TIA Centre
Tissue Image Analytics

Introduction to Computational Pathology

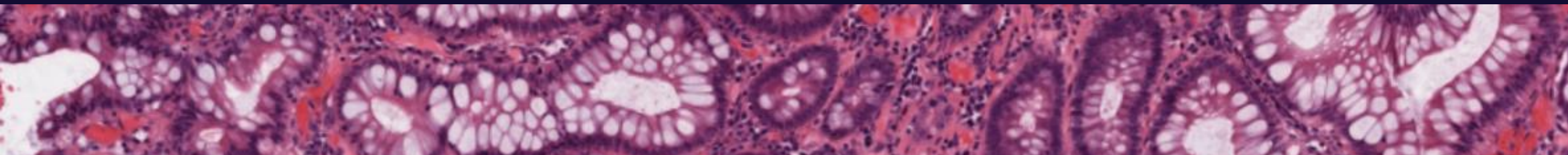
Fayyaz Minhas

Department of Computer Science

TIA Centre

PathLAKE Consortium

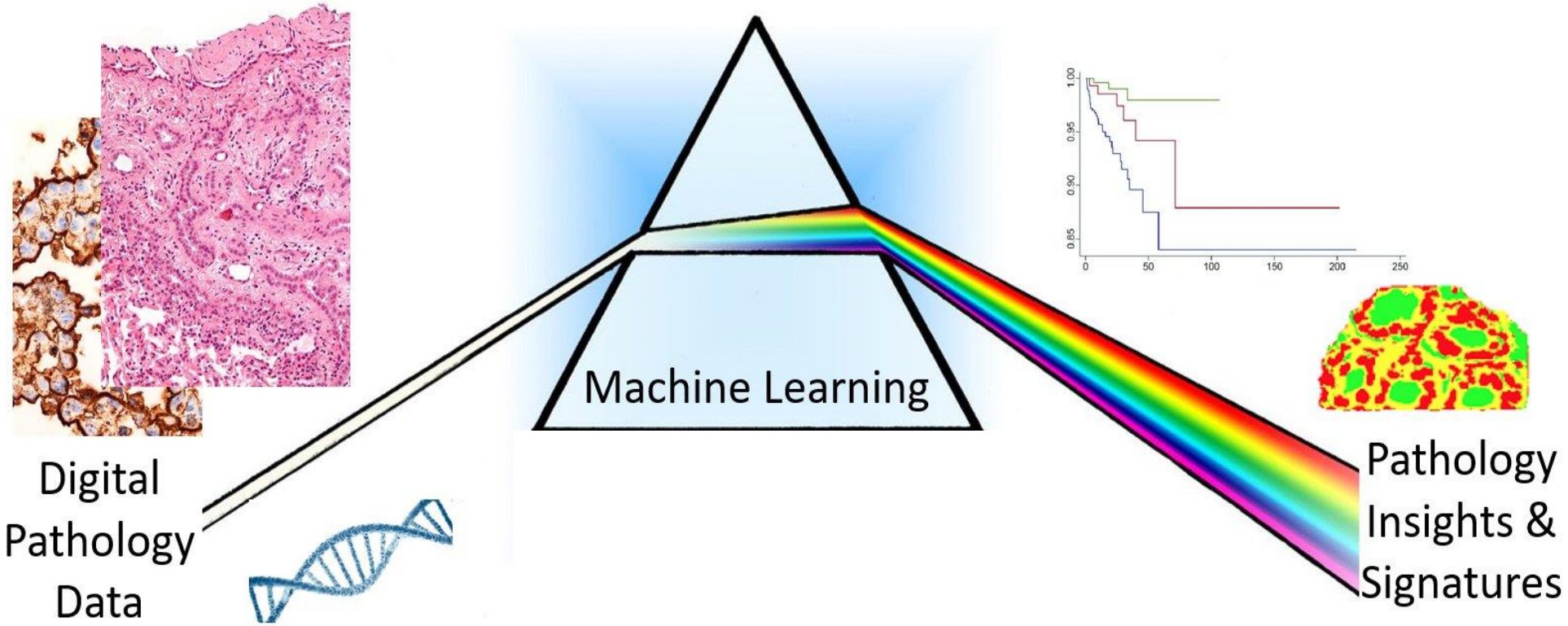
@fayyazhere



Theme of this talk

- **Workflow**
  - Digital vs. Computational Pathology
- **State of Cpath**
  - 5-star system for measuring the impact of Cpath
  - Including a taxonomy of computational problems in CPath
- **Challenges and Next Steps**

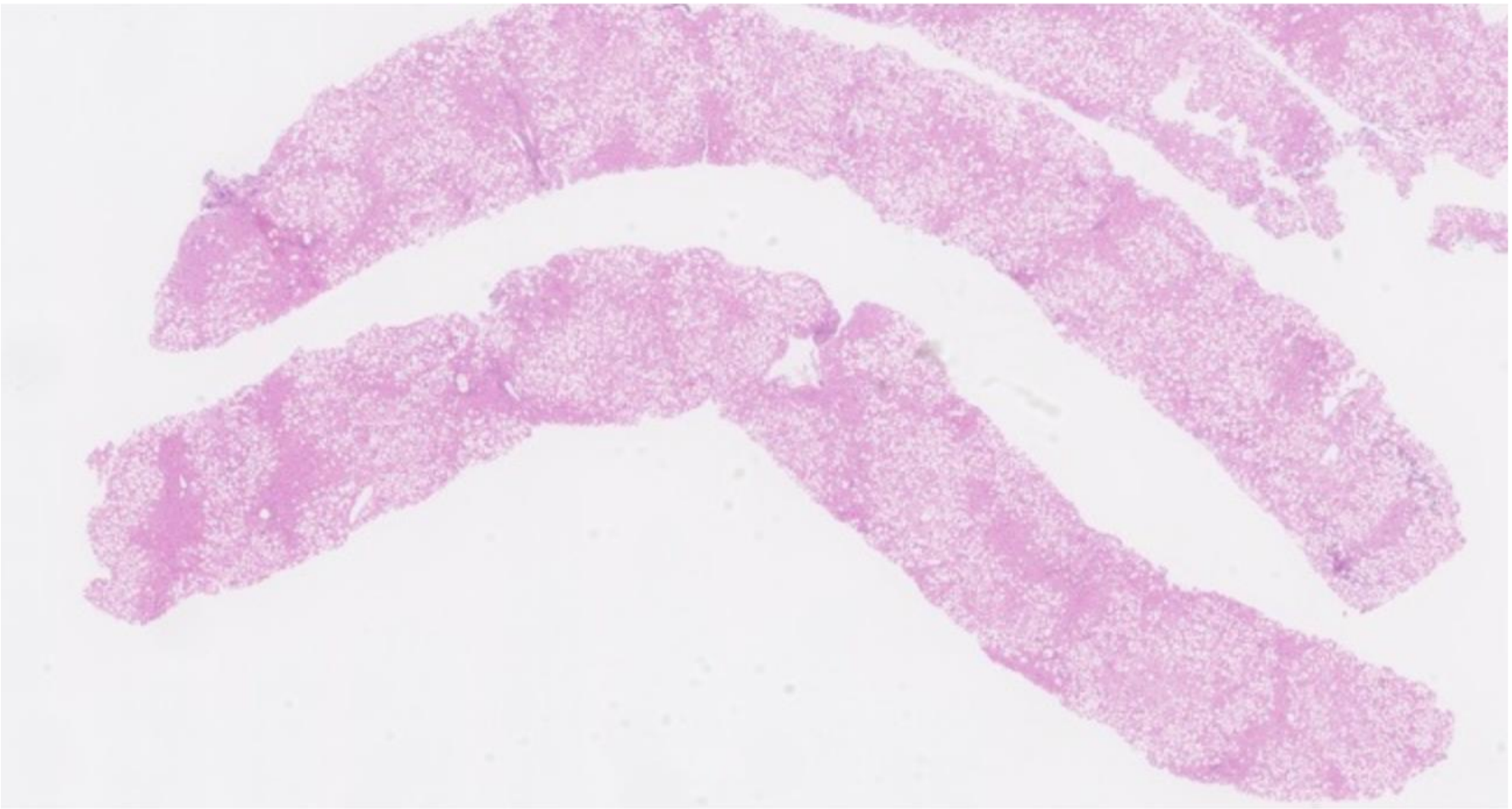
Computational Pathology



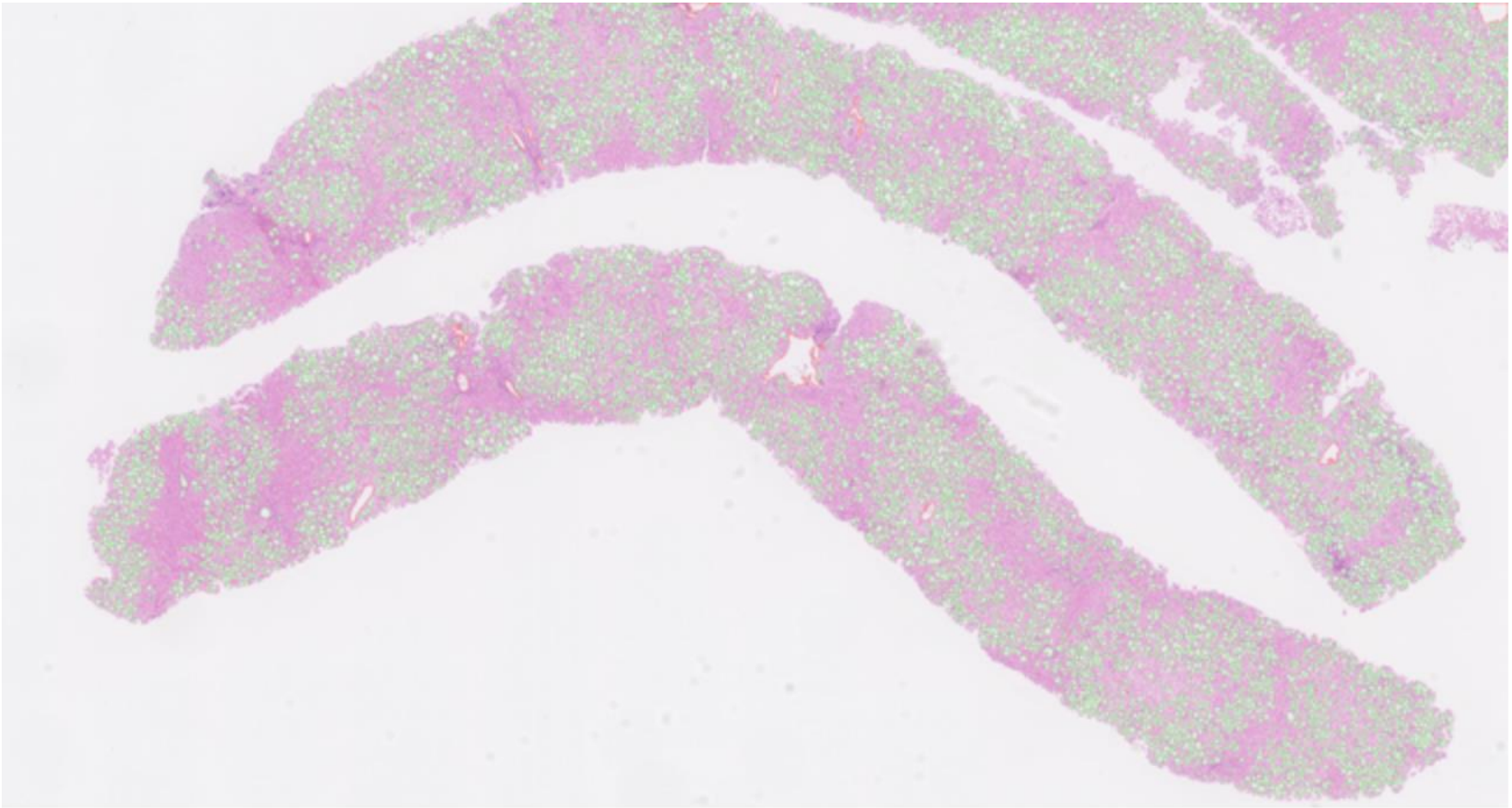
AI to help understand the cause, nature, origin and patterns of diseases for clinical decision making

Louis, David N., Georg K. Gerber, Jason M. Baron, Lyn Bry, Anand S. Dighe, Gad Getz, John M. Higgins, et al. "Computational Pathology: An Emerging Definition." *Archives of Pathology & Laboratory Medicine* 138, no. 9 (September 1, 2014): 1133–38. https://doi.org/10.5858/arpa.2014-0034-ED.

How much fat?



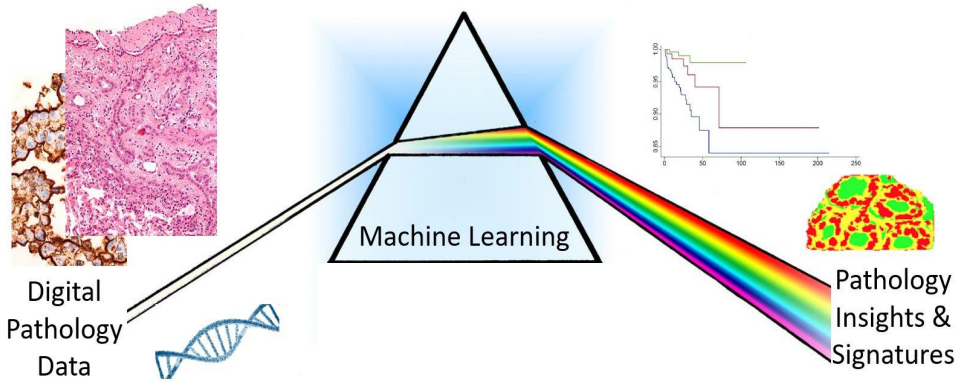
45.9%



R. Glodin, “The Clinical Challenge of Image Analysis and AI”, CRUK Early Detection Sandpit, Royal College of Pathologists, London, UK, 20 Nov. 2019

Why is applying ML to pathology a good idea?

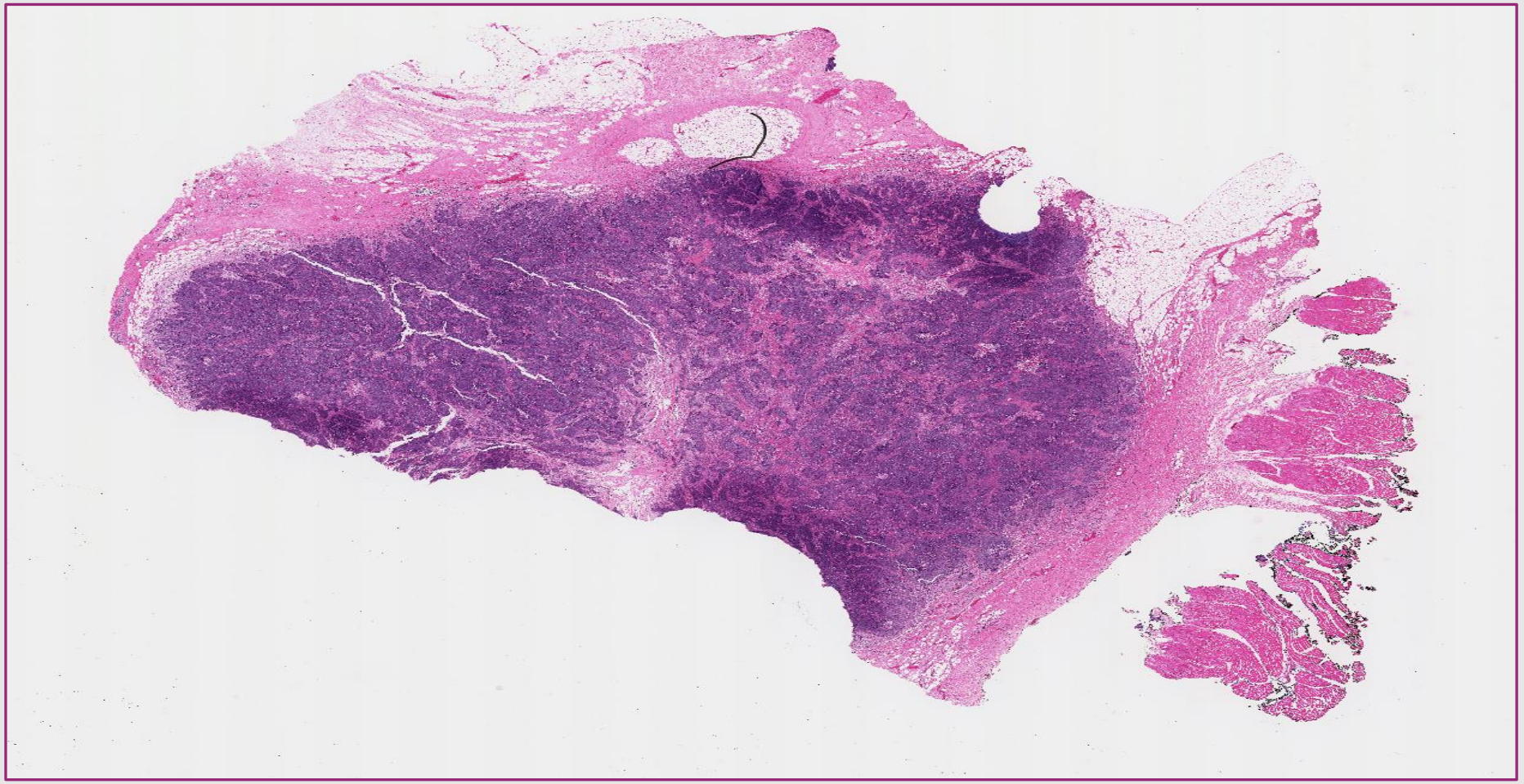
- It is a very visual discipline
- Subjectivity
- Quantification
- Digitization as an enabling technology
  - Flexible working
  - Good concordance with slide based clinical decision making
- Pathologist Recruitment and aging workforce
- Identification of discriminatory features
- Learning from “Data at Scale”
- Information Integration
- Clinical Impact
- Commercial Impact



Workflow


Glioblastoma (GBM) Samples
HE Stained FFPE Tissue Slide
Tissue Acquisition
Slide Preparation
Scanning
Quality Control
Pre-processing
Computational Algorithms

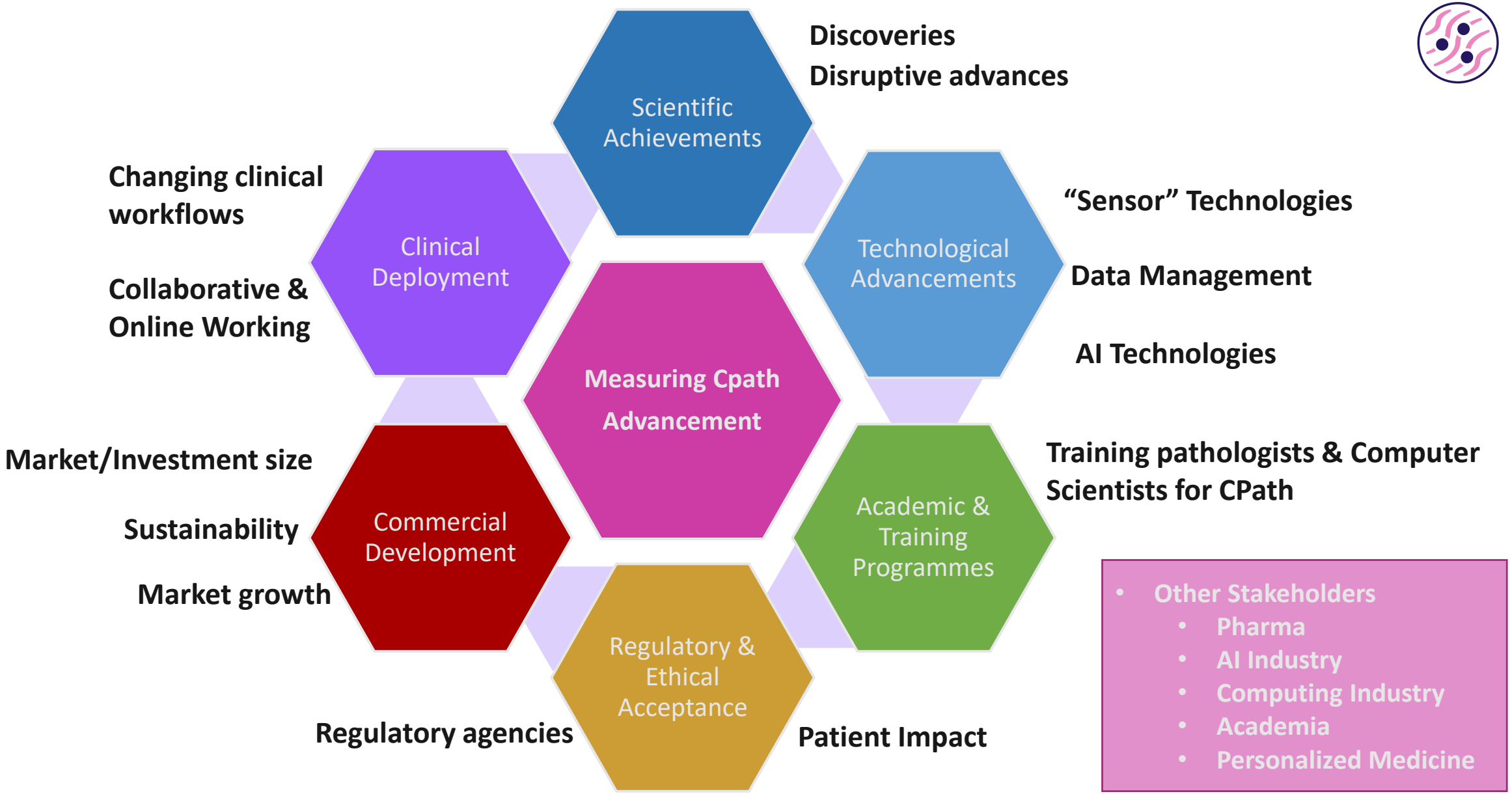
Prediction Problems in CPath



Given training data,
Can we predict:

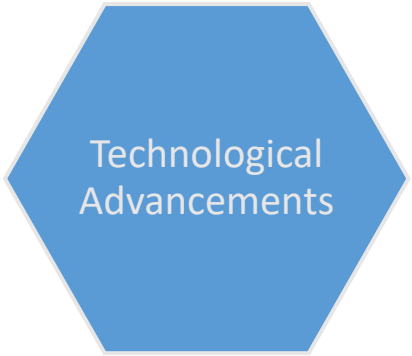
Grade?
Regression
Tumour Regions?
Segmentation
Quantifying cells
Quantification
Receptor Status?
Classification
Mutation Status?
Classification
Survival?
Drug Response?
Similar images?
Retrieval
Biomarkers?
Pattern Discovery


Scientific Achievements
Clinical Deployment
Technological Advancements
Measuring Cpath Advancement
Commercial Development
Academic & Training Programmes
Regulatory & Ethical Acceptance


Discoveries
Disruptive advances
Scientific Achievements
Changing clinical workflows
Collaborative & Online Working
Clinical Deployment
"Sensor" Technologies
Data Management
AI Technologies
Technological Advancements
Measuring Cpath Advancement
Market/Investment size
Sustainability
Market growth
Commercial Development
Training pathologists & Computer Scientists for CPath
Academic & Training Programmes
Regulatory & Ethical Acceptance
Regulatory agencies
Patient Impact
Other Stakeholders
Pharma
AI Industry
Computing Industry
Academia
Personalized Medicine

Technological Advancements in CPath

- Current Focus
  - *Image analysis with Machine Learning*



“Sensor” Technologies

Data Management

AI/ML Technologies

Cpath "Sensor" Technologies

- Whole Slide Image Scanners
  - Cost, quality and throughput
- Upcoming Technologies
  - Multiplexed Imaging
  - *In Situ* Genomics and Transcriptomics
  - Multispectral Imaging
  - Liquid Biopsies
  - 3D Imaging of WSIs
  - Virtual biopsies and autopsies: Radomics
  - Non-destructive, *In Vivo* and RT pathology
  - Microscopic imaging in the body
    - Multi-photon microendoscopy

1 to 5 min for a small biopsy to 5–20 min for a surgical specimen and 3–5 min for a liquid-based cytology smear

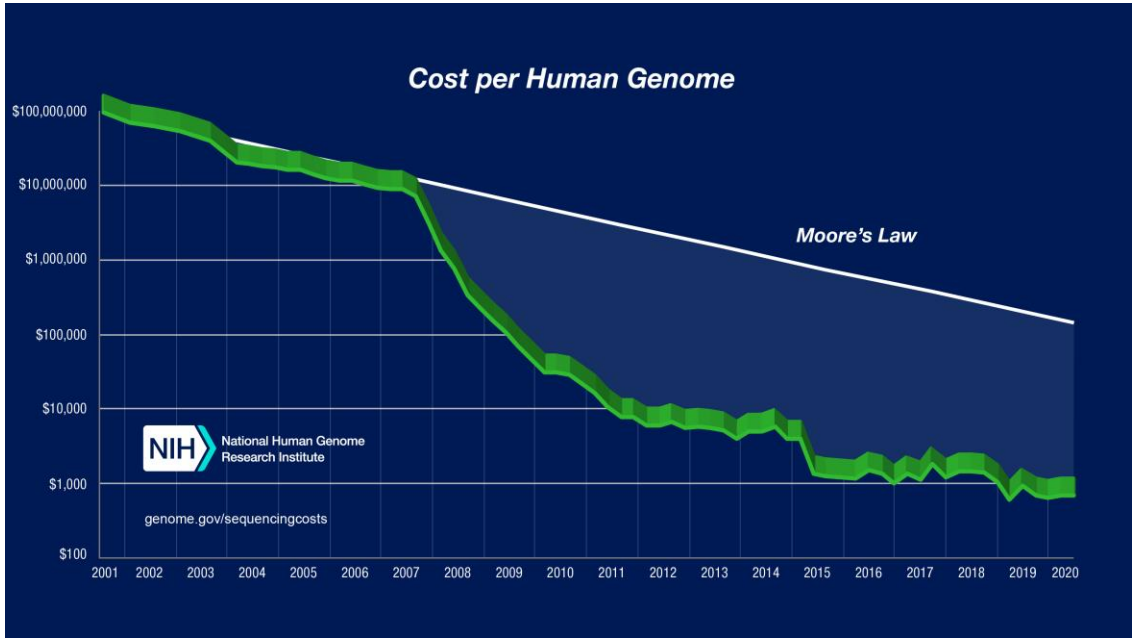
"The cost of procurement, implementation, and operational costs of WSI may be prohibitive, especially for small pathology laboratories due to huge initial cost of the scanners (US $100,000 to US $1,500,000 per piece) and additional hidden costs of training of staff and pathologists, technical support, digital slide storage systems, and regulatory or licensing costs"

"WSI images would be only as good as the original glass slide"

Kumar, Neeta, Ruchika Gupta, and Sanjay Gupta. "Whole Slide Imaging (WSI) in Pathology: Current Perspectives and Future Directions." *Journal of Digital Imaging* 33, no. 4 (August 1, 2020): 1034–40. https://doi.org/10.1007/s10278-020-00351-z.

"Sensor" technology challenges

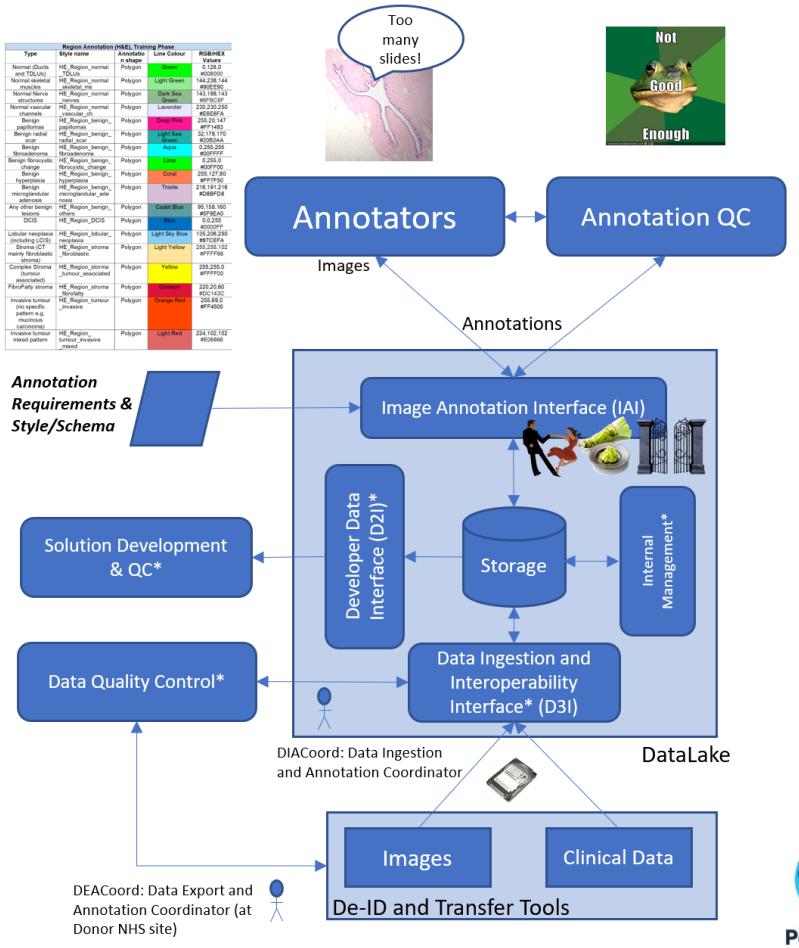
- Reduction in scanning costs
  - Does it match the decrease in sequencing or transcriptomic analysis?
- Integration of upcoming imaging technologies in clinical and scientific discovery
- Open-source hardware?
- Smartphone systems?
- Integration into LIMS



Farahani, Navid, Anil V. Parwani, and Liron Pantanowitz. "Whole Slide Imaging in Pathology: Advantages, Limitations, and Emerging Perspectives." *Pathology and Laboratory Medicine International* 7 (June 11, 2015): 23–33. https://doi.org/10.2147/PLMI.S59826.

