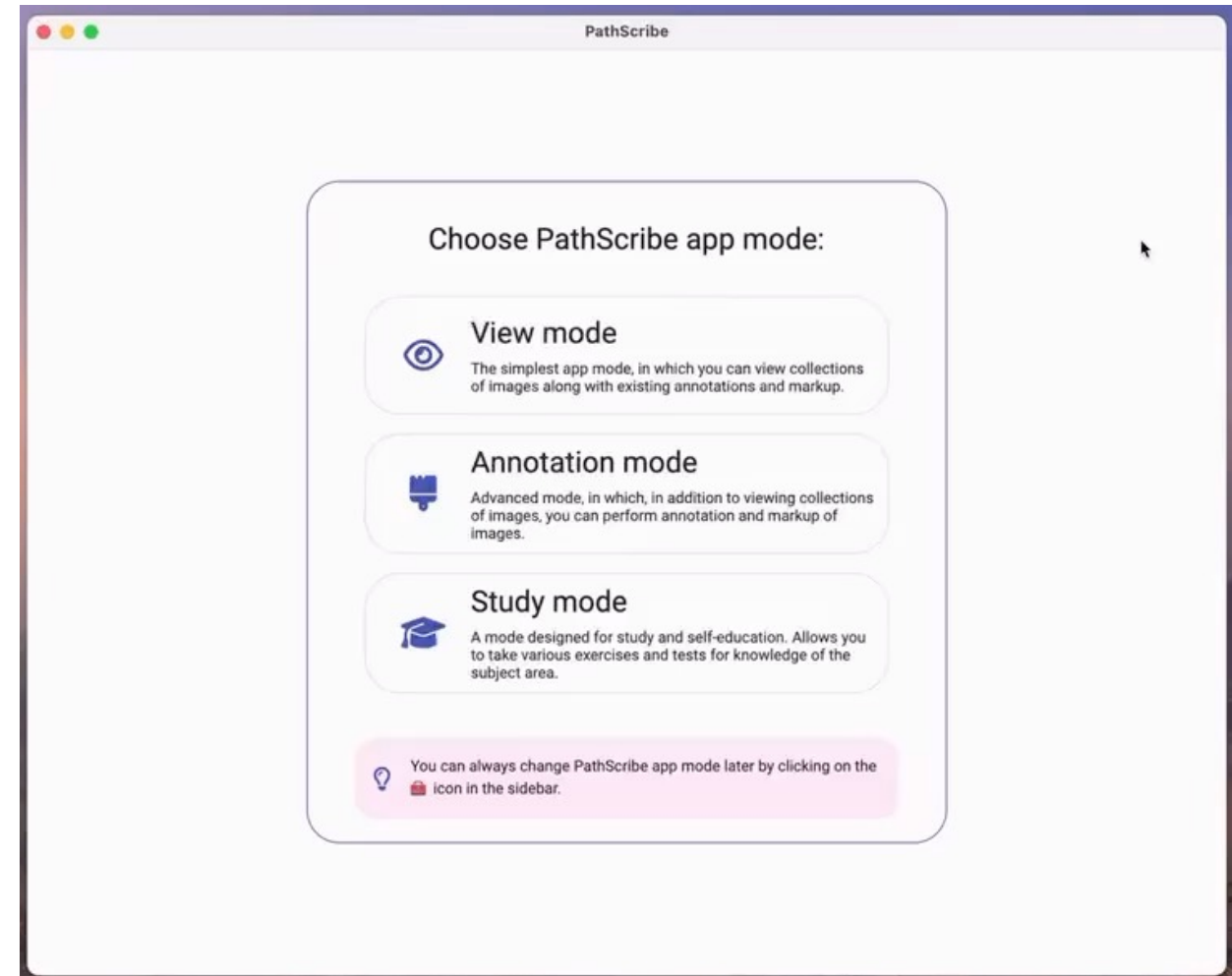

Photos of students at Lomonosov MSU studying pathology with PathScribe.

