Digital Pathology

On the intersect ion of Computer Vision and Data Science

Multimedia Laboratory
The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
University of Toronto
Toronto, ON M5S 3G4, Canada

Kostas N. Plataniotis
kostas@ece.utoronto.ca

http://www.dsp.utoronto.ca/projects/ADP/

Summer 2021 – AIDA Consortium

With input from Professor Mahdi S Hosseini

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

Who Am I

Konstantinos N (Kostas) Plataniotis, Ph.D., P.Eng. , FIEEE, FEIC, FCAE

Professor and Bell Canada Chair in Multimedia
Director, University of Toronto – Huawei Mobile AI Laboratory

The Edward S. Rogers Sr. ECE Department, University of Toronto

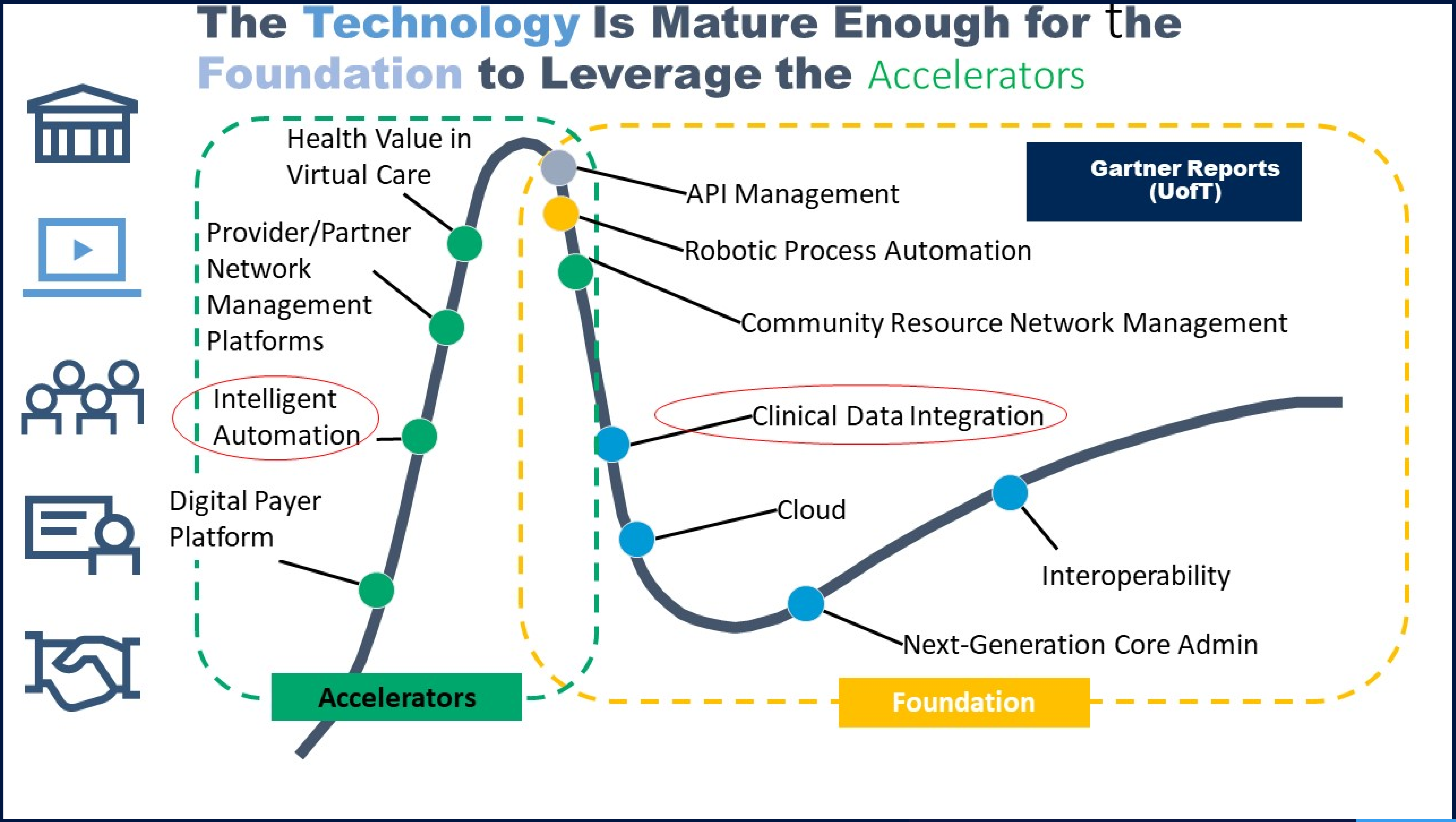
Principal Investigator
Ontario Research Fund - Research Excellence Program

`Transforming pathology using artificial intelligence to improve patient outcome and hospital efficiency'

Learning Objectives

- What is digital pathology
- Vision and data science in digital pathology
- Current state of digital pathology systems
- Limitations of digital pathology
- Opportunities in digital pathology

Digital Pathology as Health Innovation Enabler



The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

How Digital Pathology Compares to Computer Vision & Machine Learning?

- Machine Learning & Computer Vision is well matured compared to Digital Pathology in terms of data compilation

Introduction of Consumer Digital Camera

Philips intellisite Scanner Received First (Limited) FDA Approval for Clinical Pathology

AlexNet [Krizhevsky et al]

Widely Distributed Pathology Datasets

2000

2010

2012

2017

2019

More sophisticated GPUs were developed

Many CNN Classifiers are designed

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

Introduction to Clinical Pathology

Tissue Preparation

- Overall process:
  1. Embed tissue samples into paraffin wax
  2. Section into small slices using a microtome
  3. Stain tissue slices

Chemical Reagent

Staining

Biopsy

Formalin Wash

Paraffin Embedding

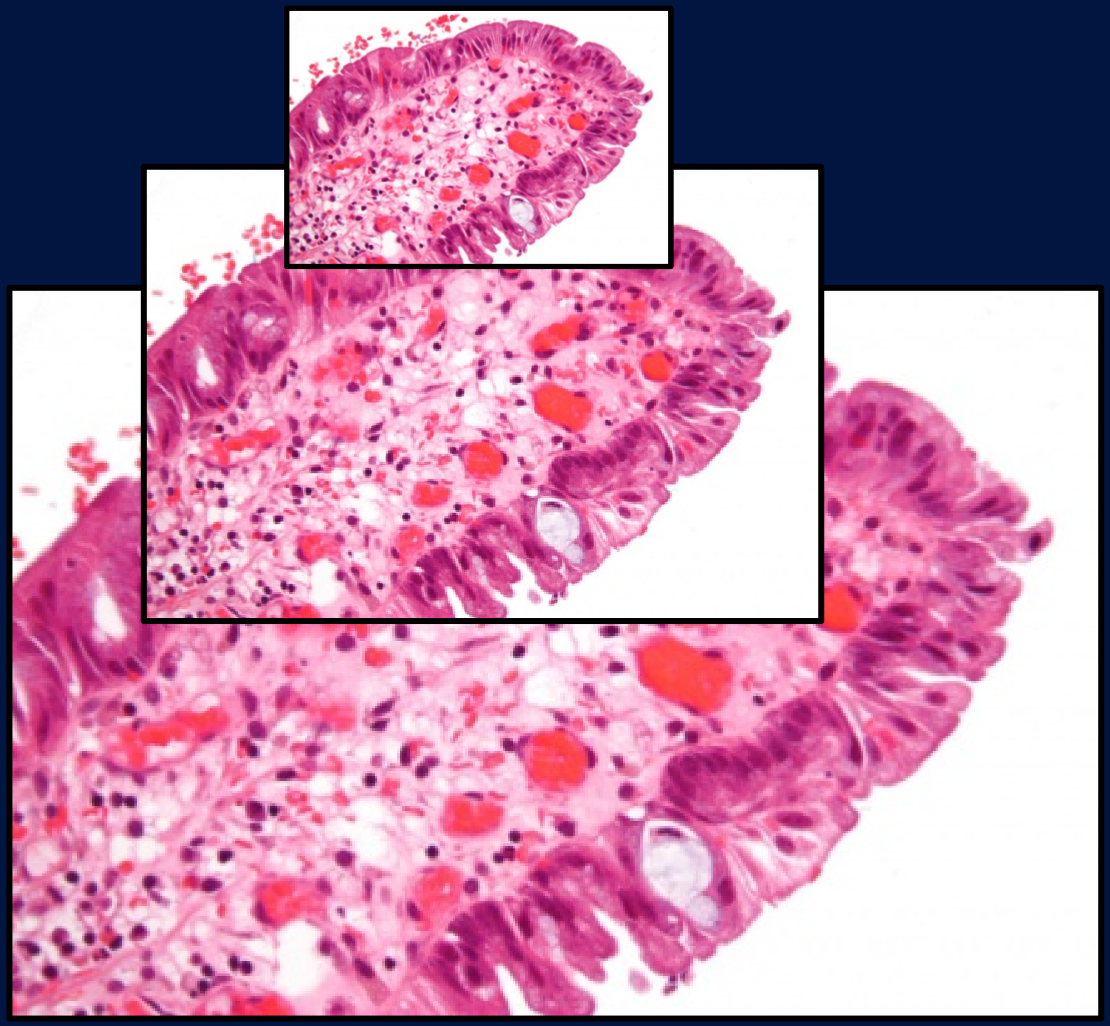
Paraffin Block

Microtome Slicing

Glass Slide Mounting

Microscopic View

- Glass slide is mounted on an optical microscope
- Different lens magnification rates i.e. 10X, 20X, 40X, 60X



Digital Pathology

Where visual input quality matters in machine learning

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

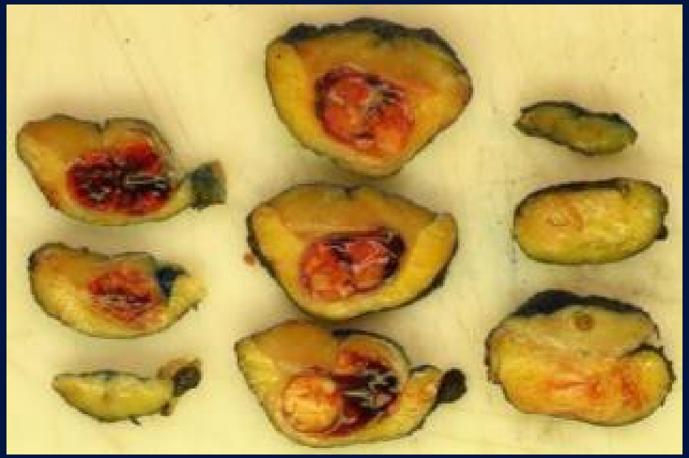
Digital Pathology - Definition

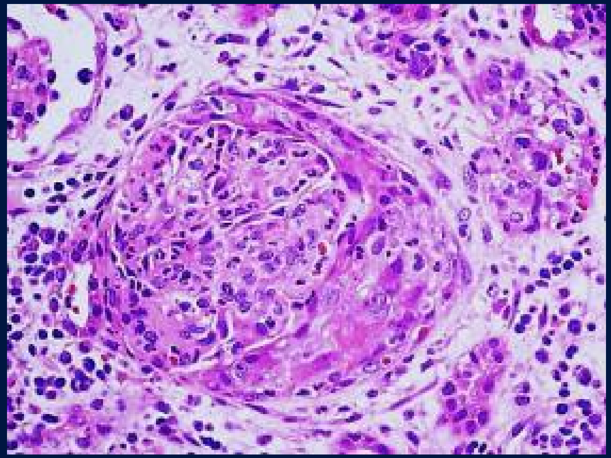
A dynamic, image-based environment that enables the acquisition, management and interpretation of pathology information generated from a digitized glass slide. Often used interchangeably with "Virtual Microscopy."

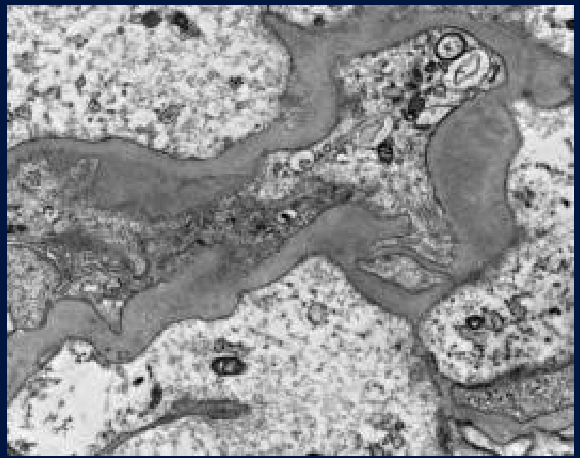
Digital Pathology Association (DPA)

Visual input in digital pathology

- Gross imaging
- Microscopic imaging
- ***Whole slide imaging (WSI)***
- Electron microscopy
- Immunofluorescence

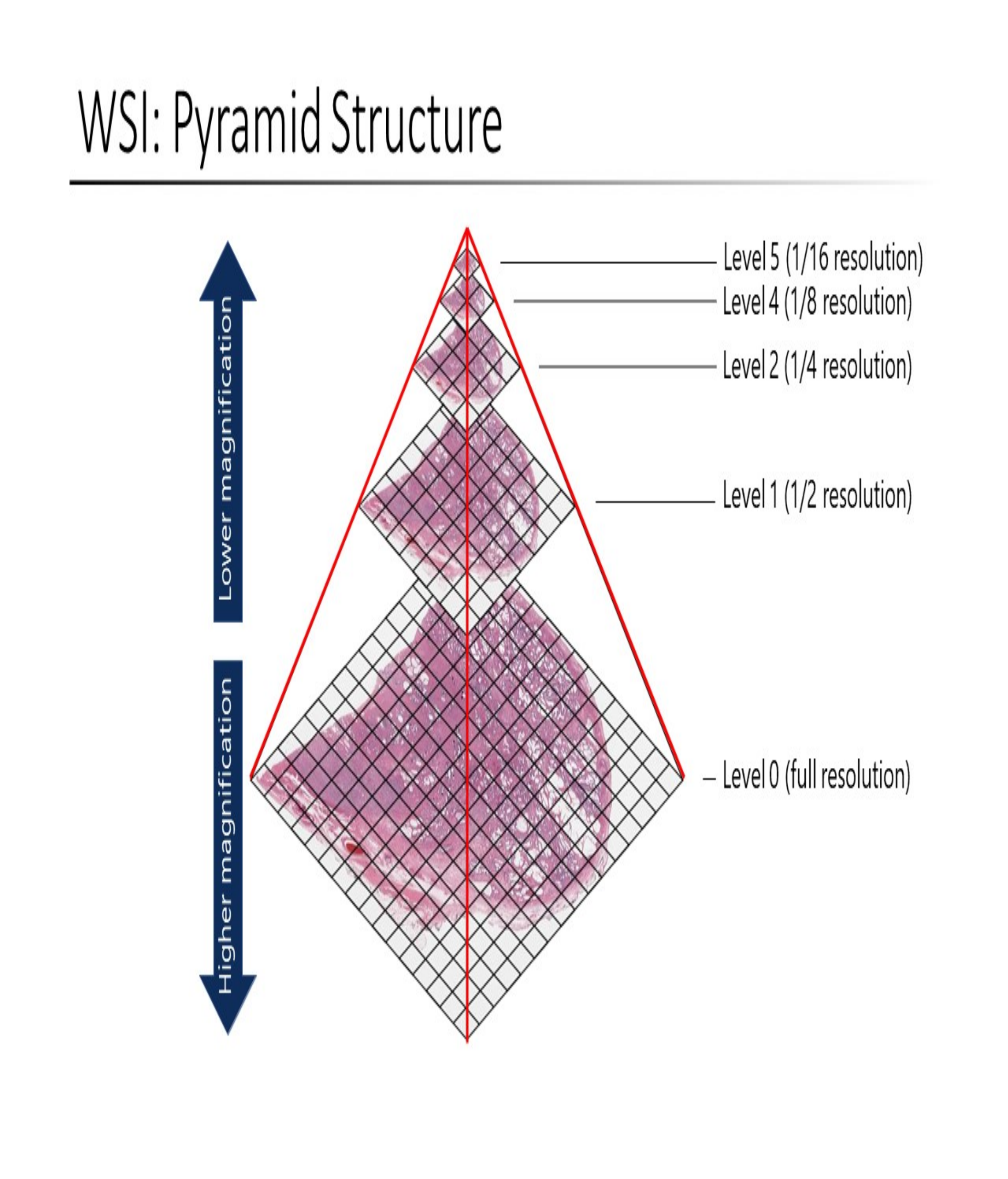


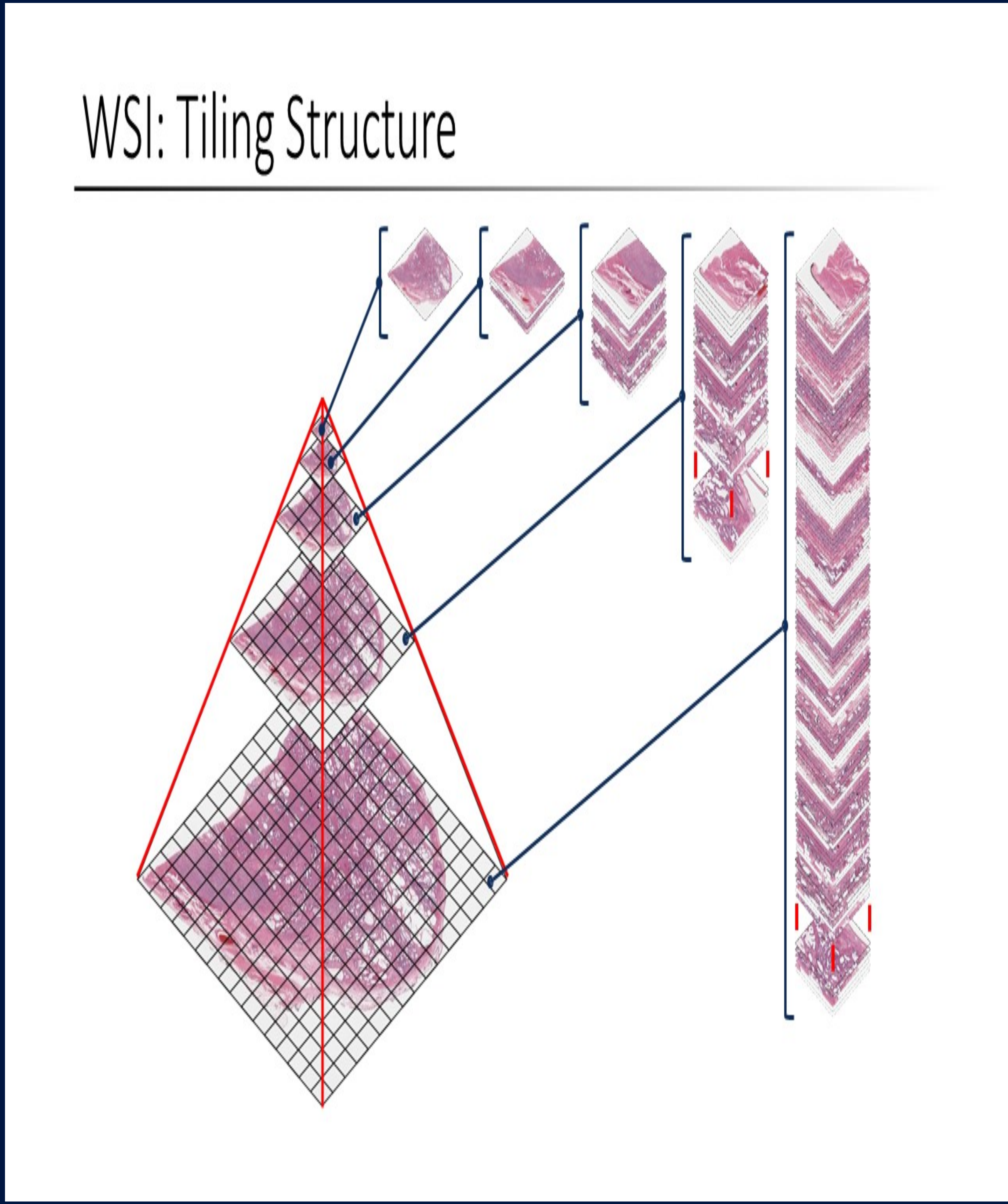




Whole Slide Imaging (WSI)

