NIDDK-CR Data Science and Meet the Expert Webinar Series

National Institute of Diabetes and Digestive and Kidney Diseases

Central Repository



National Institute of Diabetes and Digestive and Kidney Diseases

Central Repository

About the Series

- Aims to accelerate data science and AI-driven biomedical research by fostering collaboration between biomedical researchers and experts in the field
- Monthly webinar held on the **last Thursday of each month***

Upcoming Webinars

- NIDDK Early Research Partnership Engagement Challenge InfoSession – November 6th
  - Register here
- NLM SPARK dbGaP and PubMed Challenges InfoSession – December 8th
- Fusion Tool Demo – Spring 2027

Learn more about the webinar series, register for future webinars, and access past webinars materials and recordings

Join NIH Prize Competitions on the NIDDK CMP

The NIDDK Challenge Management Platform (CMP) is a centralized hub for NIH-supported prize competitions and data challenges that invite innovative ideas, methods, and partnerships to advance biomedical research.

FEATURED OPPORTUNITIES:

NIDDK Early Research Partnership Engagement Challenge

Compete for up to $1 million in total prizes by developing innovative strategies for early community engagement and building academic–community partnerships to shape future NIDDK research; register by November 30, 2026.

NLM SPARK (dbGaP and PubMed/PMC) Challenges

Compete for $500,000 in total prizes by developing AI tools to improve biomedical literature search and synthesis in PubMed/PMC and research data discovery in dbGaP; solutions due January 15, 2027.

WHY PARTICIPATE:

Advance discovery: Help solve barriers to finding biomedical knowledge and developing relevant research.

Put ideas into practice: Develop AI tools or community engagement strategies with real-world applications.

Build stronger partnerships: Bring technical, research, and community expertise together around shared goals.

Earn recognition and prizes: Compete for monetary awards and NIH recognition of winning solutions.

NIH National Institute of Diabetes and Digestive and Kidney Diseases

Central Repository

Explore current challenges and learn how to participate:

National Institute of Diabetes and Digestive and Kidney Diseases

Central Repository

The Analytics Workbench V1.0

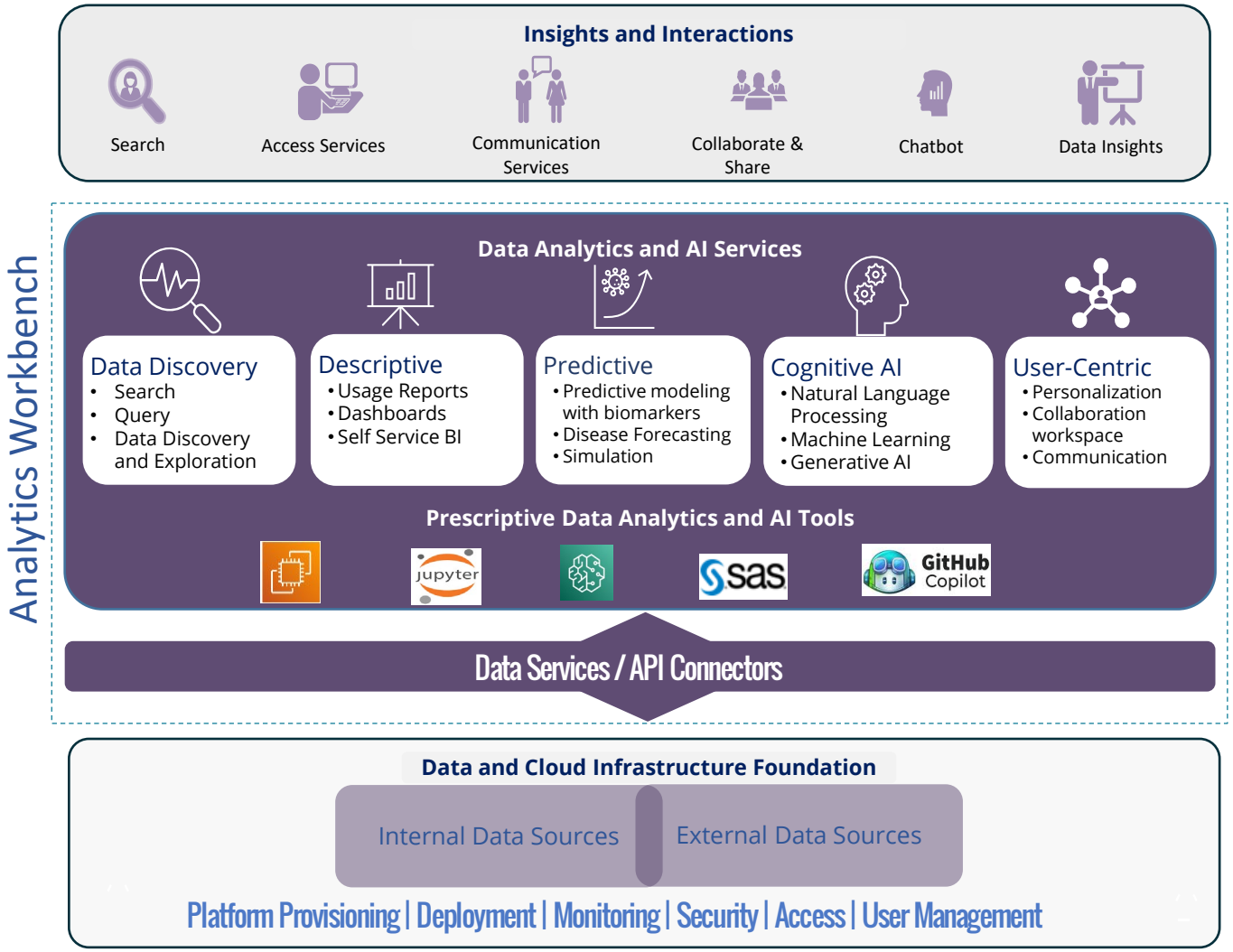
Streamlining *end-to-end data science lifecycle and discovery of data-driven biomedical insights.*

Innovation and ease of use

A cloud-based analytics environment where researchers and data scientists can access a suite of integrated analytics tools and cloud computing resources to participate in data challenges and AI innovation.