Research features

PathScribe was also conceived as a multifunctional tool for professional histologists and researchers. The research functionality includes:

▶ multiscene mode,

▶ multiple annotation spaces for each dataset,

▶ multi-user annotation within annotation space,

▶ annotation tools (brush, polygons, registration points),

▶ resulting annotations can be viewed an all devices and platforms (edited only on "big screens"),

▶ methods of semi-automatic annotations *(in development)*.


*An example of annotating WSI with brush and polygon tools.*

Research features

PathScribe was also conceived as a multifunctional tool for professional histologists and researchers. The research functionality includes:

▶ multiscene mode,

▶ multiple annotation spaces for each dataset,

▶ multi-user annotation within annotation space,

▶ annotation tools (brush, polygons, registration points),

▶ resulting annotations can be viewed an all devices and platforms (edited only on "big screens"),

▶ methods of semi-automatic annotations *(in development)*.
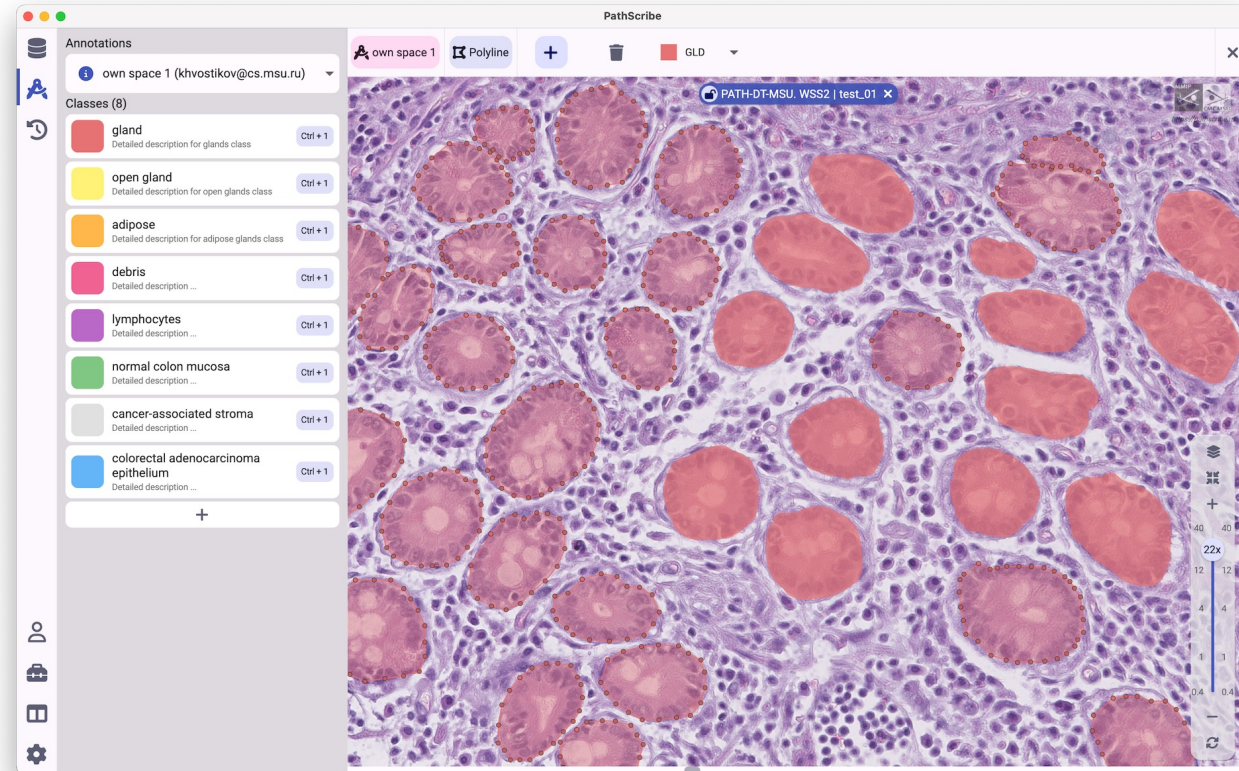
An example of WSI annotation with polygon and brush instruments in PathScribe. Displayed WSI is from PATH-DT-MSU WSS2 dataset.