CPath Data Management

- Development of large-scale openly accessible well-annotated data repositories
  - TCGA, CPTAC, PathLAKE, BIG-PICTURE
- De-identification
- Data Transfer
- Datalake design
  - Cloud based analytics or On-Prem solutions?
- Interoperability
  - Image formats: svs, DICOM, …
  - Annotations and pathology reports
- Collaboration and retrieval permissions



PathLAKE

Technology Advances in ML for CPath

- Advances
  - **Largely tracking advances in Computer Vision and Machine Learning**

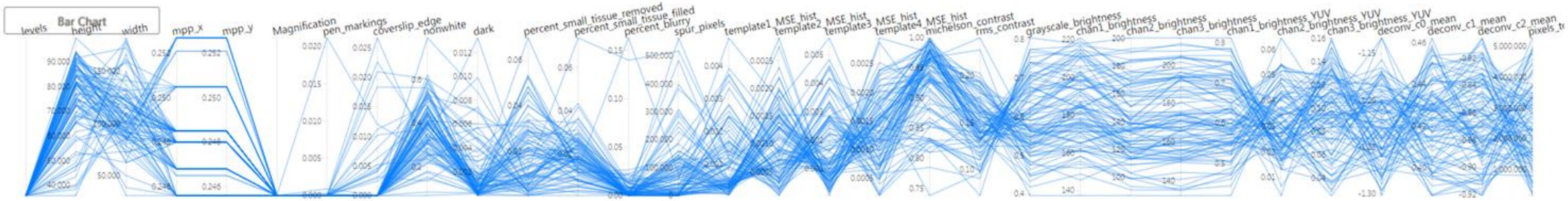
- Developed Algorithms for:
  - Pre-processing
  - Quality Control
  - Classification
  - Regression
  - Survival Prediction
  - Segmentation
  - Quantification/Counting
  - Object Detection
  - Semantic Segmentation
  - Contrastive Learning
  - Self-supervised learning
  - Active Learning
  - Registration, …
  - Generative Modelling
  - Virtual staining

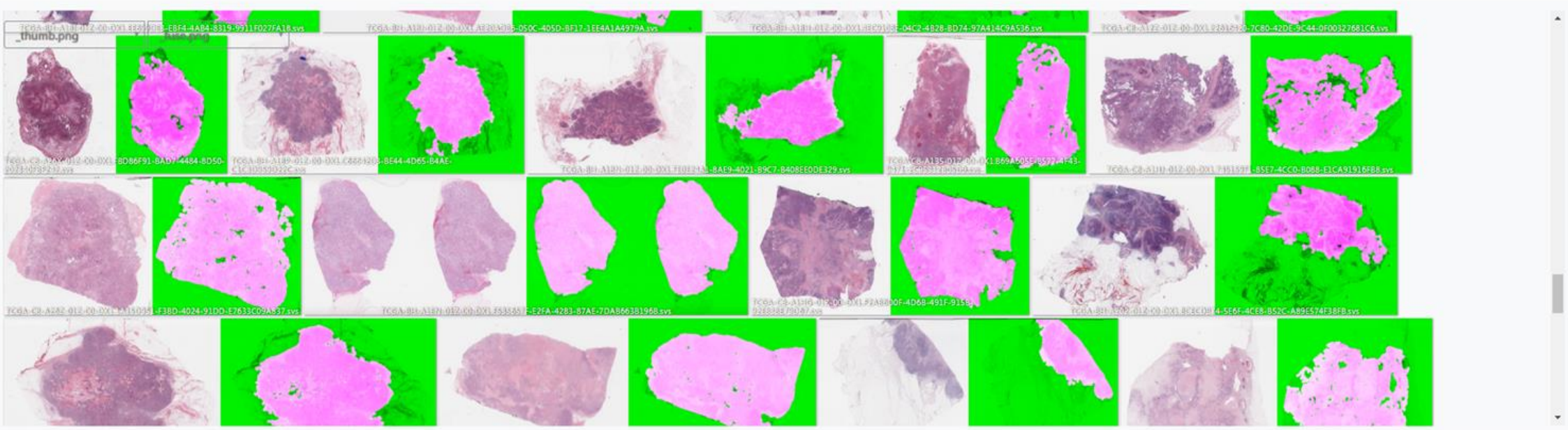
QC

HistoQc | Table | Chart | Image | About | Instruction | Reset

filename	comments	levels	height	width	mpp_x	mpp_y	Magnification	pen_markings	coverslip_edge	nonwhite	dark	percent
TCGA-D8-A1JI-01Z-00-DX2.E688E270-26C9-43EC-BC26-86BF8B74A31D.svs		4	61772	101591	0.2527	0.2527	40	0	0.010239647505201746	0.37570282730202764	0.0000585859410060789	0.02
TCGA-D8-A1JJ-01Z-00-DX1.a986b48f-b295-4d7a-b778-ce829cdf9c38.svs		4	89863	84727	0.248	0.248	40	0	0	0.40365625618646767	0.00038530934490148993	0.02
TCGA-D8-A1JJ-01Z-00-DX2.7D20F308-7DC6-4367-9459-3AC4C654E7F7.svs		4	74635	83663	0.2527	0.2527	40	0	0.02097230770845992	0.6022809295769842	0.0069632641758784015	0.02

Copy | Save | Delete | Deselect | Cols

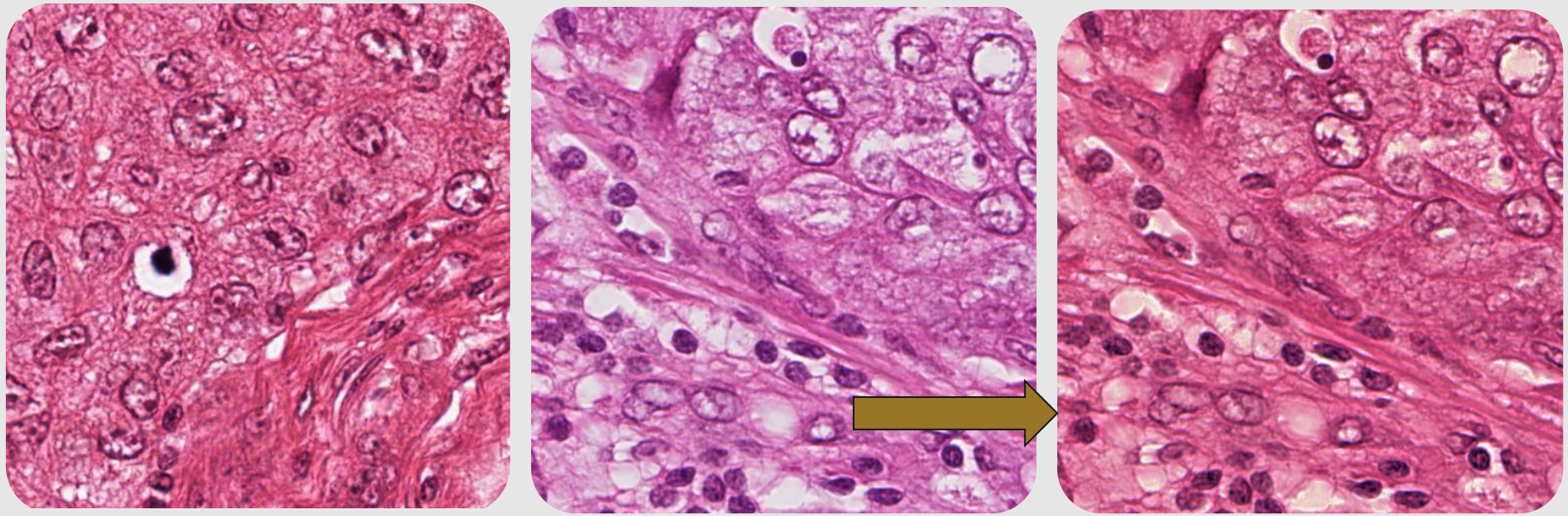




Janowczyk, Andrew, Ren Zuo, Hannah Gilmore, Michael Feldman, and Anant Madabhushi. “HistoQC: An Open-Source Quality Control Tool for Digital Pathology Slides.” *JCO Clinical Cancer Informatics* 3 (April 16, 2019). https://doi.org/10.1200/CCI.18.00157.

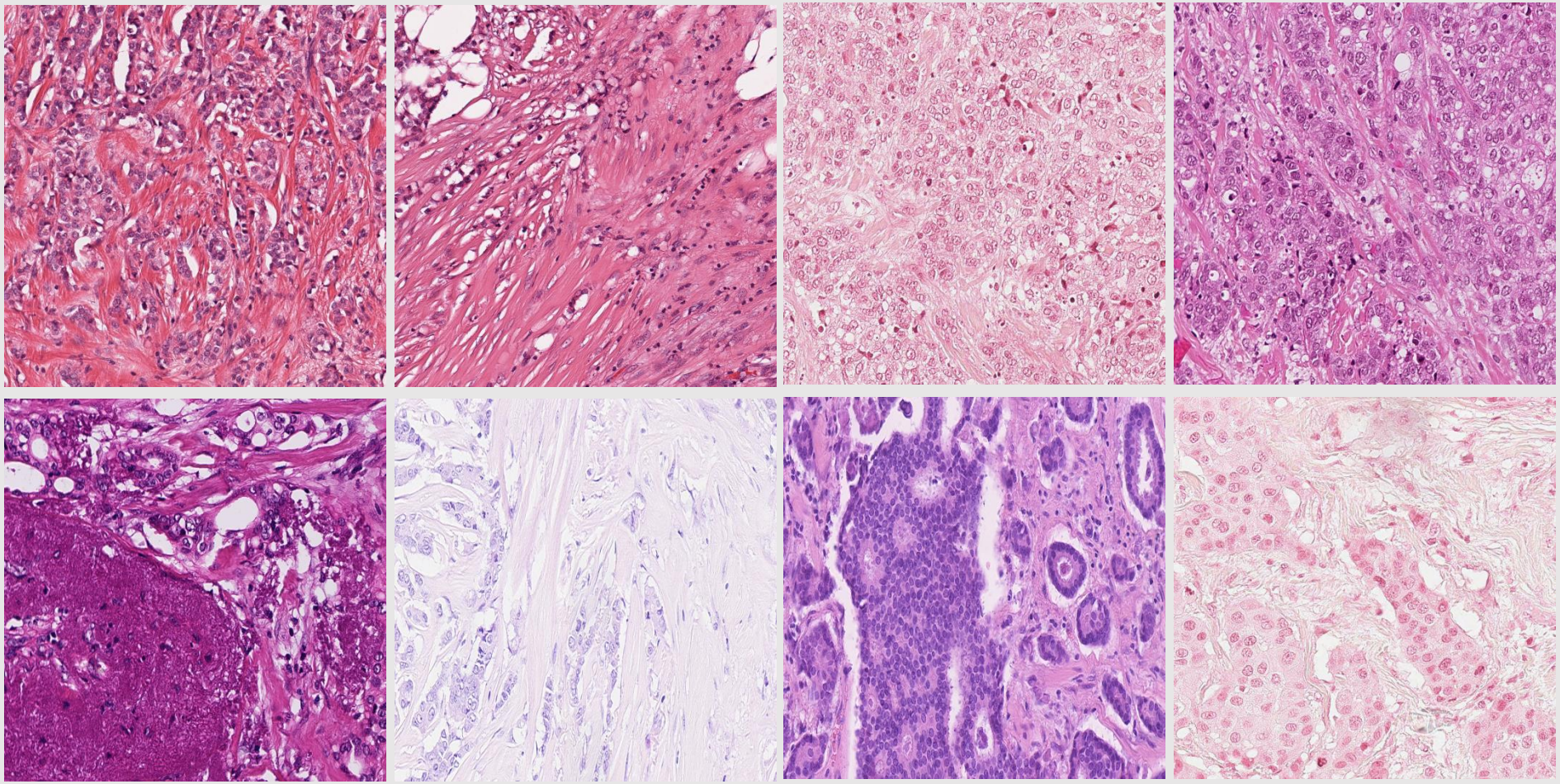
Stain Normalization

The **problem** of ‘normalizing’ the stain colour distribution according to some ‘reference’ image

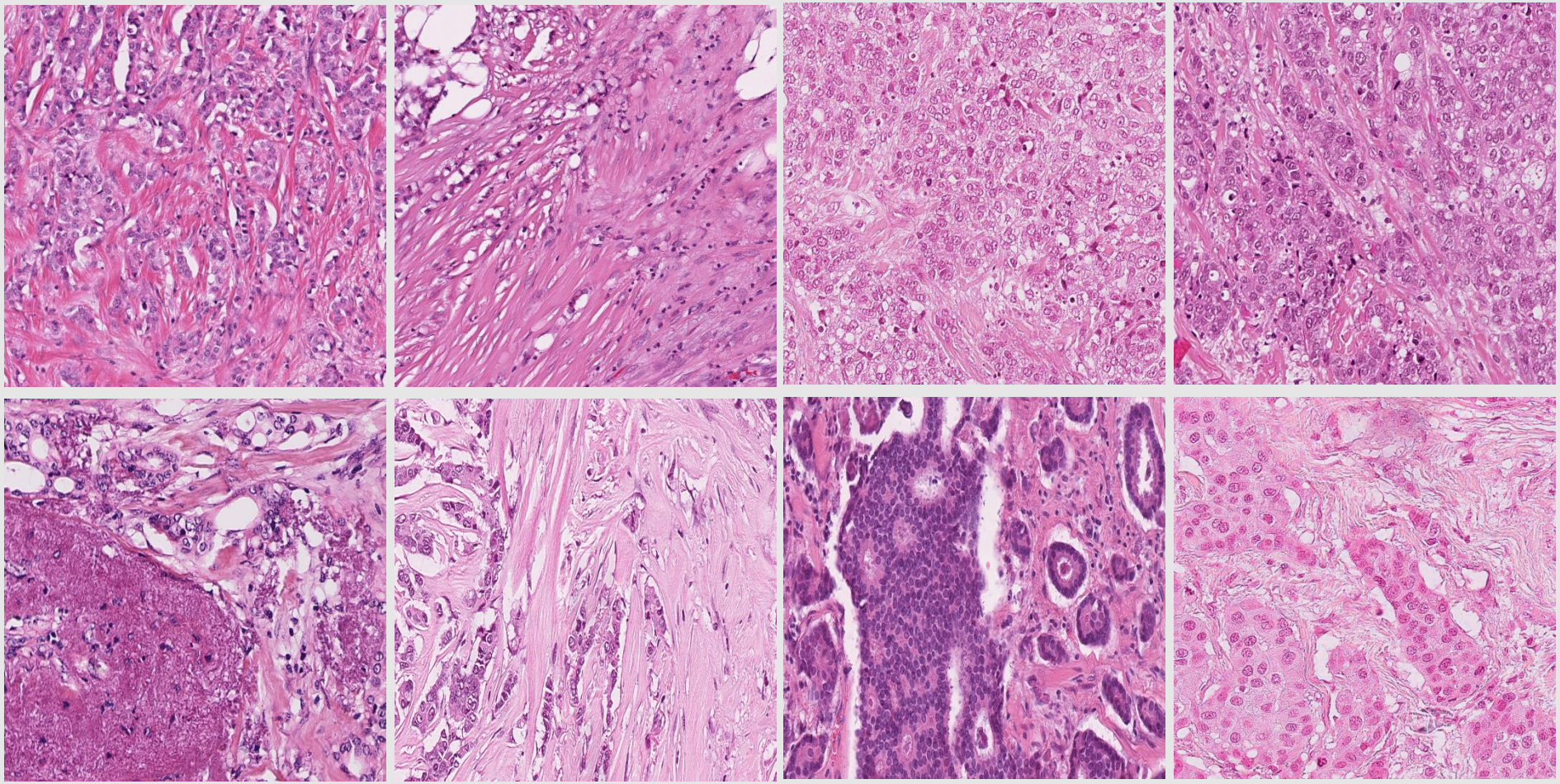


Reference Source Normalized

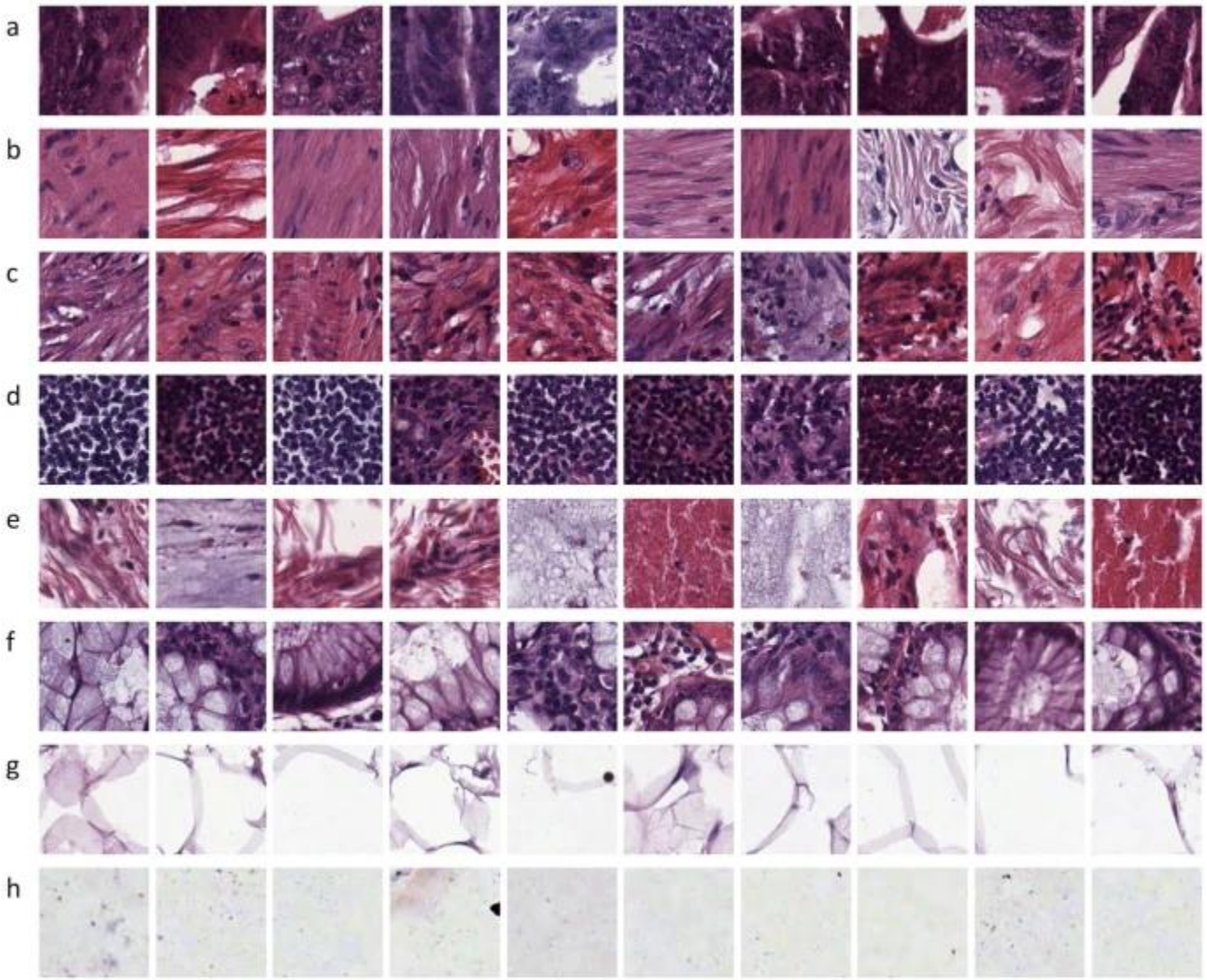
Stain Normalization



Stain Normalization



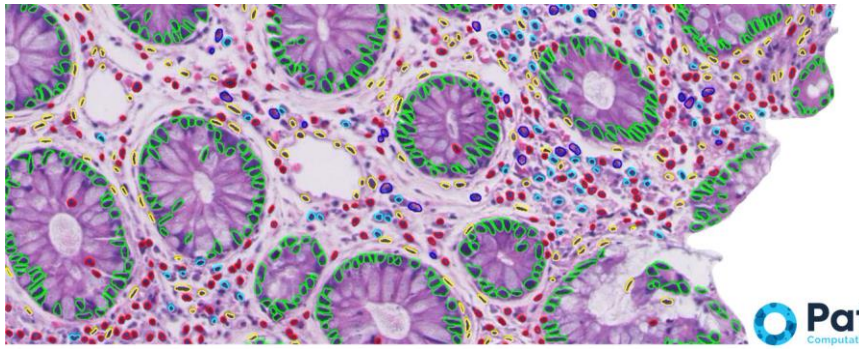
Classification



(a) tumour epithelium, (b) simple stroma, (c) complex stroma (stroma that contains single tumour cells and/or single immune cells), (d) immune cell conglomerates, (e) debris and mucus, (f) mucosal glands, (g) adipose tissue, (h) background.
(Kather et al. 2016)

Regression

- Cellular Composition Prediction: Counting how many different types of cells in a given patch



CoNiC Challenge 2022

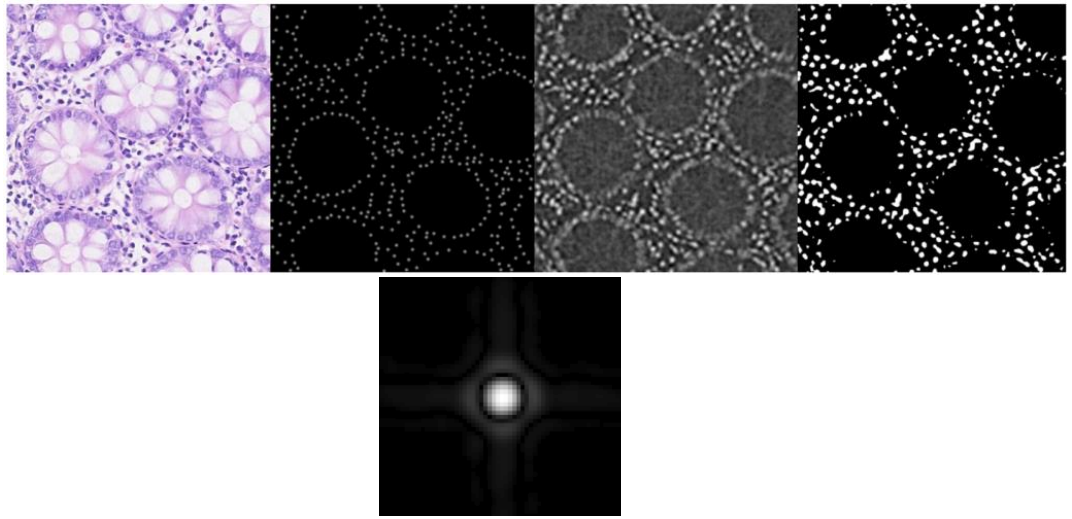
PathLAKE Computational Pathology Excellence · TIA Centre Tissue Image Analytics · NVIDIA

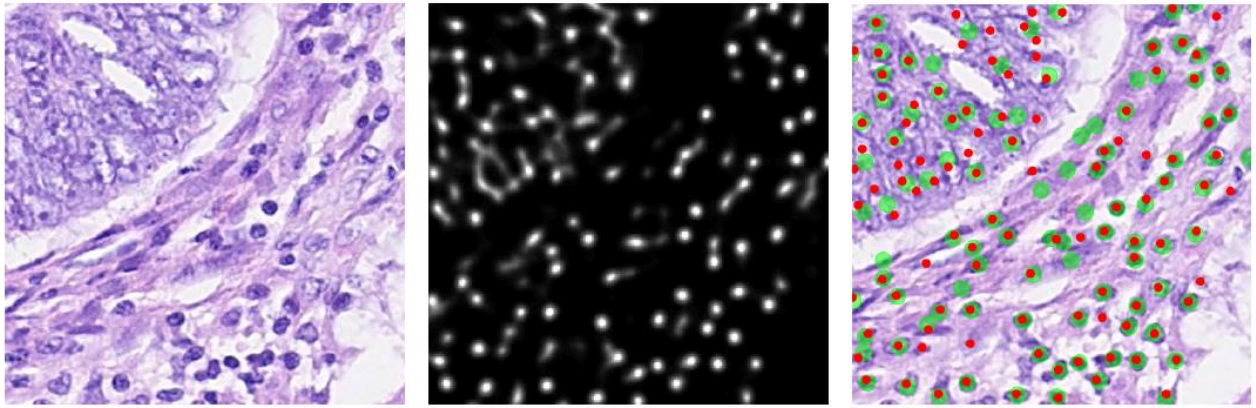
Cell type/ Method	Image 1				Image 2				Image 3				Image 4			
Neoplastic																
Inflammatory																
Connective																
Epithelial																
	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Ground Truth	0	0	4	23	85	0	6	0	0	65	7	0	7	14	3	0
HoVer-Net	1	**1**	**4**	16	93	**0**	2	**0**	**0**	7	0	71	20	**11**	1	**0**
ALBRT	**0**	**1**	5	**23**	**86**	**0**	**5**	**0**	**0**	**64**	**3**	**0**	**10**	**11**	**4**	3

Graham, Simon, Mostafa Jahanifar, Quoc Dang Vu, Giorgos Hadjigeorghiou, Thomas Leech, David Snead, Shan E. Ahmed Raza, Fayyaz Minhas, and Nasir Rajpoot. "CoNIC: Colon Nuclei Identification and Counting Challenge 2022." ArXiv:2111.14485 [Cs], November 29, 2021. http://arxiv.org/abs/2111.14485.

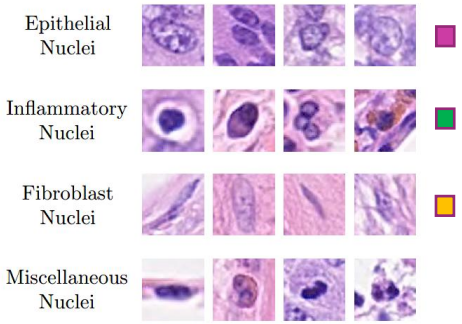
Graham, Simon, Mostafa Jahanifar, Ayesha Azam, Mohammed Nimir, Yee-Wah Tsang, Katherine Dodd, Emily Hero, et al. "**Lizard**: A Large-Scale Dataset for Colonic Nuclear Instance Segmentation and Classification." ArXiv:2108.11195 [Cs], August 25, 2021. http://arxiv.org/abs/2108.11195.

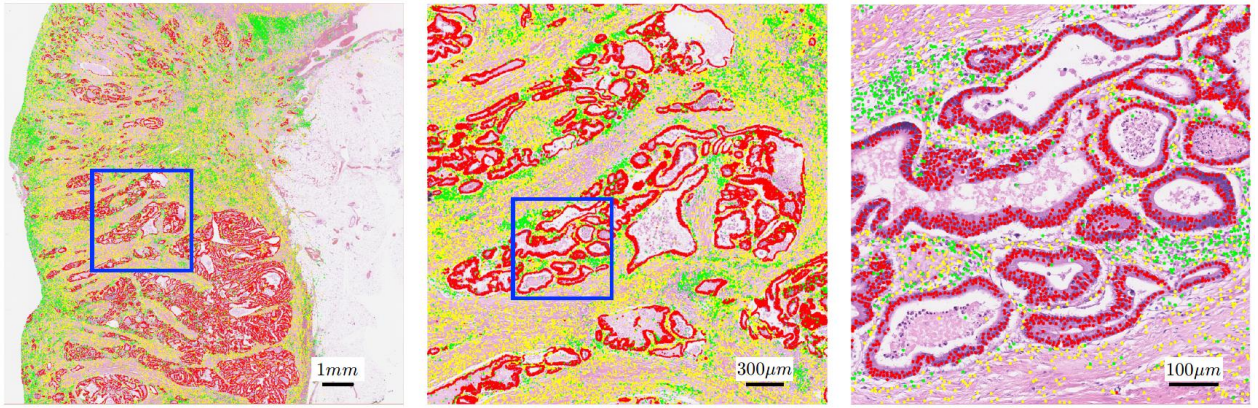
Object Detection and Classification





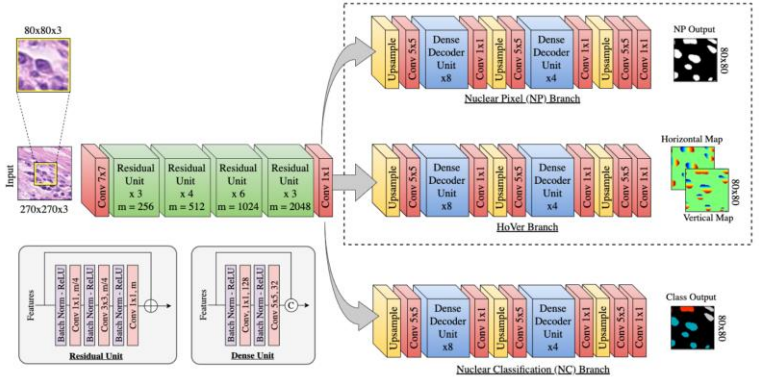
Detection approach	Precision	Recall	F1 score
Baseline	0.45	0.74	0.55
RBF Correlation Filter	**0.83**	**0.86**	**0.84**
Linear MOSSE Filter	**0.76**	**0.88**	**0.81**
SC-CNN (M = 1) [3]	0.76	0.83	0.79
SC-CNN (M = 2) [3]	0.78	0.82	0.80
CP-CNN [3]	0.70	0.69	0.69
SR-CNN [22]	0.78	0.80	0.79
SSAE [18]	0.62	0.64	0.63
LIPSyM [48]	0.72	0.52	0.60
CRImage [9]	0.66	0.46	0.54

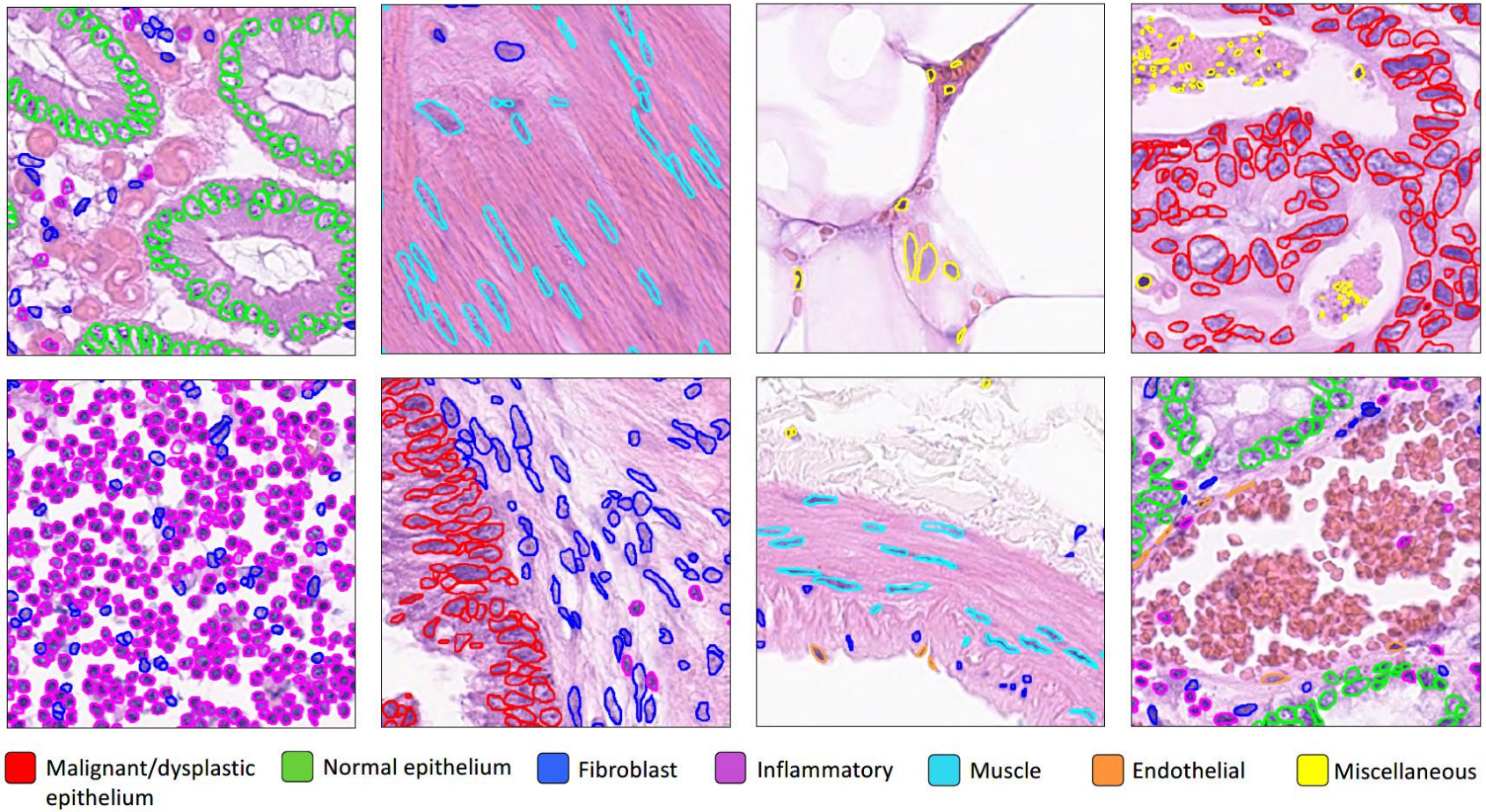




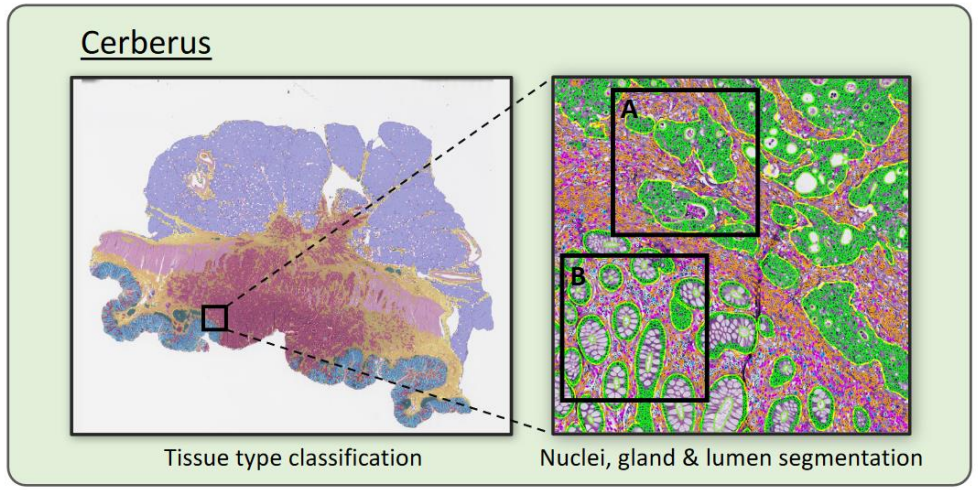
Sirinukunwattana, Korsuk, Shan E. Ahmed Raza, null Yee-Wah Tsang, David R. J. Snead, Ian A. Cree, and Nasir M. Rajpoot. "Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images." *IEEE Transactions on Medical Imaging* 35, no. 5 (May 2016): 1196–1206. https://doi.org/10.1109/TMI.2016.2525803.
Ahmad, Asif, Amina Asif, Nasir Rajpoot, Muhammad Arif, and Fayyaz Ul Amir Afsar Minhas. "Correlation Filters for Detection of Cellular Nuclei in Histopathology Images." *Journal of Medical Systems* 42, no. 1 (November 21, 2017): 7. https://doi.org/10.1007/s10916-017-0863-8.