Expected Benefits of Analytics Workbench:

- **Promote Collaboration**
- **Support AI Innovation**
- **Minimize Data Movement**
- **Improve User Experience**
- **Discover Data Insights**
- **Advance NIDDK Research Mission**



Analytics Workbench is integrated with the *NIDDK-CR R4R Platform* and *Challenge Management Platform* offerings. More information at https://repository.niddk.nih.gov/home

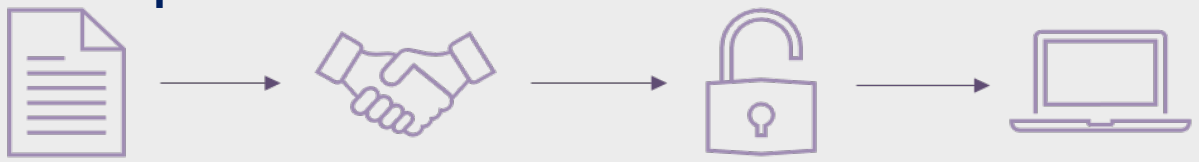
National Institute of Diabetes and Digestive and Kidney Diseases

Central Repository

Co-Analysis Pilot

A streamlined, joint access pathway unlocking cross-repository data science across metabolic, renal, digestive, and aging research

How the pilot works:



Submit one request → **Single joint DAC review** → **Access both repositories** → **Analyze in secure Workbench**

Cross-Repository Study Collections:

NIDDK CR

LookAHEAD; DPP; DPPOS; DCCT; EDIC; CRIC

NIA AgingResearchBiobank

LIFE; ENERGISE; SWAN; CALERIE; ASPREE; MOST; SOF; STRIDE; Testosterone Trials

- Larger, more diverse combined cohorts; greater statistical power and generalizability
- Well-characterized, longitudinal phenotyping linked to specimens across midlife–older age
- Enables novel cross-domain questions

Built for Co-Analysis, Not Just Co-Access

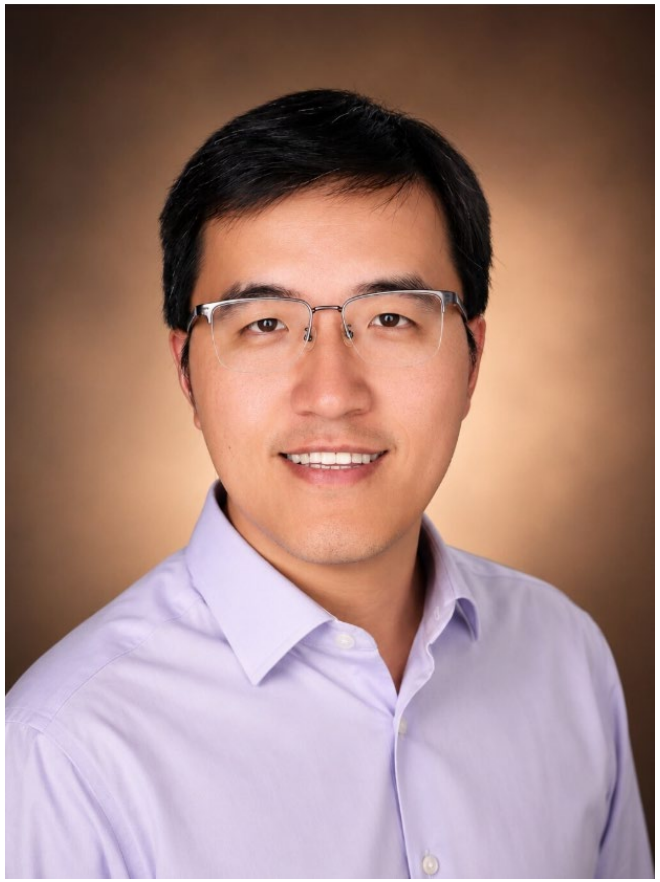
- Approved investigators can analyze data from both repositories together inside the secure NIDDK CR R4R Analytics Workbench, ***supporting integrated modeling, feature extraction, and AI/ML-driven pattern discovery without moving data out of a compliant environment.***

National Institute of Diabetes and Digestive and Kidney Diseases

Central Repository

Meet the Experts

Dr. Yuankai Huo is an Associate Professor of Computer Science, Electrical and Computer Engineering at Vanderbilt University, where he directs the Biomedical Data Representation and Learning Lab (HRLB Lab). He also holds a secondary faculty appointment in Pathology, Microbiology, and Immunology at Vanderbilt University Medical Center. His research focuses on high-dimensional multimodal data computing, generative and foundation AI, and medical computer vision. Dr. Huo has received honors including the Charles E. Ives Journal Award, the Early Career Achievement Award from the Society for Imaging Informatics in Medicine, and an NSF NAIRR Pilot Award. He is a Senior Member of IEEE and a lifetime member of SPIE, and has served as an organizing committee member and area chair for MICCAI, MIDL, and ISBI, including as Program Chair for MIDL 2026. He has served as lead PI on NIH R01 and NSF Medium projects, bridging next-generation computing and engineering.



VANDERBILT UNIVERSITY

Vanderbilt LAB FOR Immersive AI Translation

HRLB

Biomedical Data Representation and Learning Lab (HRLB)

From Pixels to Phenotypes: Scalable and Explainable Tools for Whole Slide Imaging Pathomics

NIDDK-CR Data Science and Meet the Expert Webinar Series

NIH National Institute of Diabetes and Digestive and Kidney Diseases

NIDDK Central Repository

Yuankai Huo, PhD

Associate Professor

Department of Computer Science (Primary)