Whole Slide Imaging (WSI)



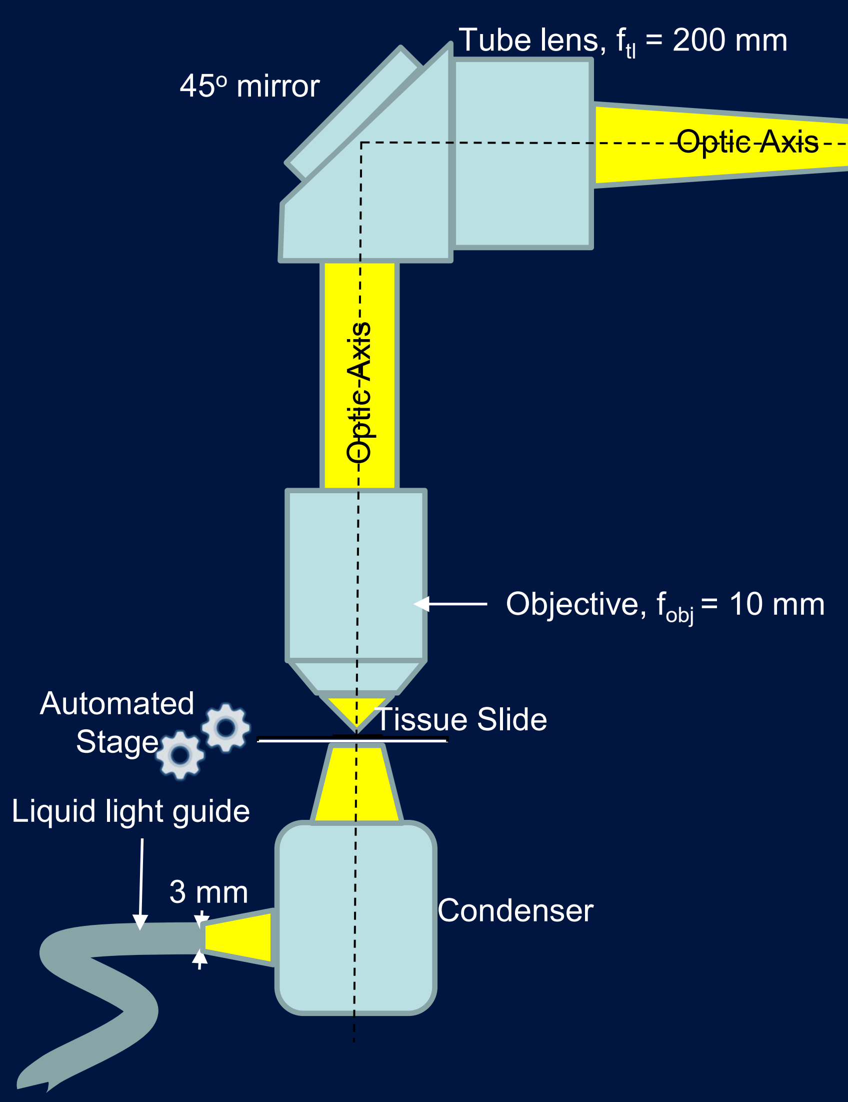


The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

TissueScope LE2.0 WSI Scanner

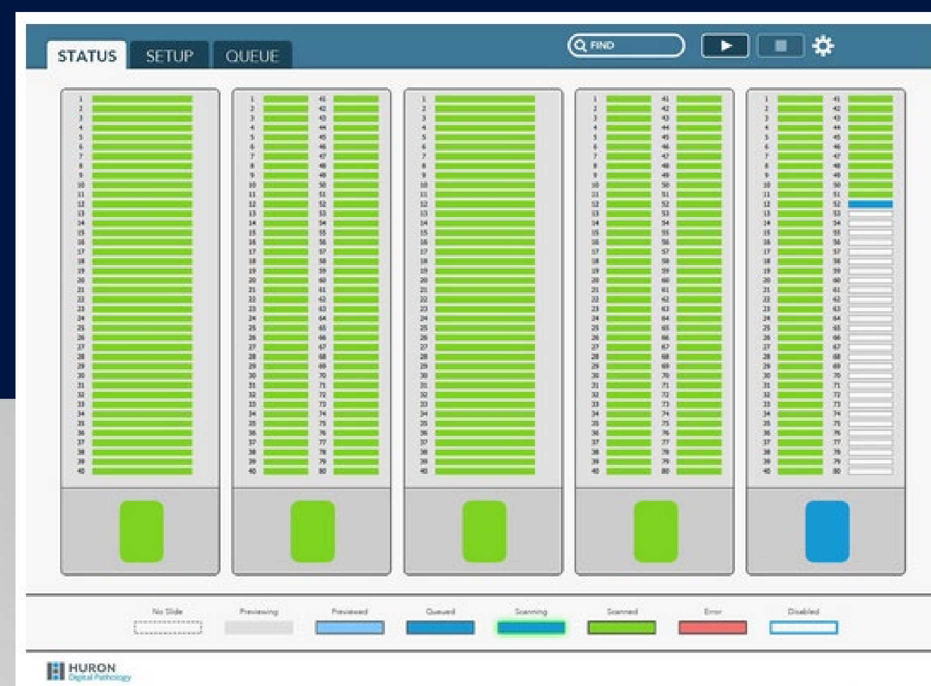
- Digital Microscopy
  - Light Condenser
  - Automated Stage
  - Objective Lens
  - Reflective Mirror
  - Tube Lens
  - Camera Sensor
- System magnification
  - f-TubeLens/f-Objective = 200mm/10mm=20

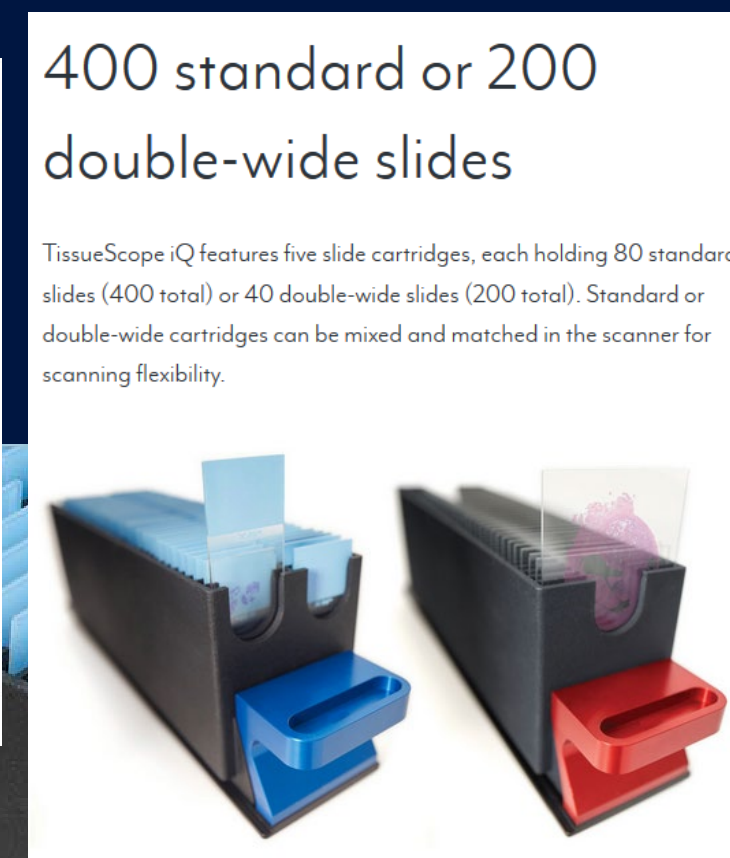


Courtesy of Huron

High-Throughput Scanner

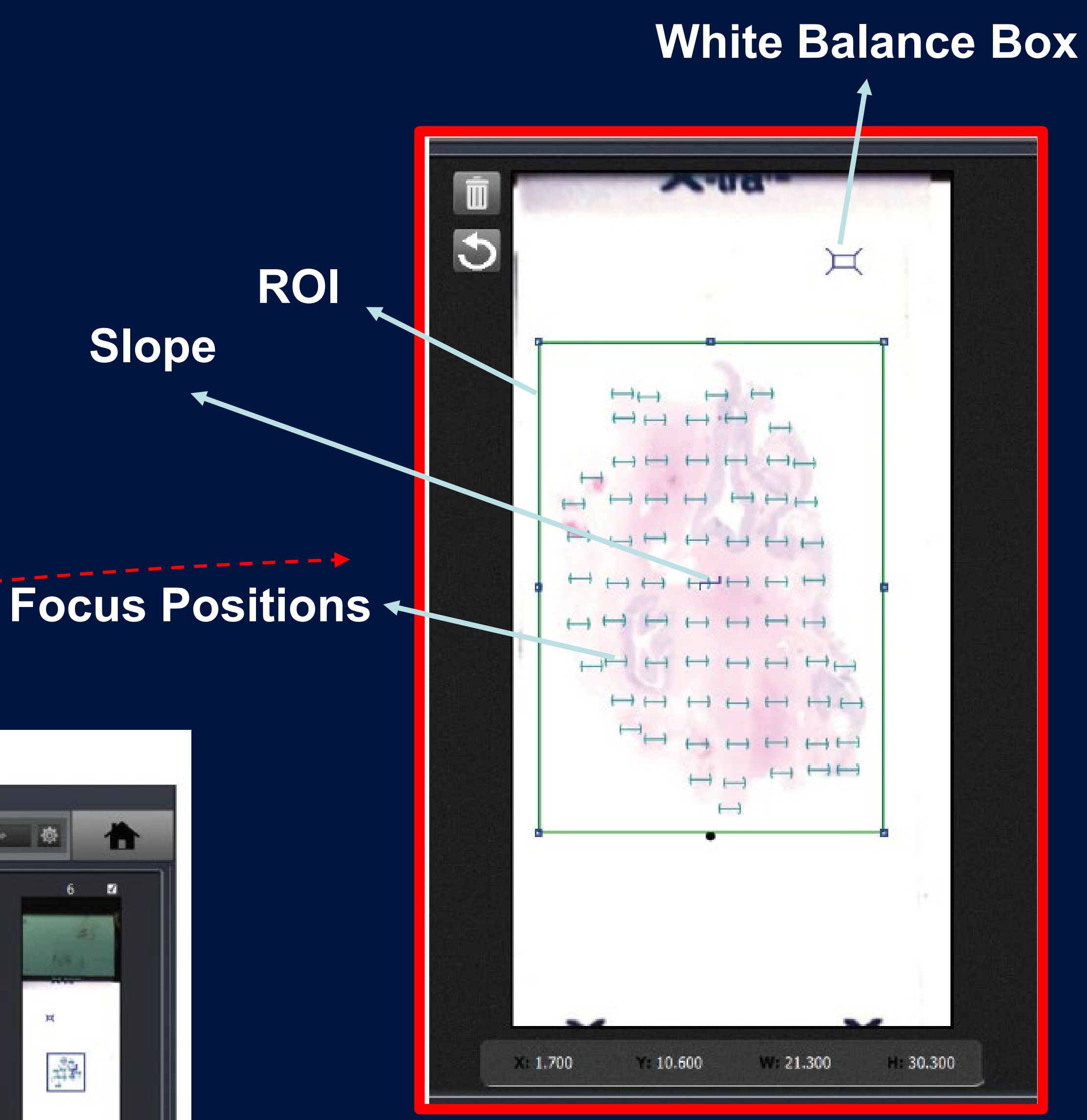
- High-volume scans e.g. 400 glass slides/200 double-wide
- Run overnight and make it ready for tomorrow for pathologists
- Semi-automated QC control





WSI Scanner Setup

- Preview camera
- Tissue Finding Software
  - Find tissue (segment)
  - Determine Region-of-Interest (ROI)
  - Allocate Focus Positions
  - Construct 3D Focus Map

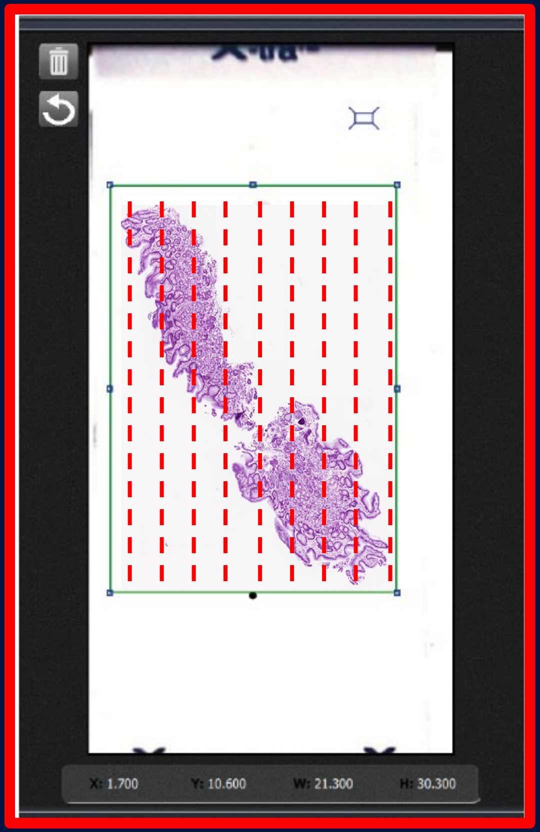


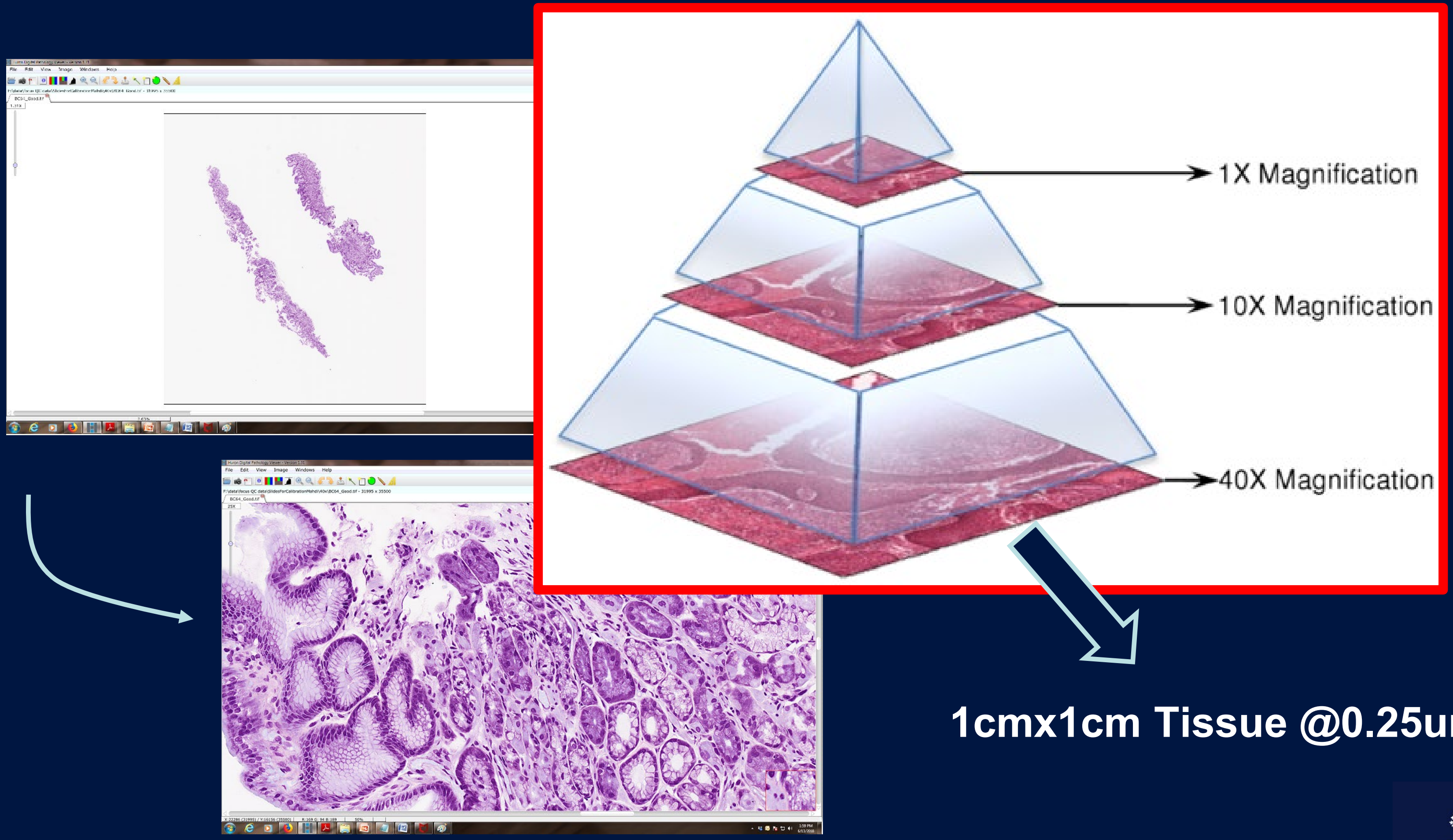
The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

WSI Scan in High-Resolution

- Scan ROI in multiple strips
- Stitch them back together to construct WSI
- Construct Pyramid on large scale TIFF image
- Pyramid emulates virtual microscopy