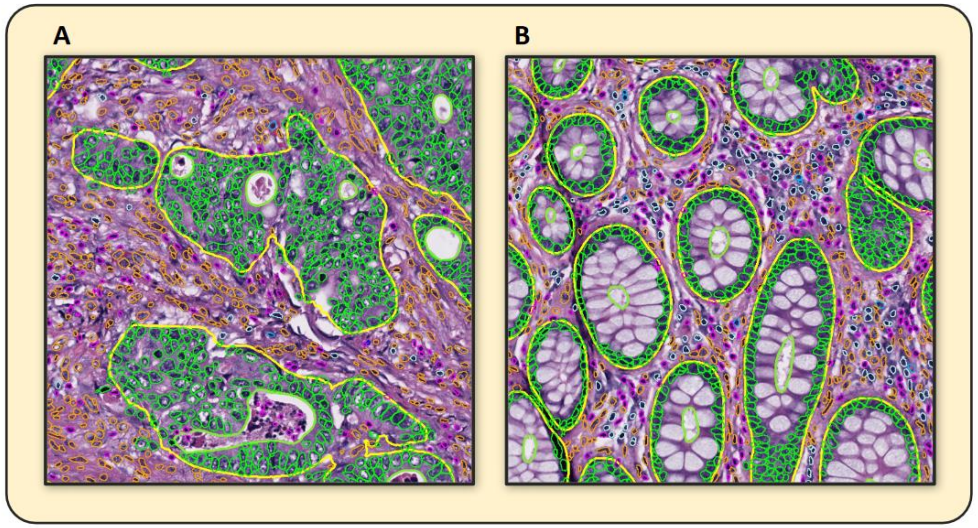
Nuclear Segmentation





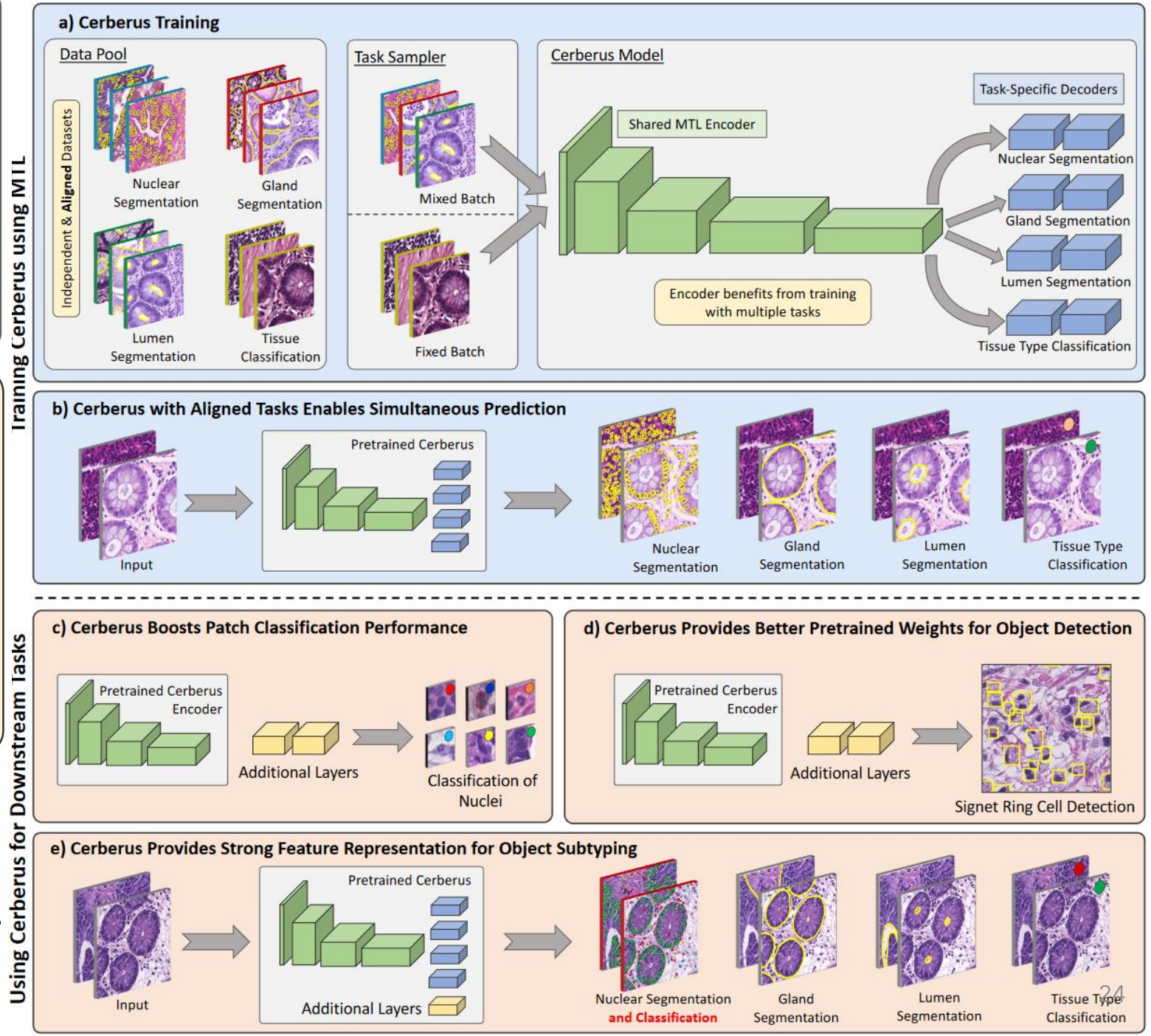
Graham, Simon, *et al.*, "Hover-Net: Simultaneous Segmentation and Classification of Nuclei in Multi-Tissue Histology Images." *Medical Image Analysis* 58 (December 1, 2019): 101563. https://doi.org/10.1016/j.media.2019.101563.





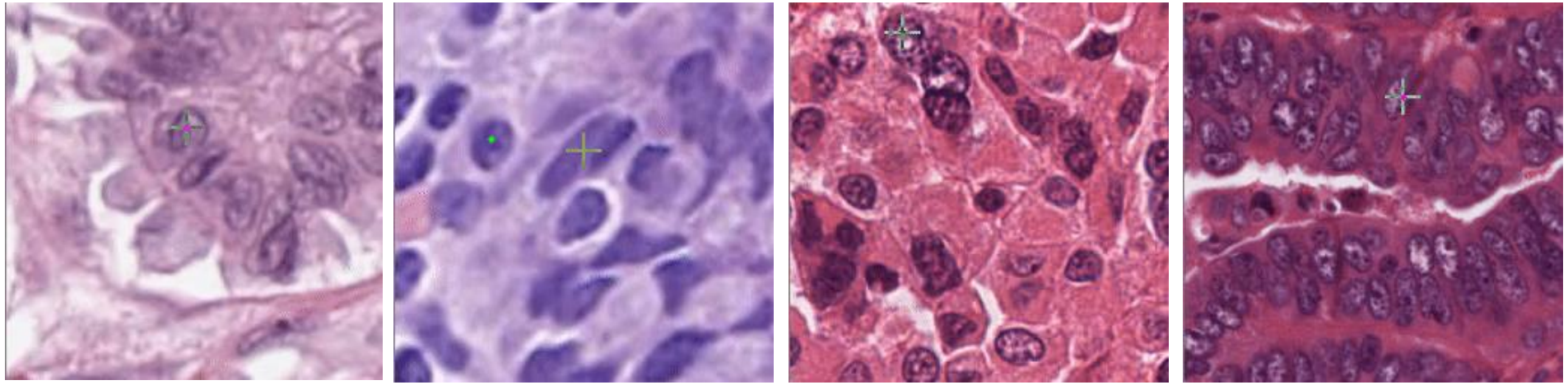
Cerberus

Graham, Simon, Quoc Vu, Mostafa Jahanifar, Fayyaz ul Amir Afsar Minhas, David Snead, and Nasir Rajpoot. One Model Is All You Need: Multi-Task Learning Enables Simultaneous Histology Image Segmentation and Classification, 2022. URL: https://arxiv.org/abs/2203.00077



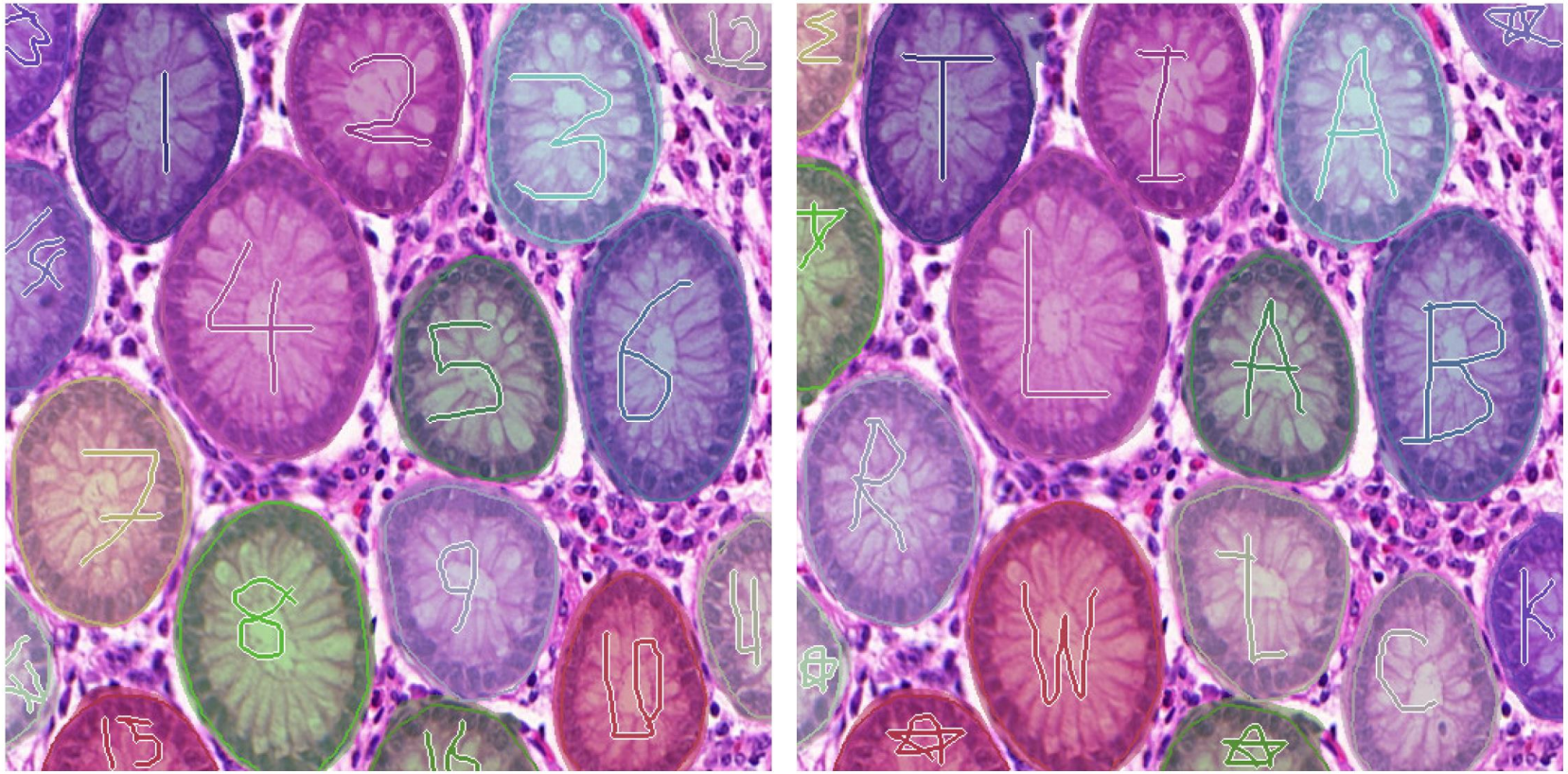
AI-Assisted Annotations

NuClick used for collecting the segmentation masks from point clicks



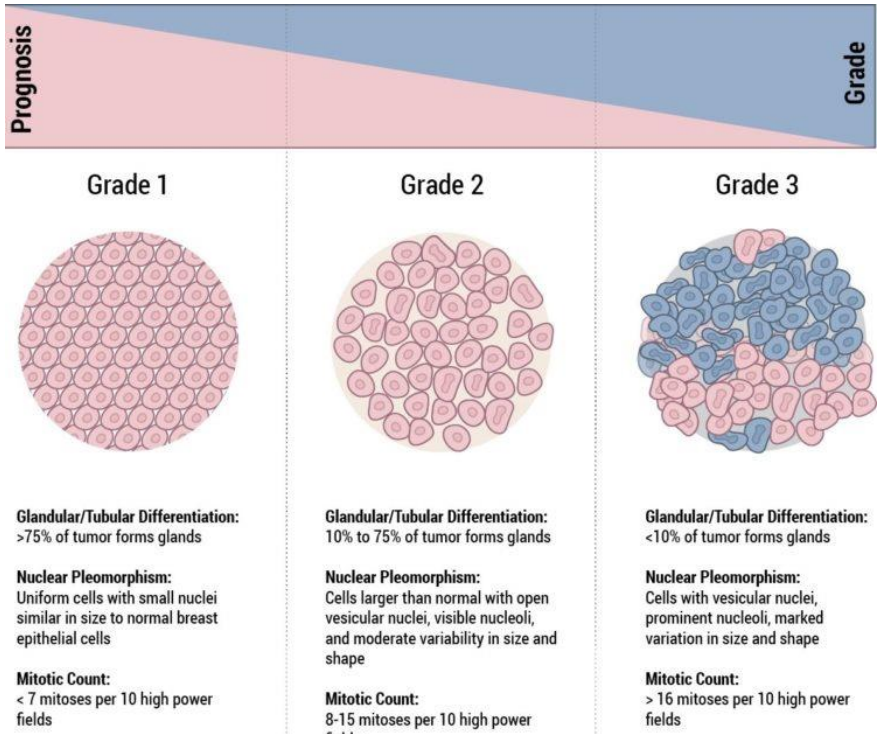
Koohbanani *et al*., **Medical Image Analysis** (Oct 2020) & Gamper et al., Arxiv (Apr 2020)

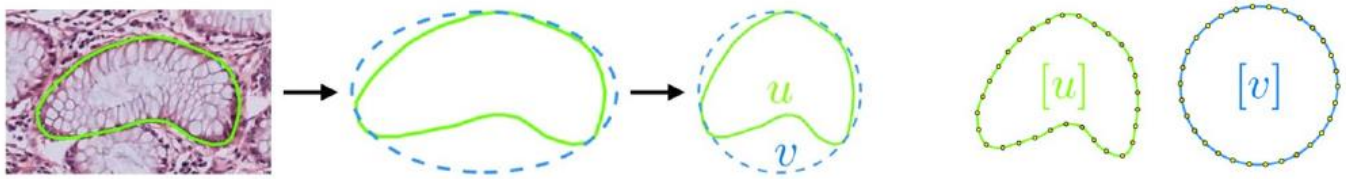
NuClick+: AI-Assisted Annotation of Glands

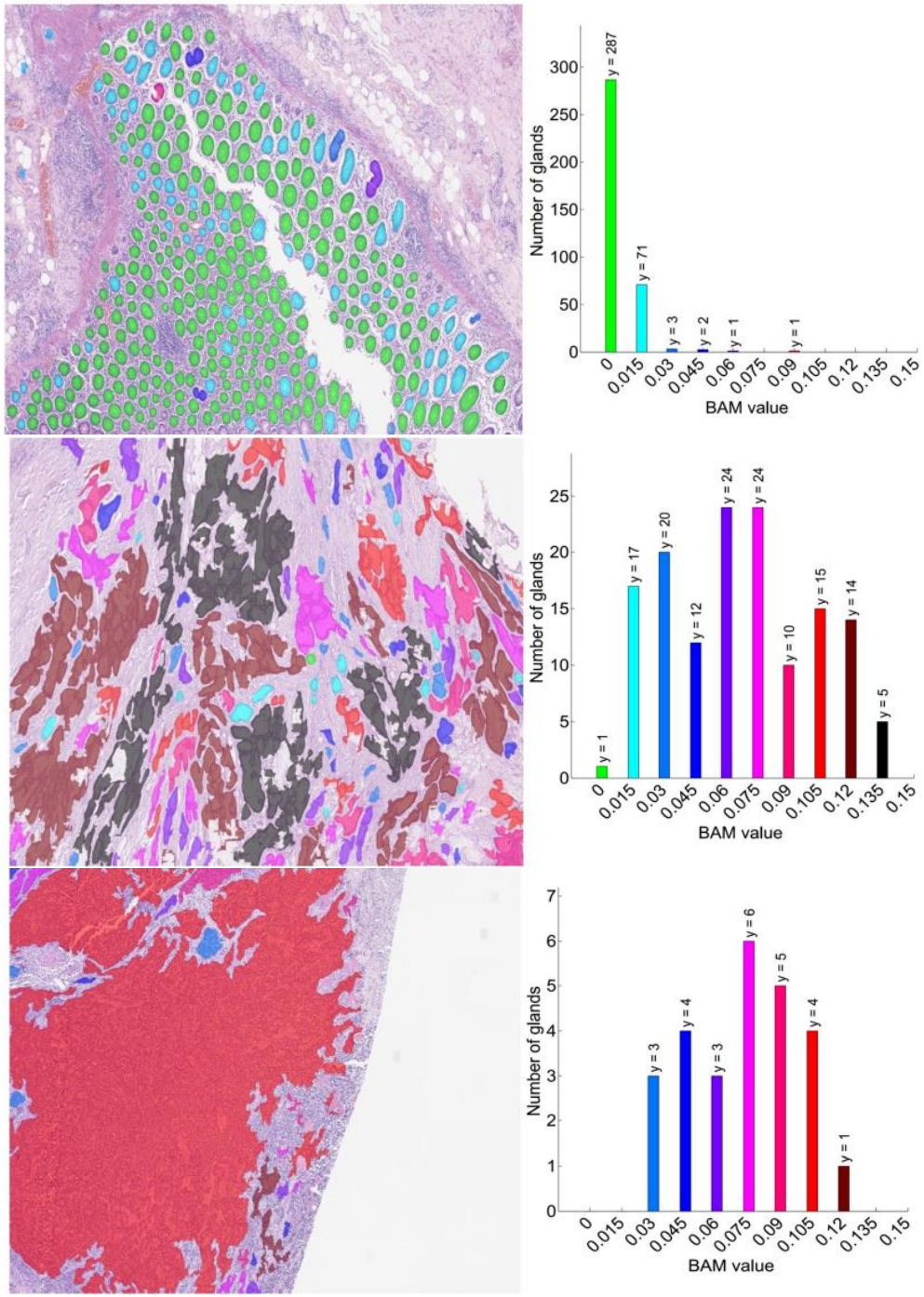


Koohbanani *et al.*, **Medical Image Analysis** (Oct 2020)

Grading problems



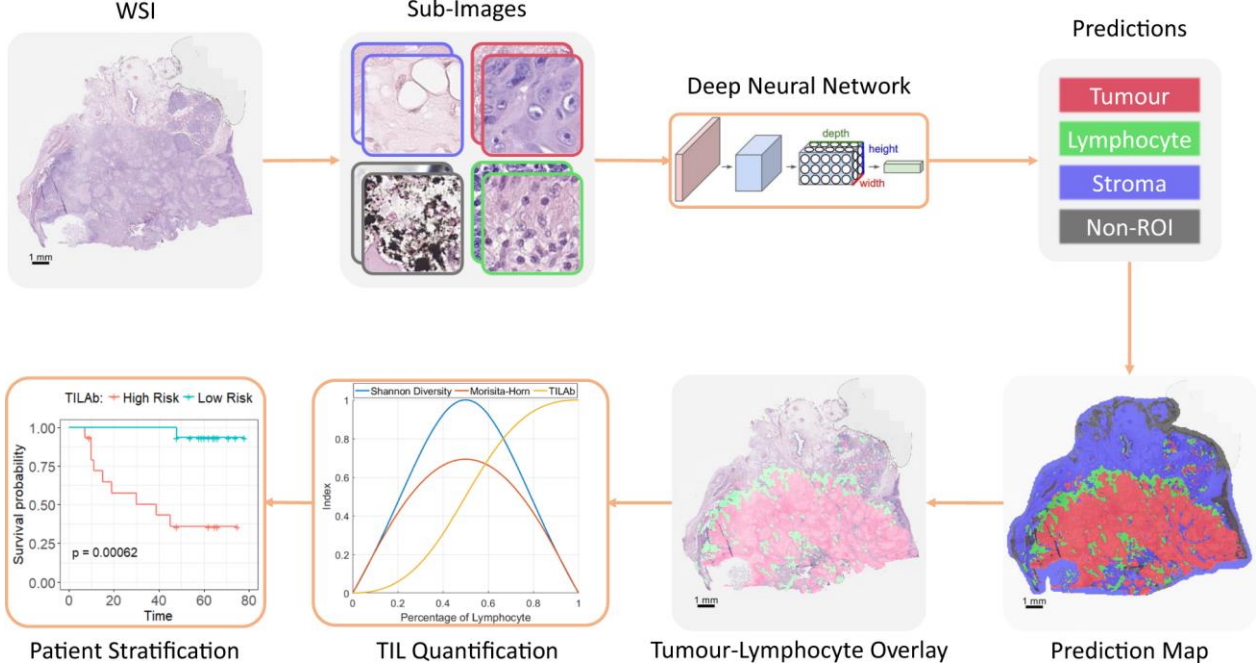


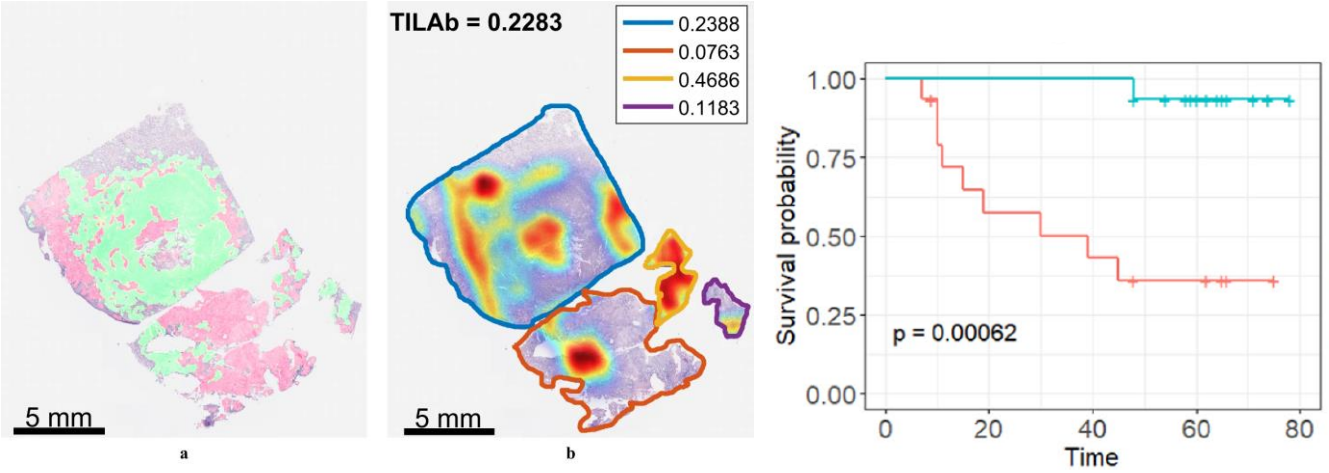


Awan, Ruqayya, Korsuk Sirinukunwattana, David Epstein, Samuel Jefferyes, Uvais Qidwai, Zia Aftab, Imaad Mujeeb, David Snead, and Nasir Rajpoot. "Glandular Morphometrics for Objective Grading of Colorectal Adenocarcinoma Histology Images." *Scientific Reports* 7, no. 1 (December 4, 2017): 16852. https://doi.org/10.1038/s41598-017-16516-w.