Department of Electrical and Computer Engineering

Data Science Institute

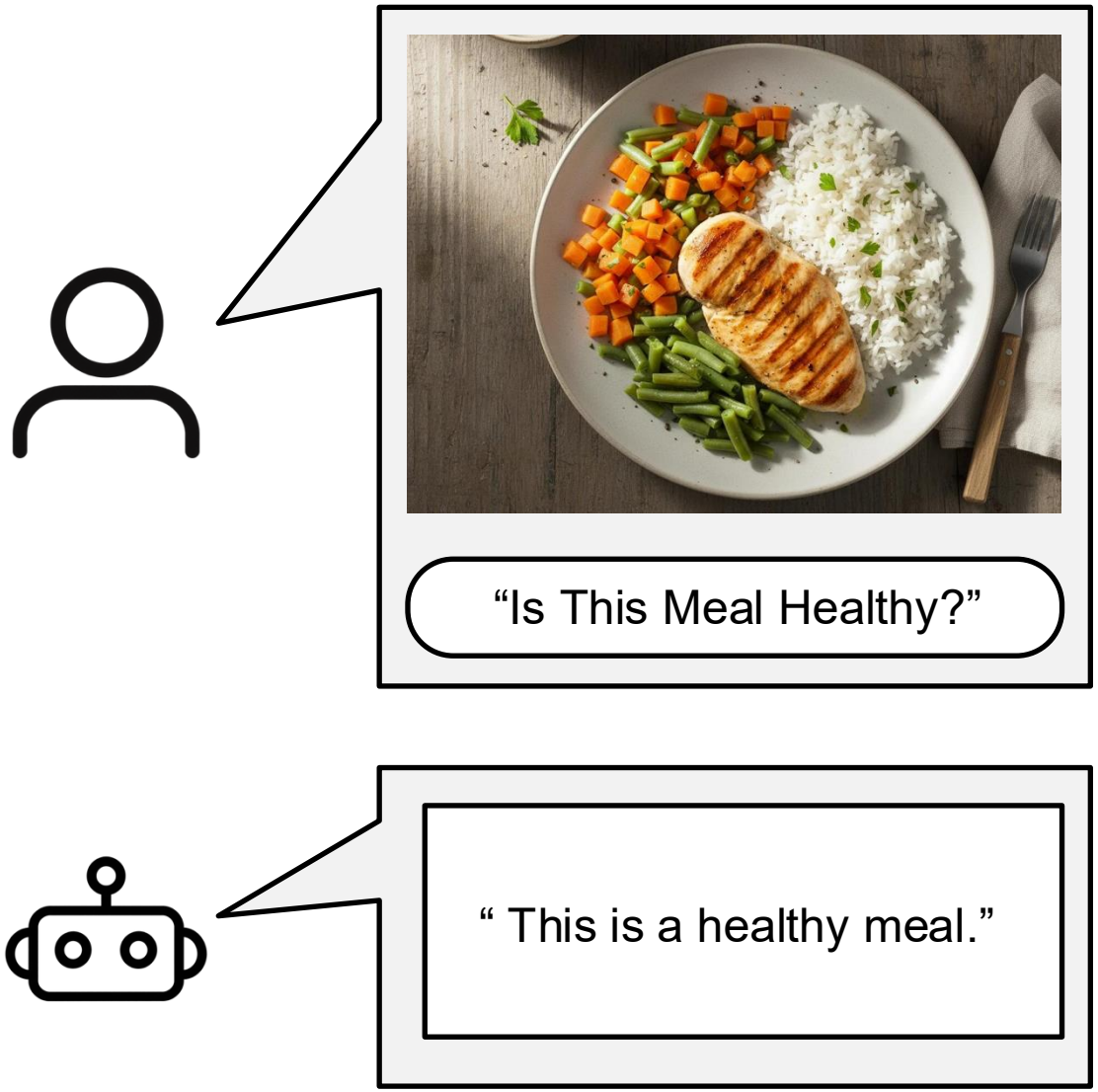
Vanderbilt University

Affiliated Faculty Member

Department of Pathology, Microbiology and Immunology

Vanderbilt University Medical Center

AI can give an answer.



Can we trust? 🤔

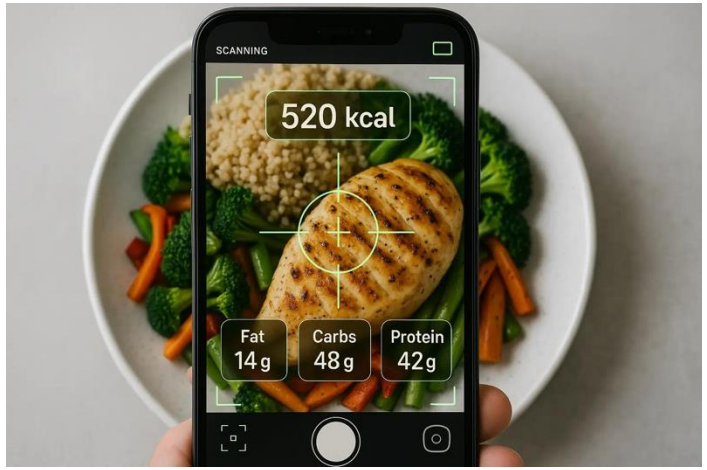
People want explainability



“Is This Meal Healthy?”

“ This is a healthy meal.”





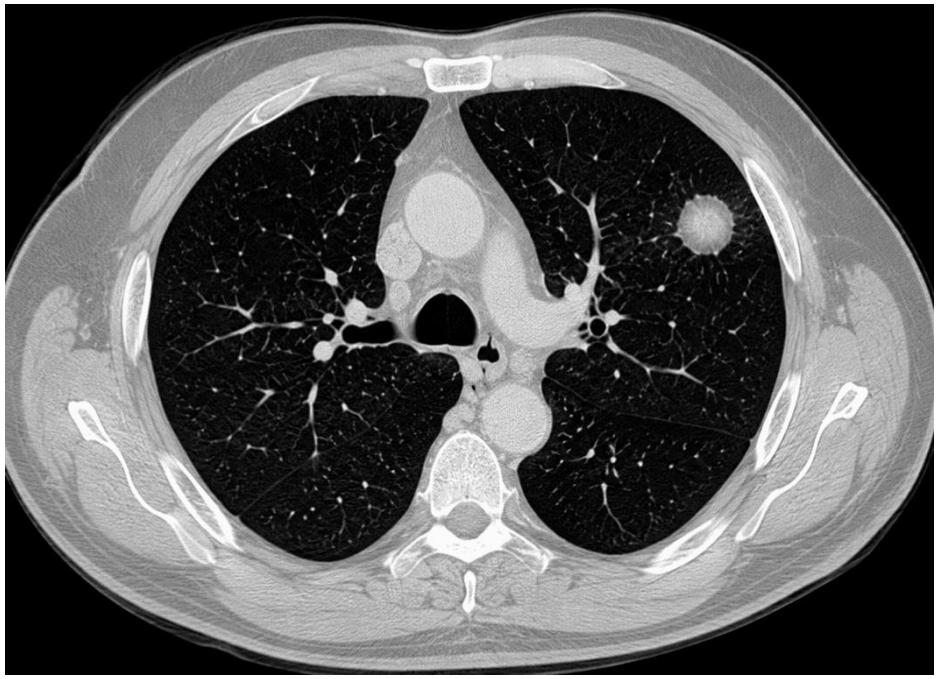
“Rice: 180 g; chicken: 140 g; vegetables: 100 g.”