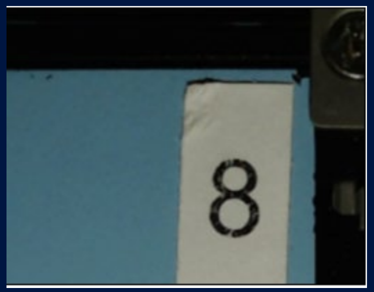


1cmx1cm Tissue @0.25um/pixel = 4.8GB (Uncompressed)

The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

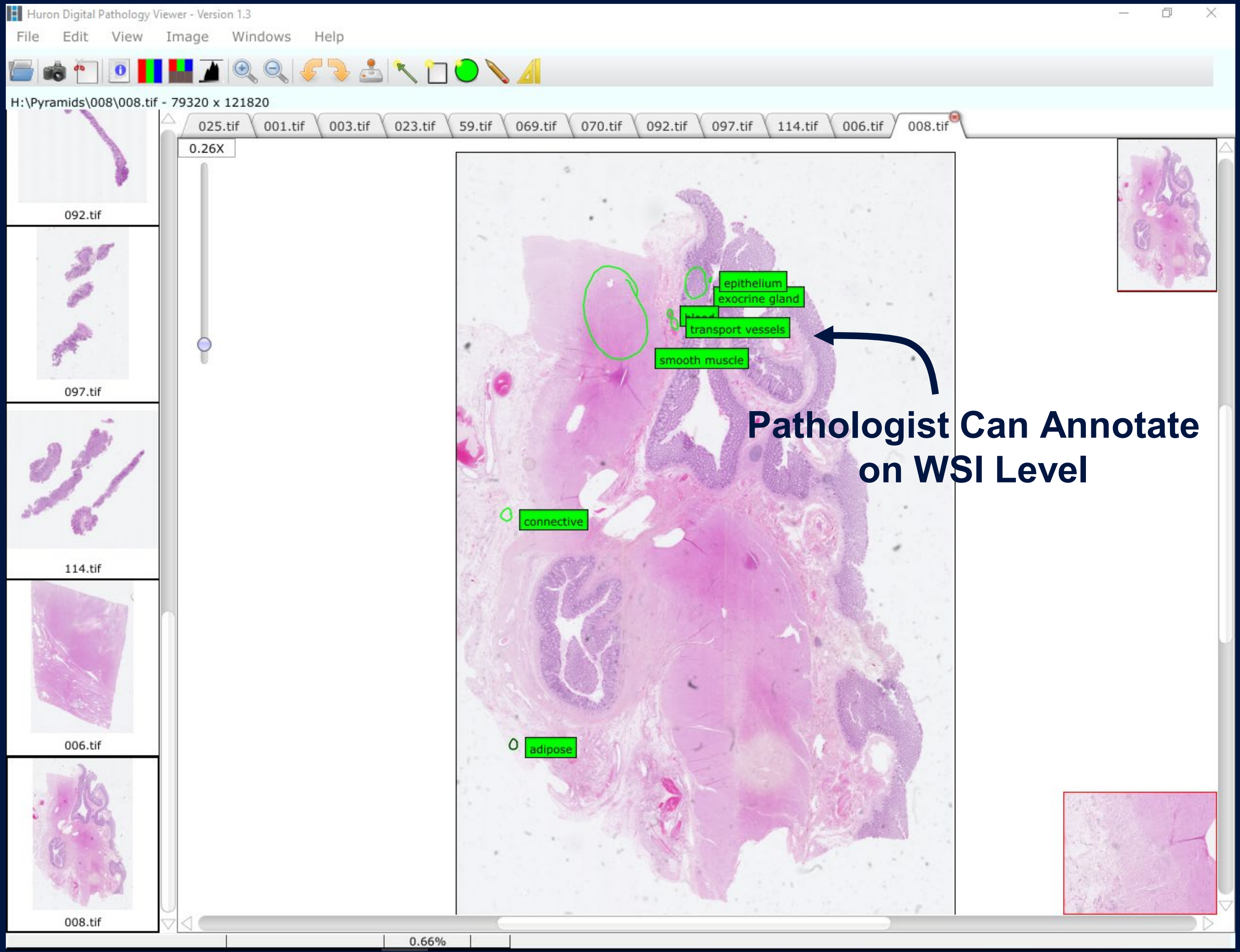
FACULTY
OF APPLIED
SCIENCE &
ENGINEERING

WSI relevant Meta Data



Slide label
(Related Patient Info)

Preview Image of
Glass Slide



Pathologist Can Annotate
on WSI Level

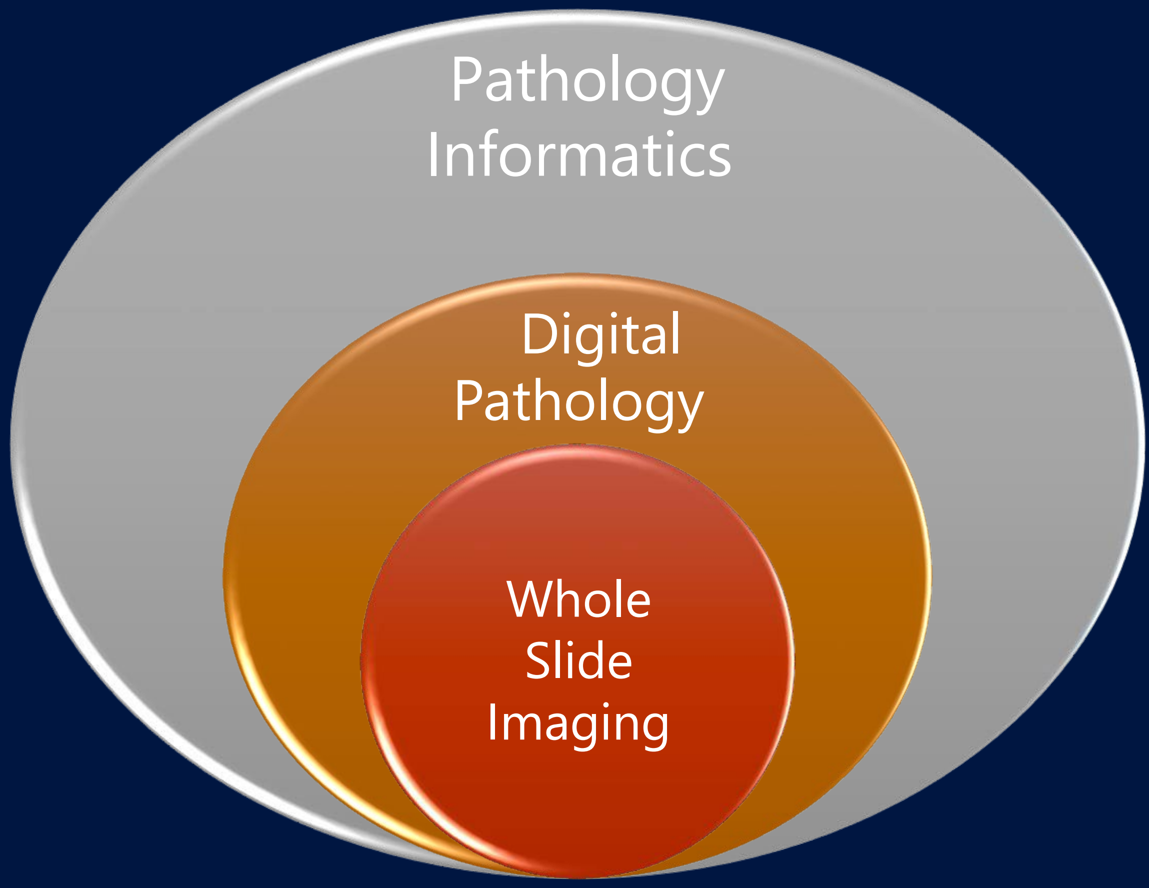
WSI Scan Info

Properties	Value
Digital Slide size	24.90 x 30.45 mm
WSI dimensions	99,600 x 122000 pixels
Scan duration	12 min 19 s
DPI	101,600
Source	Bright Field
Resolution	0.25 um
File size	39,050,676,108 bytes
Image format	24 bit RGB
Tile dimension	256 x 256 pixels

Digital Pathology vs. Pathology Informatics

- Whole Slide Imaging (WSI) is a subset of Digital Pathology (DP), which is a subset of Pathology Informatics (PI)

  - For some, PI ≅ DP
  - For many, DP ←→ WSI
  - For all, ML&CV opportunities in WSI driven DP



Digital Pathology

New Emerging Field in ML

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Digital Pathology

Data Science & Analytics

WSI Scanner

Supervised Data Compilation

Statistics

Deep Learning Model

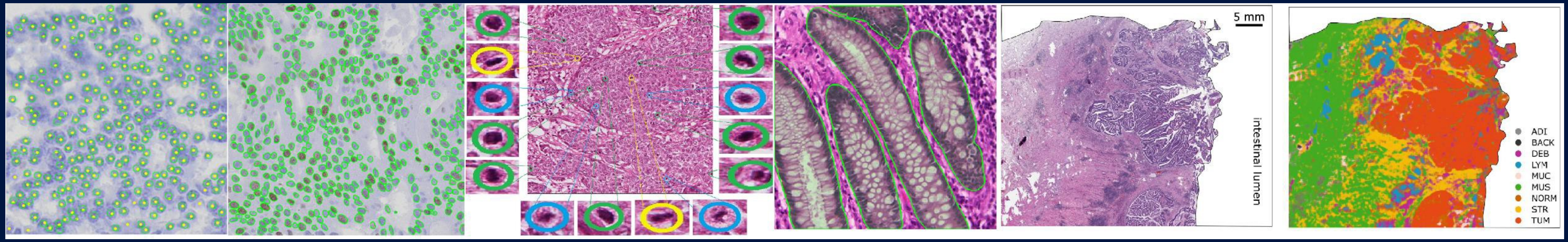
CAD Applications

The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY
OF APPLIED
SCIENCE &
ENGINEERING

Digital → Computational Pathology

- Developing Computer-Aided Diagnosis (CAD) for Assisting Clinical Pathology
  - **Detection**: cells, nuclei, mitotic
  - **Segmentation**: glands, mitosis/non-mitosis
  - **Classification (Diagnosis)**: cancer grading
  - **Regression (Prognosis)**: metastasis probability



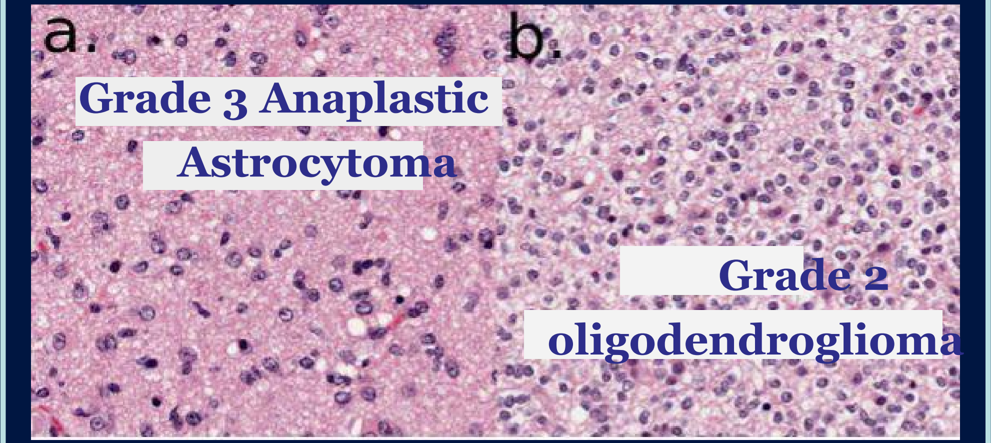
- **Detection** and **Segmentation** are well developed for quantitative analysis (recently approved by FDA for Clinical Applications)

State-Of-The-Art (SOTA)

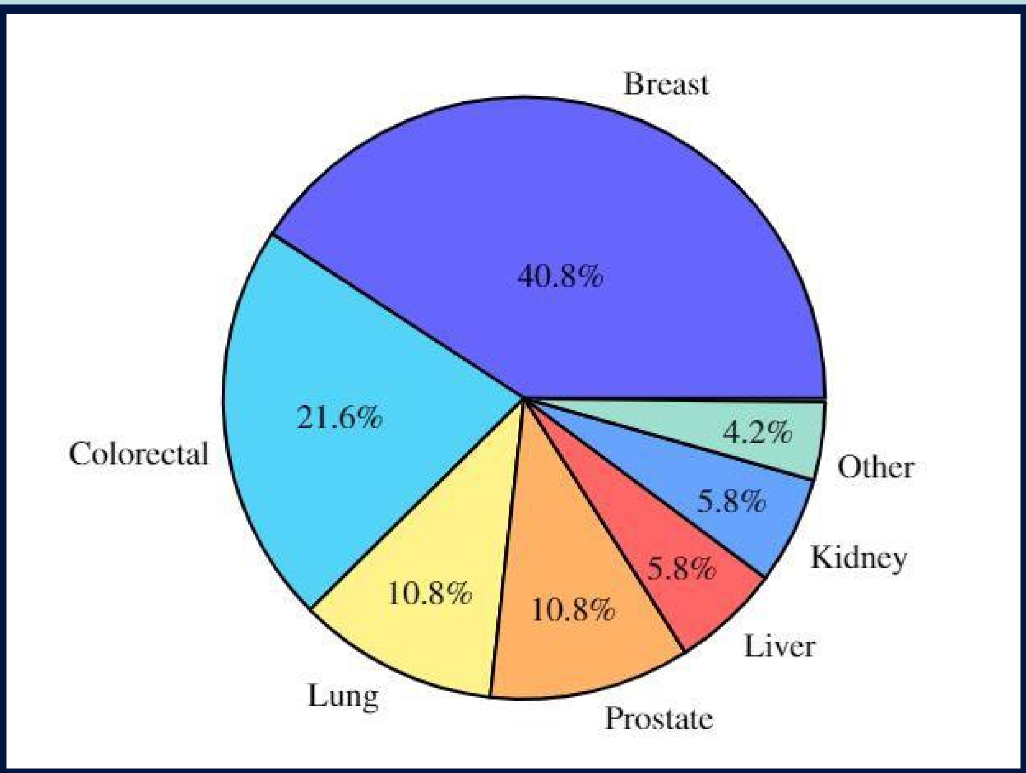
Report Survey from 150+ Papers

Centered around Organs

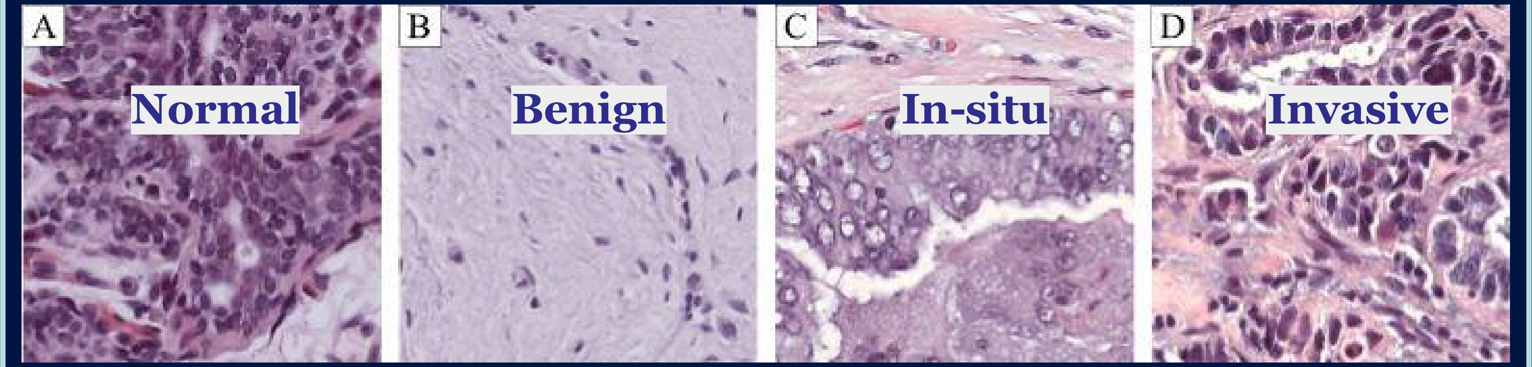
- Most funding in breast cancer research
- Breast, colorectal, prostate and lung → 48% of cancer cases[2]



Lower Grade Gliomas



Research focus distribution of 120+ papers by organ

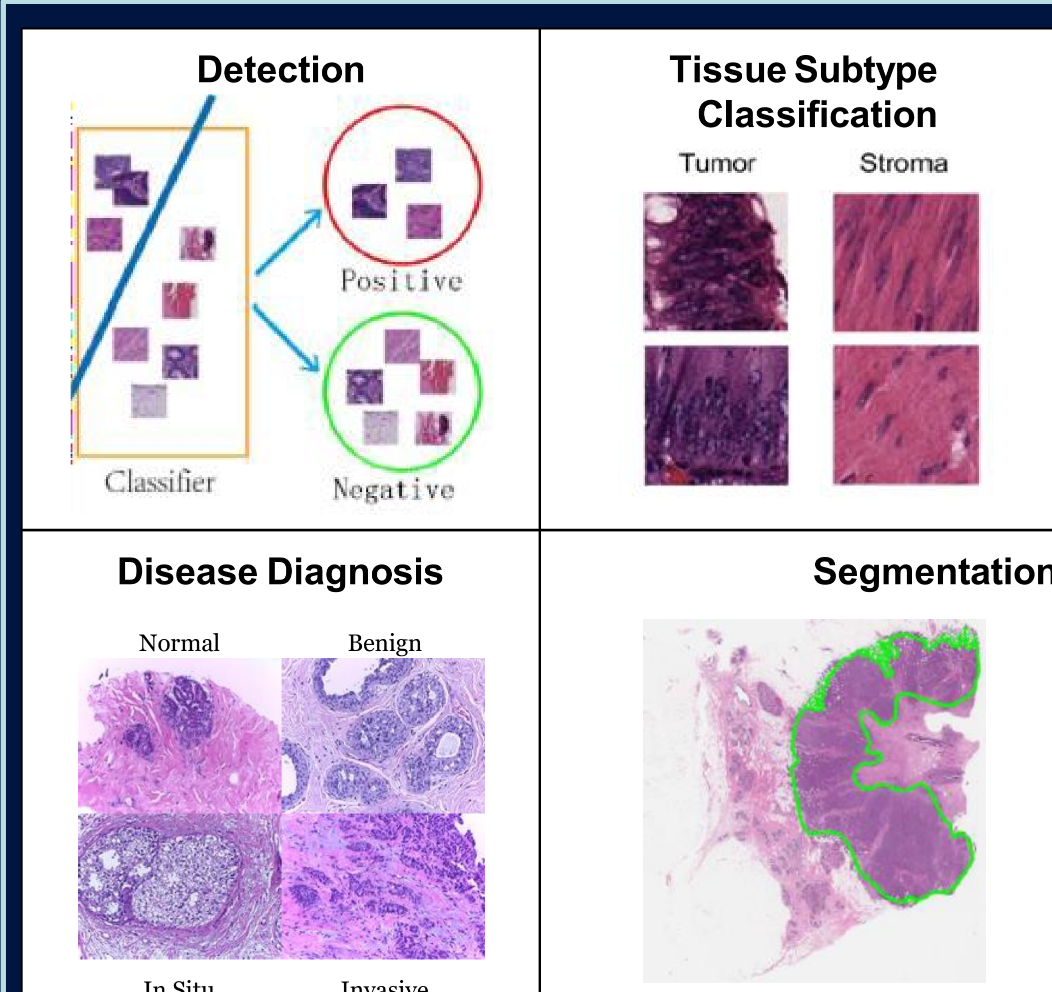


Breast cancer WSI regions (Nuclei features detect abnormality, and tissue structures classify breast carcinomas subtypes[3])