Research features
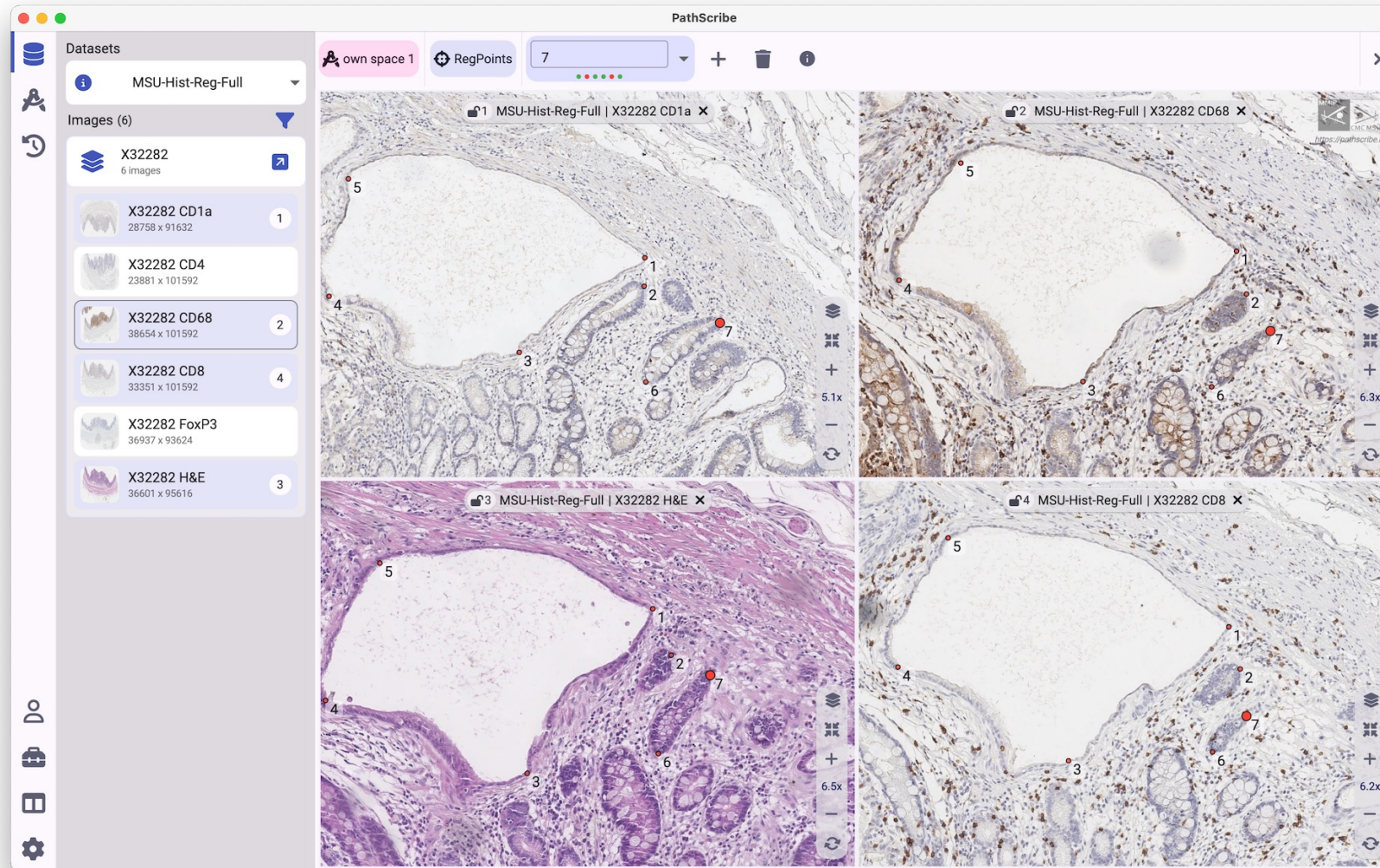
An example of adjacent WSIs annotation with registration points in PathScribe's multiscene view.