“Estimated calories: 650 kcal.”

“ This is a healthy meal.”

AI/SI in Medical Image Analysis

Recent advances in artificial intelligence (AI), also referred to as super intelligence (SI), allowing AI-assisted diagnosis from medical images.



AI/SI

"This patient has high lung cancer risk."

Can we trust that decision?

Today's AI Models are largely Black Boxes

Current



Image



Black Box

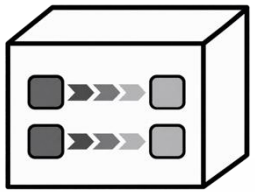
Prediction

"Lung cancer"

Ideally



Image

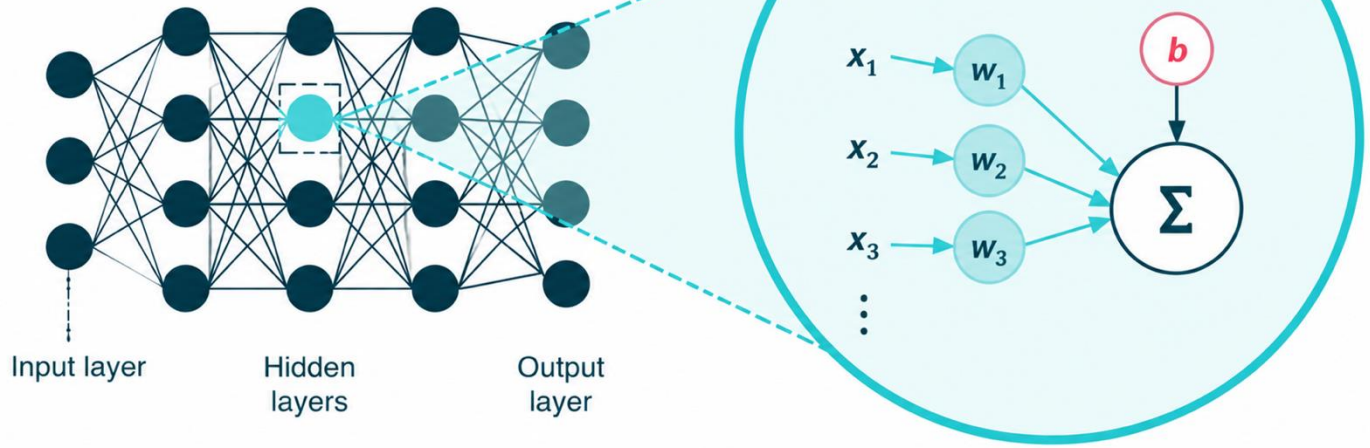


Clear Box

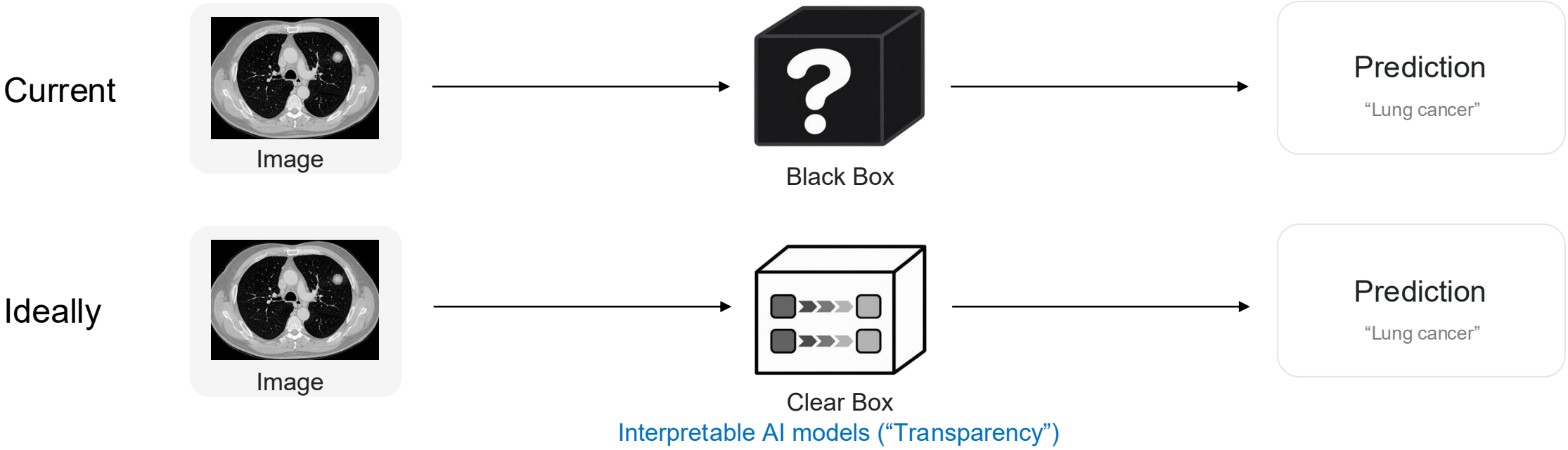
Prediction

"Lung cancer"

Interpretable AI models ("Transparency")



How to interpret?

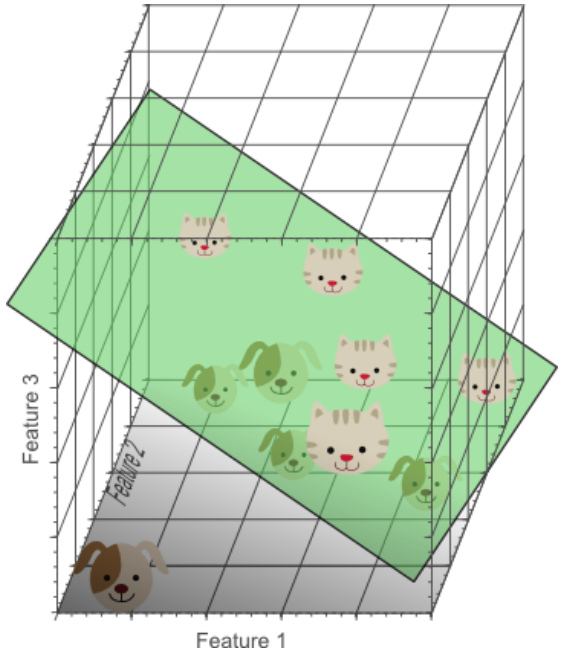


However, transparent Deep Learning is Extremely Difficult!