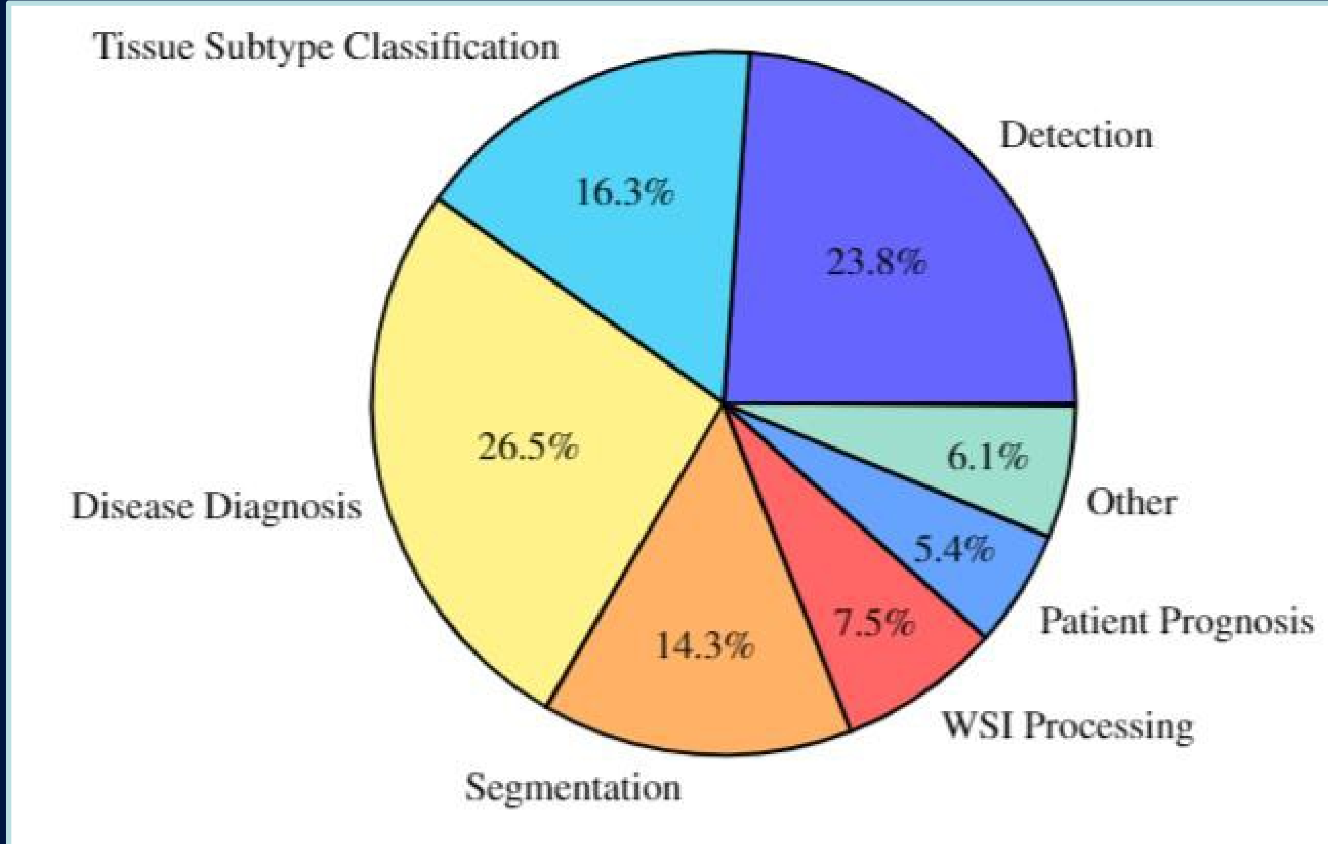
1.Barker, J., Hoogi, A., Depeursinge, A., & Rubin, D. (2016). Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles. Medical Image Analysis, 30, 60-71. https://doi.org/10.1016/j.media.2015.12.002
2. Canadian Cancer Society. (2020). Canadian Cancer Statistics 2018 [Ebook]. Retrieved 8 June 2020, from http://cancer.ca/Canadian-Cancer-Statistics-2018-EN.
3.Araújo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., & Eloy, C. et al. (2017). Classification of breast cancer histology images using Convolutional Neural Networks. PLOS ONE, 12(6), e0177544. https://doi.org/10.1371/journal.pone.0177544

Considering a multitude of tasks

- Main focus on disease diagnosis and detection tasks.
- Disease diagnosis and detection tasks primarily performed on breast cancer.



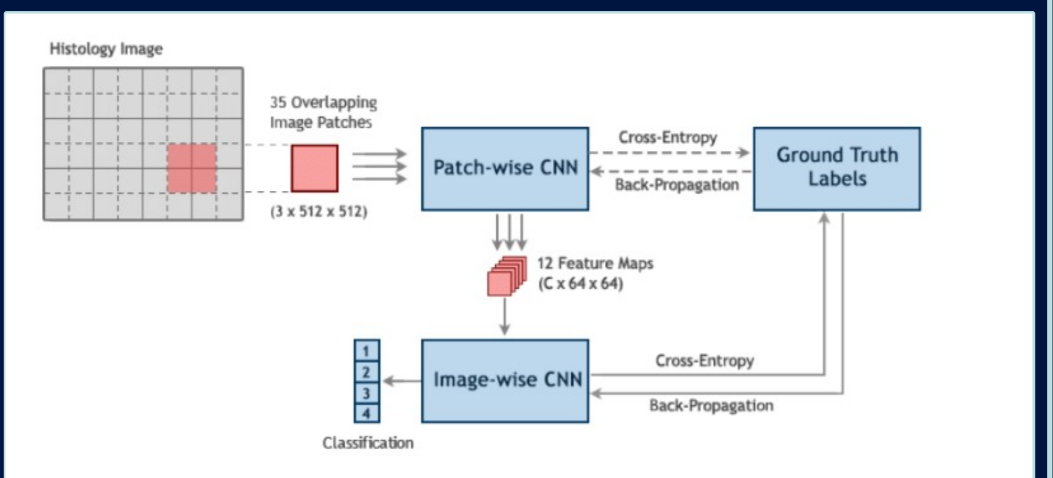
Examples of research tasks.



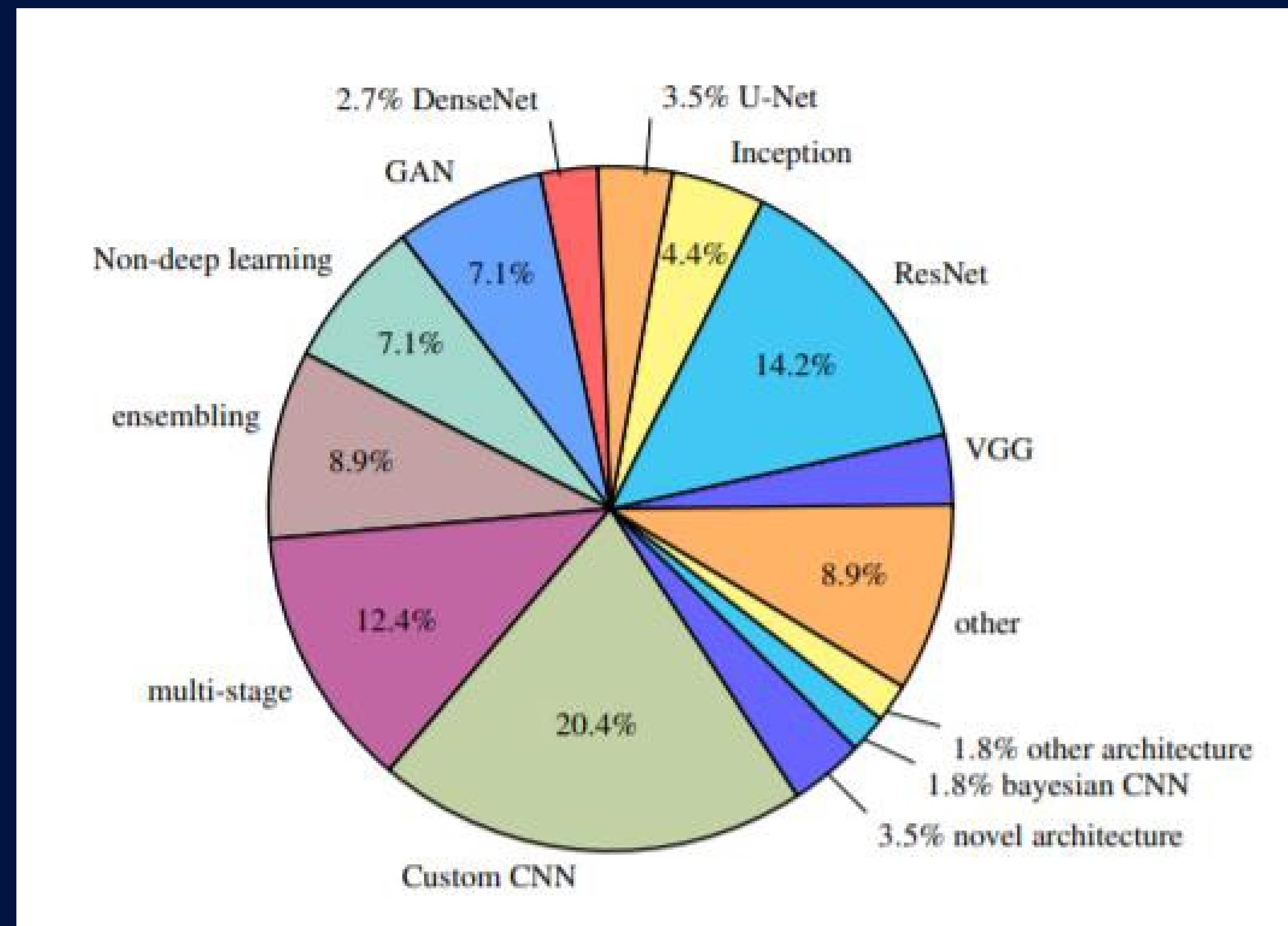
Distribution of 147 tasks over 120+ papers

Considering a multitude of models

- Most common: CNNs
- Popular approaches: Custom CNN, ResNet & Multi-Stage CNN



A Multi-Stage CNN model[1]

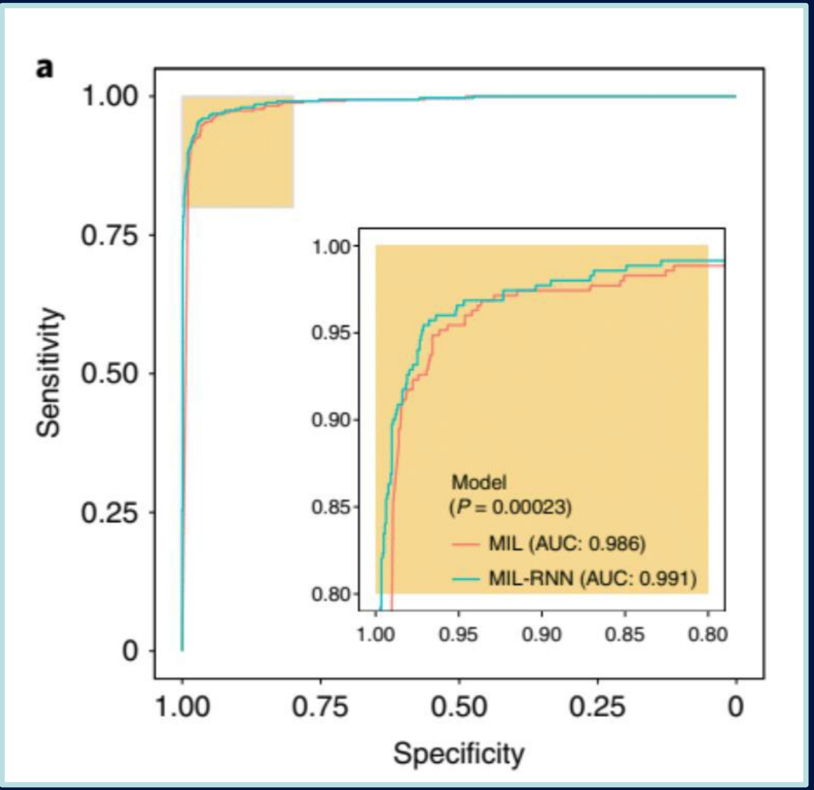


Distribution of models over 120+ papers

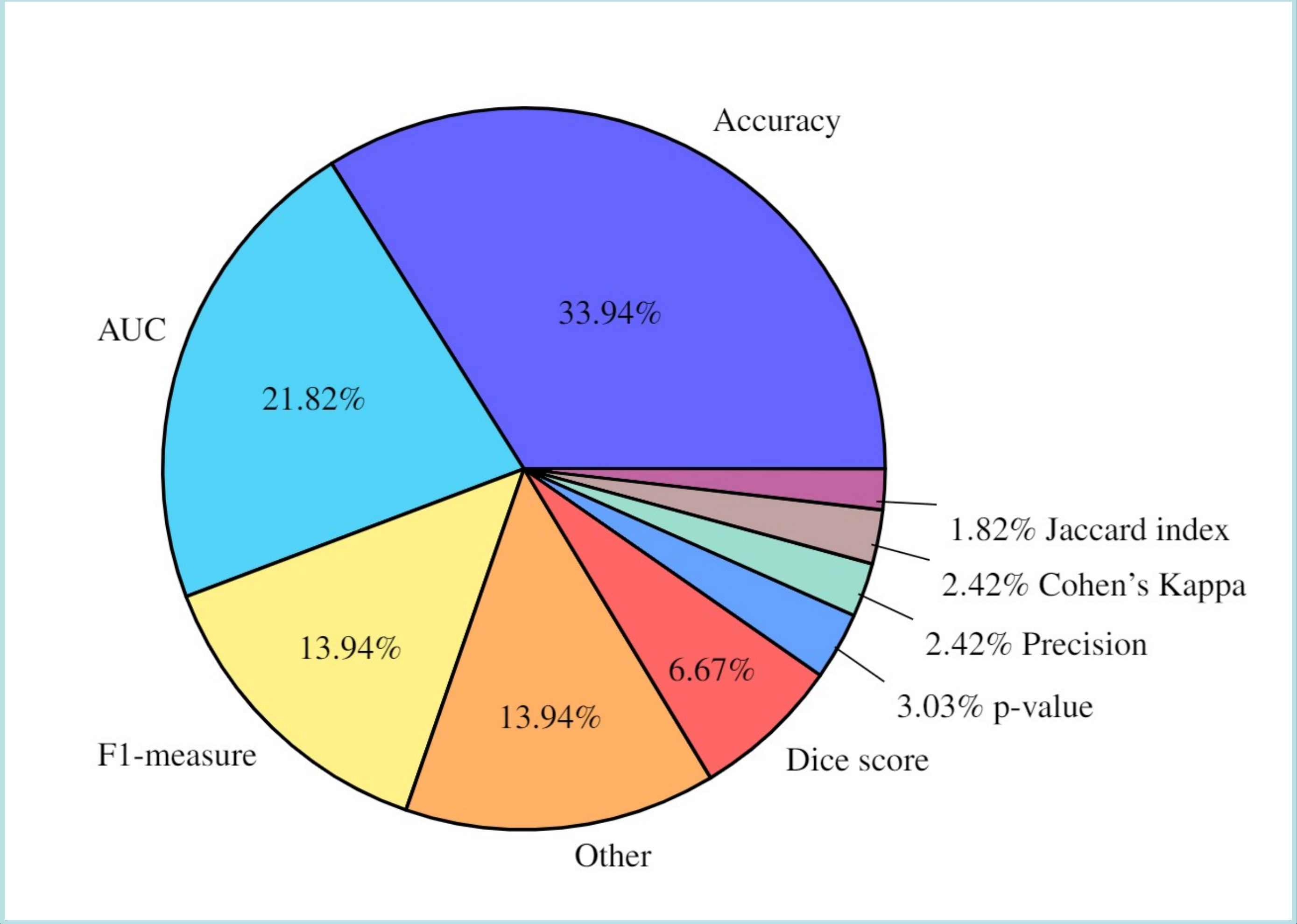
1. Nazeri, K., Aminpour, A., Ebrahimi, M.: Two-stage convolutional neural network for breast cancer histology image classification (2018). DOI 10.1007/978-3-319-93000-8 81

Multitude of Performance Metrics

- Most common metrics: Accuracy and AUC
- WSI processing and patient prognosis often use other metrics



Weakly-supervised model performance on 25k WSIs of prostate biopsies [1]

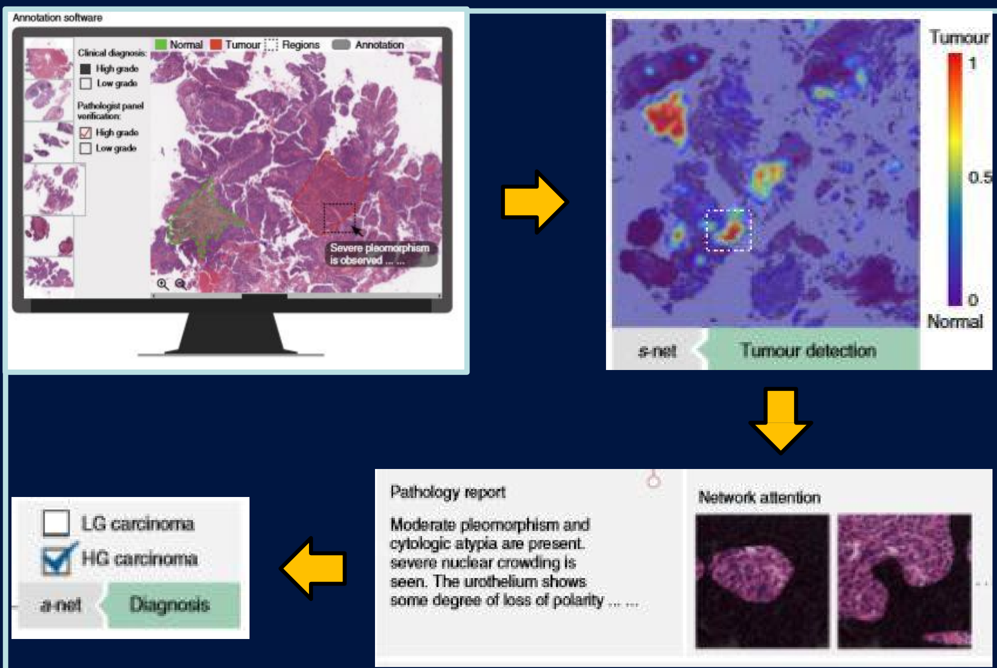


Distribution of performance metrics across 120+ papers.