Survival Prediction

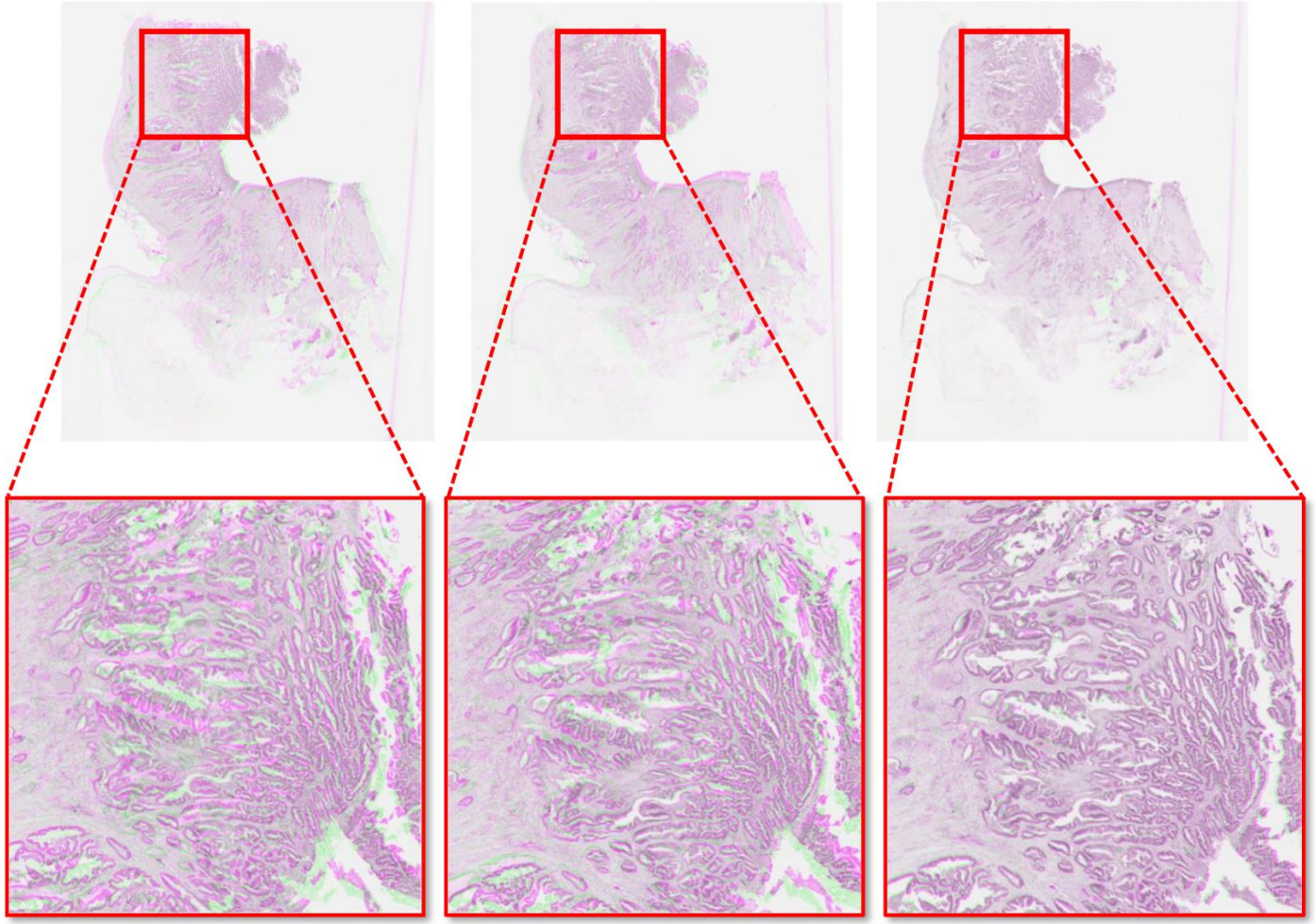
- Given
  - A set of patients with their covariates and censored survival/recurrence data
- Predict
  - Survival curves and/or identify patterns associated with survival





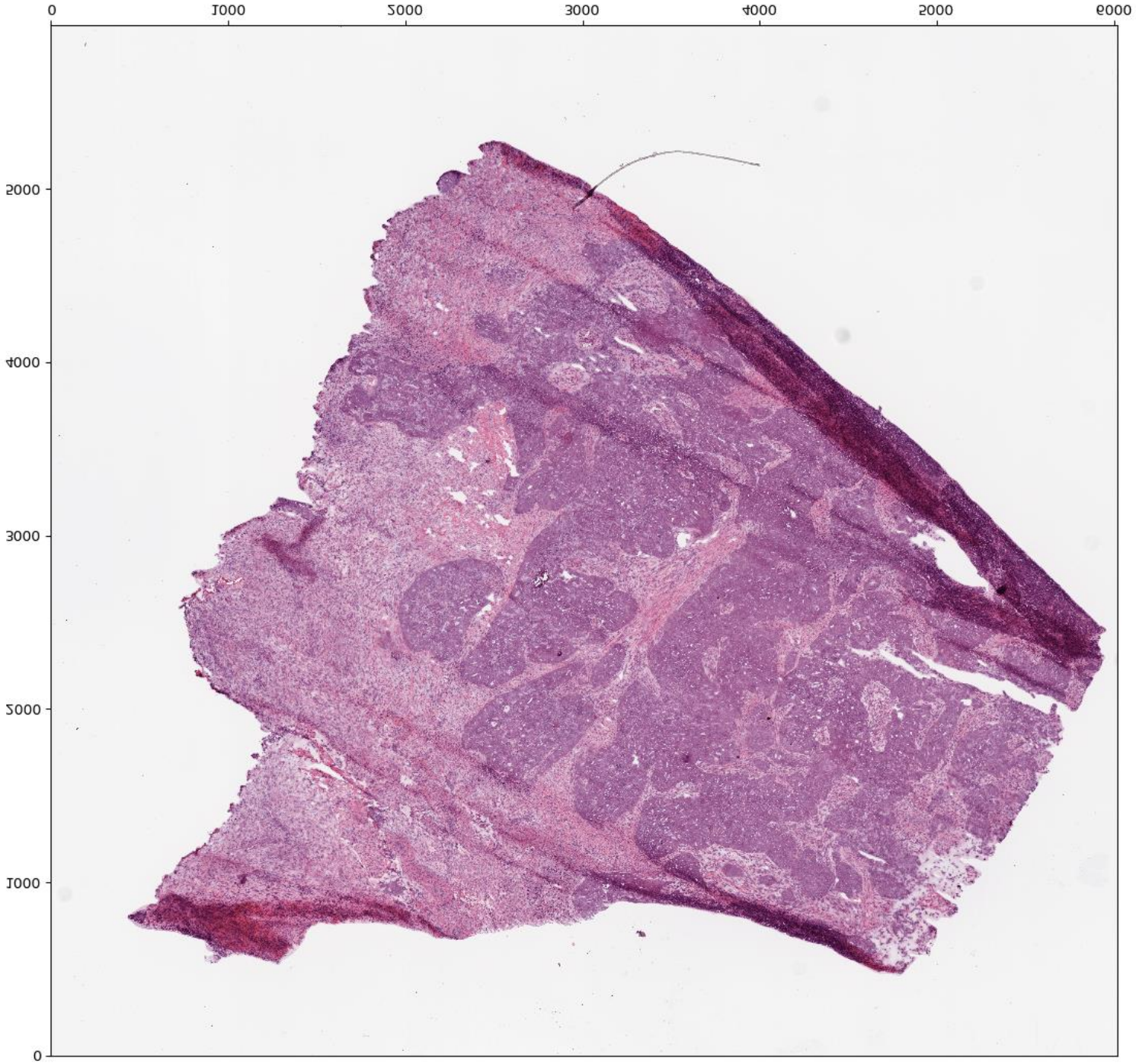
Shaban, Muhammad, Syed Ali Khurram, Muhammad Moazam Fraz, Najah Alsubaie, Iqra Masood, Sajid Mushtaq, Mariam Hassan, Asif Loya, and Nasir M. Rajpoot. "A Novel Digital Score for Abundance of Tumour Infiltrating Lymphocytes Predicts Disease Free Survival in Oral Squamous Cell Carcinoma." *Scientific Reports* 9, no. 1 (September 16, 2019): 13341. https://doi.org/10.1038/s41598-019-49710-z.
https://github.com/TIA-Lab/TILAb-Score

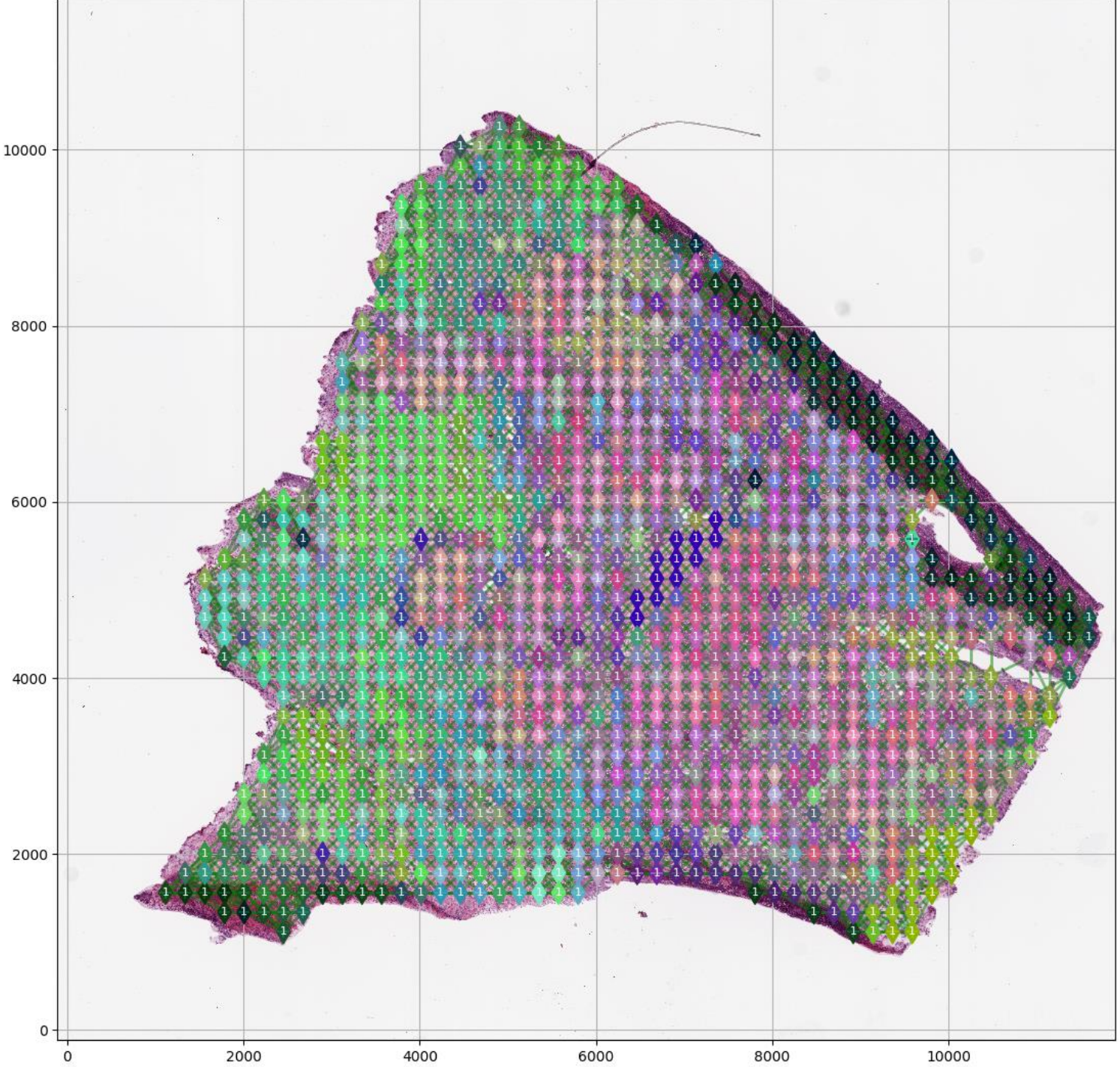
Image Registration

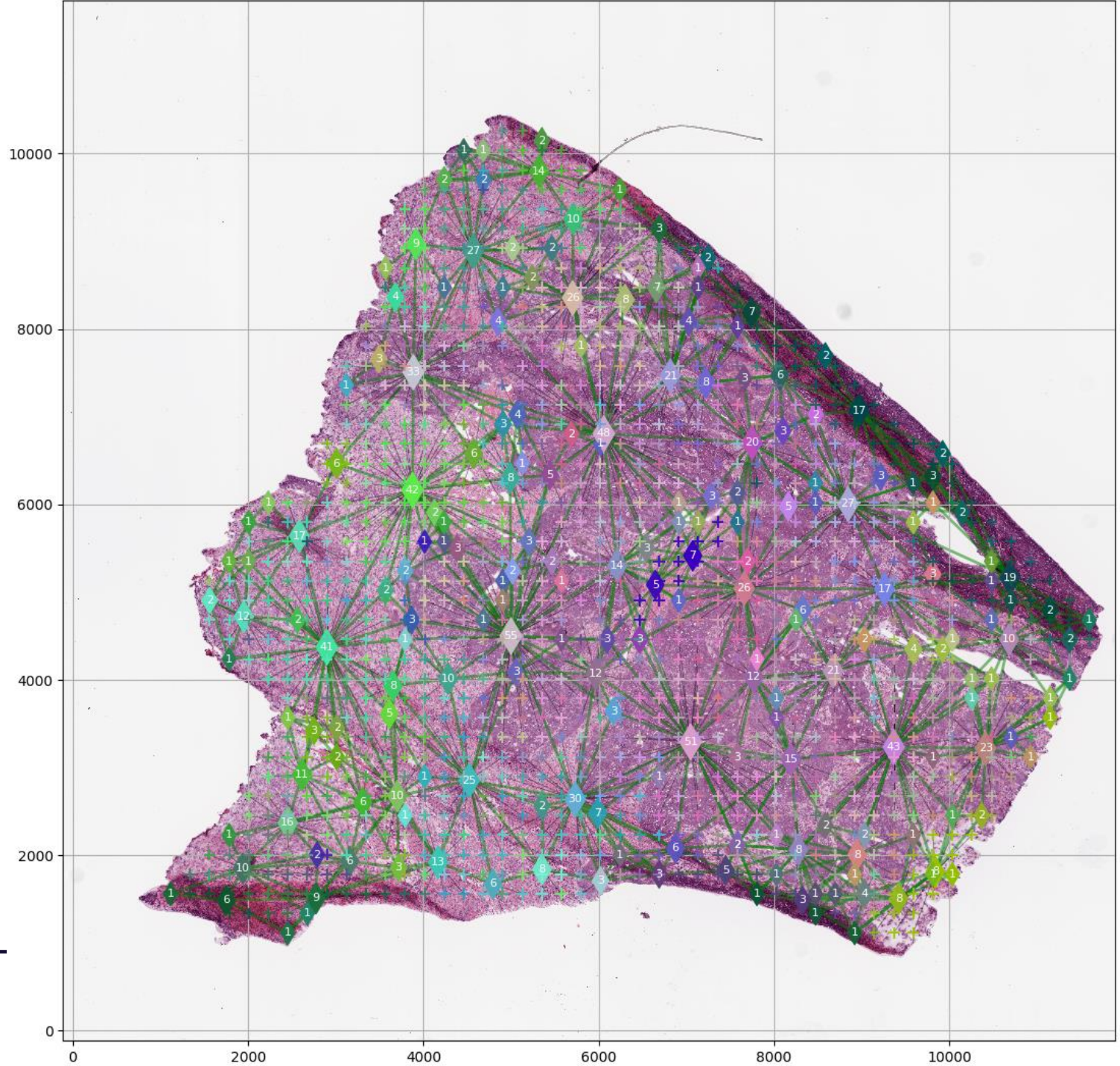


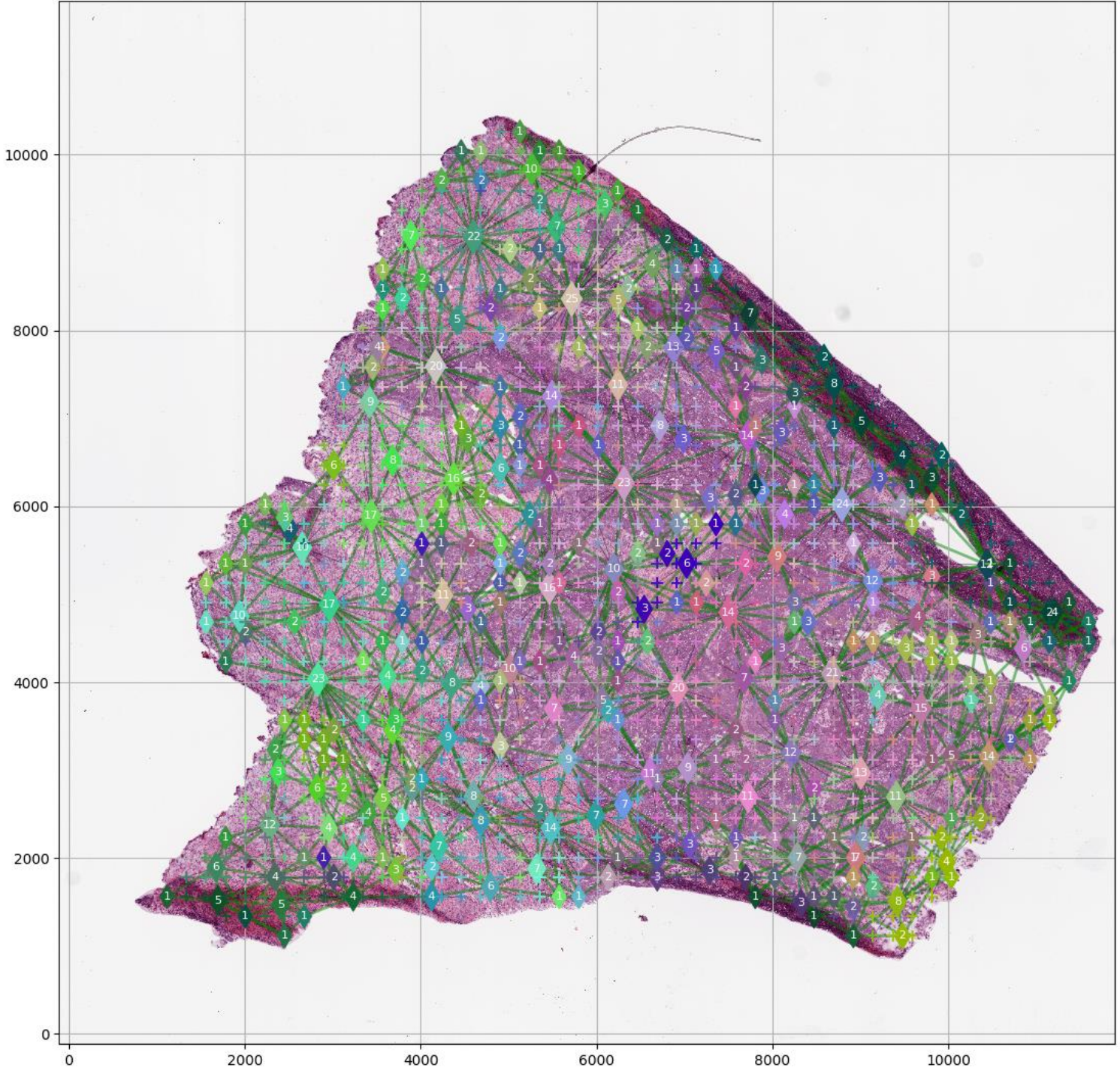
Awan, Ruqayya, Shan E. Ahmed Raza, Johannes Lotz, Nick Weiss, and Nasir M. Rajpoot. “Deep Feature Based Cross-Slide Registration.” ArXiv:2202.09971 [Cs, Eess], February 27, 2022. http://arxiv.org/abs/2202.09971.

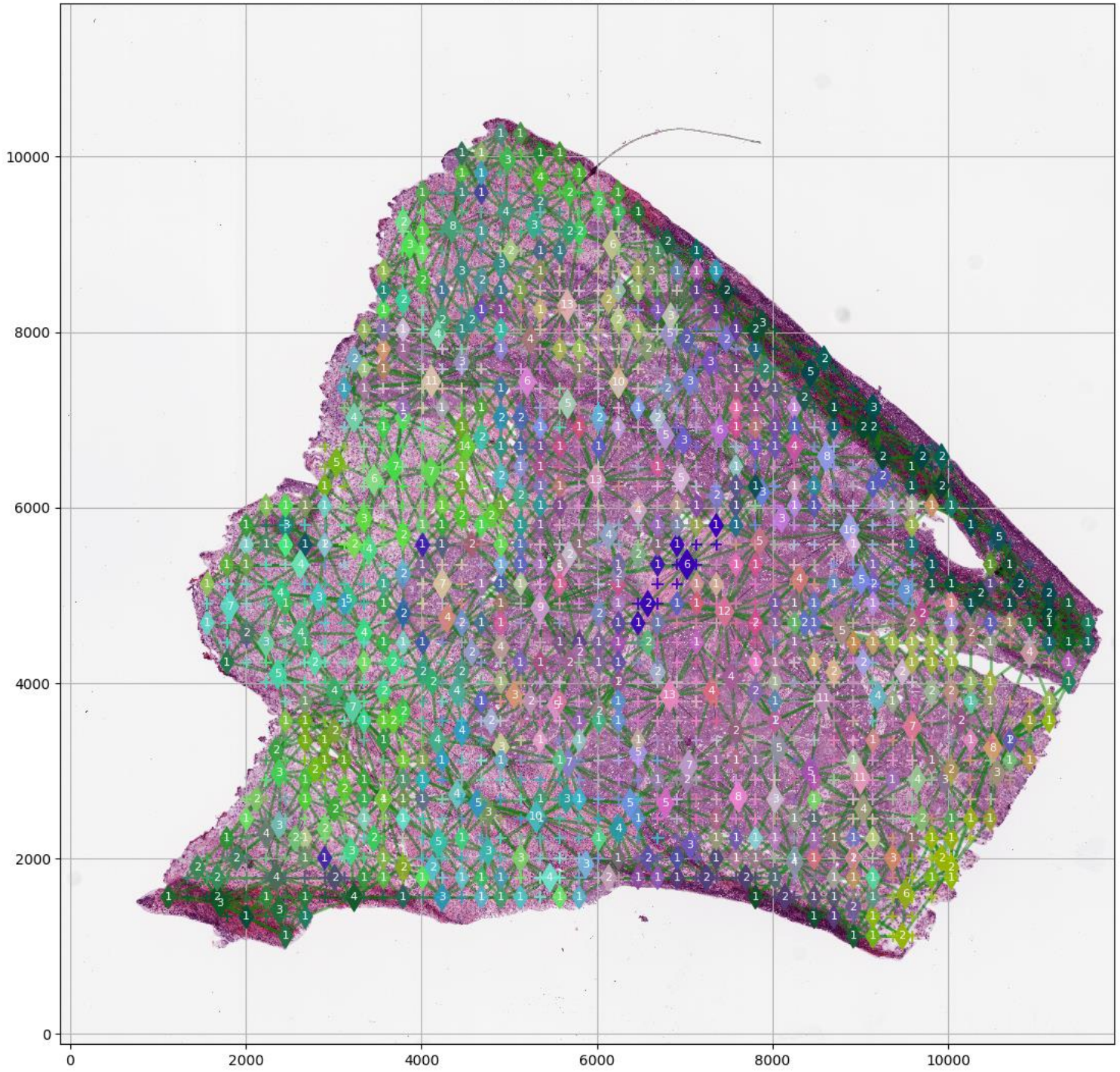
Clustering: Unsupervised Learning



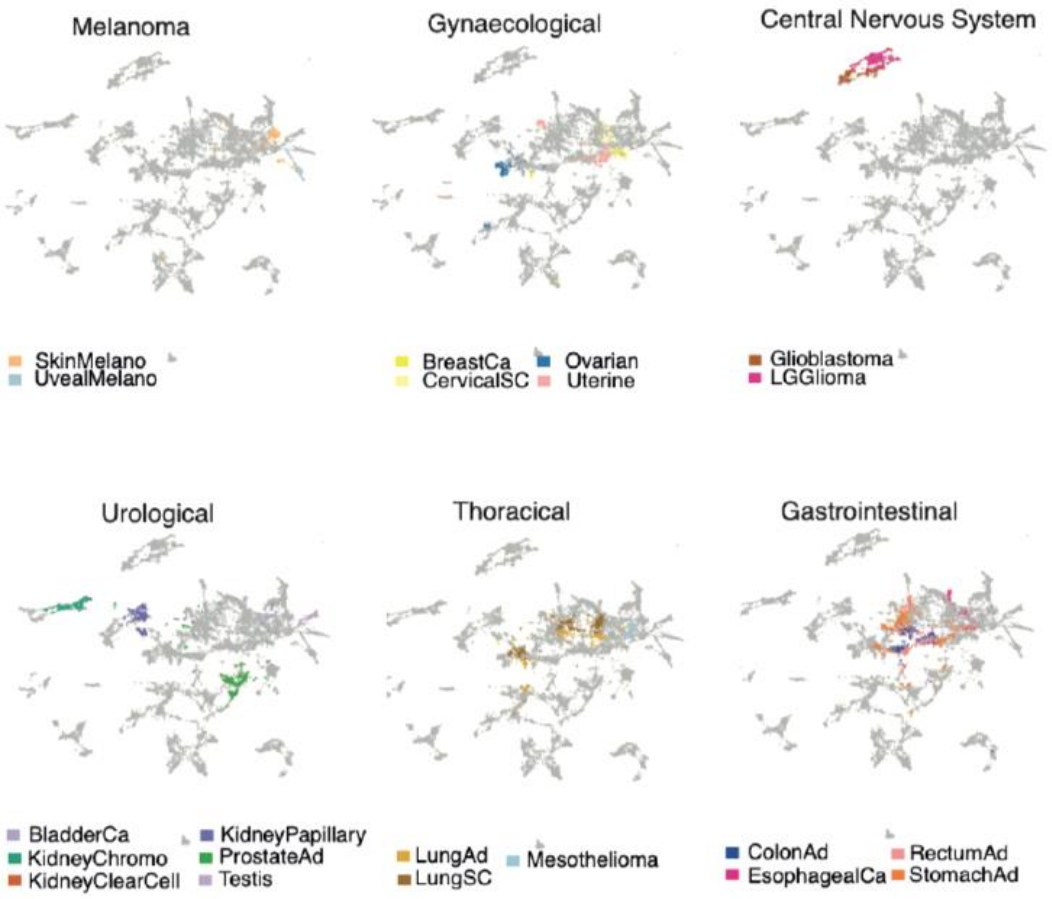

10000
8000
6000
4000
2000
0
0
2000
4000
6000
8000
10000

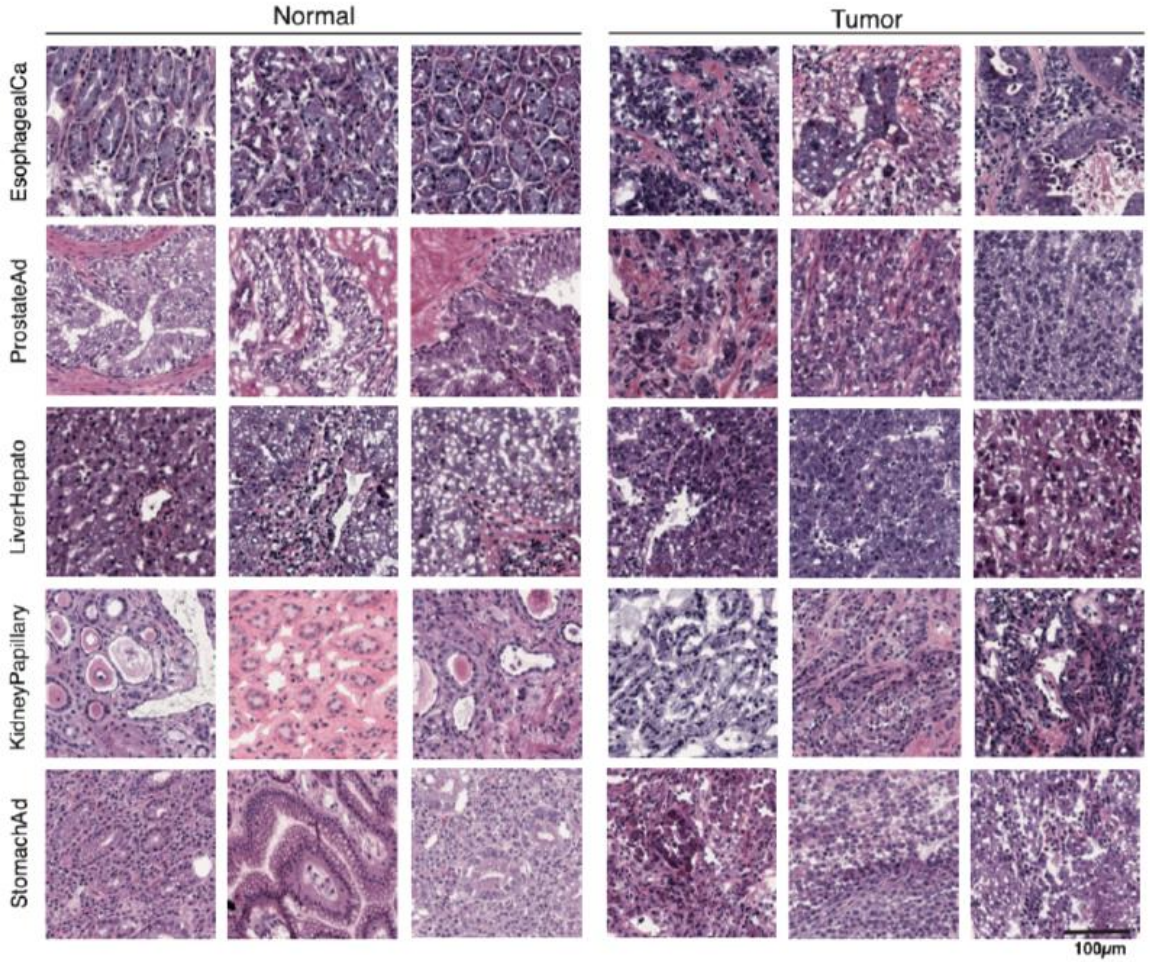

10000
8000
6000
4000
2000
0
0
2000
4000
6000
8000
10000


10000
8000
6000
4000
2000
0
0
2000
4000
6000
8000
10000

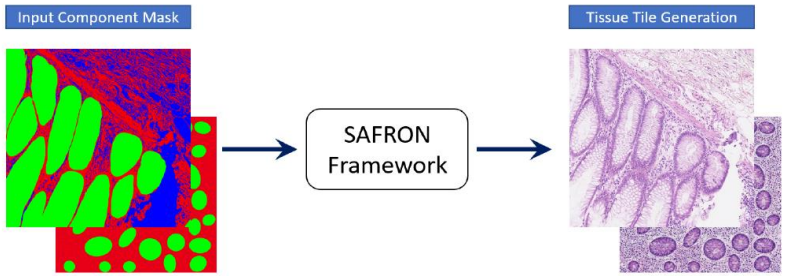

10000
8000
6000
4000
2000
0
0
2000
4000
6000
8000
10000

Discriminative vs Generative Modelling





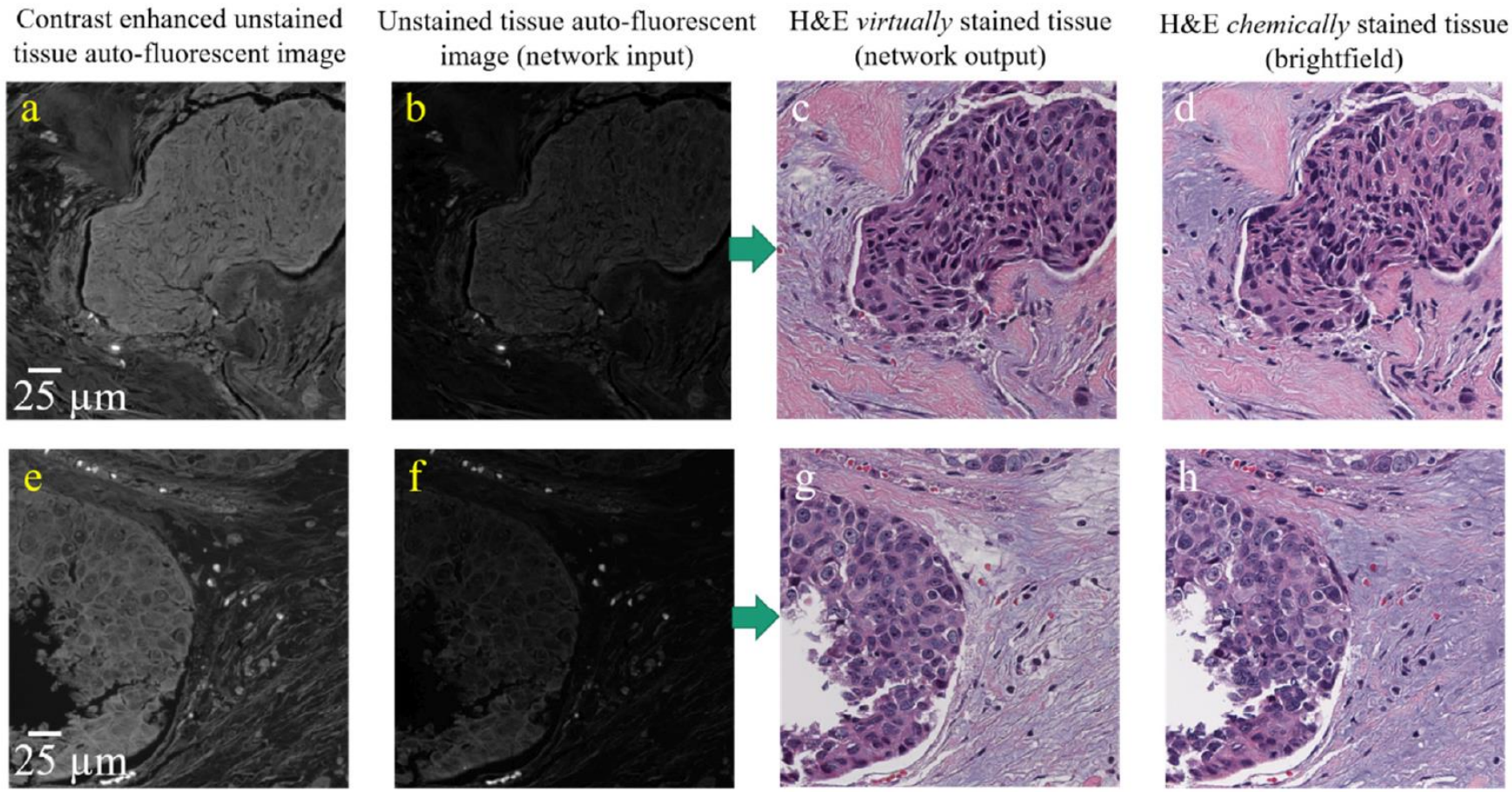
Fu, Yu, Alexander W. Jung, Ramon Viñas Torne, Santiago Gonzalez, Harald Vöhringer, Artem Shmatko, Lucy R. Yates, Mercedes Jimenez-Linan, Luiza Moore, and Moritz Gerstung. "Pan-Cancer Computational Histopathology Reveals Mutations, Tumor Composition and Prognosis." *Nature Cancer* 1, no. 8 (August 2020): 800–810. https://doi.org/10.1038/s43018-020-0085-8.