Development details

PathScribe implements cloud-based client-server architecture:

- ▶ microservices in Python (also tried Rust),
- ▶ GRPC communication with SSL encryption,
- ▶ SQLite databases,
- ▶ HDF5-based files to store images,
- ▶ caching with Redis,
- ▶ multiprocess image converter,
- ▶ Flutter client apps (native on all platforms).

Also:

- ❖ adaptive interface (looks different depending on the screen size)
- ❖ stylus support for annotation mode
- ❖ benchmark mode (for developers)
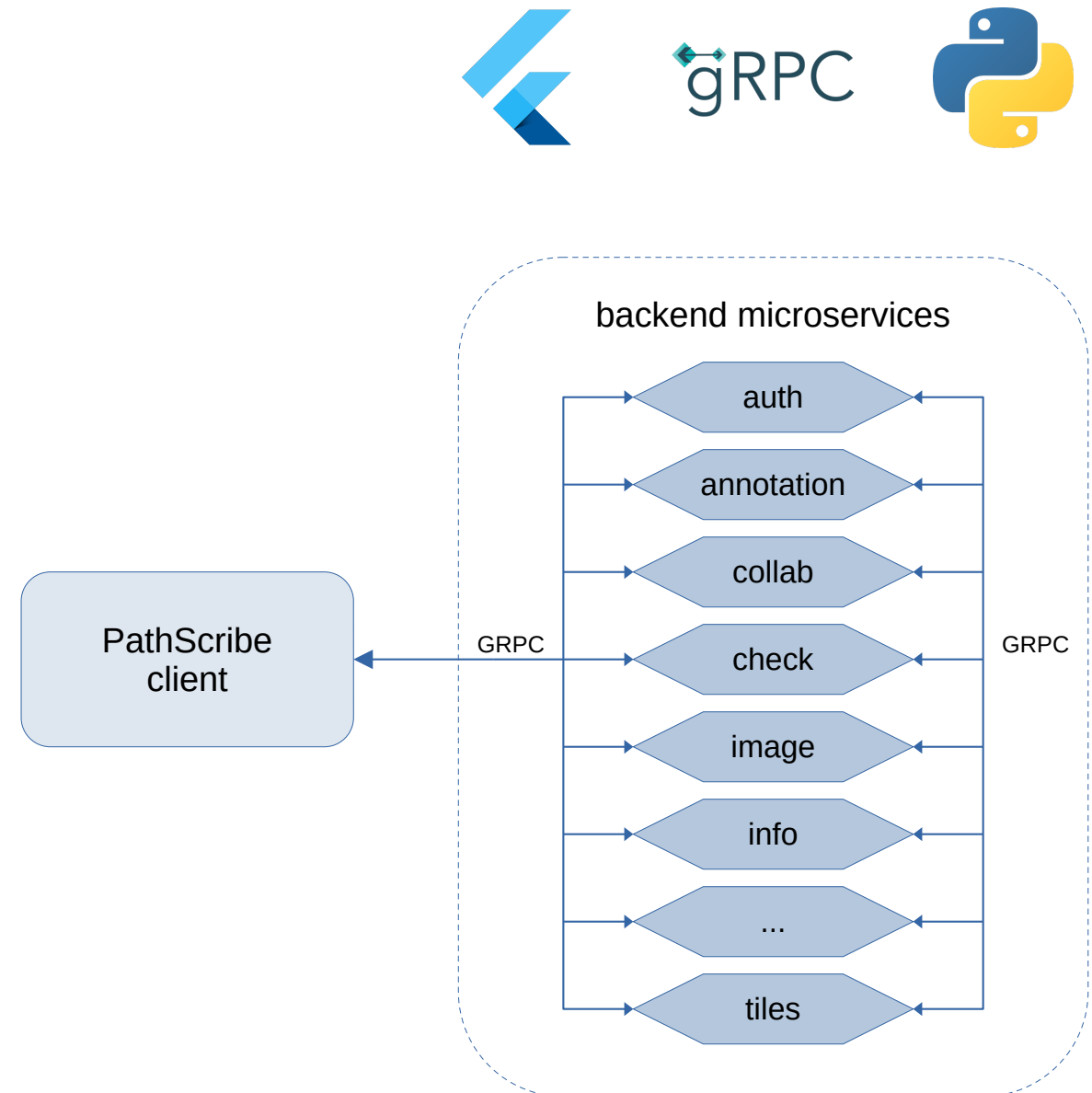
Future directions