[1] G. Campanella et al., "Clinical-grade computational pathology using weakly supervised deep learning on whole slide images," *Nature Medicine*, vol. 25, no. 8, pp. 1301-1309, 2019. Available: 10.1038/s41591-019-0508-1.

What is usually missing ? Clinical Validation

Three parts of clinical validation:
- Testing in clinical settings
- Proposed clinical application
- Concordance/discordance studies



CAD Workflow[1]:
1) WSI annotation
2) Automatic tumor detection
3) Automatic ROI description
4) Slide level diagnosis

Total # of papers: 38

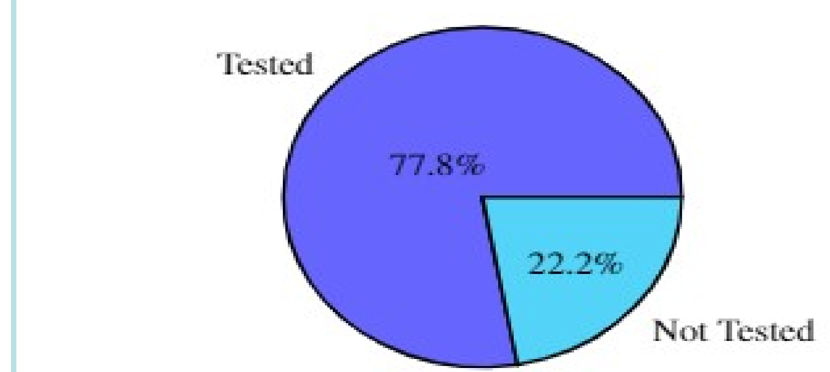


Fig. 5: Distribution of tested papers

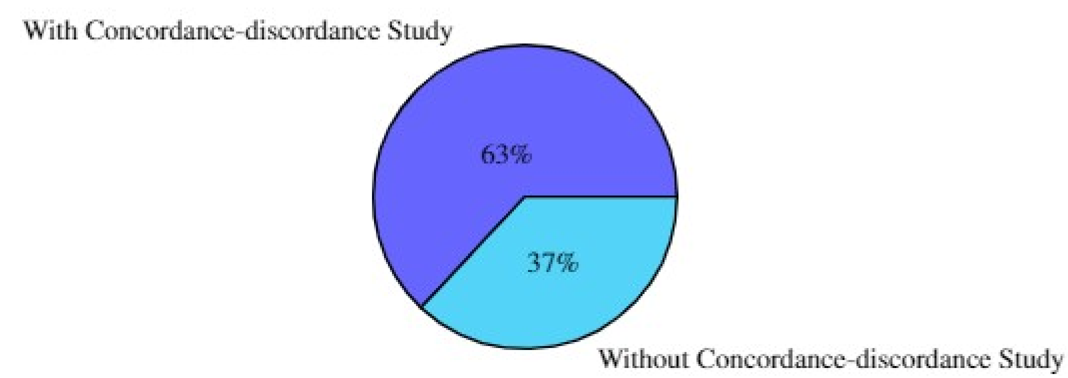


Fig. 13: Distribution of papers by Concordance-discordance Study

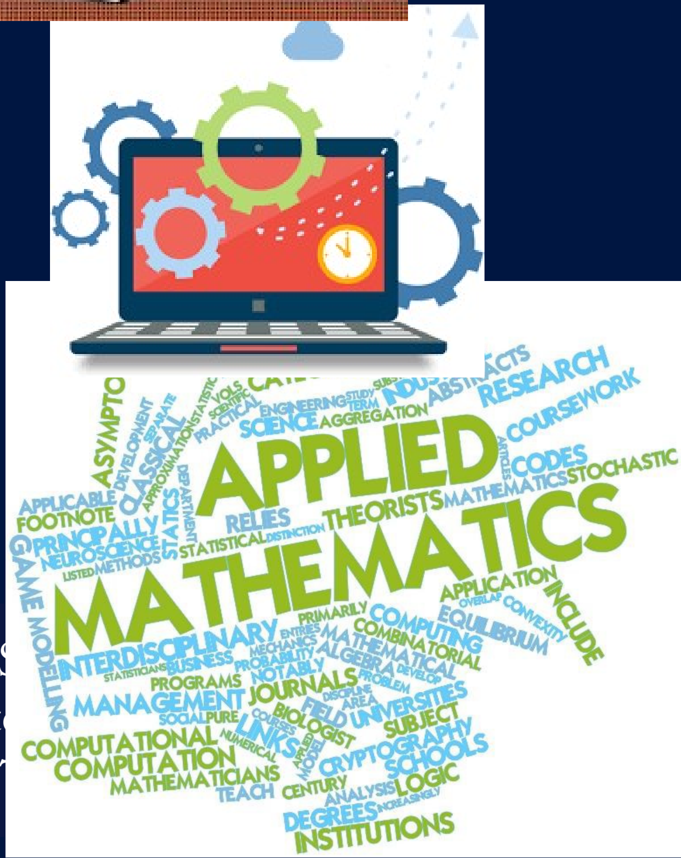
[1] Zhang, Z., Chen, P., McGough, M., Xing, F., Wang, C., & Bui, M. et al. (2019). Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. Nature Machine Intelligence, 1(5), 236-245. https://doi.org/10.1038/s42256-019-0052-1

What is the biggest problem ?

The elusive DC3

Data Collection-Compilation- Curation

What is the Efficient Data Compilation in Digital Pathology?



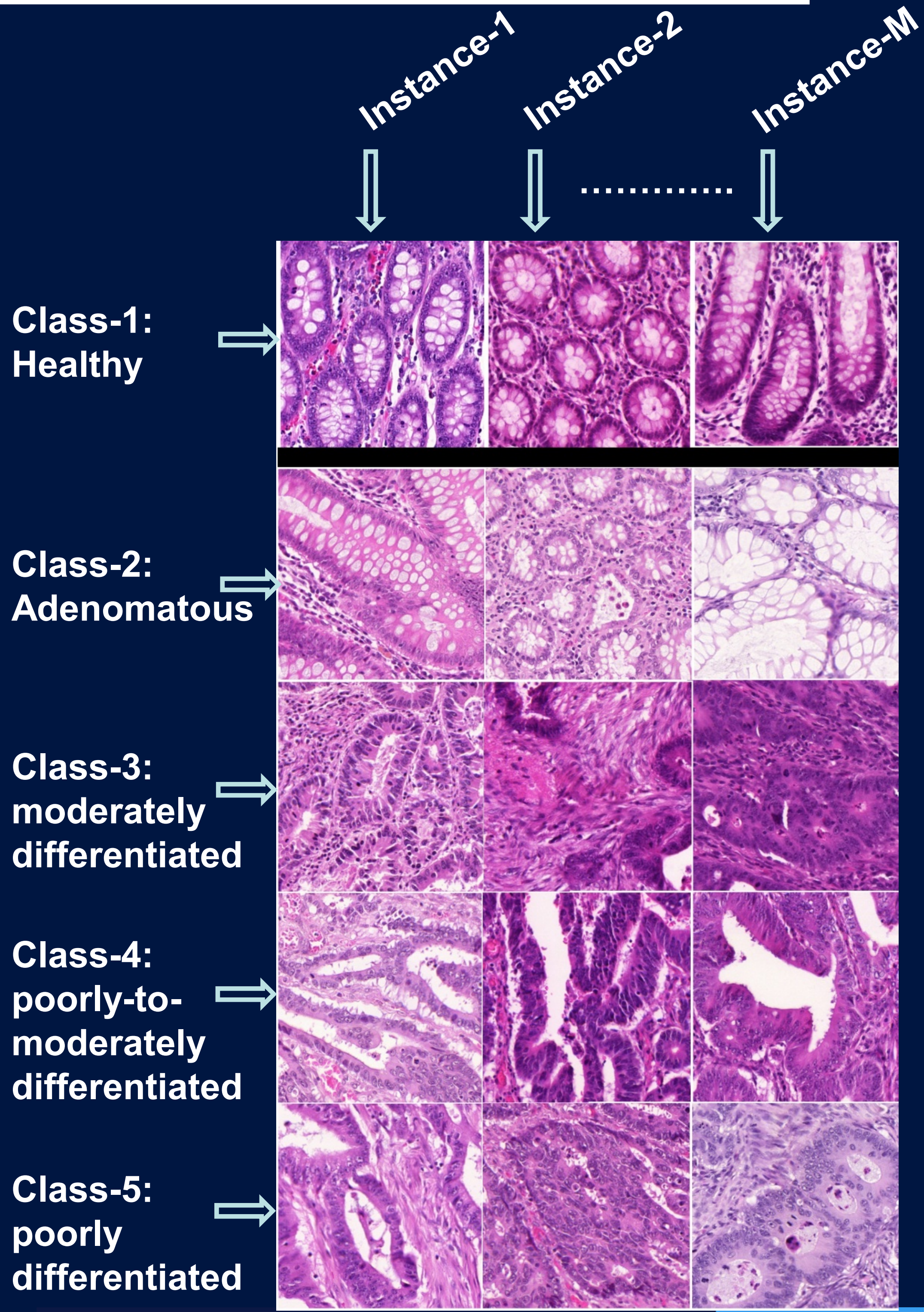
The Edward S. Rogers S
of Electrical & Comput
UNIVERSITY OF T

No Imagenet in Digital Pathology

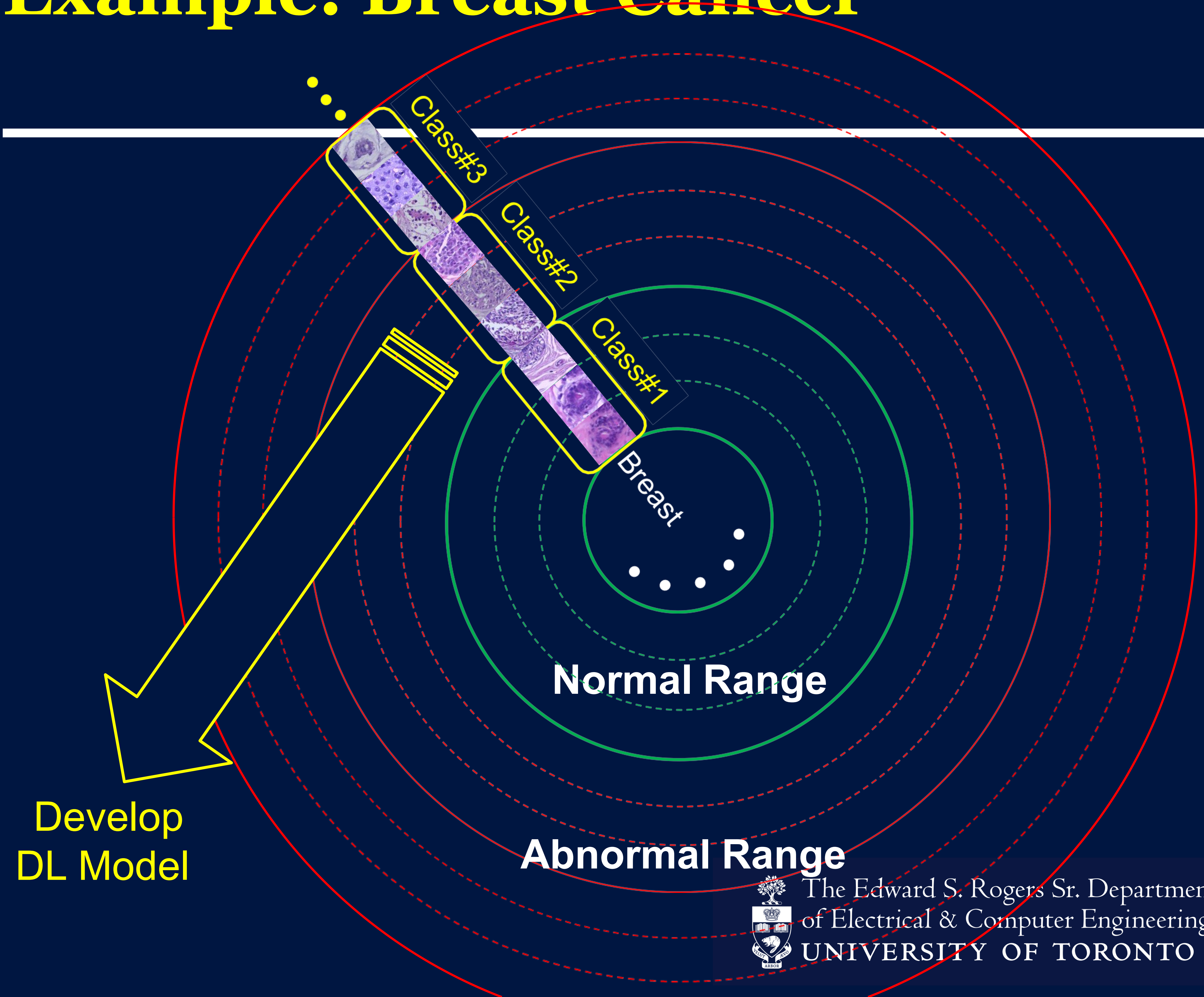
- ImageNet covers 1000 object classes across 1M images

How Digital Pathology Datasets Compare to ImageNet?

- There are limited number of classes to train
- Classes are mixtures of healthy and diseased instances
- Compared to Healthy tissues, the number of disease categories (classes) are much higher
- Each disease category varies extensively in terms of representation complexity
- There is No Consensus on how computational pathology database should look like for CAD and ML developments that can generalize for pathology field (similar to what ImageNet did for computer vision)



Example: Breast Cancer



The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

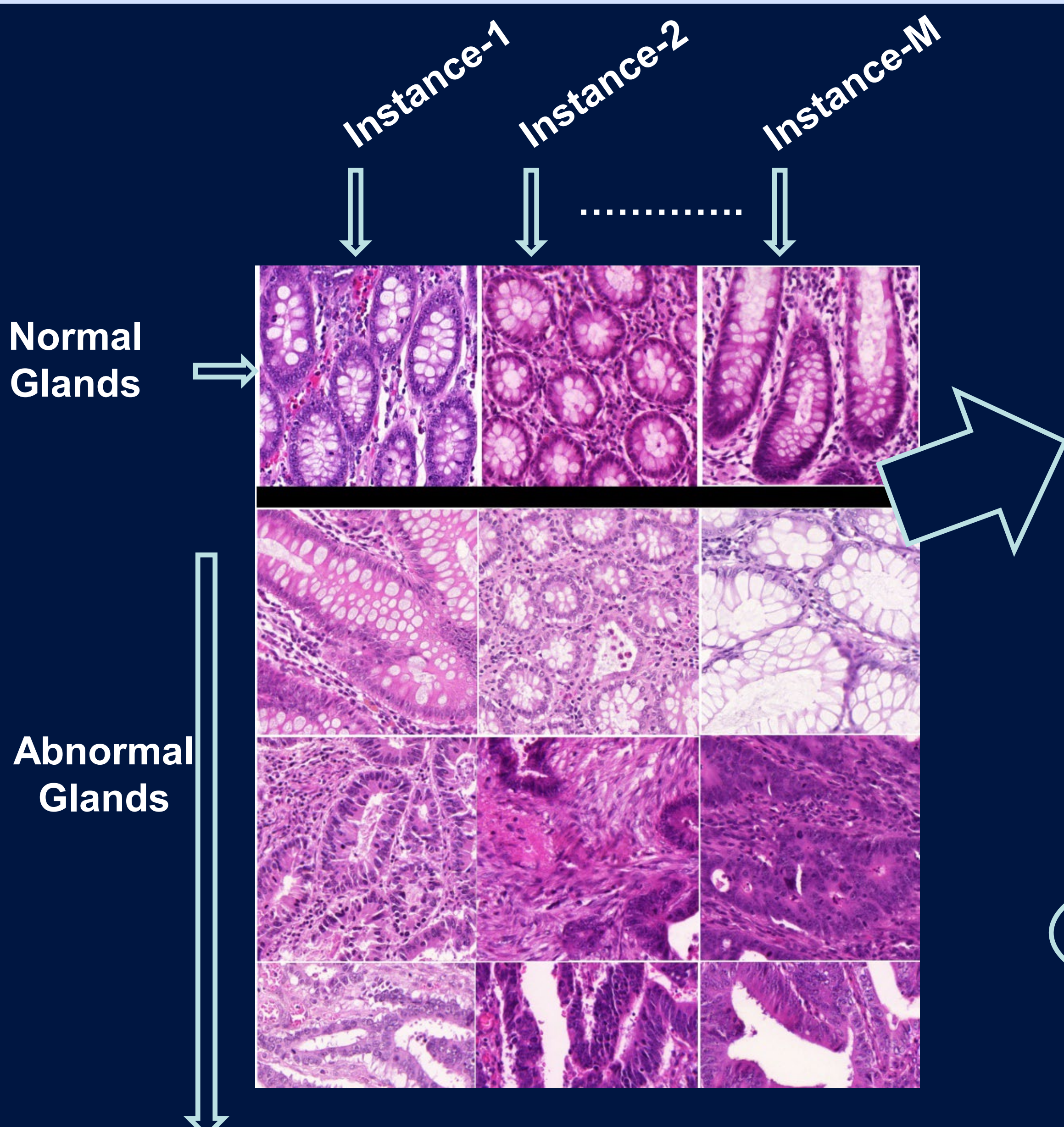
FACULTY
OF APPLIED
SCIENCE &
ENGINEERING

Why Histology Matters?

- Pathology residences study histology (the normal structure) to better understand a pathological (abnormal) change and the consequences of that change.