6Kx5K

Deshpande, Srijay, Fayyaz Minhas, Simon Graham, and Nasir Rajpoot. “SAFRON: Stitching Across the Frontier for Generating Colorectal Cancer Histology Images.” *ArXiv:2008.04526 [Cs, Eess]*, March 26, 2021. http://arxiv.org/abs/2008.04526.

Virtual Staining



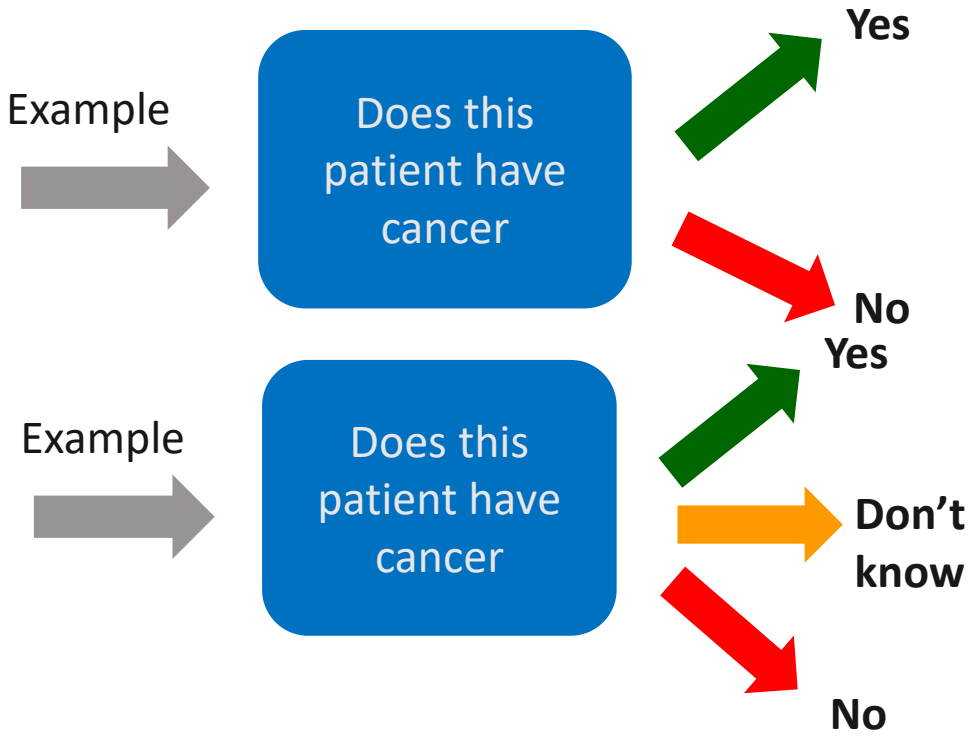
Rivenson, Yair, Hongda Wang, Zhensong Wei, Kevin de Haan, Yibo Zhang, Yichen Wu, Harun Günaydın, et al. "Virtual Histological Staining of Unlabelled Tissue-Autofluorescence Images via Deep Learning." *Nature Biomedical Engineering* 3, no. 6 (June 2019): 466–77. https://doi.org/10.1038/s41551-019-0362-y.

Advanced Machine Learning in CPath

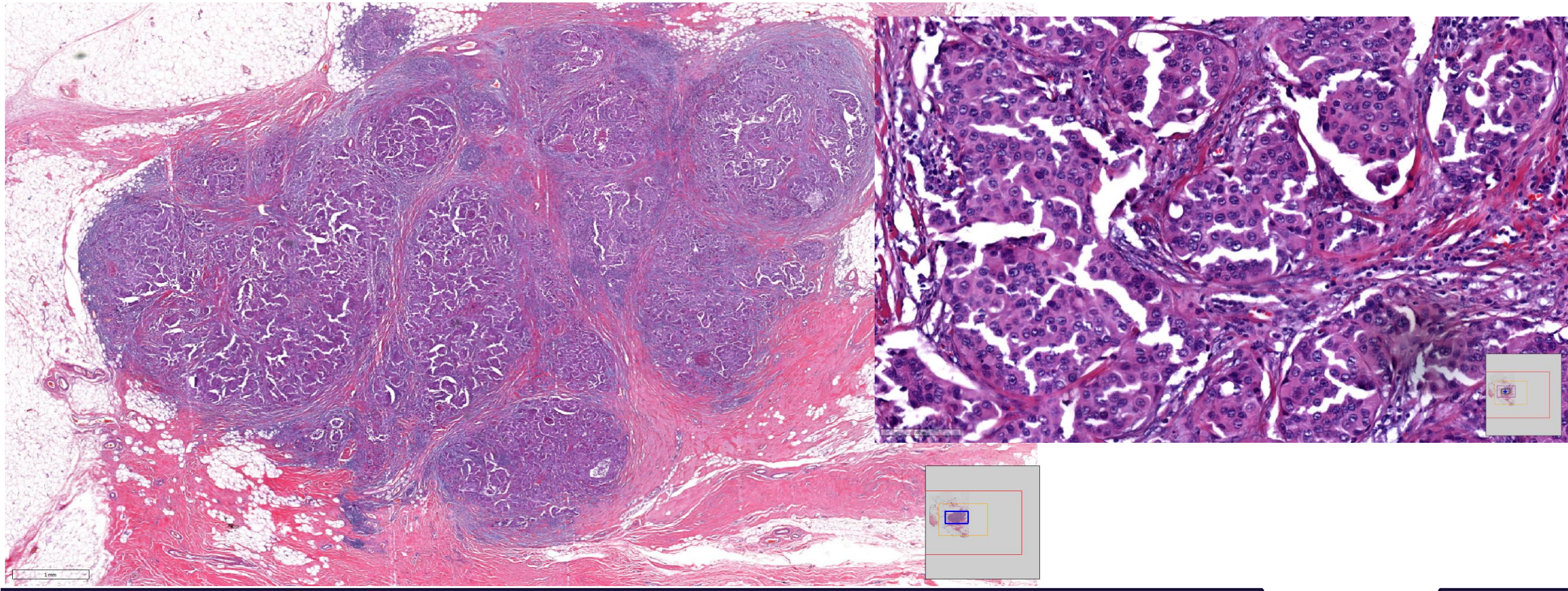
- Contrastive Representation Learning
- Active Learning
- Self-Supervised Learning
- Graph Neural Networks
- Community Detection
- Attention-based modelling
- Super-resolution imaging
- …

ML Technology in CPath

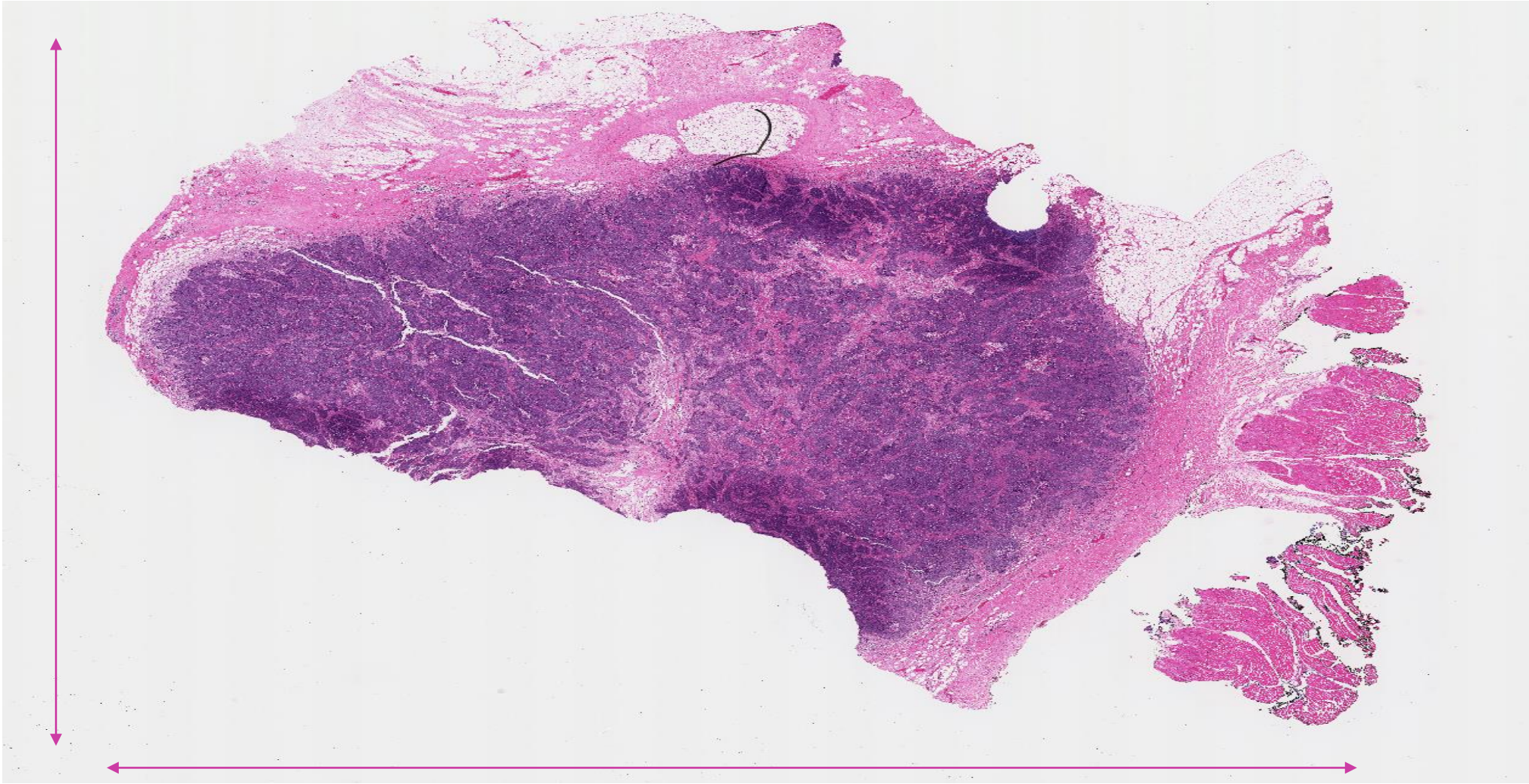
- Issues
  - Pre-processing and data quality management
  - Aggregation and information integration
  - Annotation Efficiency
  - Robustness
  - Explainability
  - Confidence prediction and "Rejection"
  - "Auto-ML" in Cpath
  - Generalization and reproducibility



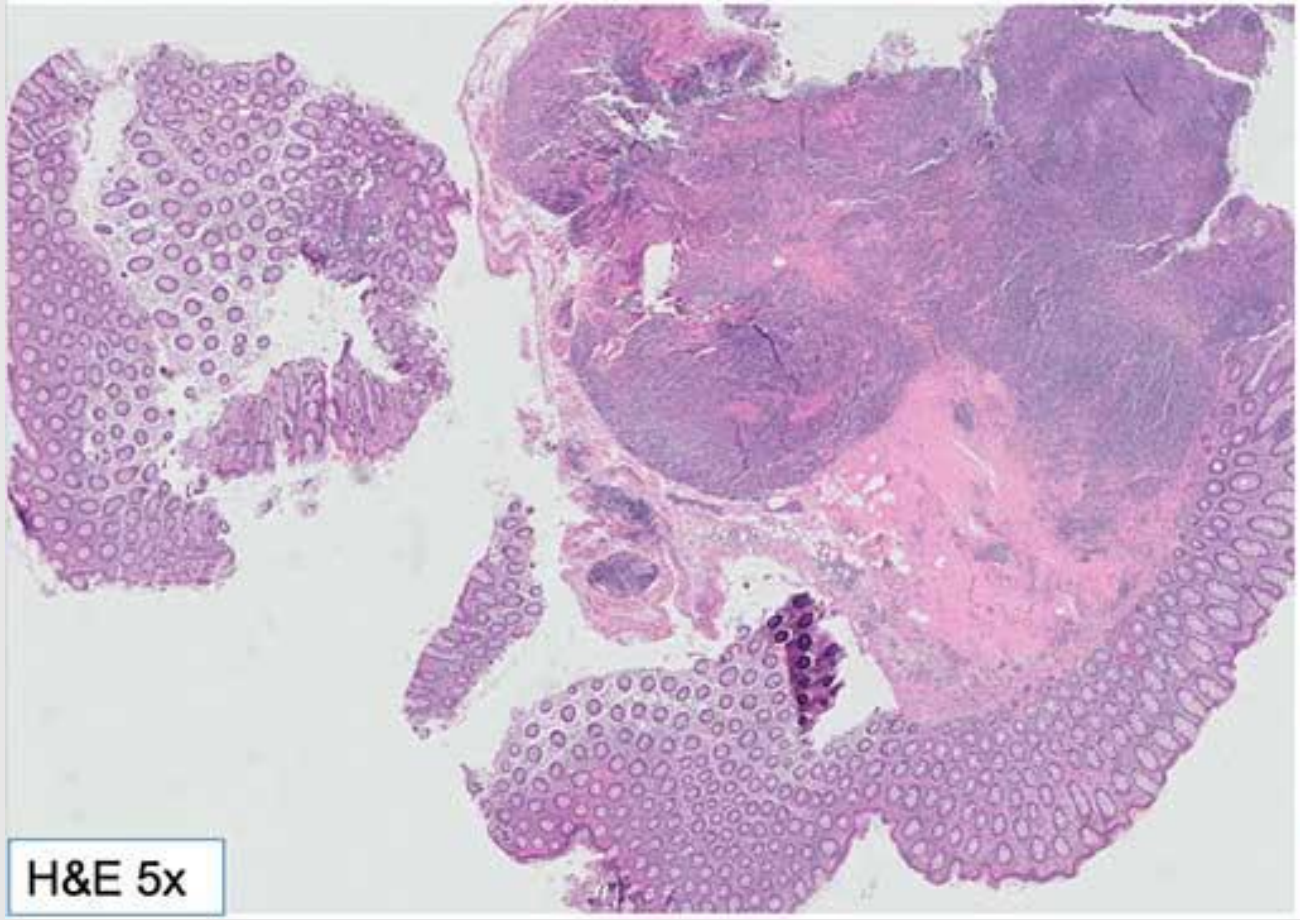
ML Problem(s) with WSIs: Multiresolution

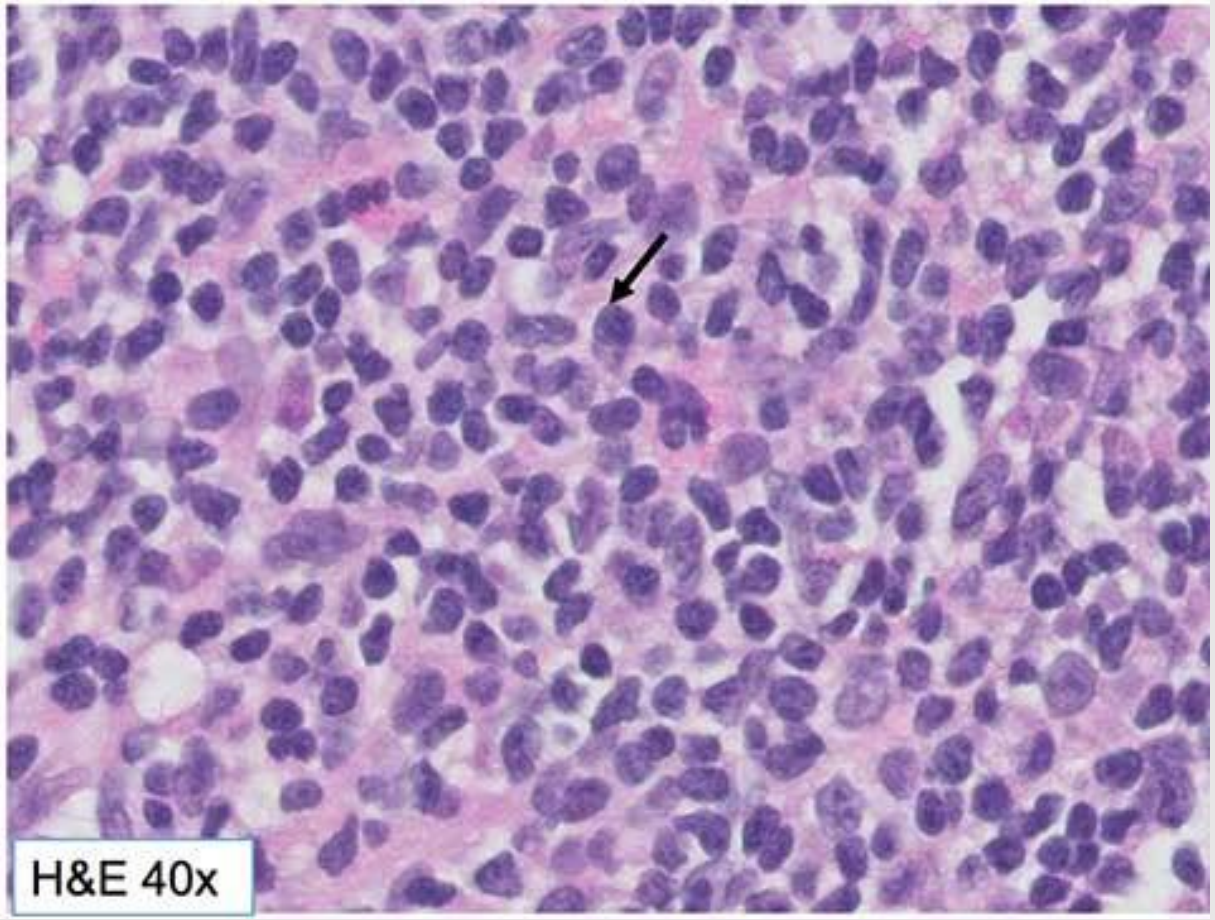


ML Problem(s) with WSIs: Size



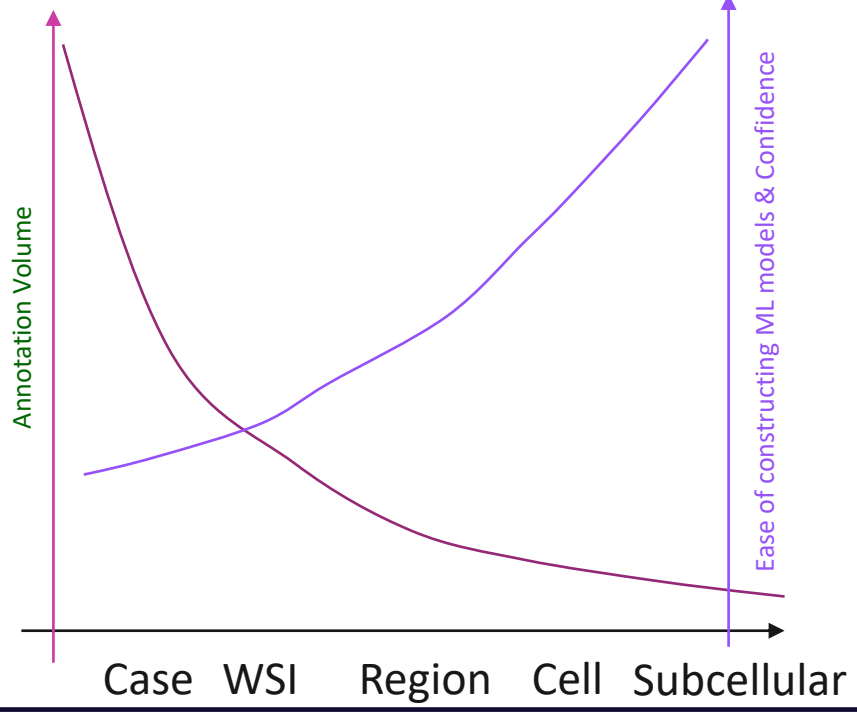
ML Problem(s) with WSIs: Spatial Context Matters

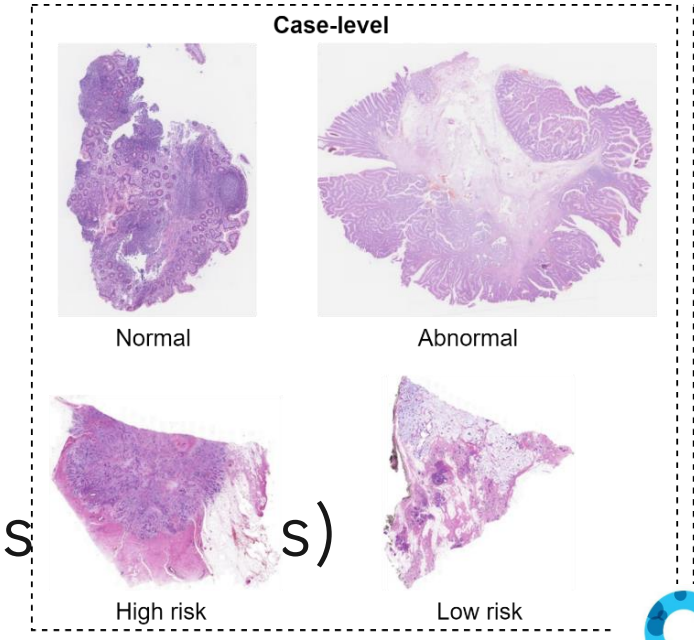


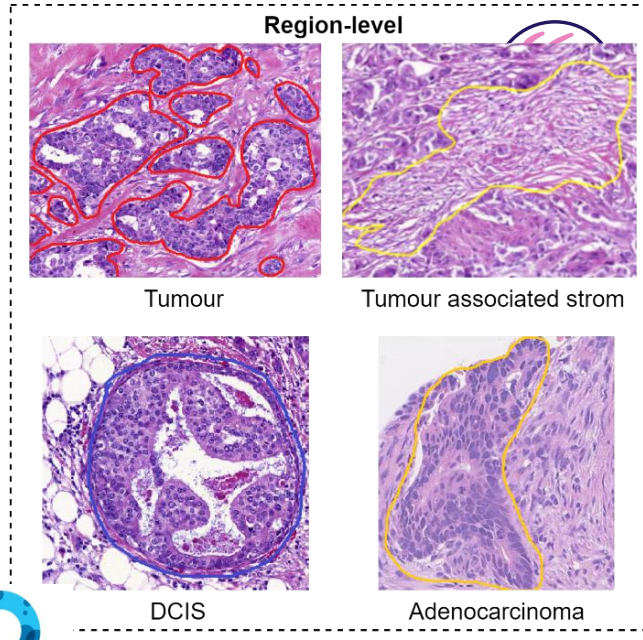


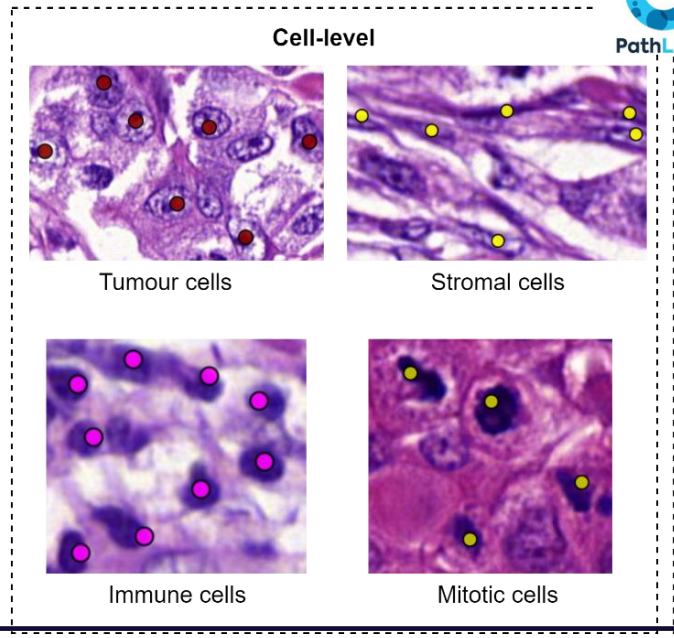
ML Problem(s) with WSIs: Training Labels

- Case or WSI-level labels
- Detailed annotations (Regions s)

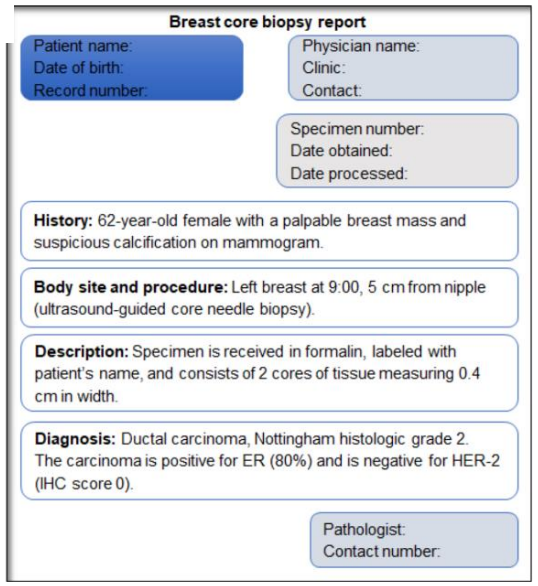






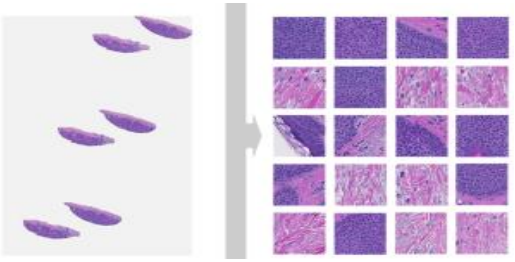


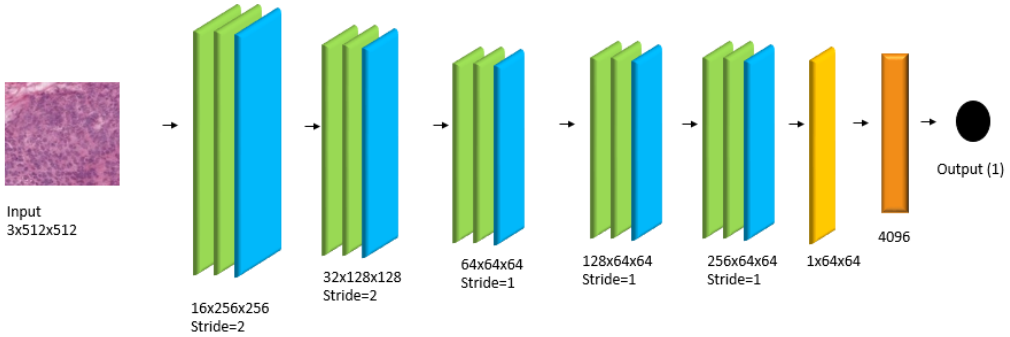
PathLAKE



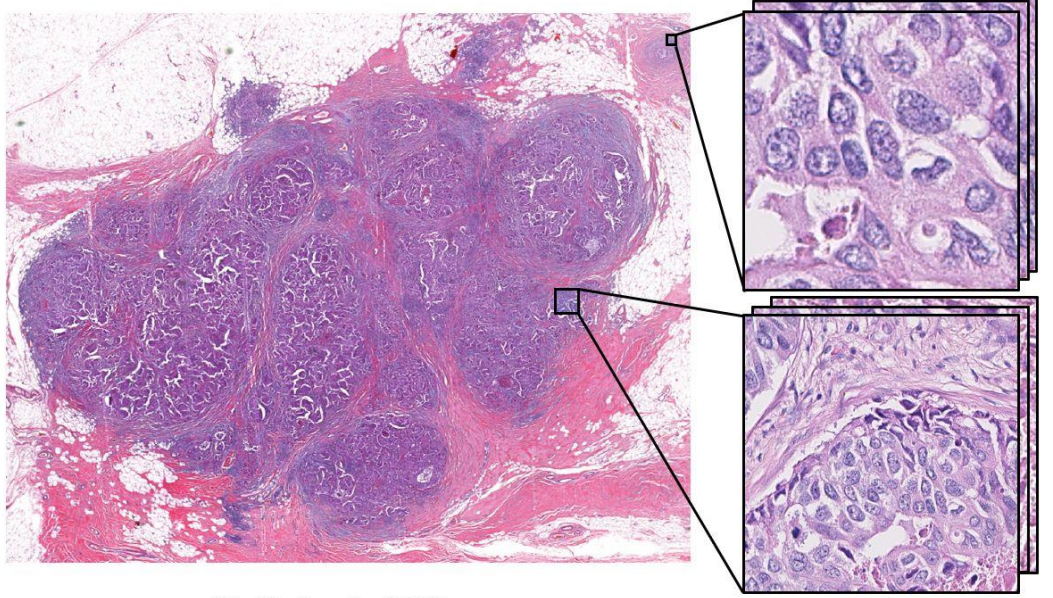
Biopsy report

Aggregation Problem



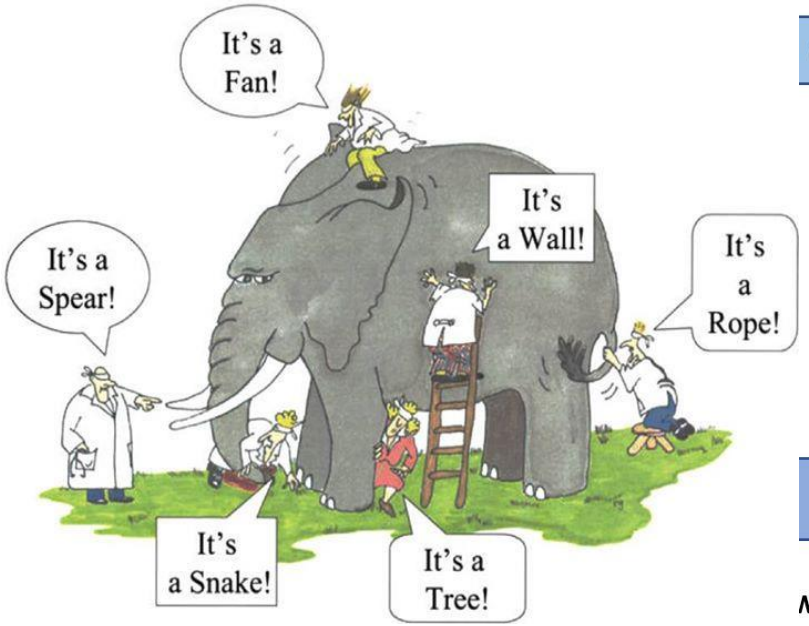


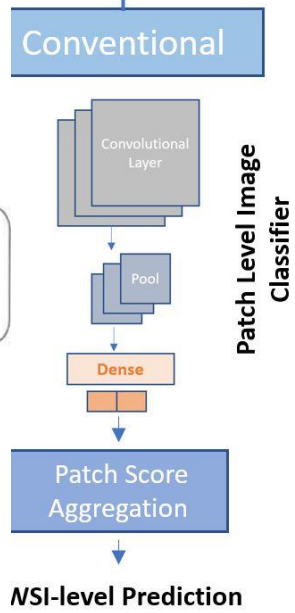
- Aggregation: Moving from patches to slide level predictions
  - Weak supervision approaches such as Averaging, Multiple Instance Learning, Attention Mechanisms LSTMs…