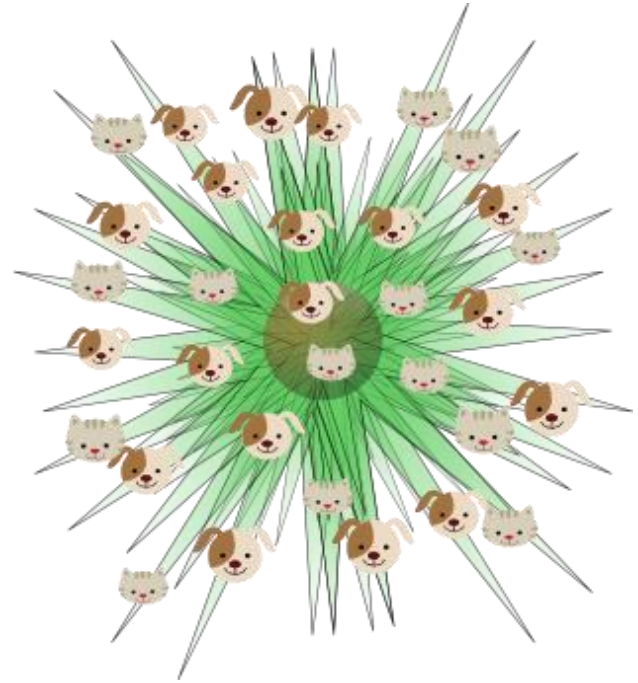
Curse of Dimensionality



Richard Bellman 1957



Dimension = 3



Dimension = billion?

The **curse of dimensionality** is one of the fundamental challenges to making AI transparent and explainable.

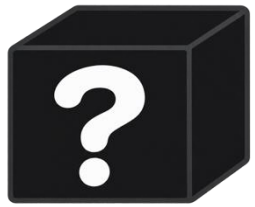
How Can We Get the Explainability



Current



Image



Black Box

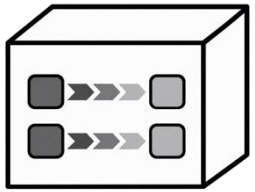
Prediction

"Lung cancer"

Ideally



Image



Clear Box

Interpretable AI models ("Transparency")

Prediction

"Lung cancer"

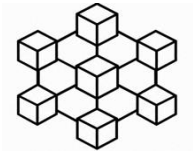
In the middle



Image



Still Black Box Model



Evidence

Interpretable evidence ("Explainability")

Prediction

"Lung cancer"

Evidence-based AI



"Which Product Is Better?"

"Suitcase A is better for your trip."



"Both are segmented as the red boxes"

"Left one has dimensions: 55 × 35 × 23 cm; capacity: 38 L."

"Suitcase A is better for your trip."

Evidence is Essential

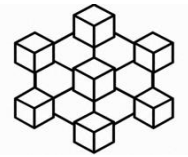




Still Black Box Model

Segmentation AI

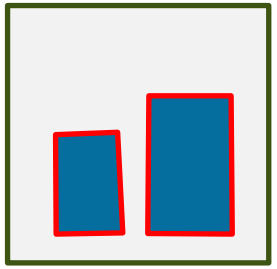


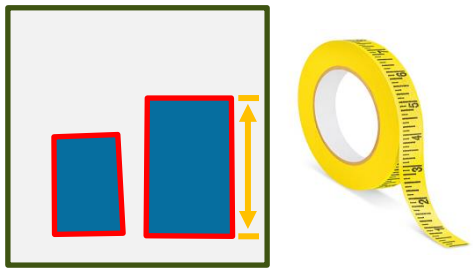


Evidence

dimensions: 55 × 35 × 23 cm
capacity: 38 L

"Suitcase A is better for your trip.."

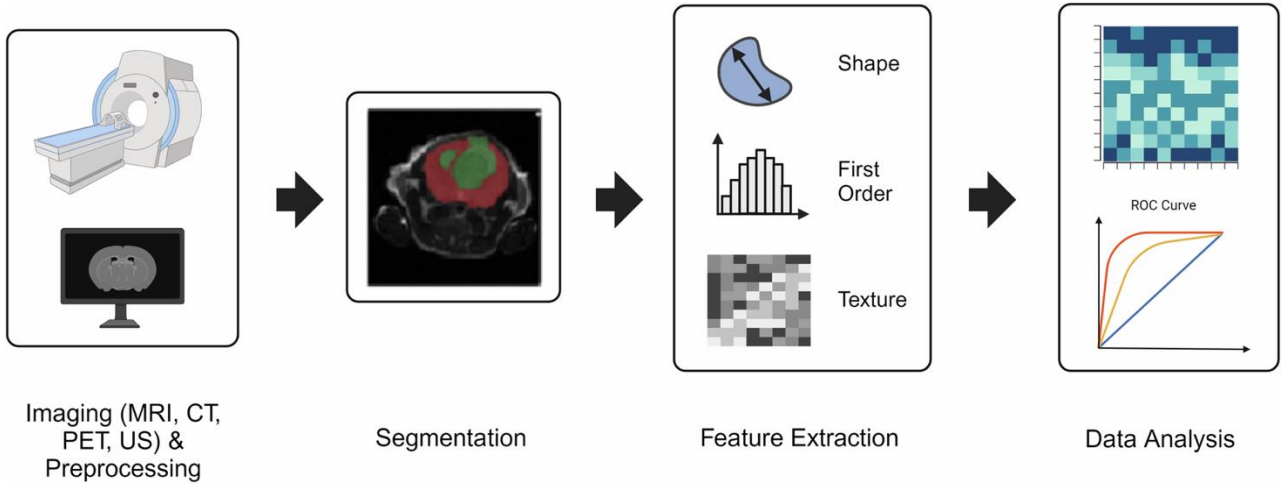




Evidence-based Decision

In Healthcare?