In Pathology Medicine

In Machine Learning

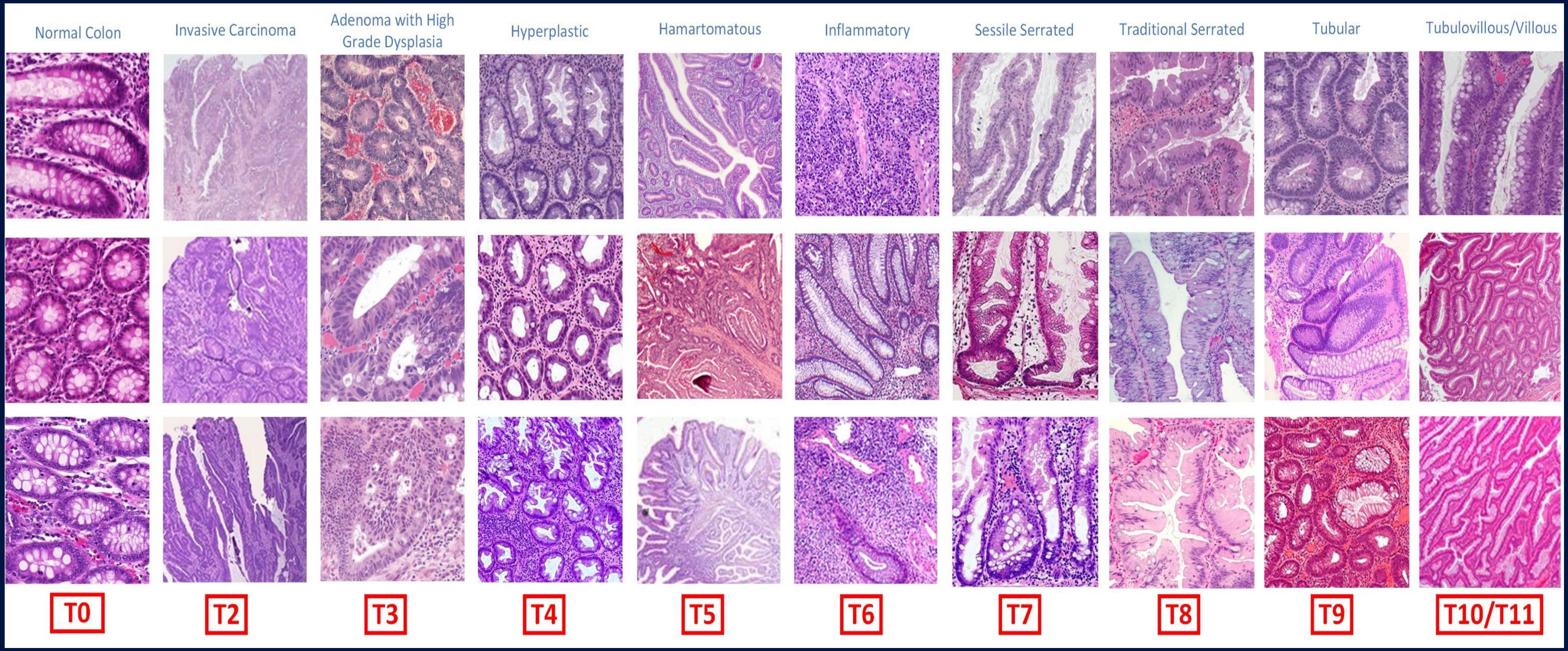


If I can learn the "Normal" Looking tissues, then I can tell which ones are "Abnormal" without learning them!!!

There is some consensus on histology image representation differentiated by tissue category compared to diseased tissue

Tolstoy on Happy/Unhappy Families

All happy families are alike; each unhappy family is unhappy in its own way



How is it really effecting ML development?

The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Existing Limitations of Current DC3 approaches

- Here is the ultimate question: Is there any way to compile a database to generalize for more conditions?
- Downside: It will probably take for ever… to compile more organ types and diseases for the purpose of generalization!

A radical new approach:

Atlas of Digital Pathology (ADP)

ADP Database—Leave No Tissue Behind

Identified all Histological Tissue Types (HTTs) in Hierarchical Taxonomy

- Multiple Organs

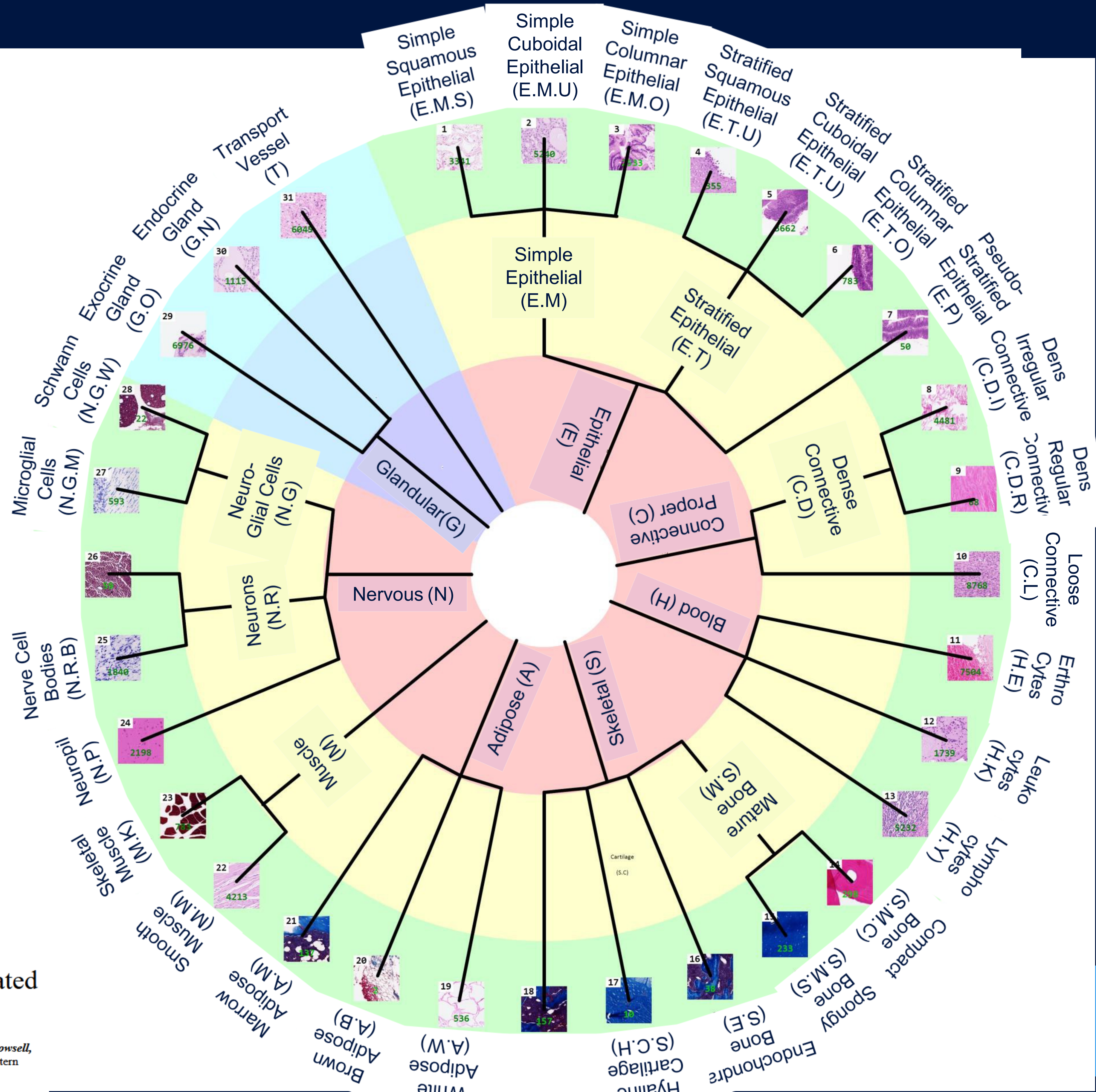
-Selections are primarily from Healthy tissue range

Level1: 36 HTTs

Level2: 23 HTTs

Level1: 9 HTTs

Multi-Label per patch



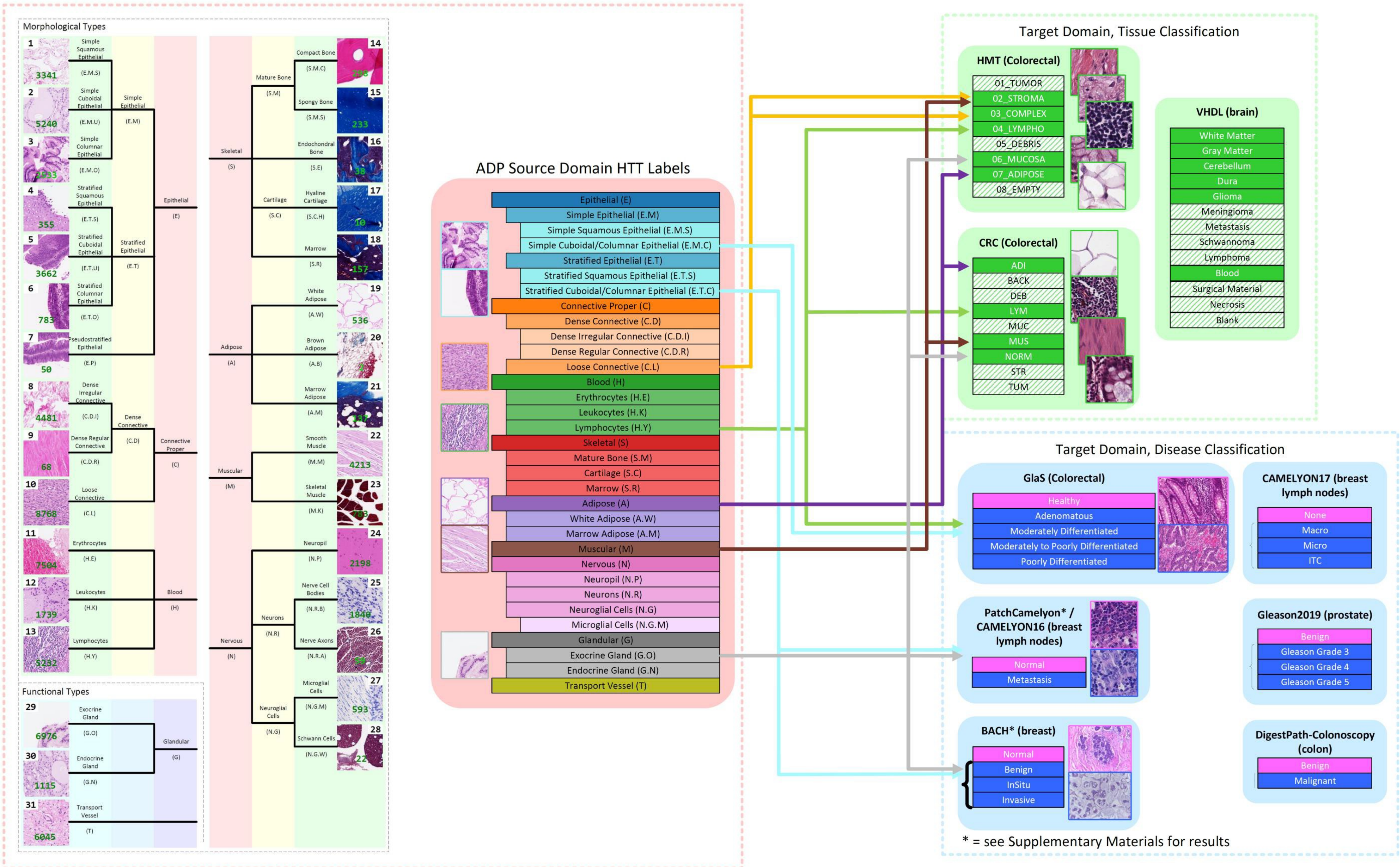
CVPR 2019 open access

These CVPR 2019 papers are the Open Access versions, provided by the Computer Vision Foundation.
Except for the watermark, they are identical to the accepted versions; the final published version of the proceedings is available on IEEE Xplore.

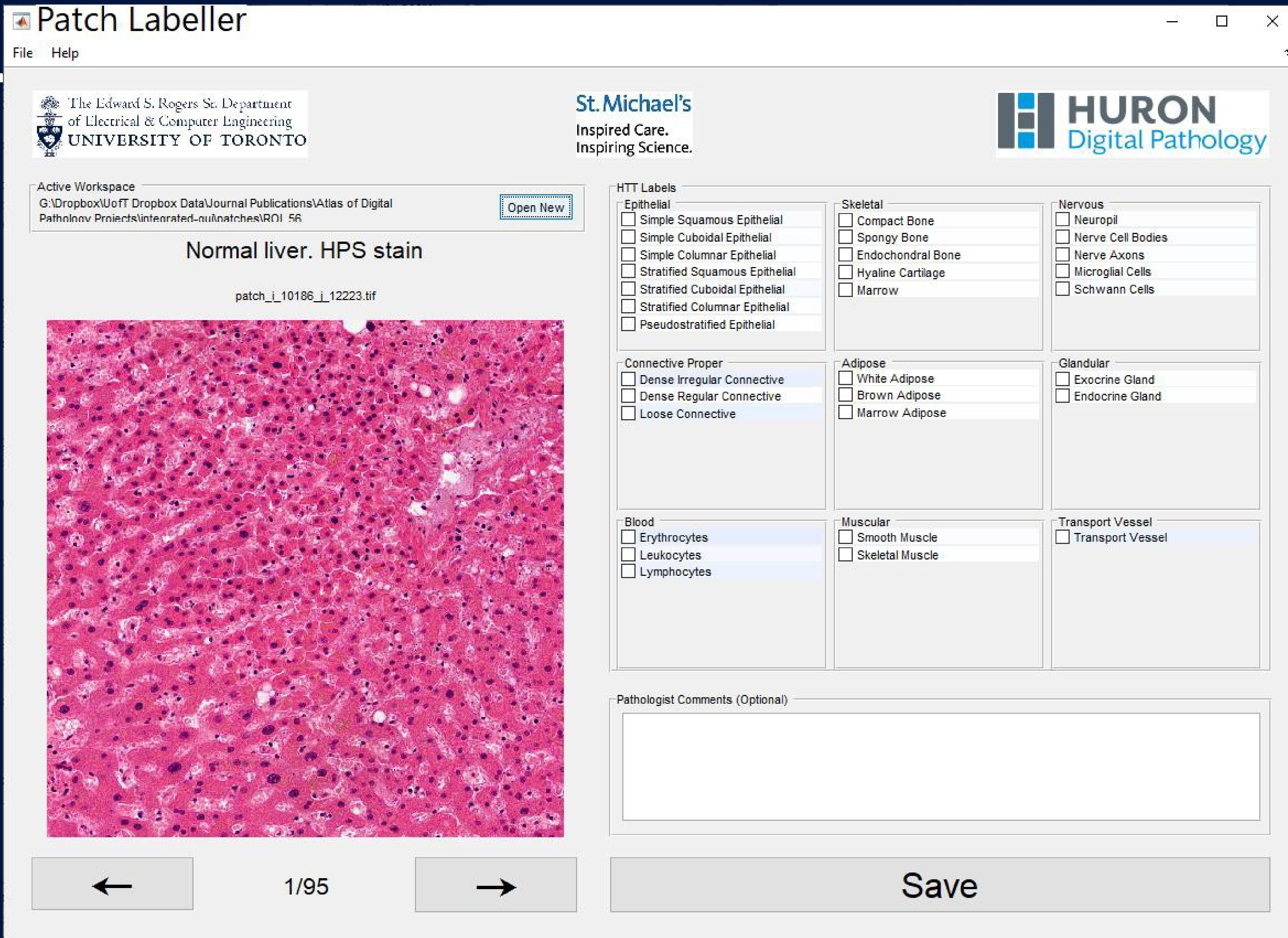
Atlas of Digital Pathology: A Generalized Hierarchical Histological Tissue Type-Annotated Database for Deep Learning

Mahdi S. Hosseini, Lyndon Chan, Gabriel Tse, Michael Tang, Jun Deng, Sajad Norouzi, Corwyn Rowsell, Konstantinos N. Plataniotis, Savvas Damaskinos; The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 11747-11756

ADP is Union of Multiple Datasets in Digital Pathology



ADP GUI for Patch Labeling



FACULTY OF APPLIED SCIENCE & ENGINEERING

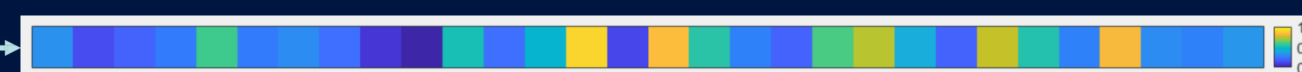
ADP is Comprehensive Supervised Database for CNN Training