Multi-gigapixel WSI

Patches



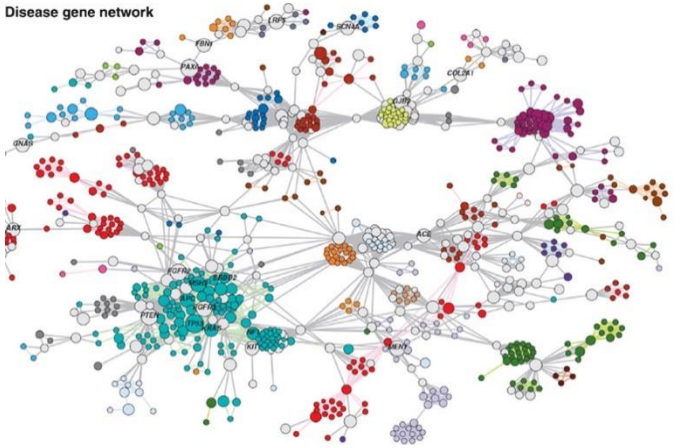


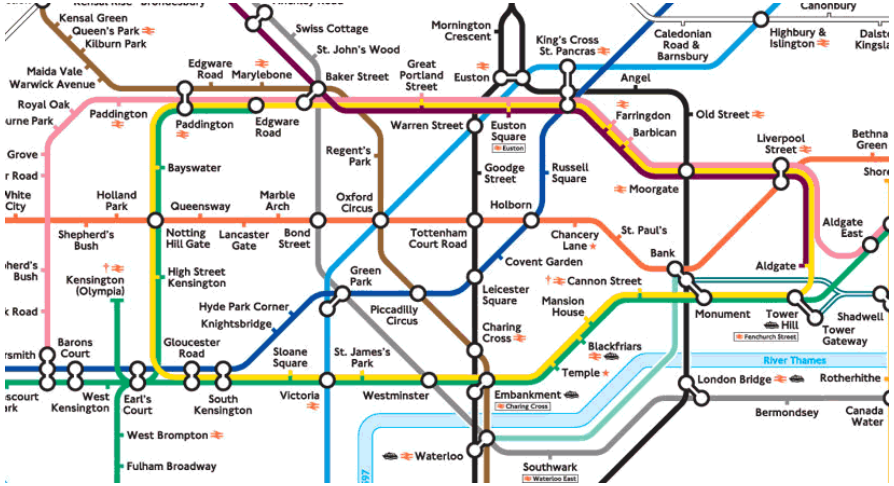
Weakly Supervised Learning Example

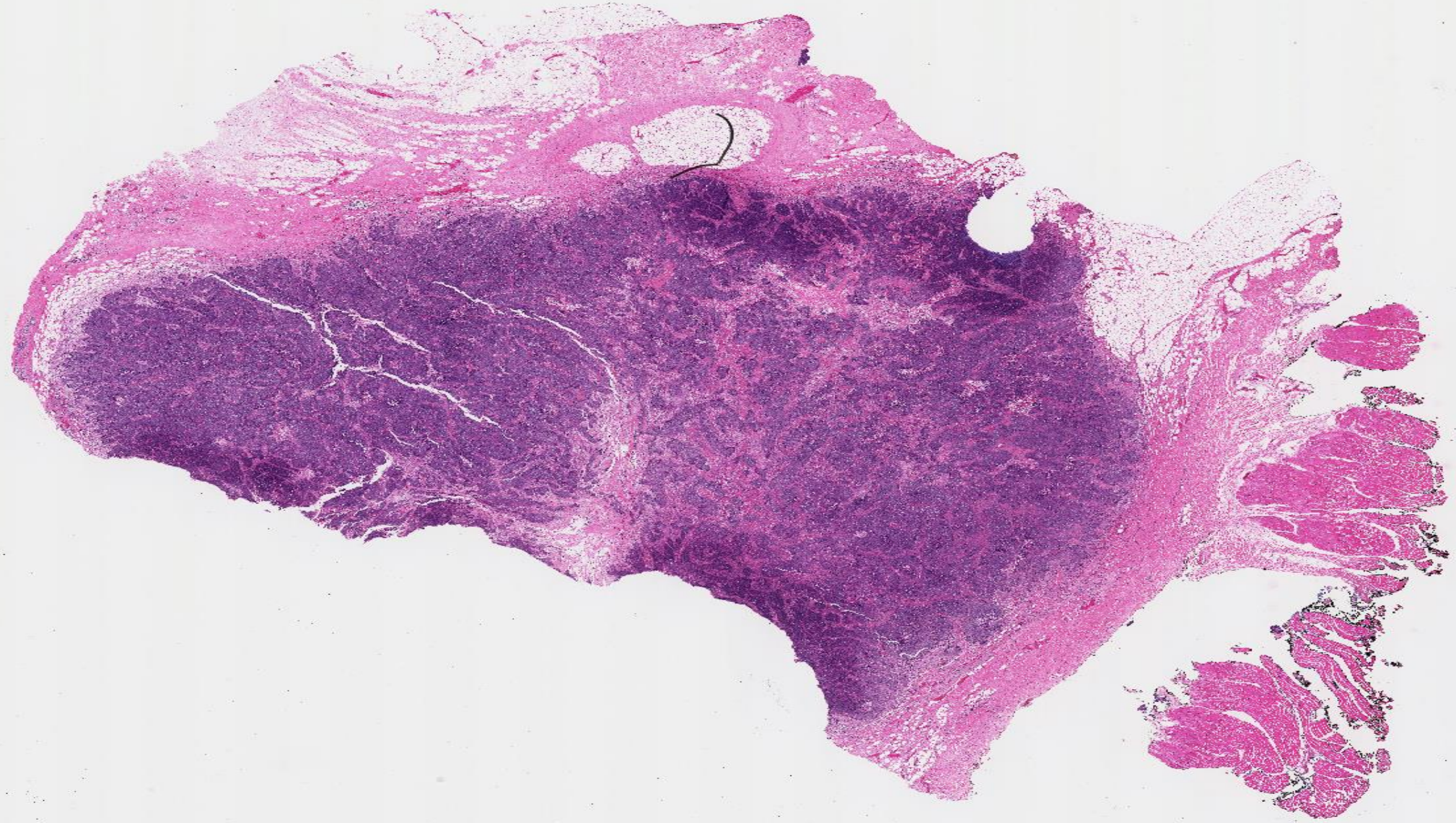
- Given:
  - Patches from Tumor Slides
  - Patches from Non-tumor slides
- Tumor Slides will have non-tumor patches
- Predict
  - Given a patch – is it a tumor patch?
  - Given a slide – does it have a tumor patch? (Is tumorous?)

An alternate modelling strategy: Graphs

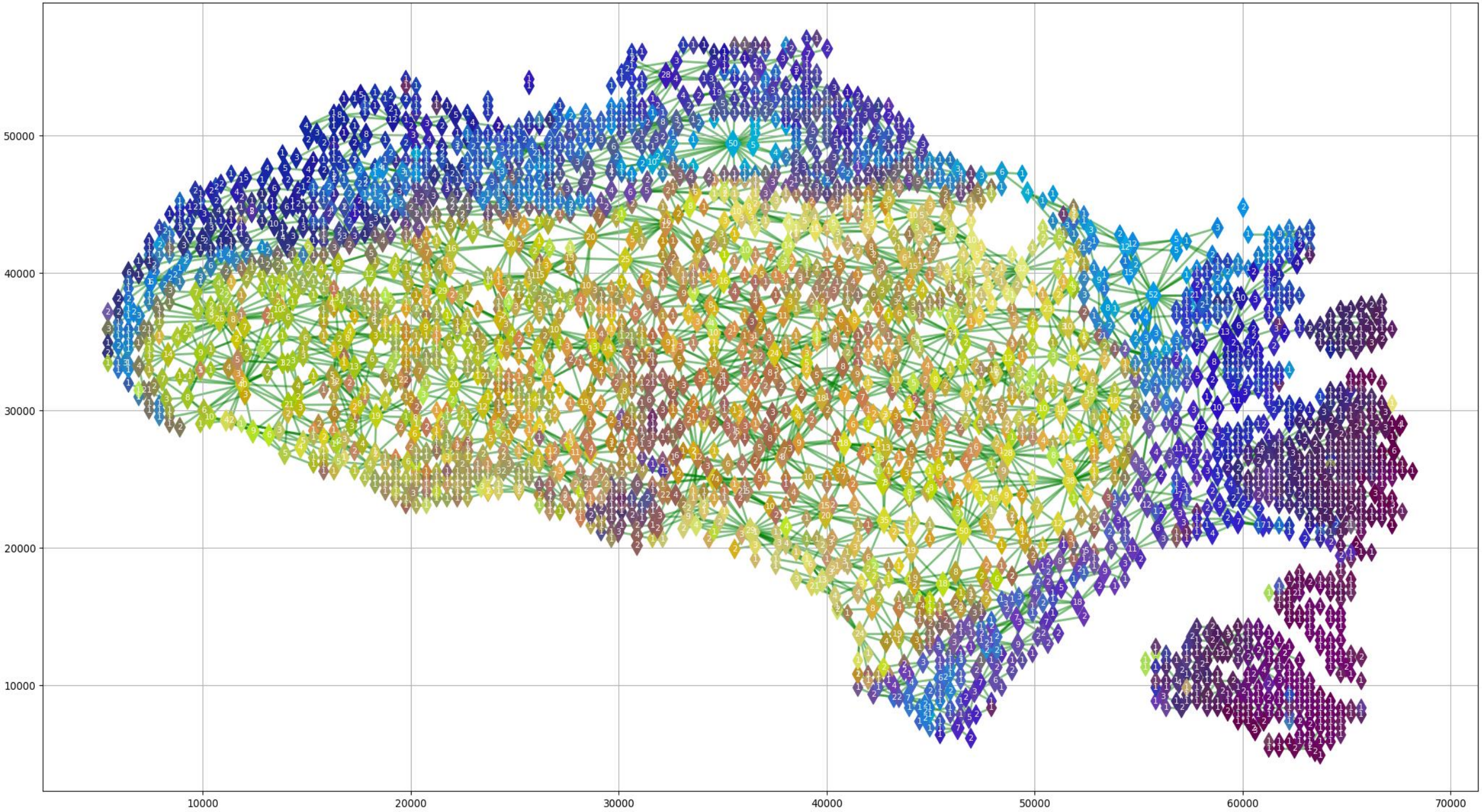
- What are graphs?
  - Sets of nodes
    - Each node has a set of features
  - Links or Edges
    - Each edge can have a set of descriptors



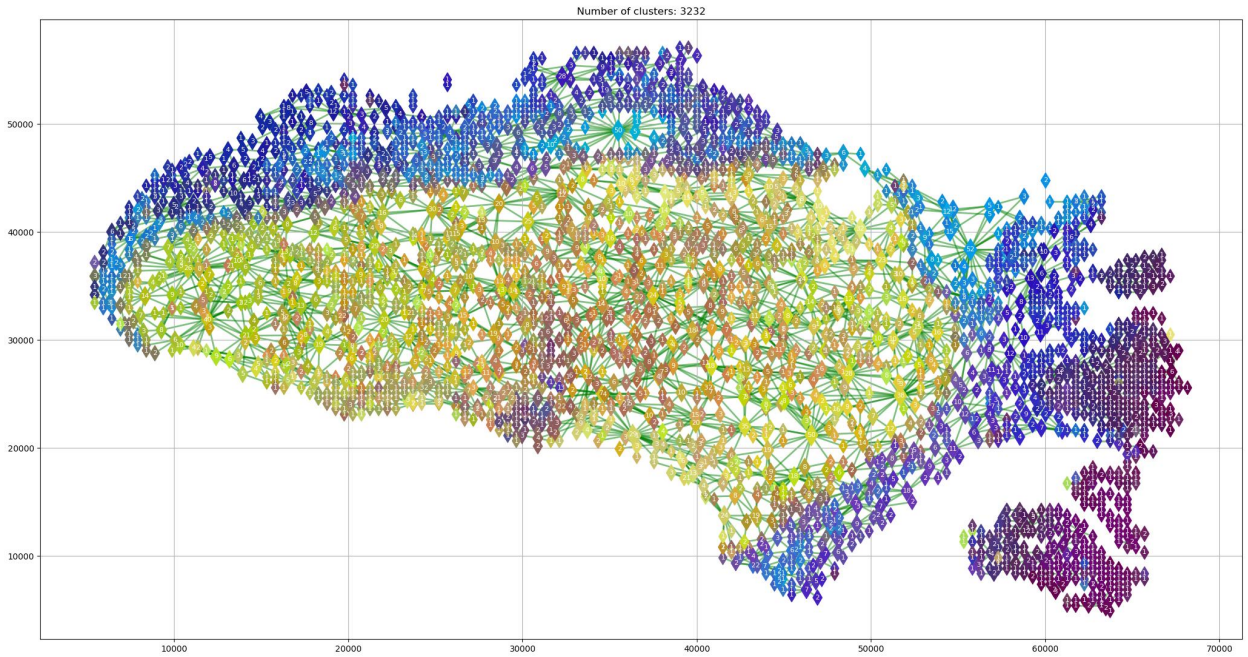


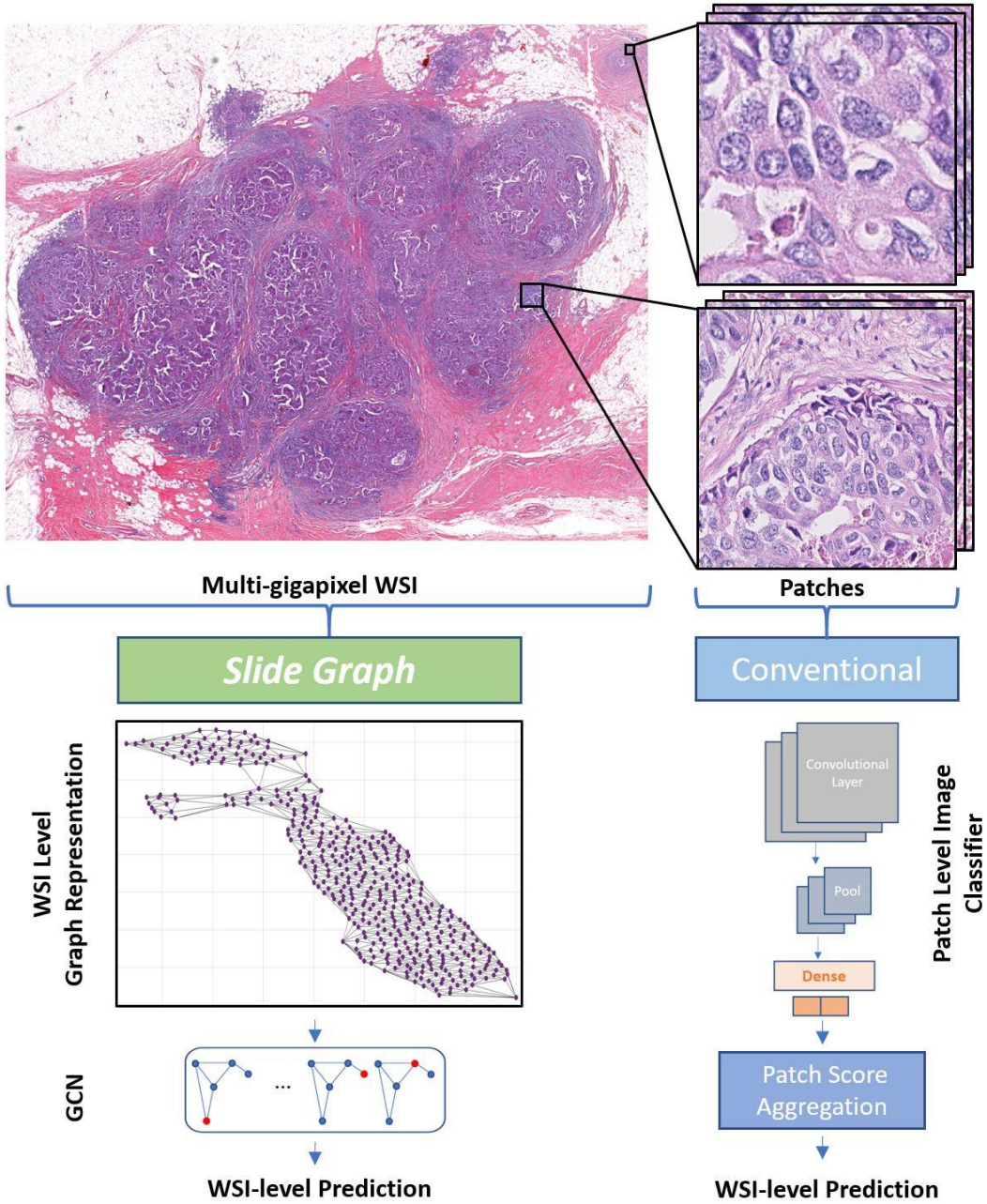


Number of clusters: 3232



SlideGraph+



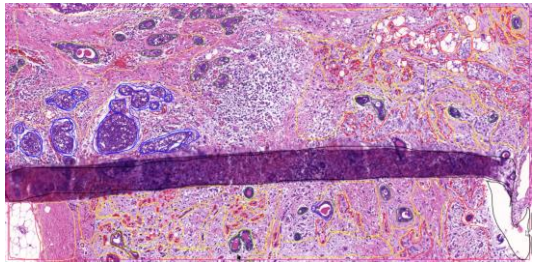


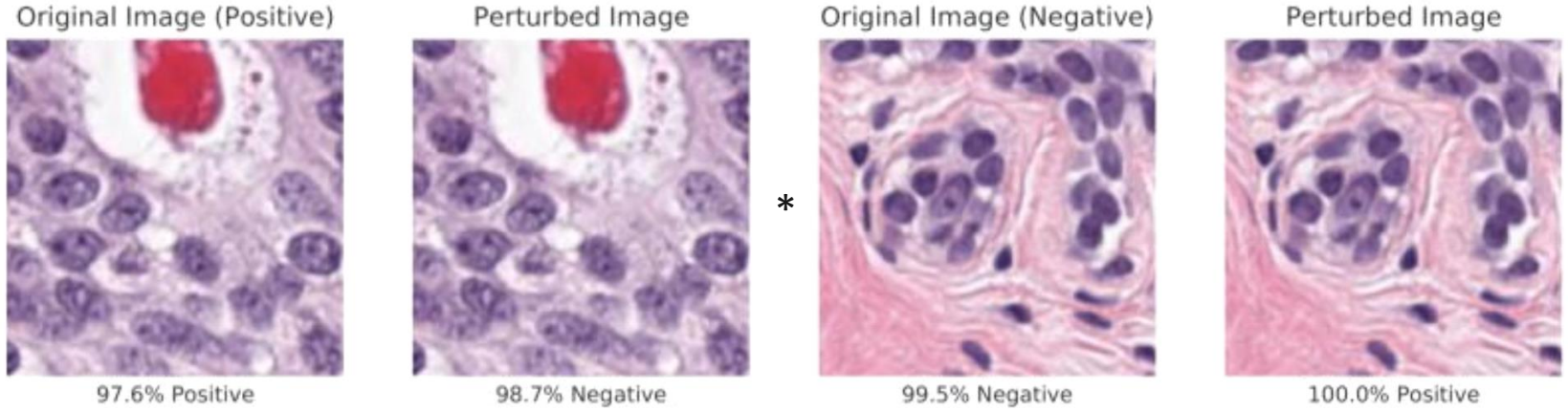
- Lu, Wenqi, Simon Graham, Mohsin Bilal, Nasir Rajpoot, and Fayyaz Minhas. "Capturing Cellular Topology in Multi-Gigapixel Pathology Images." In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1049–58, 2020. https://doi.org/10.1109/CVPRW50498.2020.00138.

- Lu, Wenqi, Michael Toss, Muhammad Dawood, Emad Rakha, Nasir Rajpoot, and Fayyaz Minhas. "SlideGraph+: Whole Slide Image Level Graphs to Predict HER2 Status in Breast Cancer." *Medical Image Analysis*, May 25, 2022, 102486. https://doi.org/10.1016/j.media.2022.102486.

Generalization and Reproducibility

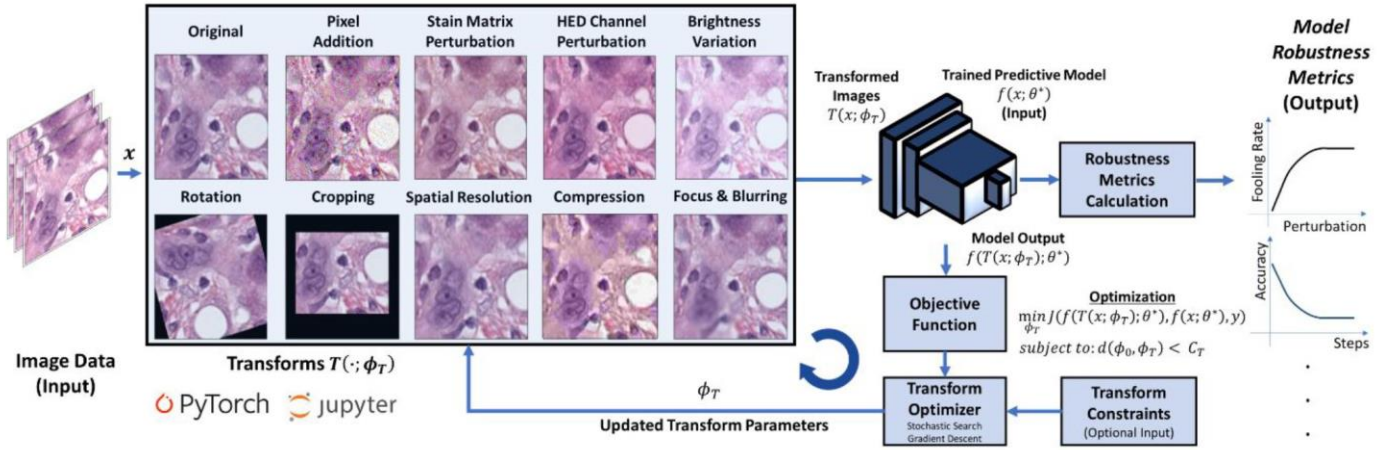
- Sources of variation
  - Inherent differences in populations
  - Difference in disease pathology
  - Tissue Processing and Handling
  - Scanner Changes
  - Pre-processing and QC
  - Changes in population
    - Due to treatment
    - Aging
- Possible Solution(s)
  - Publishing code and data
  - Preventing over-engineering
  - Large-scale validation
    - PathLAKE
  - Competitions
    - Mitosis detection
    - MSI prediction
  - Friendly "Adversarial" attacks
  - Peer-review process
  - Making solutions available to pathologists for trials (webservers)

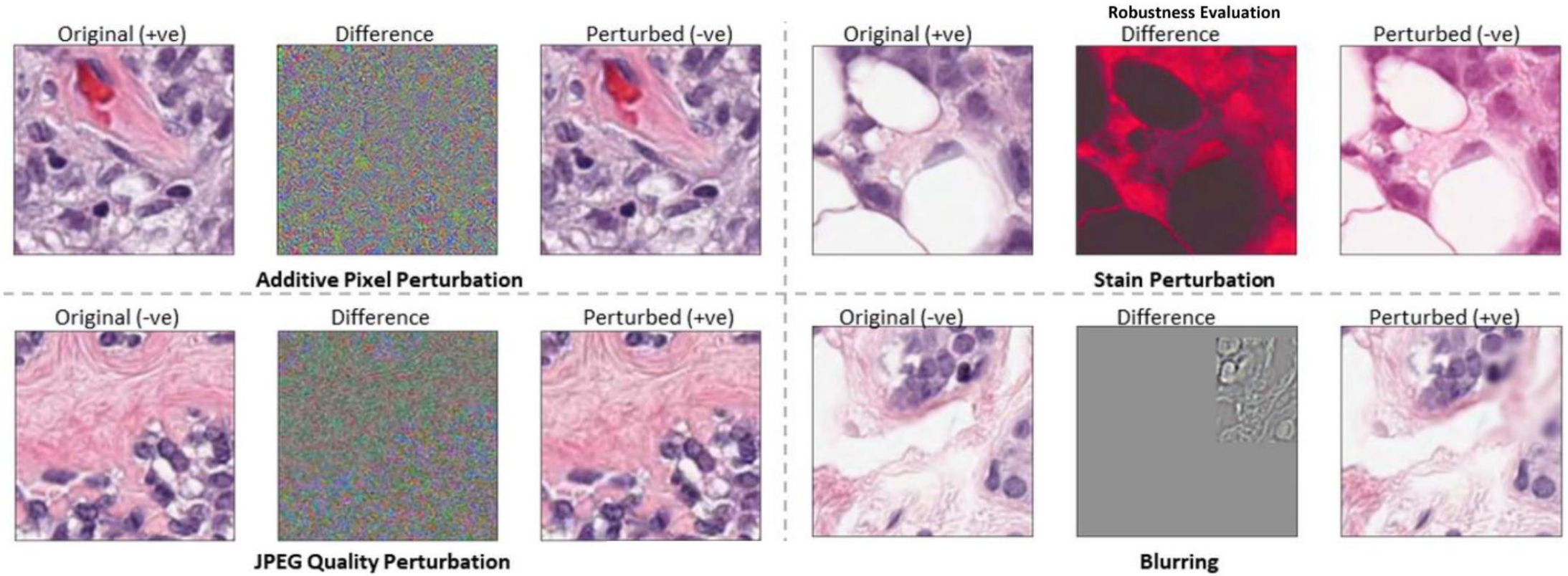




REET

Training	Accuracy (%)	Fooling Rate (%)				
		Stain Mixing	JPEG Quality	Rotate	Blur	Average
Standard	94.1 ± 0.5	64.3 ± 6.8	25.2 ± 2.5	21.7 ± 1.3	5.6 ± 1.1	29.2 ± 2.9
Adversarial	95.0 ± 0.8	2.2 ± 0.6	3.0 ± 0.4	6.7 ± 0.8	2.3 ± 0.6	3.6 ± 0.6





If you can make a CPath Model, We can break it!

Foote, Alex, Amina Asif, Nasir Rajpoot, and Fayyaz Minhas. "REET: Robustness Evaluation and Enhancement Toolbox for Computational Pathology." *Bioinformatics*, May 9, 2022, btac315. https://doi.org/10.1093/bioinformatics/btac315.

	TIA Toolbox	HistoCarto Graphy[11]	HEAL[10]	QuPath[7]	PathML[13]	CLAM[12]	Multi_Scale_Tools[14]	stainlib†	IBM CODAIT deep-histopath‡	HistoQC[2]
Platform Compatibility	Windows, Linux, Mac	?	Docker (Linux)	Windows, Linux, Mac	?	?	?	?	?	?
Language	Python	Python	Python	Java (+Groovy)	Java +Python	Python	Python	Python	Python	Python
Modular Usage	✓	✓	✗	✓	✓	✓	✗	✓	✗	✗
Unit Test Coverage	>99%	88%	0%	?	87%	0%	0%	0%	0%	?
Read Visual Fields (PNG, JPEG)	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗
Read WSI TIFF (SVS, SCN, MRXS)	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓
Read Generic Tiled TIFF	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓
Read JP2 (Omnyx)	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
Read OME-TIFF	✓	✗	✗	✓	✓	✗	✗	✗	✗	✗
Physical Unit Based Reads	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
Tissue Masking / Detection	✓	✓	✗	✓	✓	✓	✗	✗	✓	✓
Patch Extraction	✓	✗	✓	✗	✗	✓	✓	✗	✓	✗
Stain Separation	✓	✓	✗	✓	✓	✗	✗	✓	✗	✗
Stain Normalization	✓	✓	✓	✗	✓	✗	✗	✓	✗	✗
Stain Augmentation	✓	✗	✗	✗	✗	✗	✗	✓	✗	✗
Patch Classification	✓	✗	✗	✓	✗	✗	✓	✗	!	✗
Tissue Semantic Segmentation	✓	✗	✗	!	✓	✗	!	✗	✗	✗
Nucleus Detection	✓	✓	✗	✓	✓	✗	✗	✗	✗	✗
Nucleus Segmentation	✓	✗	✗	✓	✓	✗	✗	✗	✗	✗
Nucleus Classification	✓	✗	✗	✓	✓	✗	✗	✗	✗	✗
Deep Feature Extraction	✓	✓	✗	✗	✗	✓	✗	✗	✗	✗
Graph Learning	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗
Whole Slide Classification	✓	✓	✓	✗	✗	✓	✗	✗	✗	✗

Pocock, Johnathan, Simon Graham, Quoc Dang Vu, Mostafa Jahanifar, Srijay Deshpande, Giorgos Hadjigeorghiou, Adam Shephard, et al. "TIAToolbox: An End-to-End Toolbox for Advanced Tissue Image Analytics." bioRxiv, December 24, 2021. https://doi.org/10.1101/2021.12.23.474029.