Radiomics



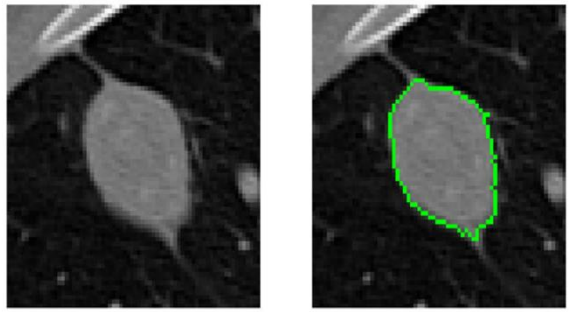
Riberdy et al.
NPJ Imaging 2025



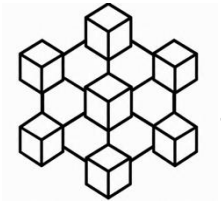
Image



Still Black Box Model



Seeing is believing



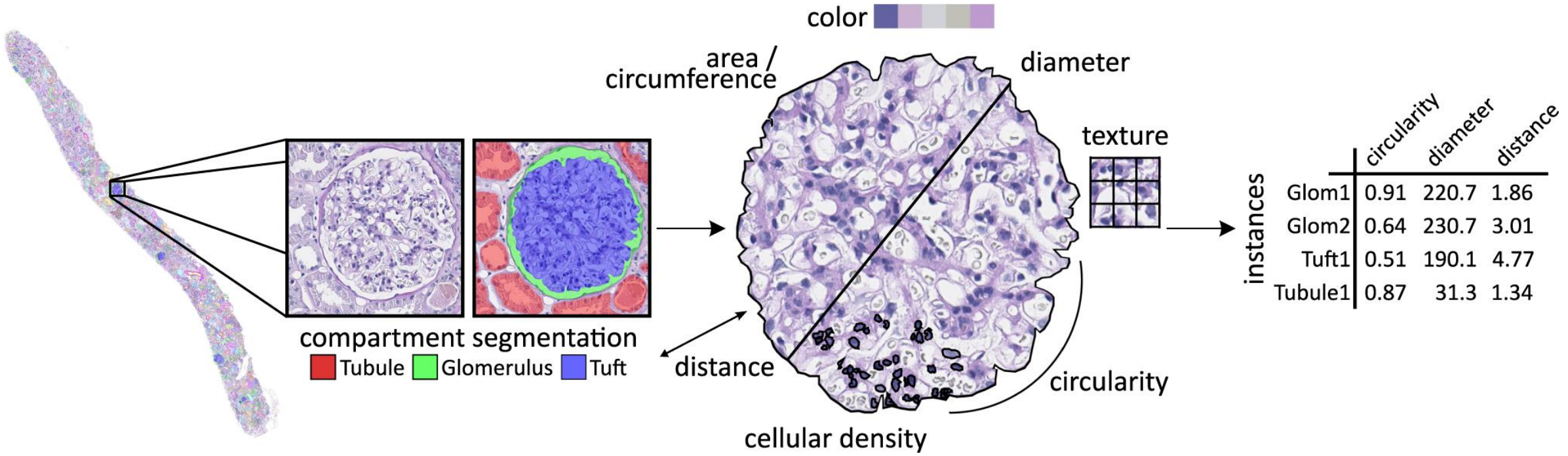
Evidence

Evidence-based
Clinical Decision

Segmentation AI

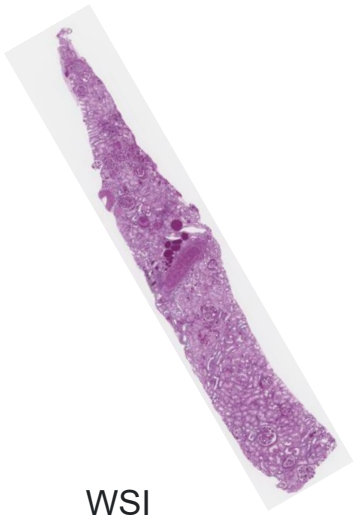
Nodule diameter 12 mm
Nodule volume 650 mm³
Volume growth +35%