Step-1: Train CNN on ADP database

Step2: Predict multiple HTTs on Novel patch

Pre-Trained CNN on ADP

HTT Confidence Scores



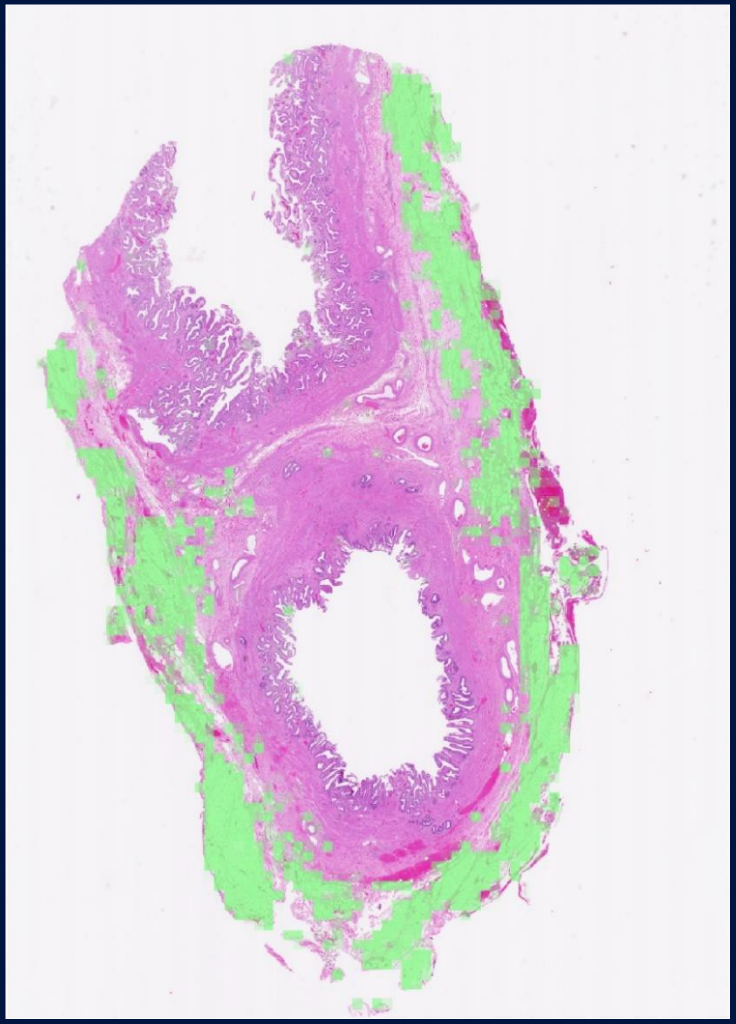
Code of 36 Values

HTT scores indicate Confidence of existence of tissue type in patch

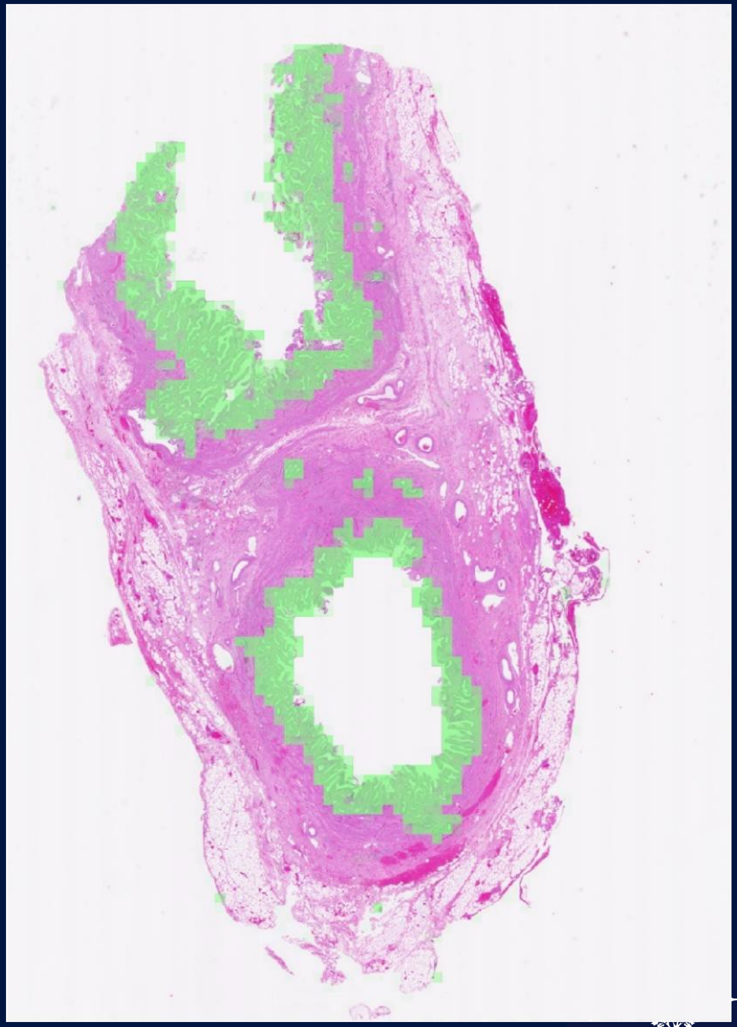
Digital Pathology-1: Visual Aiding

Visual Aid heatmaps can guide through educational pathology

White Adipose Tissues

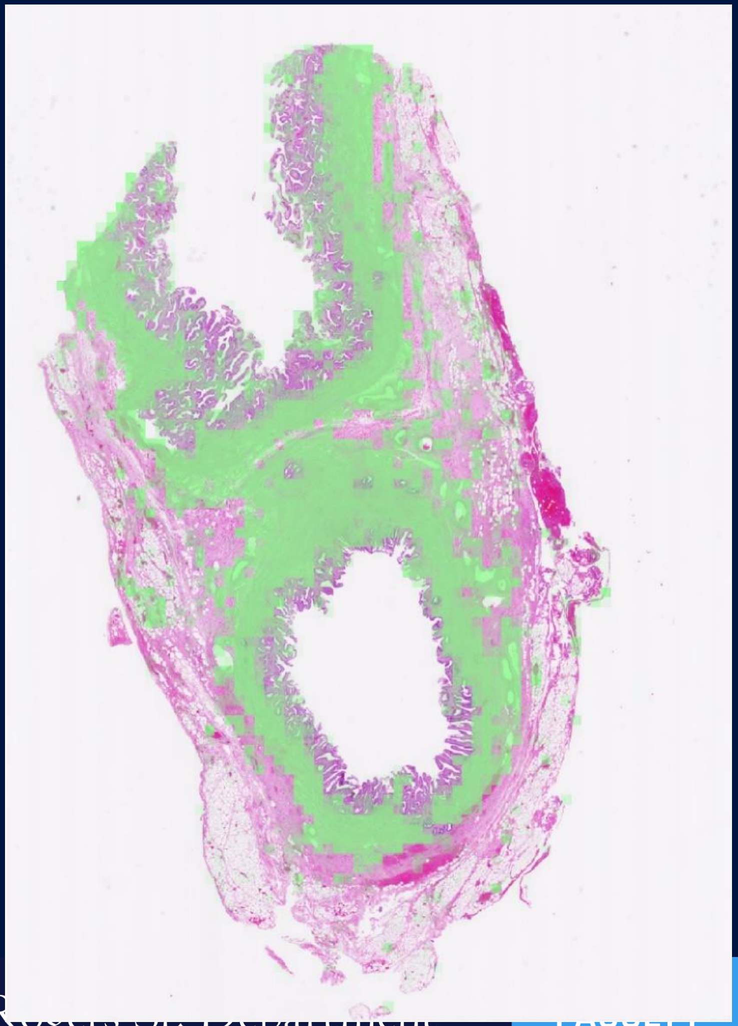


Exocrine Gland Tissues



....

Smooth Muscle Tissues



Digital Pathology 2: Pixel-Level Classification

Developing Weakly-Supervised Semantic Segmentation on ADP to infer Pixel-Level class from Patch-Level

ICCV 2019 Seoul, Korea

ICCV 2019 open access

These ICCV 2019 papers are the Open Access versions, provided by the Computer Vision Foundation.
Except for the watermark, they are identical to the accepted versions; the final published version of the proceedings is available on IEEE Xplore.

HistoSegNet: Semantic Segmentation of Histological Tissue Type in Whole Slide Images

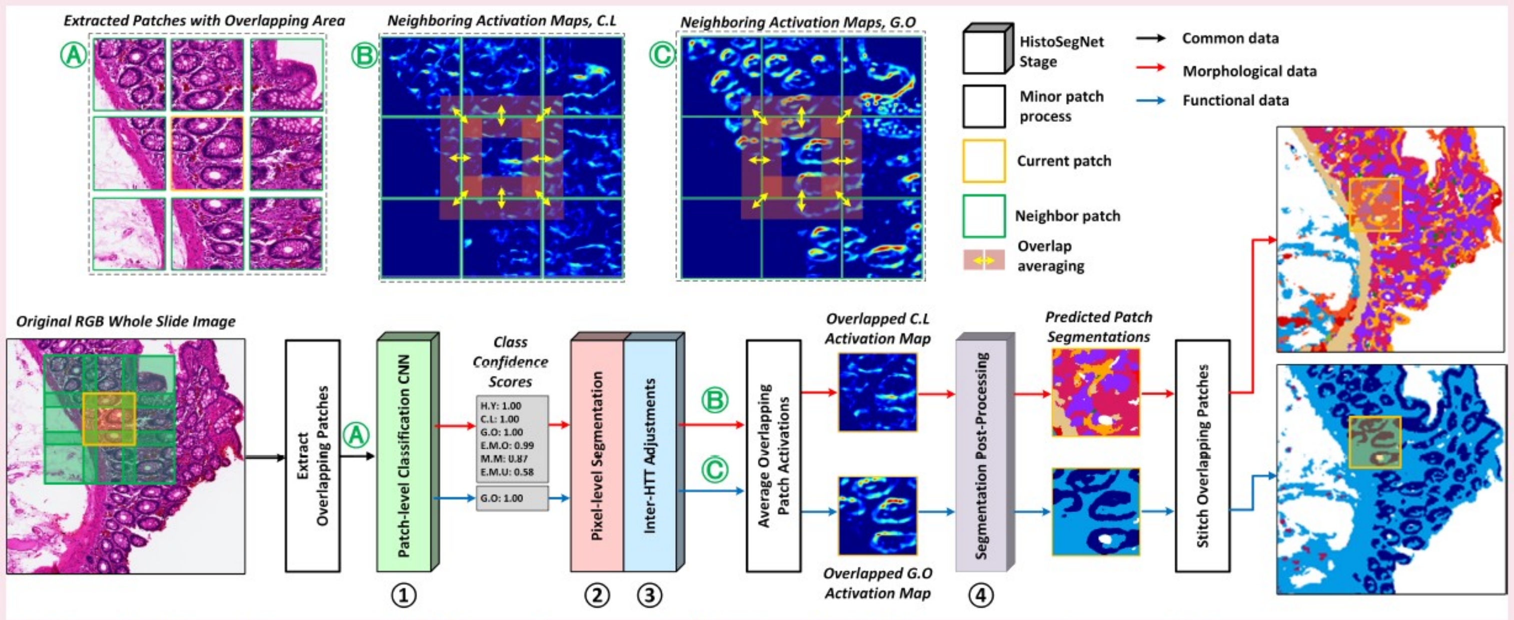
Lyndon Chan, Mahdi S. Hosseini, Corwyn Rowsell, Konstantinos N. Plataniotis, Savvas Damaskinos; The IEEE International Conference on Computer Vision (ICCV), 2019, pp. 10662-10671

International Journal of Computer Vision (2020)

A Comprehensive Analysis of Weakly-Supervised Semantic Segmentation in Different Image Domains

Lyndon Chan · Mahdi S. Hosseini · Konstantinos N. Plataniotis

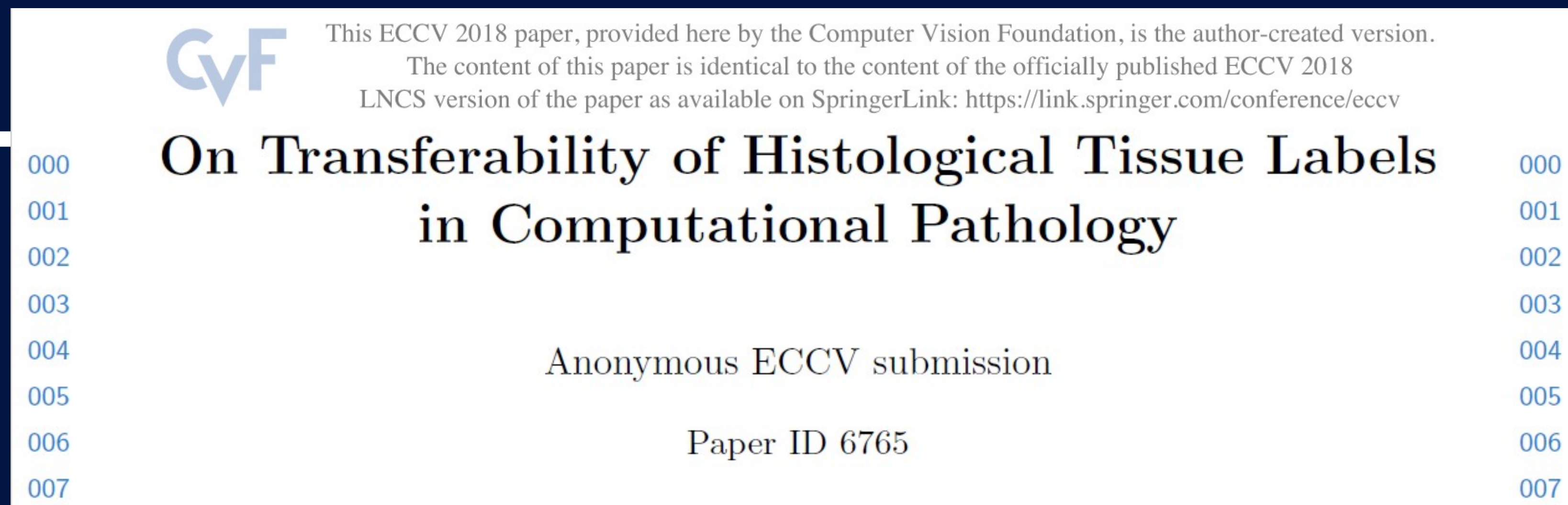
Methodology (HistoSegNet)



❶ **Patch-level Classification CNN**: predict patch HTTs with CNN pre-trained on ADP
❷ **Pixel-level Segmentation**: Grad-CAM applied on CNN to give score-scaled class activation maps
❸ **Inter-HTT Adjustments**: generate *background*/ *other* maps, subtract maps from each other
❹ **Segmentation Post-Processing**: apply dense CRF to improve segmentation

Code: `https://github.com/lyndonchan/hsn_v1`

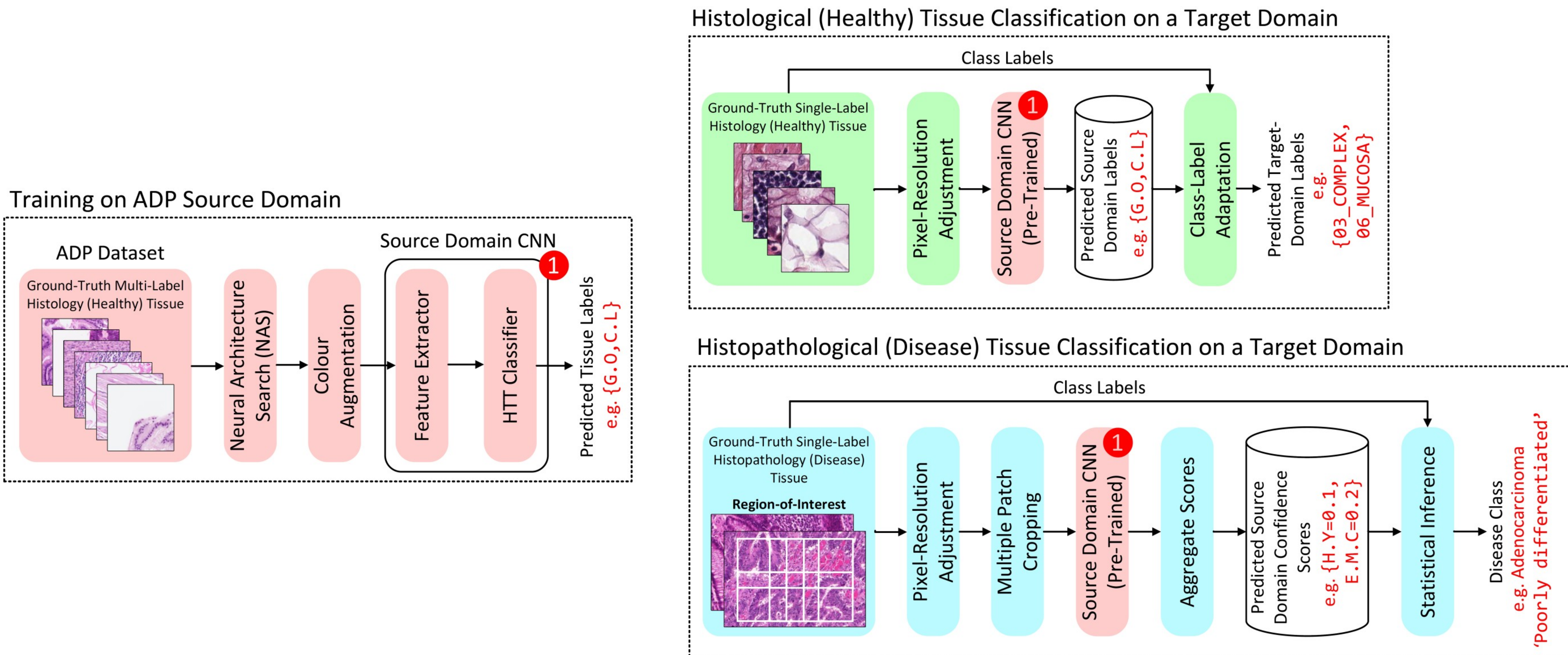
Digital Pathology-3: Cancer Diagnosis


This ECCV 2018 paper, provided here by the Computer Vision Foundation, is the author-created version. The content of this paper is identical to the content of the officially published ECCV 2018 LNCS version of the paper as available on SpringerLink: https://link.springer.com/conference/eccv