PathScribe is actively developed, and a lot of new features are coming out soon. We plan to release the following new functionality in the next versions of PathScribe:

▶ collab mode (conference mode),

▶ multi-user collaborative annotations;

▶ interactive tests for students focused on testing and improving their knowledge within the course of pathology;

▶ integration a set of modern **algorithms for segmentation**, classification and analysis of histological images (developed earlier and being under development by our team);

▶ web version of PathScribe,

▶ data upload mode – a mode in which users can upload local images from their devices to PathScribe server (with automatic conversion), form datasets and image groups.
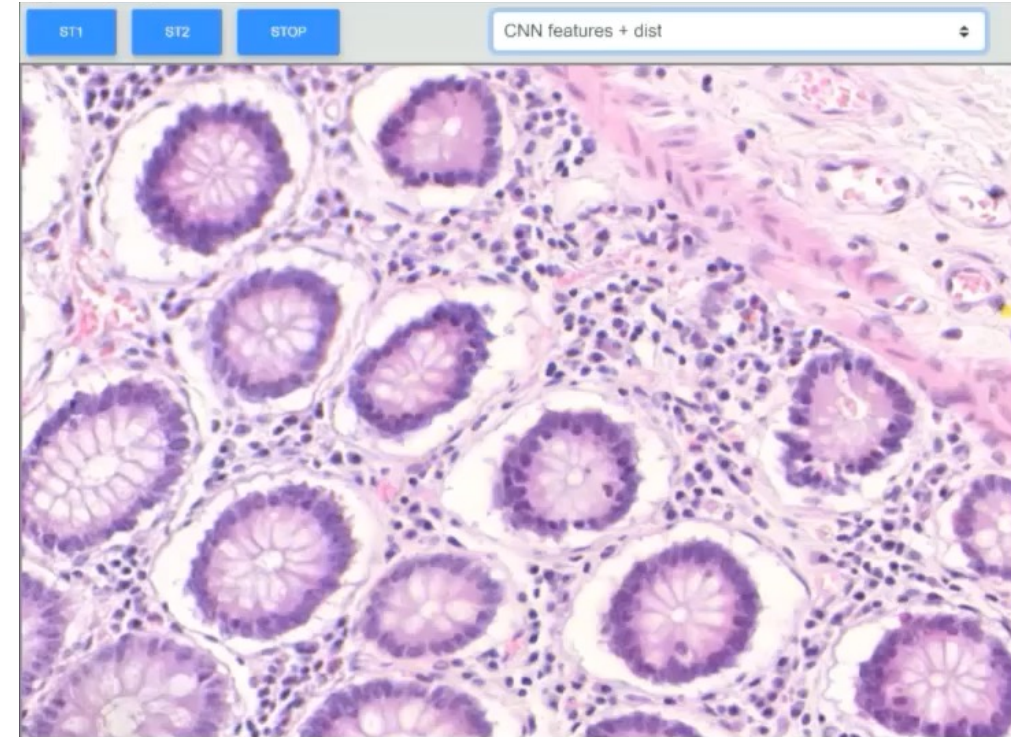
Our prototype of an interactive segmentation method (10x speedup compared to brushes and polygons).

PathScribe. Try it yourself!

❖ Everyone is free to download the latest PathScribe distribution from official website www.pathscribe.ru;

❖ There are 4 public datasets available within PathScribe;

❖ To sign up and start using PathScribe you should get an invite code through telegram bot **@PathScribeInvitesBot** (link available at the project website);

www.pathscribe.ru

Thank you for your attention!

For more information visit:

https://pathscribe.ru

https://imaging.cs.msu.ru