Pathomics

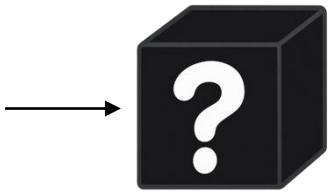


Bülow et al.
NPJ Systems Biology and Applications 2023

Pathomics for Kidney Pathology

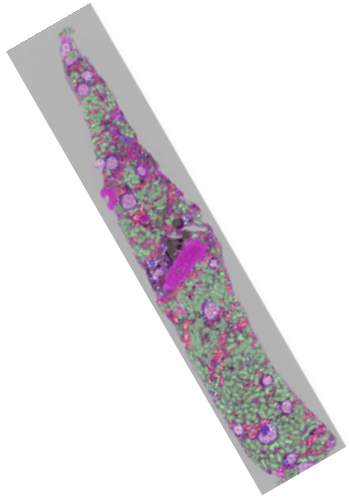


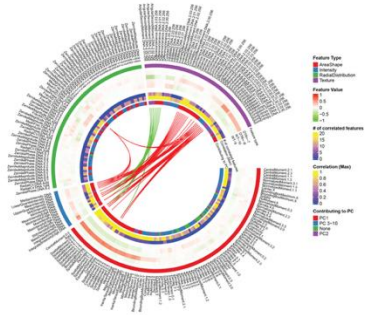
WSI



Still Black Box Model

Segmentation AI





Evidence

Pathomics



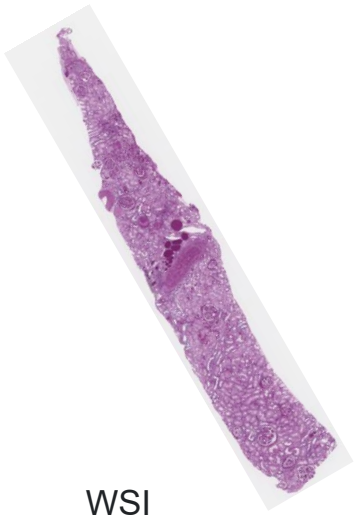
Analyze

Challenge 1

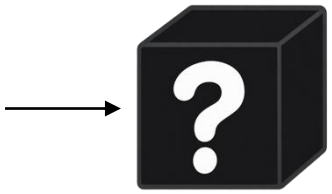
Challenge 2

Challenge 3

Kidney Pathology

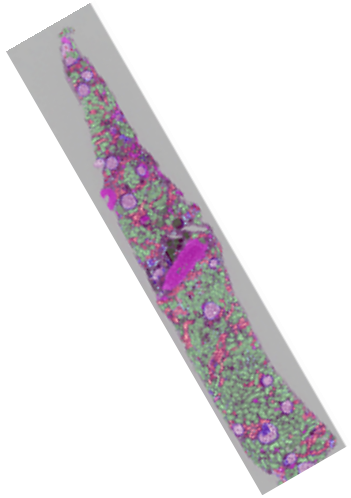


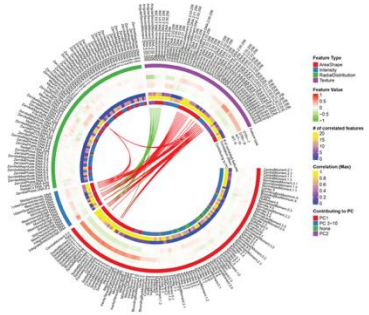
WSI



Still Black Box Model

Segmentation AI





Evidence

Pathomics



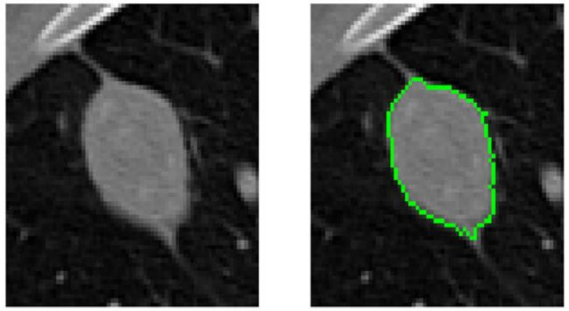