On Transferability of Histological Tissue Labels in Computational Pathology

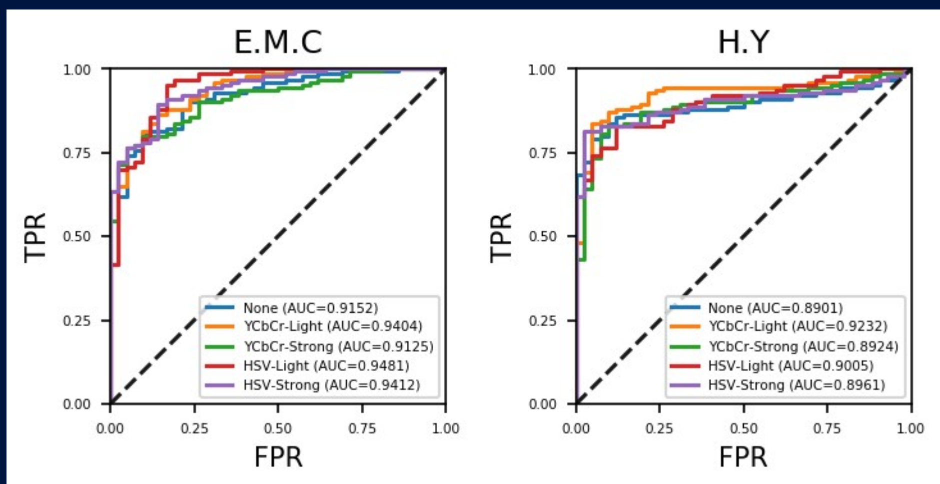
Anonymous ECCV submission

Paper ID 6765

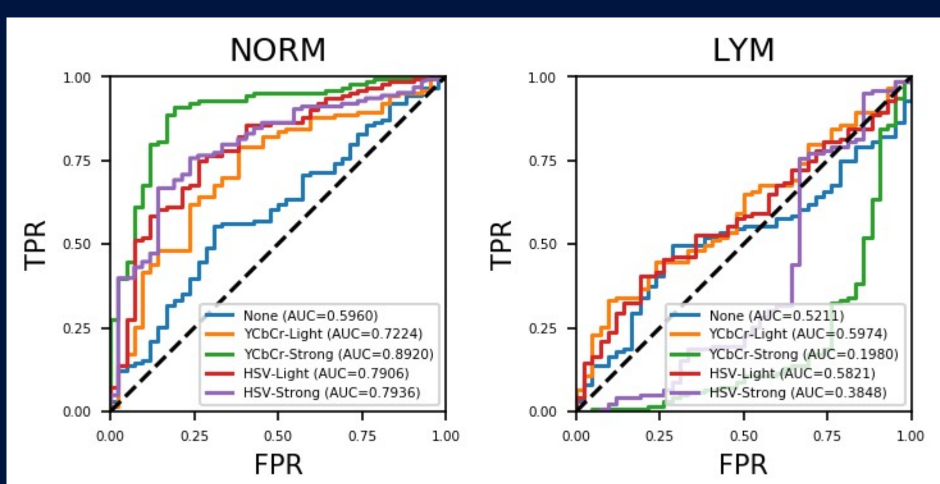


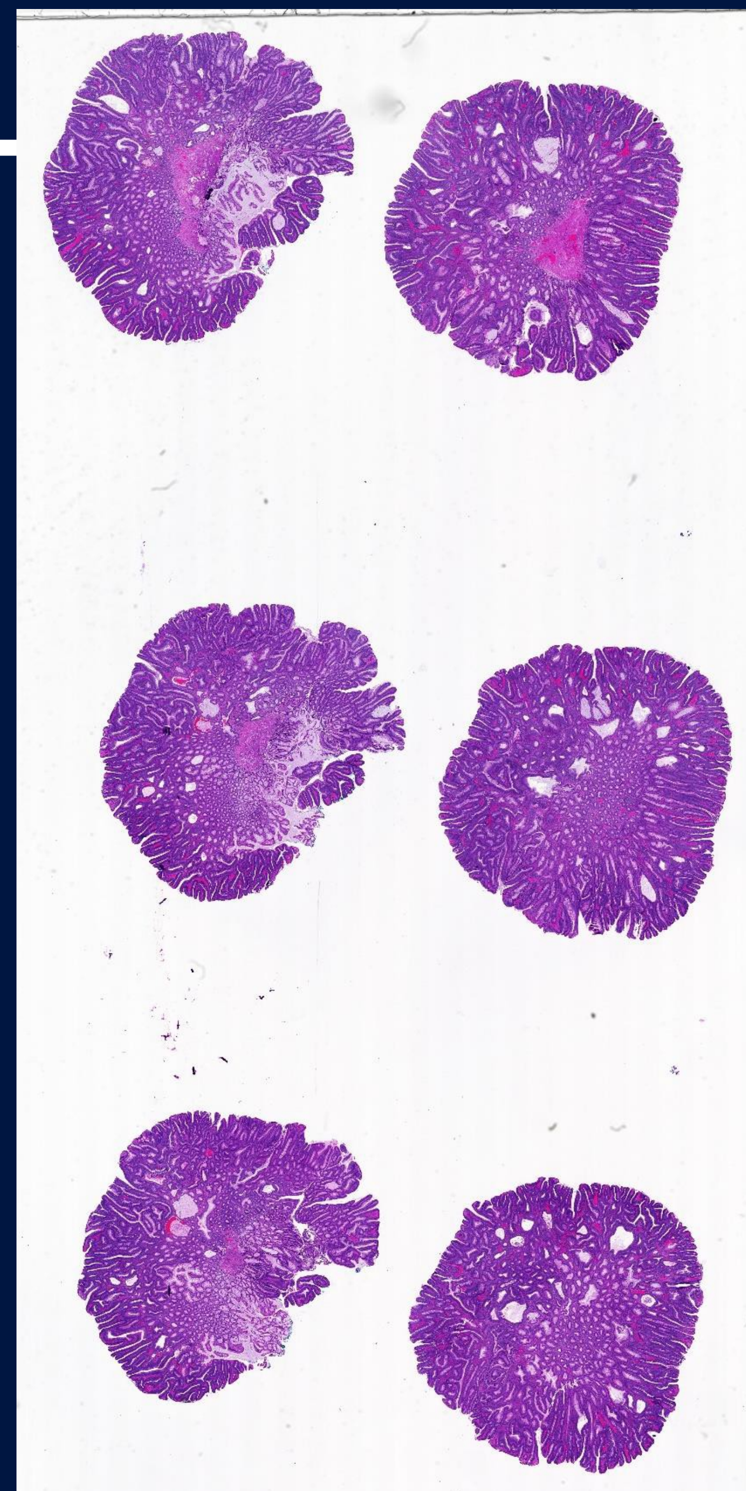
Example Application-3: Colon Polyps

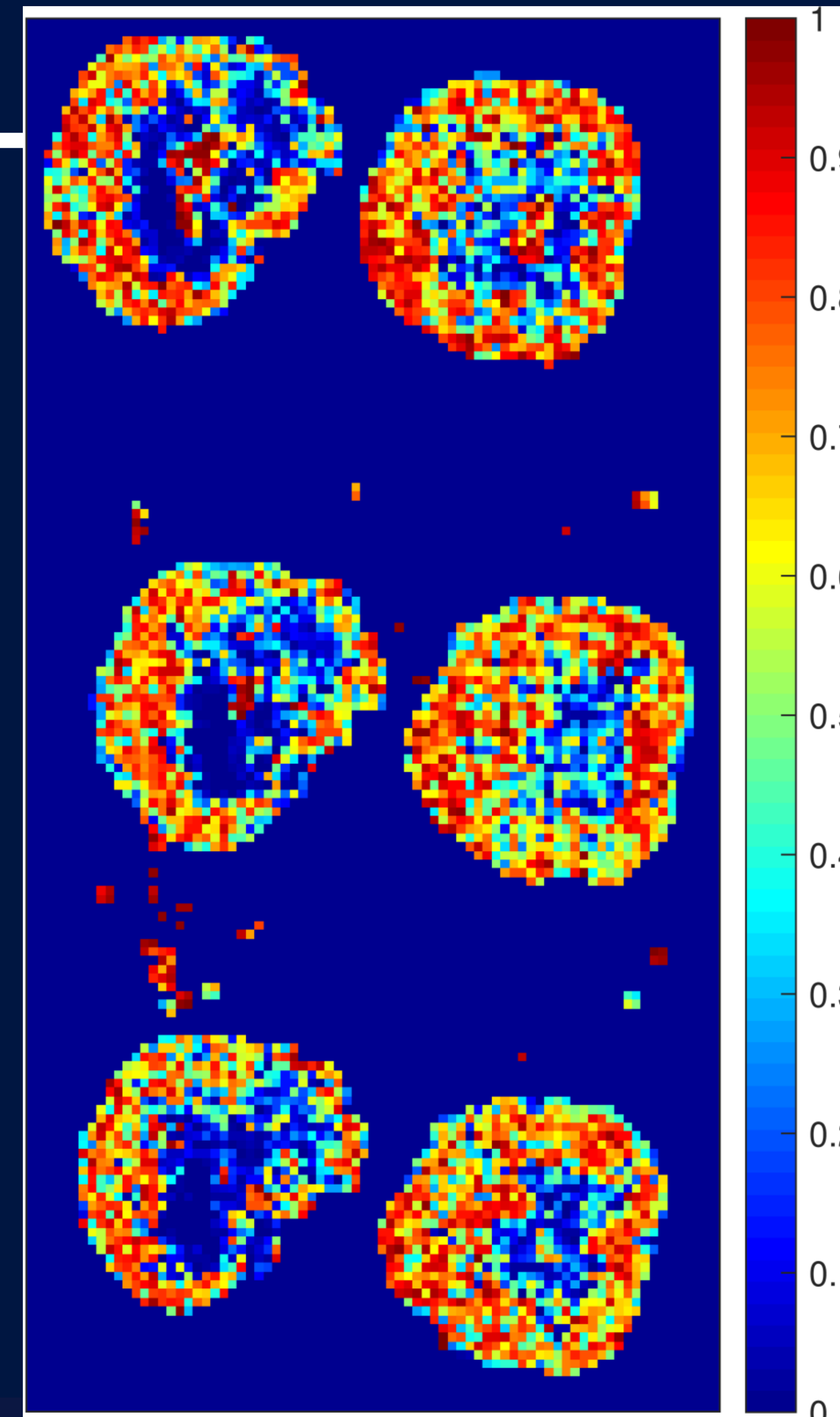
Transferring from ADP



Transferring from Other Data



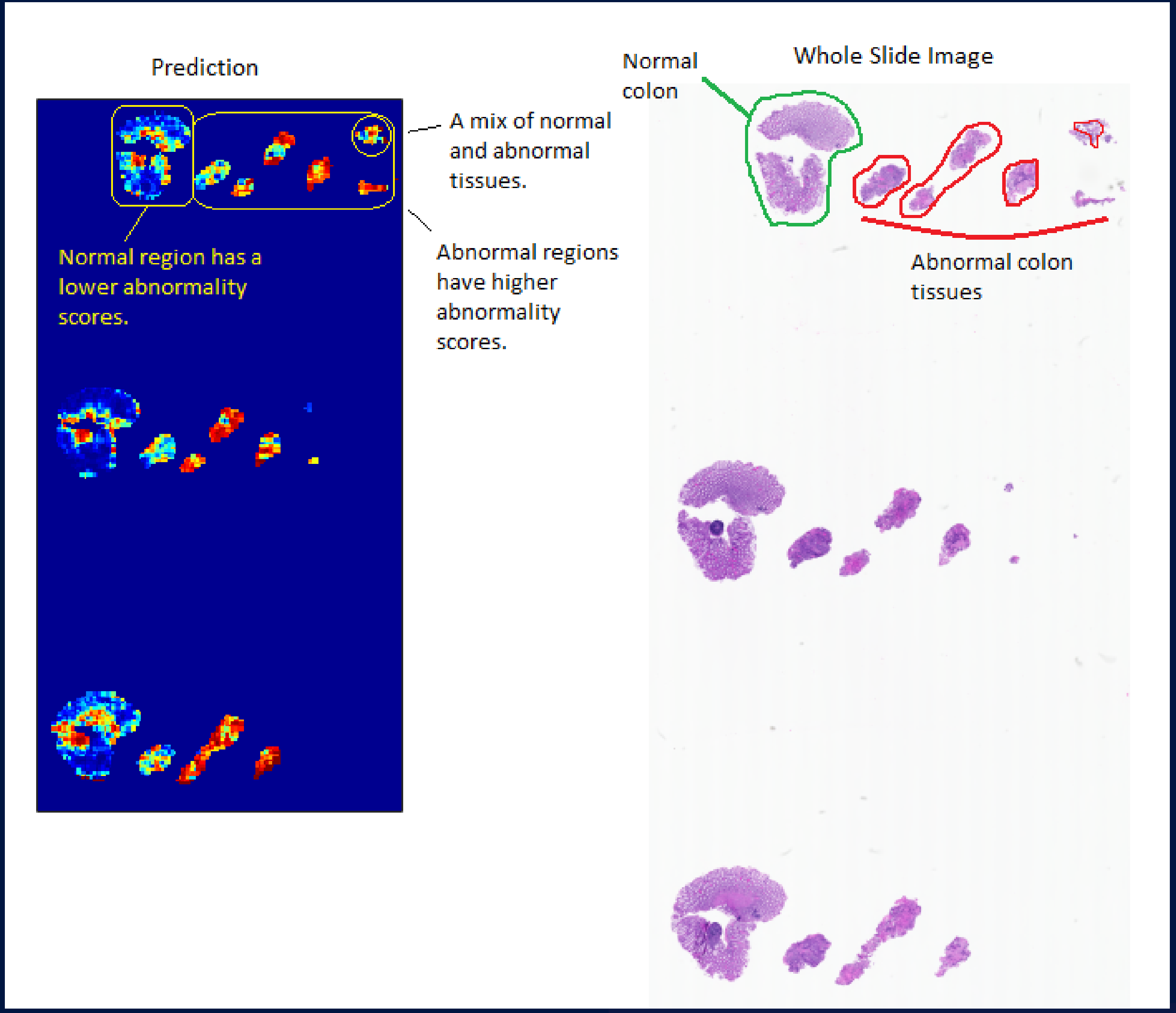




Example Application-4: Abnormality Detection

- The confidence level of CNN prediction associates with healthiness of tissue, because CNN is trained on ADP (histological tissues)
- Low confidence correlates with abnormalities e.g. disease/cancer

Performance on Unseen Colon Tissue



Digital Pathology
Clinical Integration

Coming Soon

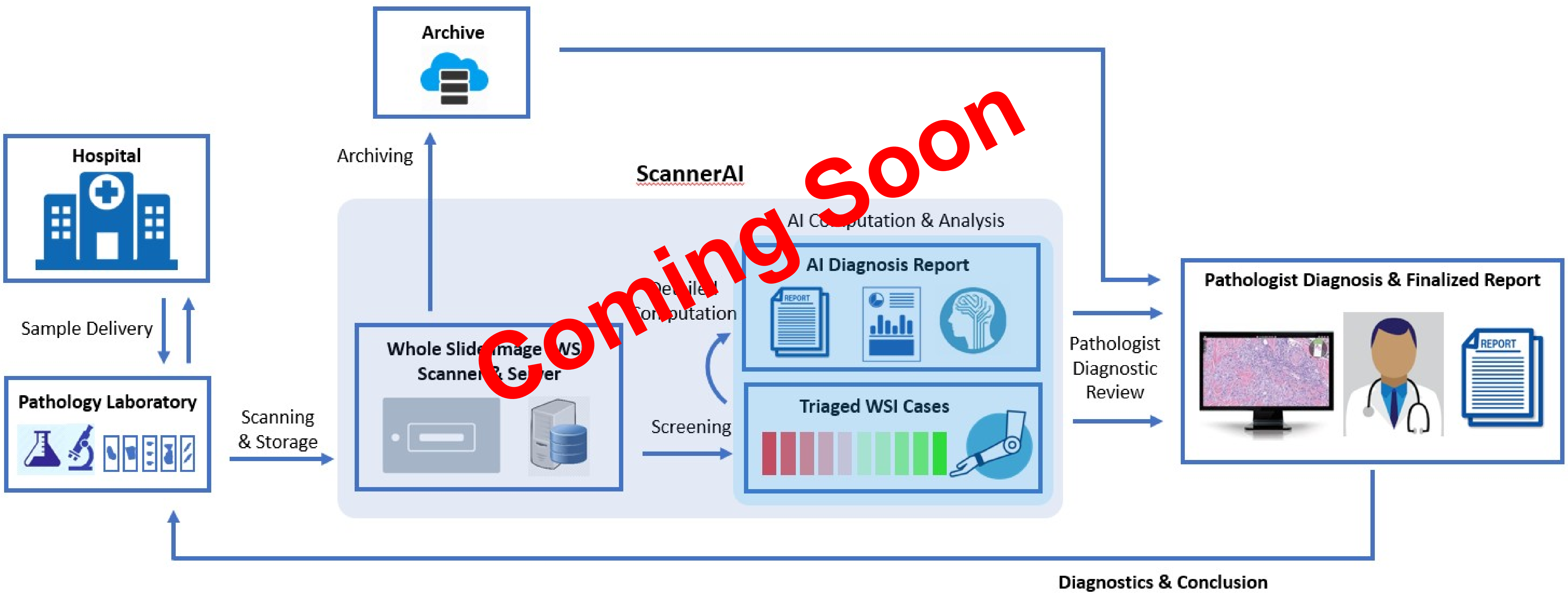
Digital Pathology
Clinical Integration

Current Monolithic IT Is a Major Barrier

Source: Alex Motoc, Unsplash.com

Gartner Reports (UofT)

Digital Pathology
Clinical Integration



The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY
OF APPLIED
SCIENCE &
ENGINEERING

The ATLAS team

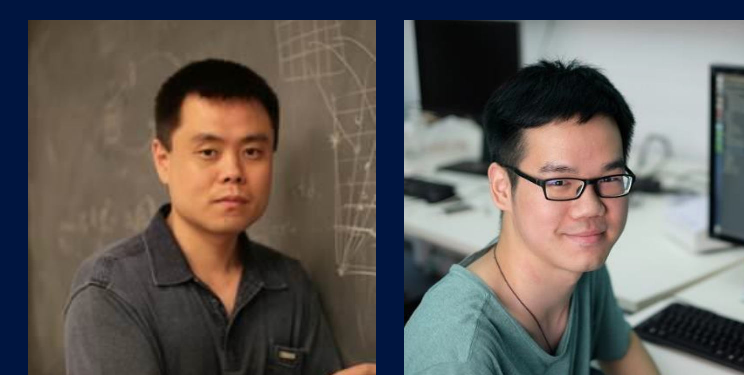
Multimedia Lab – UofT

St. Michael Hospital & UofT

Kingston Health Science Center & Queens University

Huron Digital Pathology

University of Waterloo



The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE & ENGINEERING

Reading Material (Optional)

1. A comprehensive analysis of weakly-supervised semantic segmentation in different image domains, Lyndon Chan, Mahdi S Hosseini, Konstantinos N Plataniotis, International Journal of Computer Vision, also available: https://arxiv.org/pdf/1912.11186.pdf.
2. On Transferability of Histological Tissue Labels in Computational Pathology, Mahdi S Hosseini, Lyndon Chan, Weimin Huang, Yichen Wang, Danial Hasan, Corwyn Rowsell, Savvas Damaskinos, Konstantinos N Plataniotis, European Conference on Computer Vision, 2020, https://link.springer.com/chapter/10.1007/978-3-030-58526-6_27.
3. HistoSegNet: Semantic Segmentation of Histological Tissue Type in Whole Slide Images, Savvas Damaskinos Lyndon Chan, Mahdi S. Hosseini, Corwyn Rowsell, Konstantinos N. Plataniotis, The IEEE International Conference on Computer Vision (ICCV), 2019, http://openaccess.thecvf.com/content_ICCV_2019/html/Chan_HistoSegNet_Semantic_Segmentation_of_Histological_Tissue_Type_in_Whole_Slide_ICCV_2019_paper.htm.
4. Atlas of Digital Pathology: A Generalized Hierarchical Histological Tissue Type-Annotated Database for Deep Learning, Savvas Damaskinos Mahdi S Hosseini, Lyndon Chan, Gabriel Tse, Michael Tang, Jun Deng, Sajad Norouzi, Corwyn Rowsell, Konstantinos N Plataniotis, CVPR 2019,, https://openaccess.thecvf.com/content_CVPR_2019/html/Hosseini_Atlas_of_Digital_Pathology_A_Generalized_Hierarchical_Histological_Tissue_Type-Annotated_CVPR_2019_paper.html

Additional Information:

- K.N. Plataniotis Google Scholar: https://scholar.google.com/citations?user=W-4N_2gAAAAJ&hl=en

Thank You!

Q&A

www.dsp.utoronto.ca

http://www.dsp.utoronto.ca/projects/ADP/

kostas@ece.utoronto.ca