Future Technological Advances in CPath

- Overcoming previously stated challenges
- Incorporating Natural Language Processing
- Focus on information fusion

- Are pathologist labels noisy? (Invalidate Turing tests)

Scientific Achievements of CPath

- **5 Pathologist Nobel Prize Winners [1]**
  - **Johannes Fibiger** received the Nobel Prize in Physiology or Medicine in 1926 for the experiments in which he produced gastric carcinoma in rats by feeding them *Spiroptera*-infected cockroaches.
  - **George Whipple** was awarded the 1934 Nobel Prize in Physiology or Medicine for his discovery that a diet rich in liver cured pernicious anemia.
  - **Renato Dulbecco** was awarded the 1975 Nobel Prize for discovering by molecular techniques that the genetic material of viruses was incorporated into the genetic material of the transformed cells and represented the first phase in carcinogenesis.
  - **Baruj Benacerraf** received the 1980 Nobel Prize in Physiology or Medicine for his discovery of immune response genes.
  - **John Warren** received the Nobel Prize in Physiology or Medicine in 2005 in recognition of this discovery of the role of *H. Pylori* in ulcers.

"Rutherford used to divide science into physics and stamp collecting.[2]"

[1] http://www.annclinlabsci.org/content/39/2/196.long
[2] Bernal, J *The Social Importance of Science*1939 p. 9

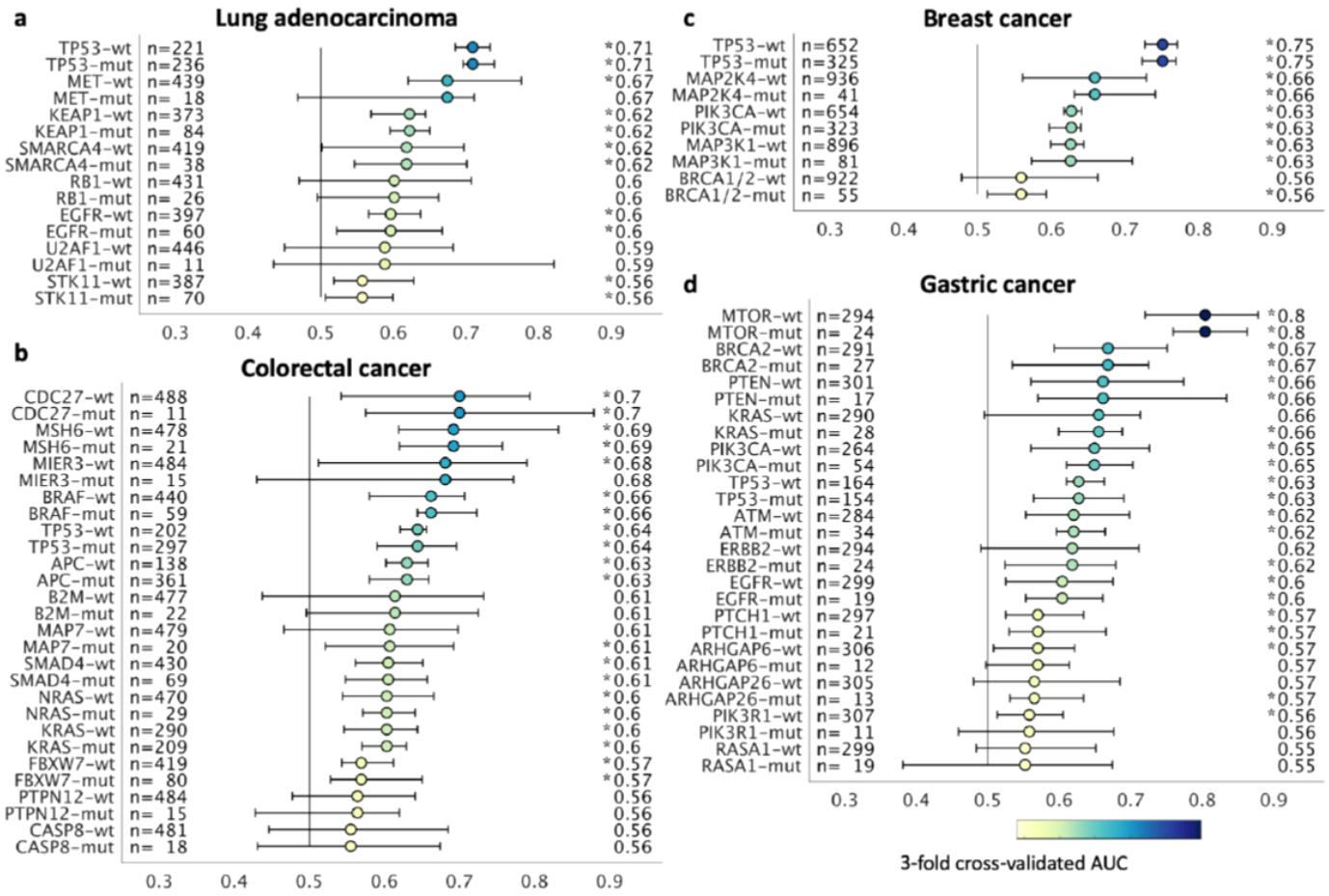
Scientific Discoveries of CPath

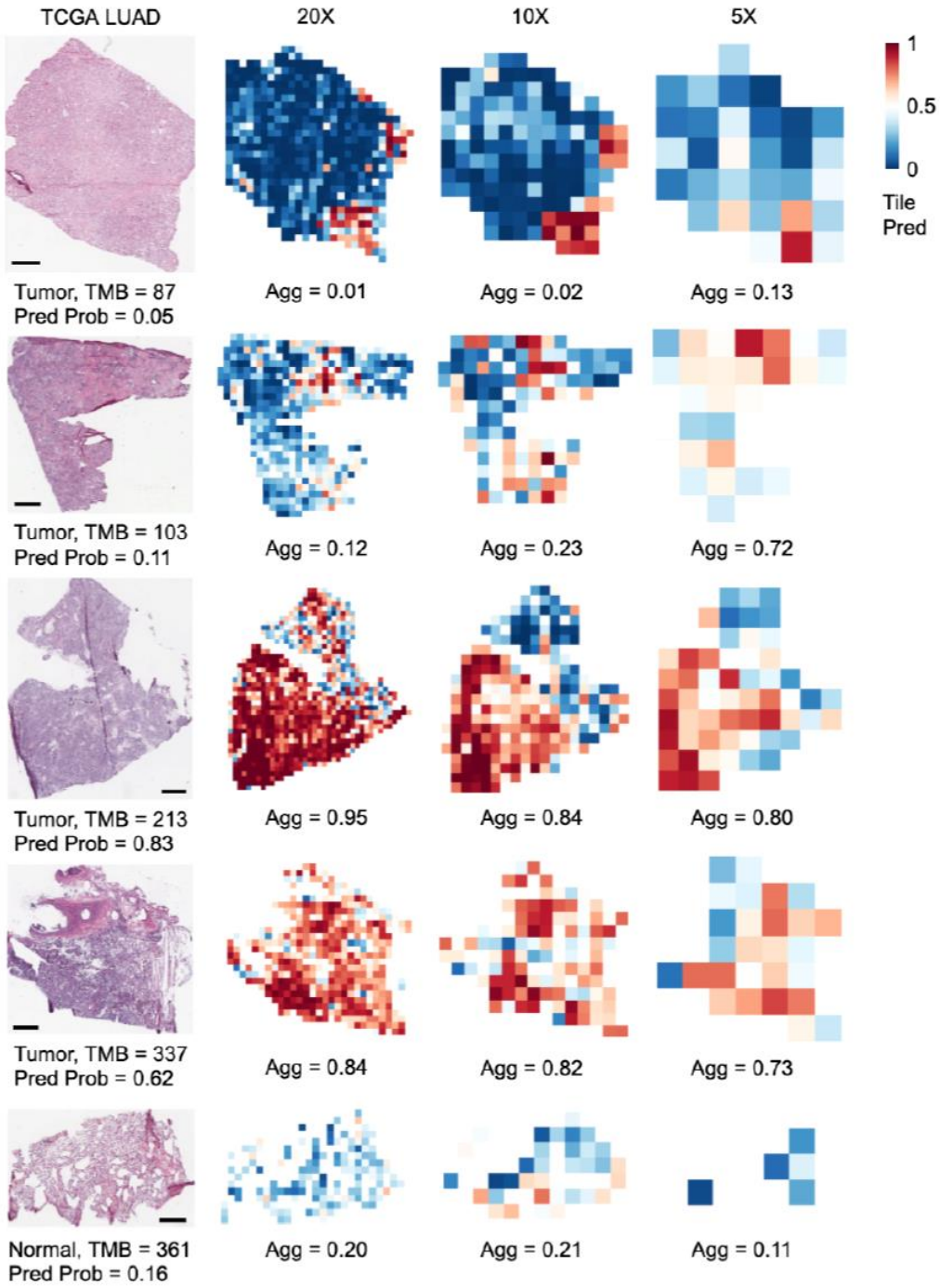
- What has CPath help discover?
  - Current focus
    - Identification of associations between various clinical parameters and Cpath Imaging data (visual and/or subvisual)
  - It is possible to predict (from WSIs)
    - Mutations
    - Gene Expression Patterns
    - Source of cancers of unknown primaries
    - Microsatellite Instability
    - Receptor Status in BrCa
    - Chemo-response/drug response in OvCa
    - Recurrence

Some Literature

- Kather, Jakob Nikolas, Lara R. Heij, Heike I. Grabsch, Chiara Loeffler, Amelie Echle, Hannah Sophie Muti, Jeremias Krause, et al. "Pan-Cancer Image-Based Detection of Clinically Actionable Genetic Alterations." *Nature Cancer*, July 27, 2020, 1–11. https://doi.org/10.1038/s43018-020-0087-6.
- Fu, Yu, Alexander W. Jung, Ramon Viñas Torne, Santiago Gonzalez, Harald Vöhringer, Artem Shmatko, Lucy R. Yates, Mercedes Jimenez-Linan, Luiza Moore, and Moritz Gerstung. "Pan-Cancer Computational Histopathology Reveals Mutations, Tumor Composition and Prognosis." *Nature Cancer* 1, no. 8 (August 2020): 800–810. https://doi.org/10.1038/s43018-020-0085-8.
- Jain, Mika S., and Tarik F. Massoud. "Predicting Tumour Mutational Burden from Histopathological Images Using Multiscale Deep Learning." *BioRxiv*, June 16, 2020, 2020.06.15.153379. https://doi.org/10.1101/2020.06.15.153379.
- Bilal, Mohsin, Shan E. Ahmed Raza, Ayesha Azam, Simon Graham, Muhammad Ilyas, Ian A. Cree, David Snead, Fayyaz Minhas, and Nasir M. Rajpoot. "Novel Deep Learning Algorithm Predicts the Status of Molecular Pathways and Key Mutations in Colorectal Cancer from Routine Histology Images." *MedRxiv*, February 4, 2021, 2021.01.19.21250122. https://doi.org/10.1101/2021.01.19.21250122. Echle, Amelie, Heike Irmgard Grabsch, Philip Quirke, Piet A. van den Brandt, Nicholas P. West, Gordon G. A. Hutchins, Lara R. Heij, et al. "Clinical-Grade Detection of Microsatellite Instability in Colorectal Tumors by Deep Learning." *Gastroenterology*, June 17, 2020. https://doi.org/10.1053/j.gastro.2020.06.021.
- Echle, Amelie, Heike Irmgard Grabsch, Philip Quirke, Piet A. van den Brandt, Nicholas P. West, Gordon G. A. Hutchins, Lara R. Heij, et al. "Clinical-Grade Detection of Microsatellite Instability in Colorectal Tumors by Deep Learning." *Gastroenterology*, June 17, 2020. https://doi.org/10.1053/j.gastro.2020.06.021.
- Lu, Ming Y., Melissa Zhao, Maha Shady, Jana Lipkova, Tiffany Y. Chen, Drew F. K. Williamson, and Faisal Mahmood. "Deep Learning-Based Computational Pathology Predicts Origins for Cancers of Unknown Primary." *ArXiv:2006.13932 [Cs, q-Bio]*, June 28, 2020. http://arxiv.org/abs/2006.13932.
- …

Mutation Prediction





Kather, Jakob Nikolas, Lara R. Heij, Heike I. Grabsch, Chiara Loeffler, Amelie Echle, Hannah Sophie Muti, Jeremias Krause, et al. "Pan-Cancer Image-Based Detection of Clinically Actionable Genetic Alterations." *Nature Cancer* 1, no. 8 (August 2020): 789–99. https://doi.org/10.1038/s43018-020-0087-6.

Jain, Mika S., and Tarik F. Massoud. "Predicting Tumour Mutational Burden from Histopathological Images Using Multiscale Deep Learning." *Nature Machine Intelligence* 2, no. 6 (June 2020): 356–62. https://doi.org/10.1038/s42256-020-0190-5.

Pan-Cancer Computational Histopathology Reveals Mutations, Tumor Composition and Prognosis." Nature Cancer 1, no. 8 (August 2020): 800–810. https://doi.org/10.1038/s43018-020-0085-8.

Microsatellite Instability

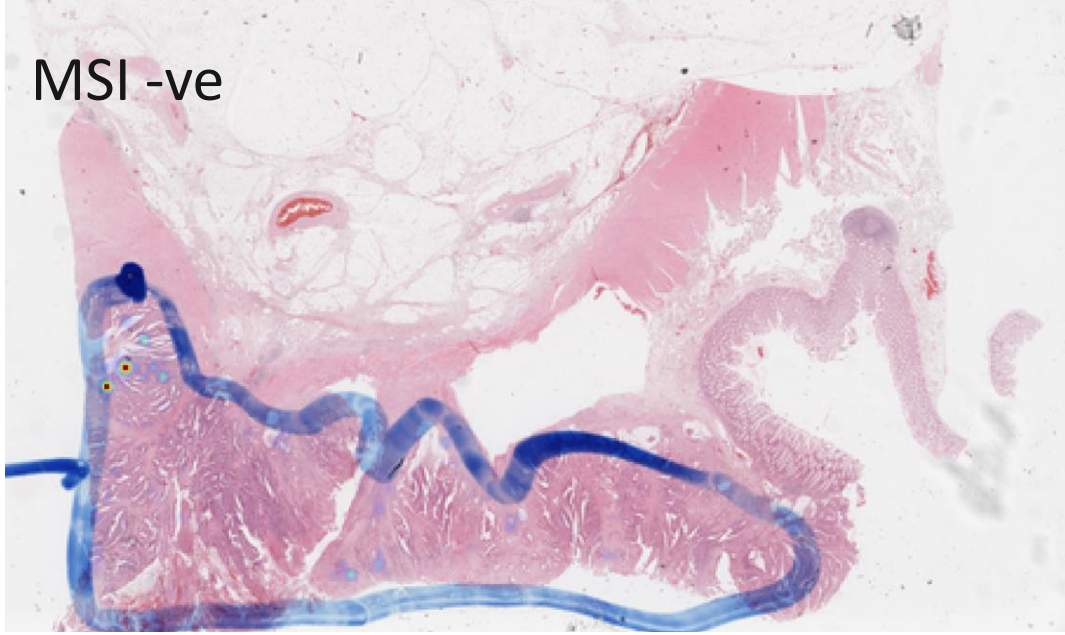
Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer

Jakob Nikolas Kather[1,2,3,4,5*], Alexander T. Pearson[4], Niels Halama[2,5,6], Dirk Jäger[2,3,5], Jeremias Krause[1], Sven H. Loosen[1], Alexander Marx[7], Peter Boor[8], Frank Tacke[9], Ulf Peter Neumann[10], Heike I. Grabsch[11,12], Takaki Yoshikawa[13,14], Hermann Brenner[2,15,16], Jenny Chang-Claude[17,18], Michael Hoffmeister[15], Christian Trautwein[1] and Tom Luedde[1*]

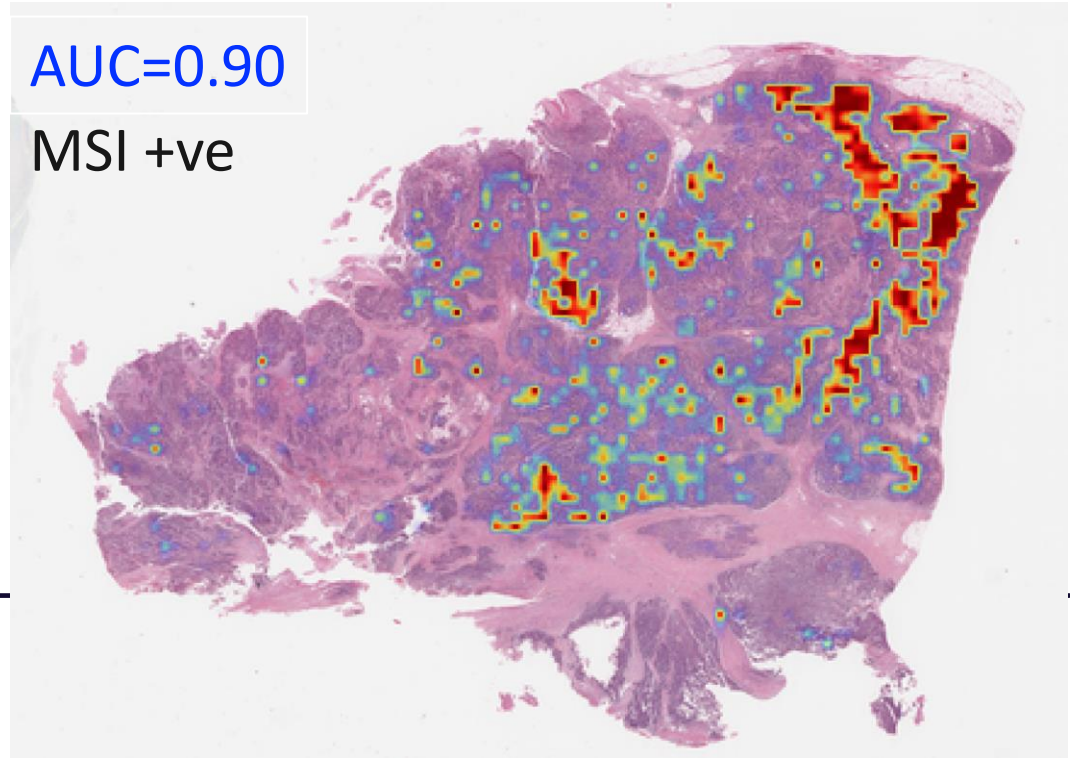
Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study

Rikiya Yamashita, MD • Jin Long, PhD • Prof Teri Longacre, MD • Lan Peng, MD • Prof Gerald Berry, MD • Brock Martin, MD • et al. Show all authors • Show footnotes

Published: January, 2021 • DOI: https://doi.org/10.1016/S1470-2045(20)30535-0 •

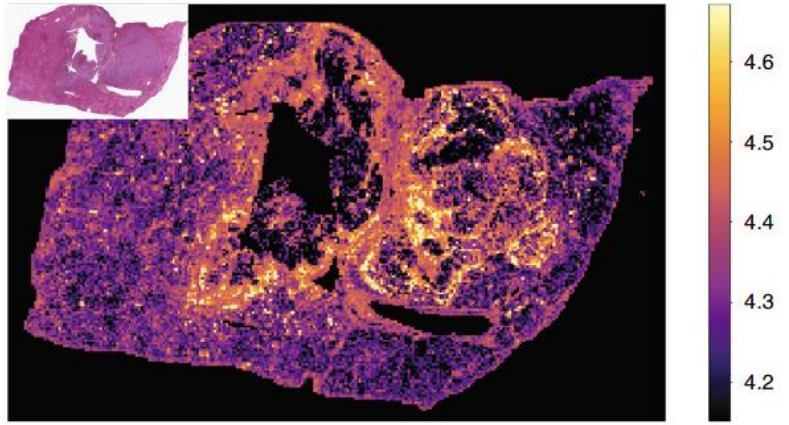


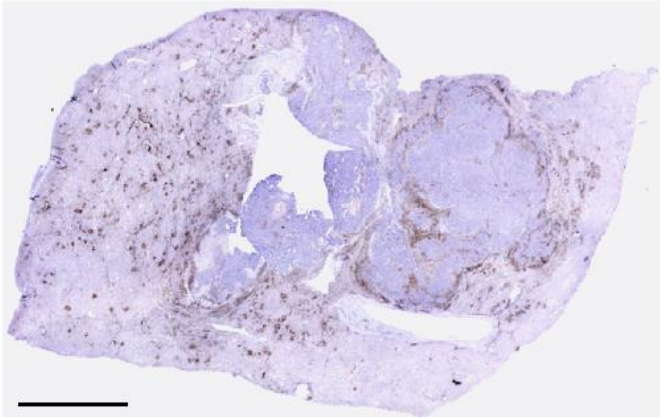
Bilal et al. (*under review*)

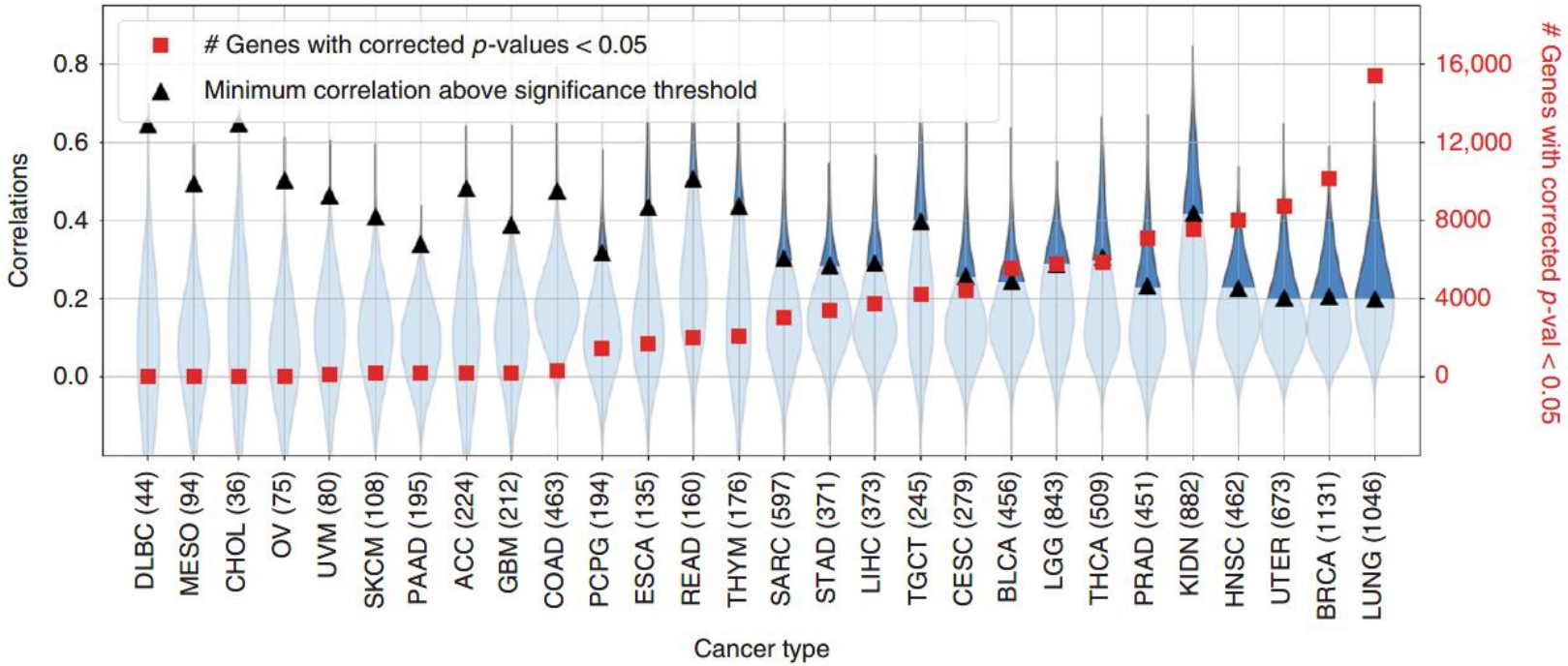


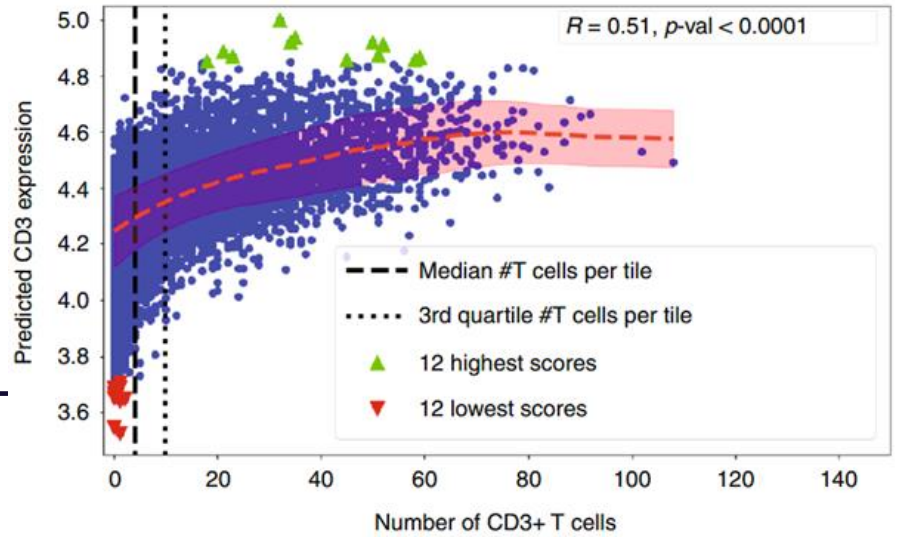
Bilal, Mohsin, Shan E. Ahmed Raza, Ayesha Azam, Simon Graham, Muhammad Ilyas, Ian A. Cree, David Snead, Fayyaz Minhas, and Nasir M. Rajpoot. "Novel Deep Learning Algorithm Predicts the Status of Molecular Pathways and Key Mutations in Colorectal Cancer from Routine Histology Images." *MedRxiv*, February 4, 2021, 2021.01.19.21250122. https://doi.org/10.1101/2021.01.19.21250122.

Gene Expression Patterns



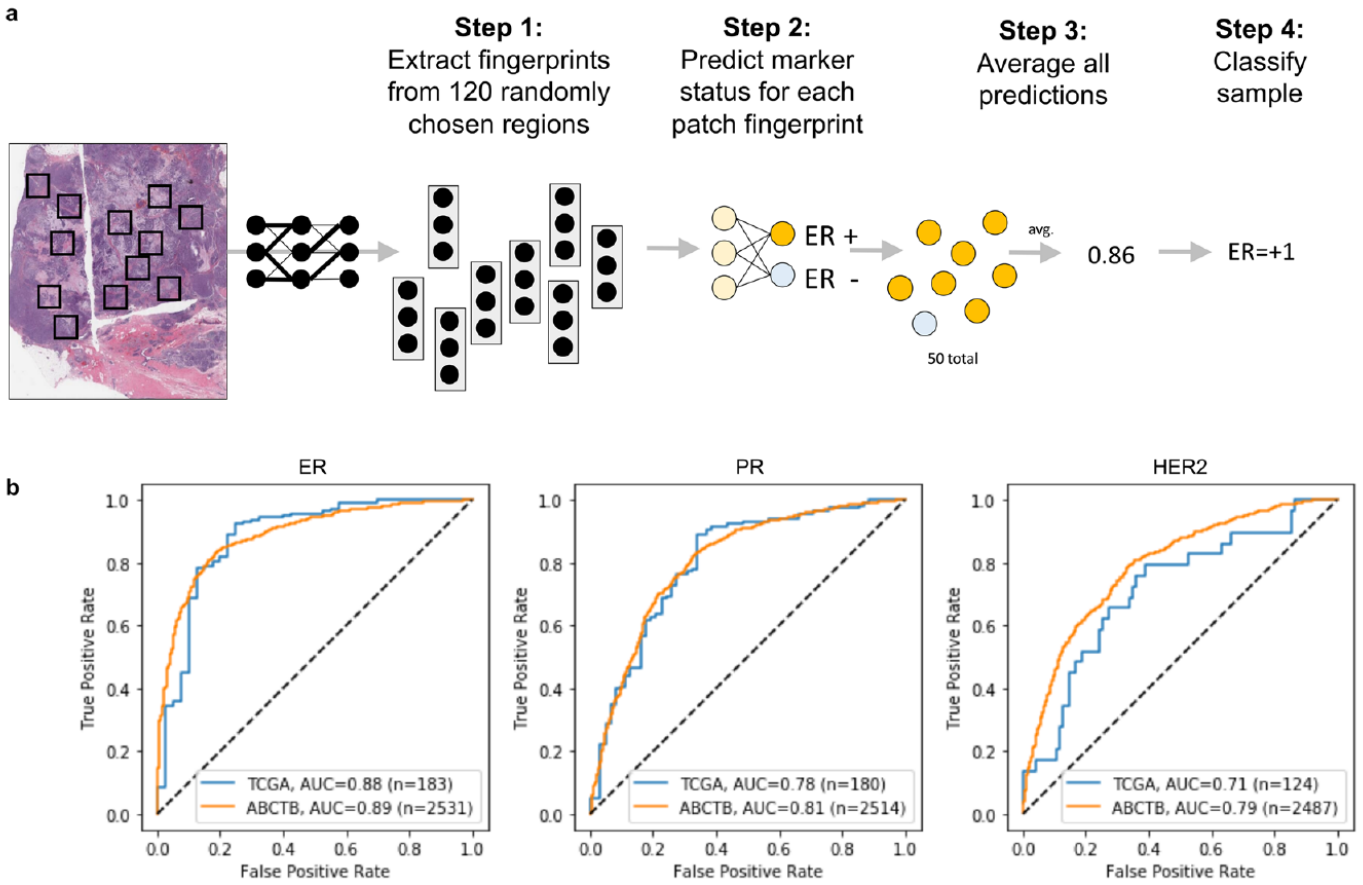






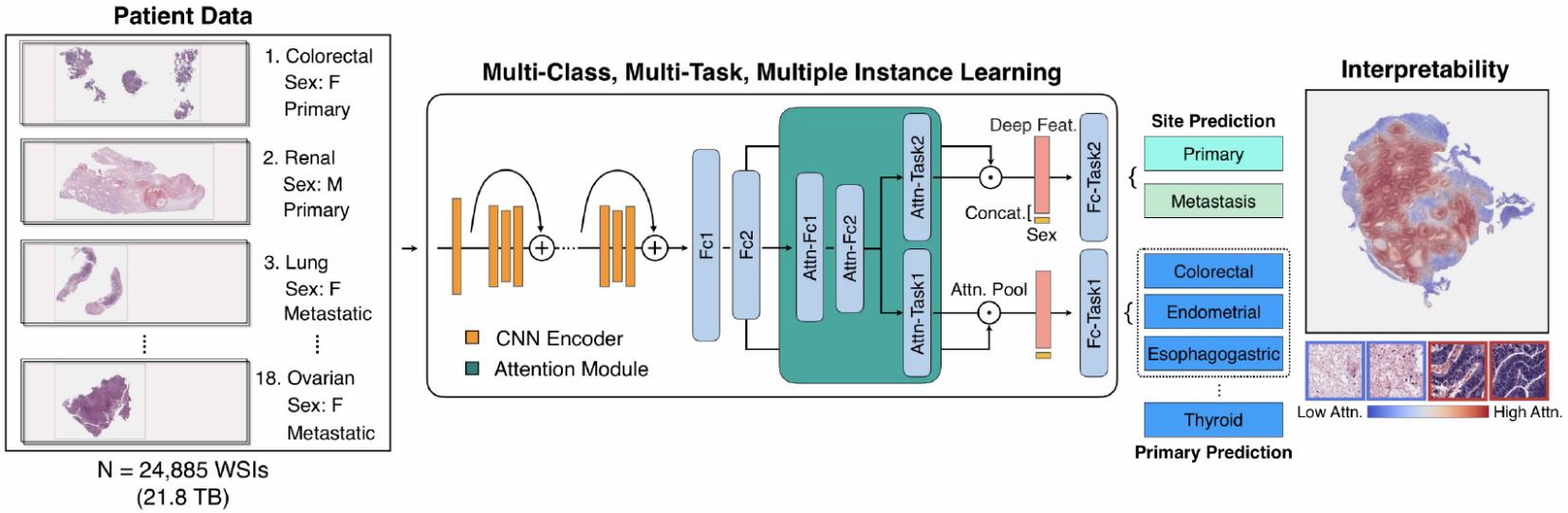
Schmauch, Benoît, Alberto Romagnoni, Elodie Pronier, Charlie Saillard, Pascale Maillé, Julien Calderaro, Aurélie Kamoun, et al. "A Deep Learning Model to Predict RNA-Seq Expression of Tumours from Whole Slide Images." *Nature Communications* 11, no. 1 (August 3, 2020): 3877. https://doi.org/10.1038/s41467-020-17678-4.

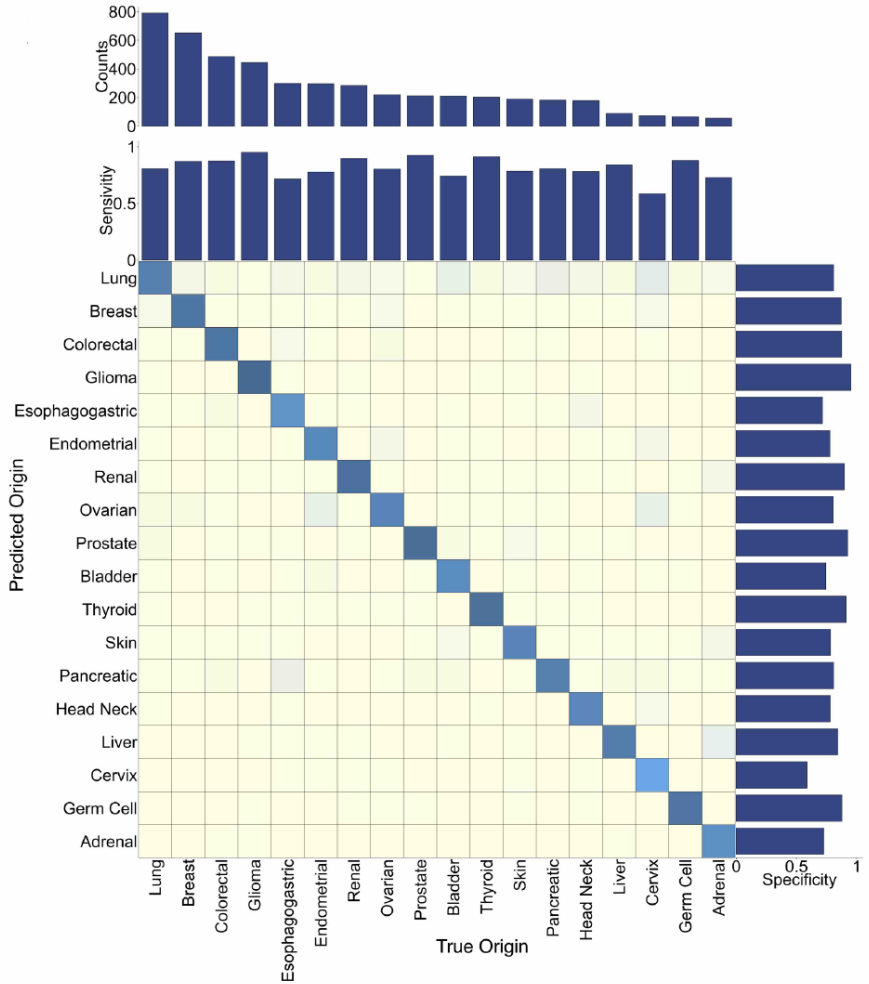
Receptor Status in BrCa



Rawat, Rishi R., Itzel Ortega, Preeyam Roy, Fei Sha, Darryl Shibata, Daniel Ruderman, and David B. Agus. "Deep Learned Tissue 'Fingerprints' Classify Breast Cancers by ER/PR/Her2 Status from H&E Images." *Scientific Reports* 10, no. 1 (April 29, 2020): 1–13. https://doi.org/10.1038/s41598-020-64156-4.

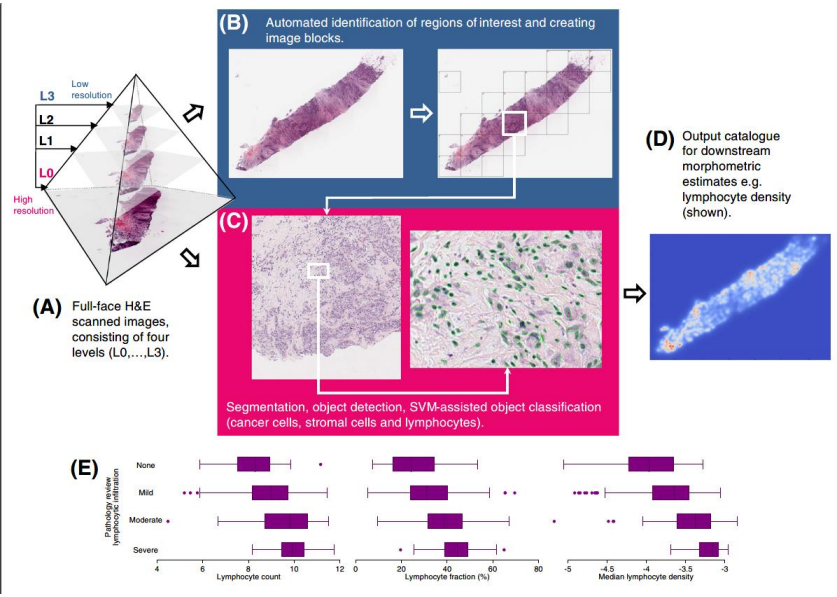
Cancers of Unknown Primaries

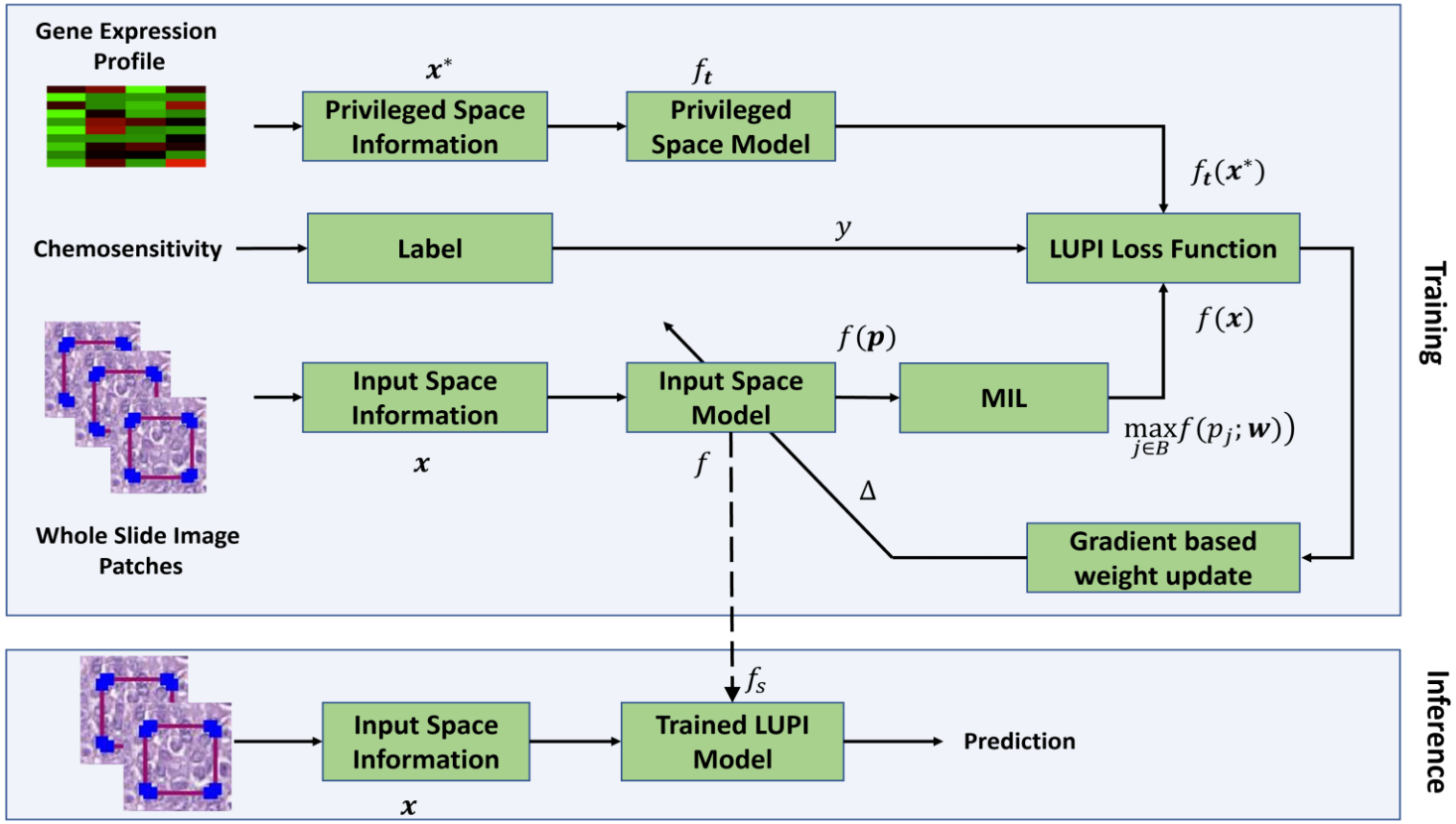




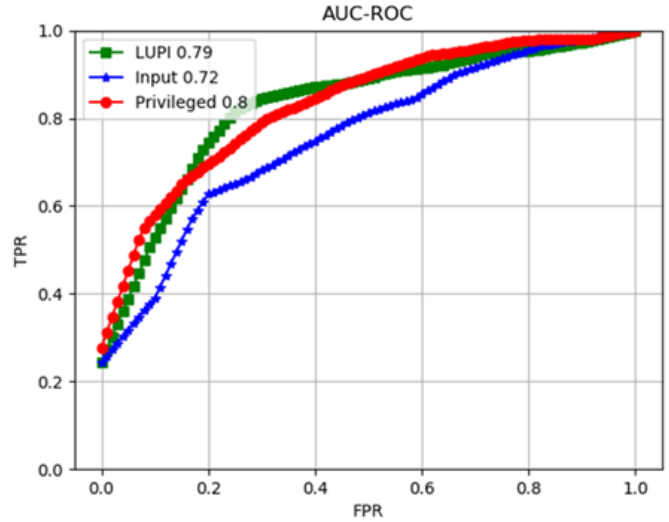
Lu, Ming Y., Melissa Zhao, Maha Shady, Jana Lipkova, Tiffany Y. Chen, Drew F. K. Williamson, and Faisal Mahmood. "Deep Learning-Based Computational Pathology Predicts Origins for Cancers of Unknown Primary." ArXiv:2006.13932 [Cs, q-Bio], June 28, 2020. http://arxiv.org/abs/2006.13932.

Chemotherapy Response Prediction





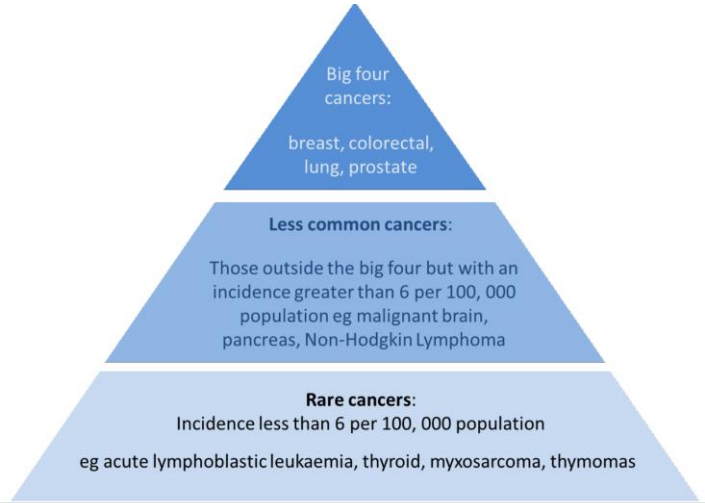
		Univariate				Multivariate			
Variable	Categories	Odds ratio	95 % CI	*P* value	Observations	Odds ratio	95 % CI	*P* value	Observations
Age	Continuous	0.99	0.97, 1.02	0.6	755	0.99	0.96-1.02	0.44	406
Tumour size	≤50 mm, >50 mm	0.89	0.54, 1.48	0.66	755	0.68	0.27-1.75	0.43	406
Node status	Negative, positive	0.8	0.54, 1.18	0.25	755	0.56	0.31-1.00	0.05	406
Grade	1, 2, 3	3.98	2.38, 6.67	<0.00001	588	3.62	1.74-7.52	0.0006	406
ER status	Negative, positive	0.26	0.17, 0.38	<0.00001	755	0.3	0.17-0.53	0.00004	406
HER2 status	Negative, positive	1.6	1.03, 2.50	0.04	665	1.81	0.97-3.37	0.06	406
Median lymphocyte density	Continuous	4.46	2.34, 8.50	<0.00001	614	2.42	1.08-5.40	0.03	406



Ali, H. Raza, Aliakbar Dariush, Elena Provenzano, Helen Bardwell, Jean E. Abraham, Mahesh Iddawela, Anne-Laure Vallier, et al. "Computational Pathology of Pre-Treatment Biopsies Identifies Lymphocyte Density as a Predictor of Response to Neoadjuvant Chemotherapy in Breast Cancer." *Breast Cancer Research* 18, no. 1 (February 16, 2016): 21. https://doi.org/10.1186/s13058-016-0682-8.
Yaar, Asfand, Amina Asif, Shan E. Ahmed Raza, Nasir Rajpoot, and Fayyaz Minhas. "Cross-Domain Knowledge Transfer for Prediction of Chemosensitivity in Ovarian Cancer Patients," 928–29, 2020. https://openaccess.thecvf.com/content_CVPRW_2020/html/w54/Yaar_Cross-Domain_Knowledge_Transfer_for_Prediction_of_Chemosensitivity_in_Ovarian_Cancer_CVPRW_2020_paper.html.

Future Scientific Discoveries of CPath?

- Previously undiscovered but clinically relevant patterns in data by integrating multimodal and multisource data
- Raising interesting biological questions
- Identification of novel causes of treatment effectiveness or resistance
- Leading to discovery of novel therapies and drugs
- Understanding evolutionary mechanisms at play that result in mutations
- Using Cpath as a tool by Pathologists
  - Current ML technologies are good at learning patterns
  - However, they do not have clinical and biological understanding of disease phenomenon
- Rare cancers



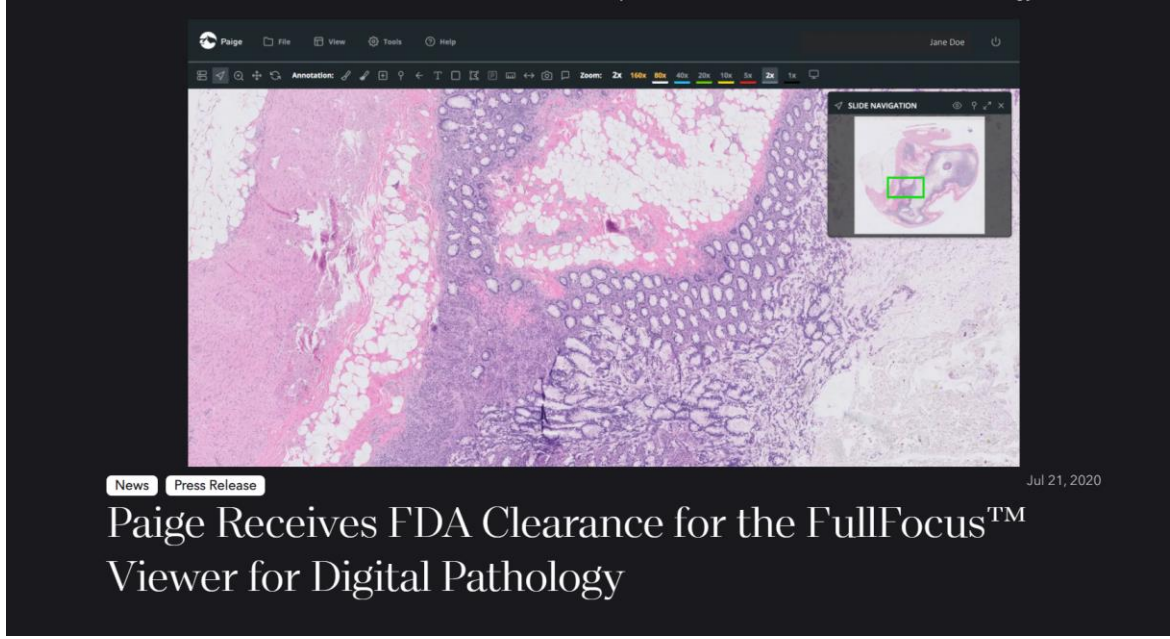
"By definition, you are less likely to be diagnosed with a less common or rare cancer. The real issue is that if you do, you are more likely to die from it than from a common cancer."

The Guardian

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6852275/

CCR Deployment

- How many commercial Cpath solutions have been clinically deployed (and are being used) after regulatory approval?
- Measuring Improvement in Clinical workflows and efficiency
- Closing the gap between volume of scientific publications and clinical deployment



Paige Prostate Clinical (CE-IVD)

A prostate cancer solution that detects, grades and quantifies areas suspicious for cancer.

Paige Breast Clinical (CE-IVD)

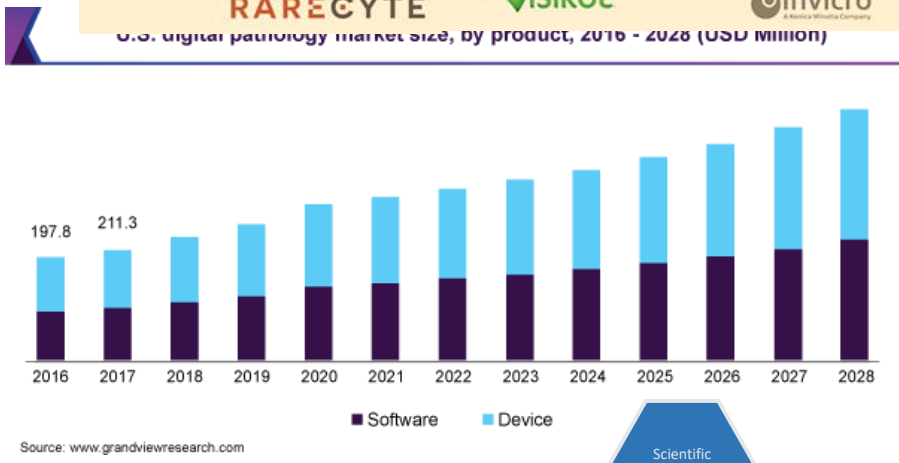
A breast cancer solution that detects areas suspicious for cancer.

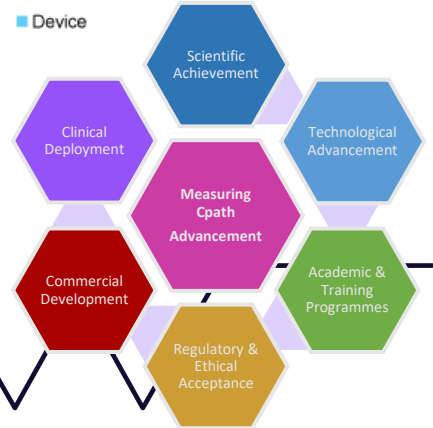
A 510(K) is a premarket submission made to FDA to demonstrate that the device to be marketed is as safe and effective, that is, substantially equivalent, to a legally marketed device (section 513(i)(1)(A) FD&C Act) that is not subject to premarket approval.

Commercial Development

- Interesting questions: What business models work in Cpath?
  - Data to compute
  - Compute to data
  - Integrated frameworks vs. distributed tools
  - Public data lakes vs. proprietary data banks
- Estimating market size

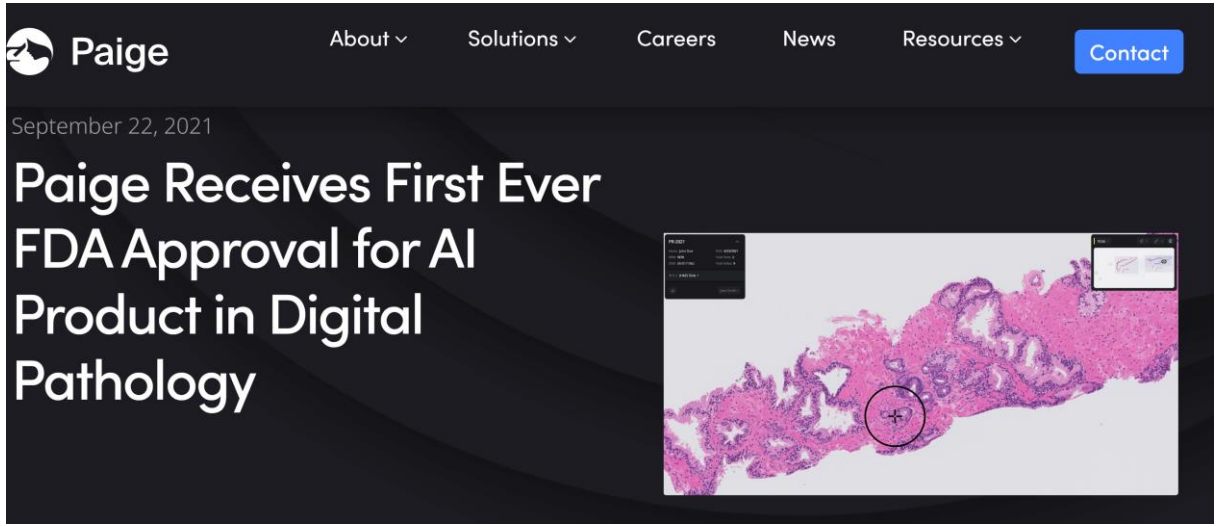
Market size value in 2021	USD 382.3 million
Revenue forecast in 2028	USD 593.9 million





"North America Digital Pathology Market Report, 2021-2028." Accessed May 31, 2021. https://www.grandviewresearch.com/industry-analysis/north-america-digital-pathology-market.

Regulatory and Ethical Approval



- More data vs privacy
- "Black-box"
  - Bias?
- Data Repurposing?
- Competing interests?
  - Using data to sell a product?
- Rapidly changing regulatory landscape
  - European regulation (in vitro diagnostic medical device regulation) 2017/746 after May 26, 2024

Sorell, Tom, Nasir Rajpoot, and Clare Verrill. "Ethical Issues in Computational Pathology." *Journal of Medical Ethics*, March 3, 2021. https://doi.org/10.1136/medethics-2020-107024.
Garcia-Rojo, Marcial, David de Mena, Pedro Muriel-Cueto, Lidia Atienza-Cuevas, Manuel Domínguez-Gómez, and Gloria Bueno. "New European Union Regulations Related to Whole Slide Image Scanners and Image Analysis Software." *Journal of Pathology Informatics* 10 (January 24, 2019): 2. https://doi.org/10.4103/jpi.jpi_33_18.

Education and Training Programs

- How has Cpath improved training of pathologists?
- Improving recruitment into pathology as a discipline

PATHCOMP

Lectures + Hands-on Workshops

Mon – Wed

Hackathon

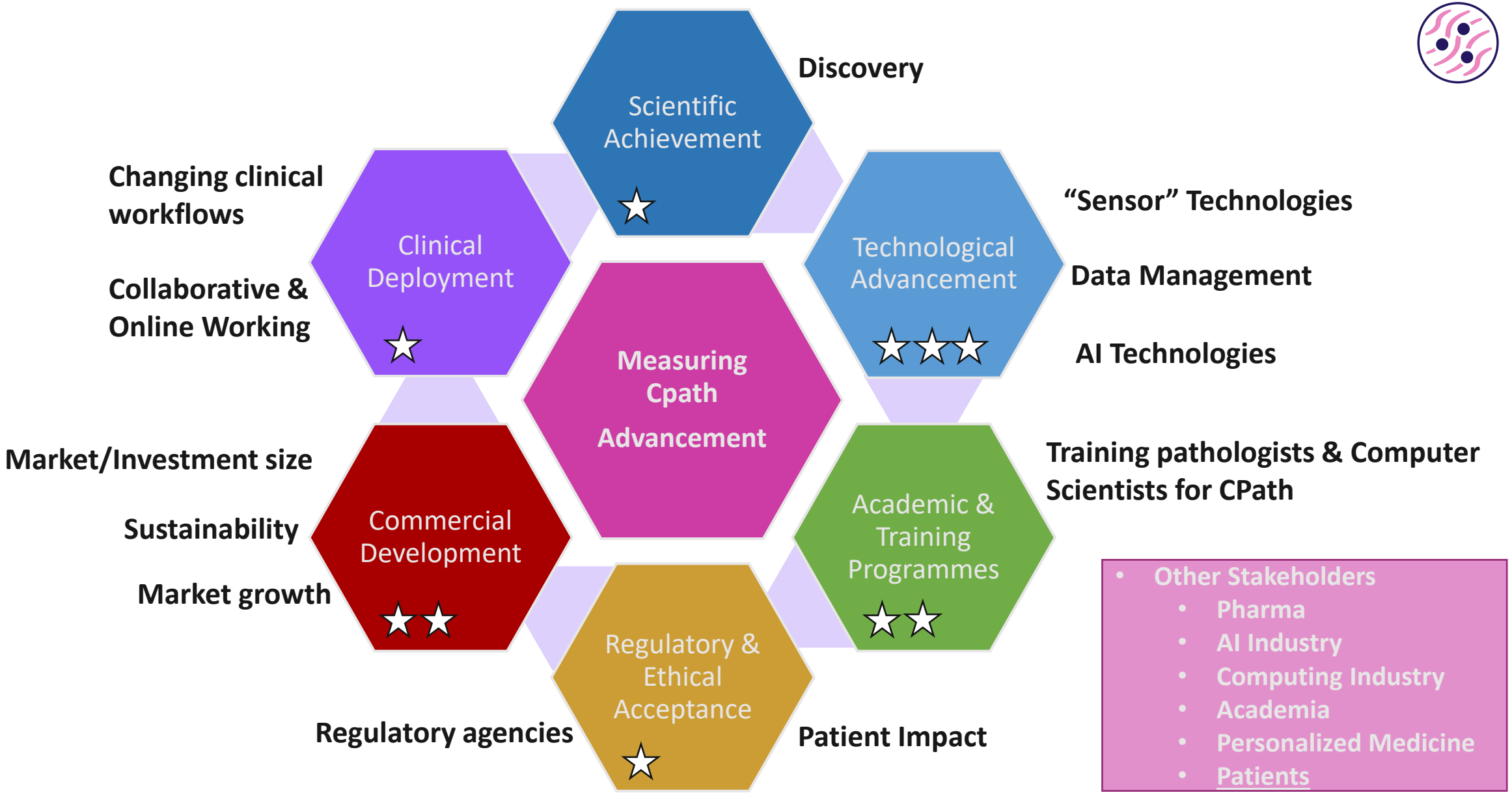
NVIDIA

5pm Wed – 4pm Fri

Short Talks + Hands-on Workshops

Panel Discussion

Industry Exhibitions


Discovery
Scientific Achievement
Changing clinical workflows
Clinical Deployment
Collaborative & Online Working
"Sensor" Technologies
Technological Advancement
Data Management
AI Technologies
Measuring Cpath Advancement
Market/Investment size
Commercial Development
Sustainability
Market growth
Training pathologists & Computer Scientists for CPath
Academic & Training Programmes
Regulatory & Ethical Acceptance
Regulatory agencies
Patient Impact
Other Stakeholders
Pharma
AI Industry
Computing Industry
Academia
Personalized Medicine
Patients