Analyze

Challenge 1

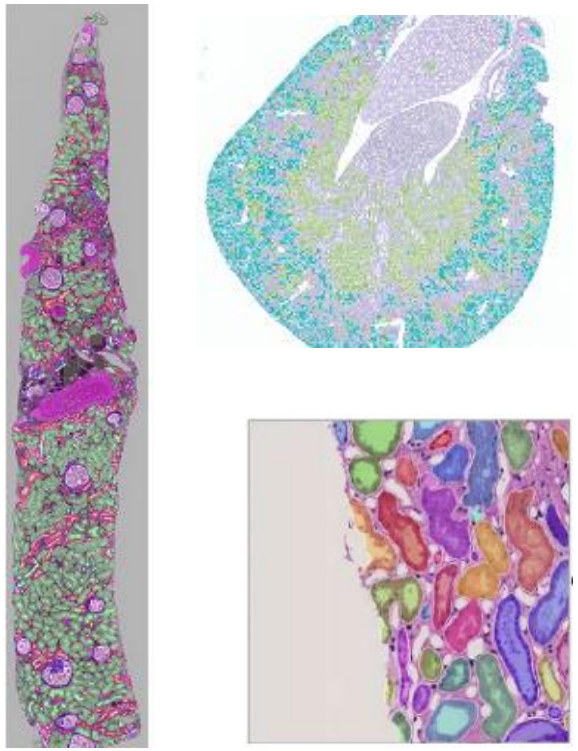
Challenge 2

Challenge 3

Challenge 1: N x M times more complicated



Radiomics:
Single Object Single Type



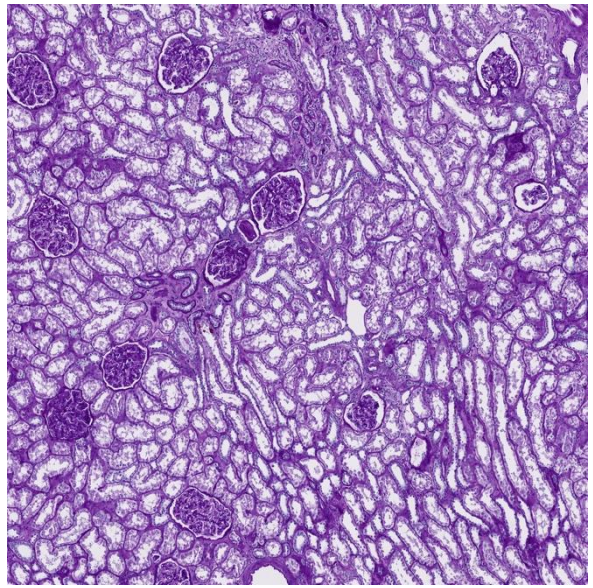
Pathomics:
N Objects M Types

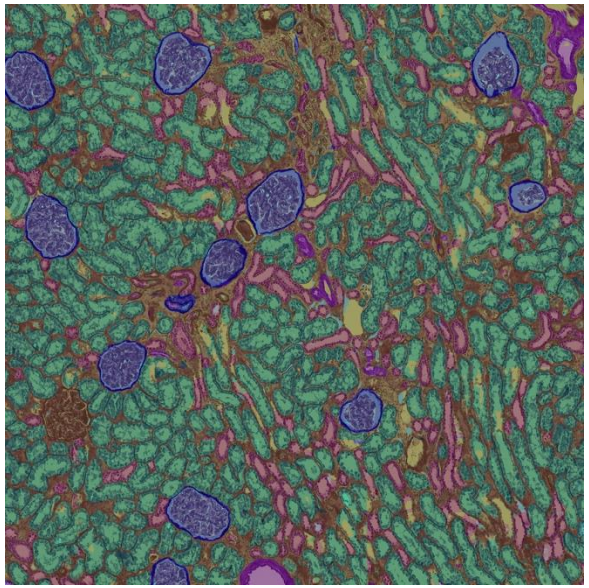
Challenge 1: tools for segmentation

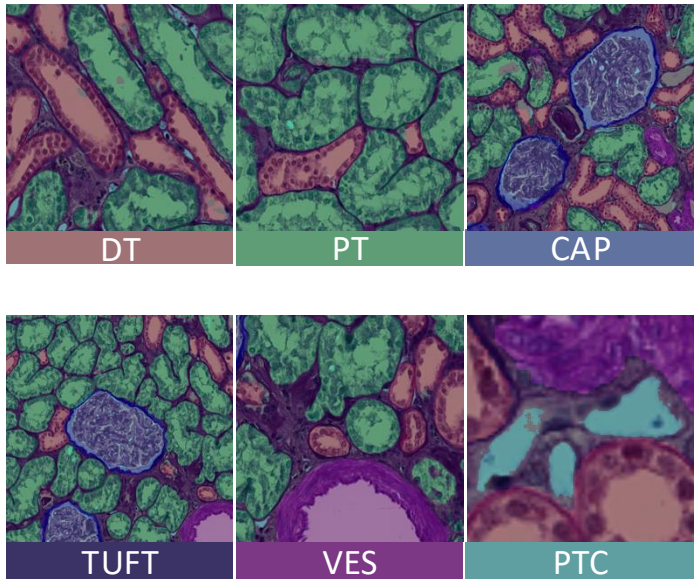
Densely segmentation
(Testing stage)

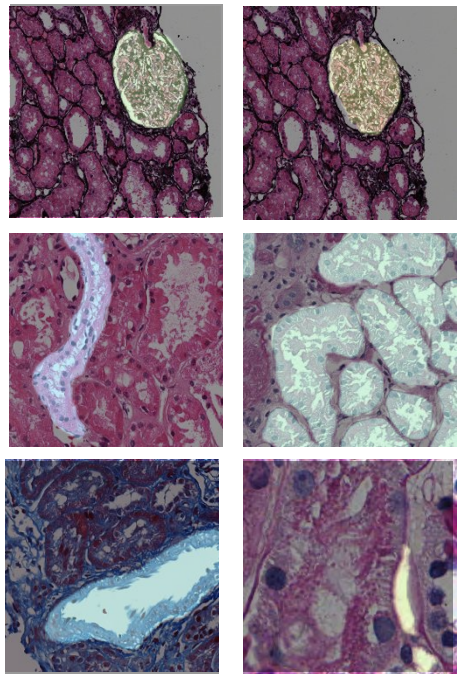
Partially labelled
(Training stage)

Pathological image	Aggregation	Aggregation patches



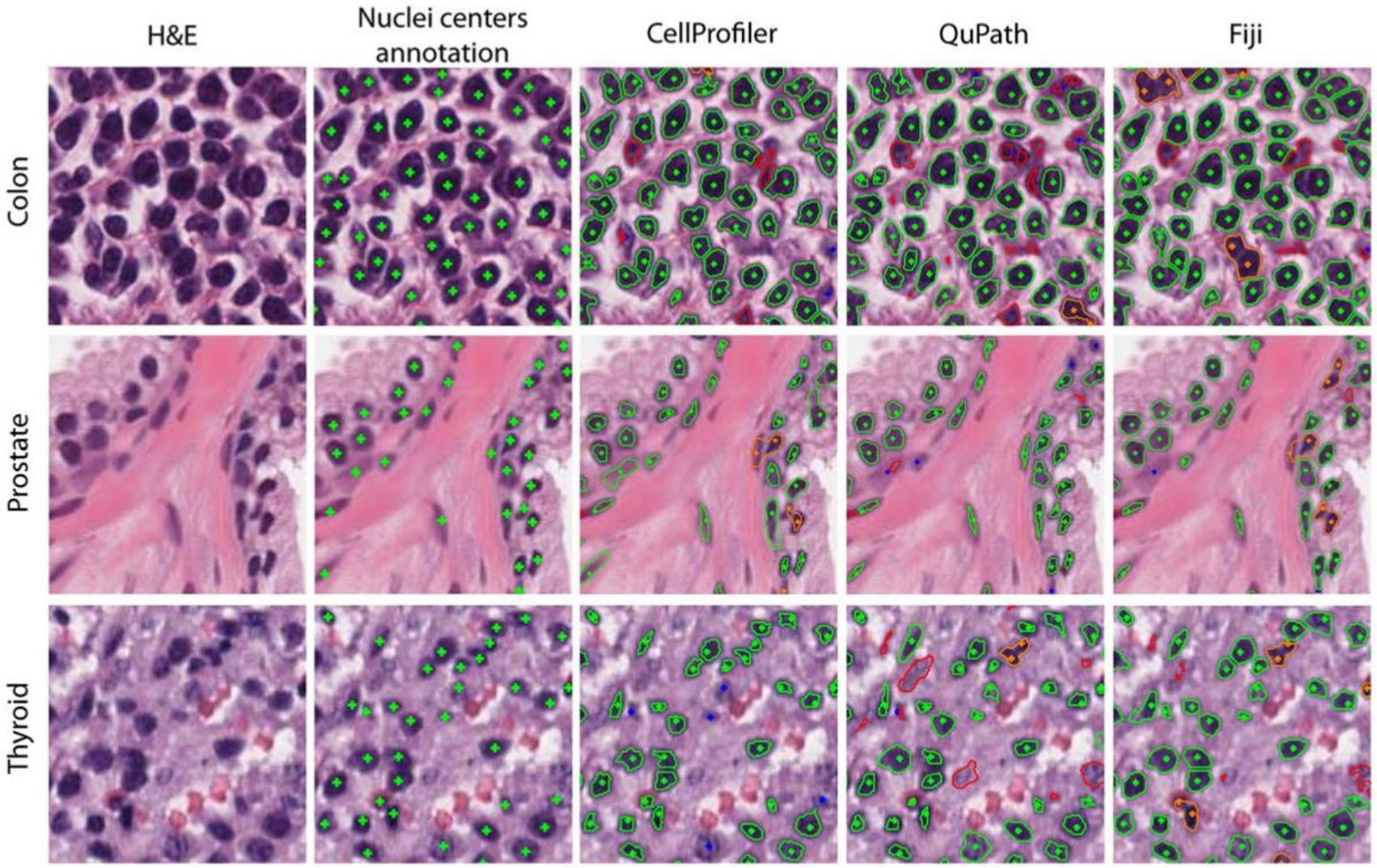






"Dense" tissue segmentation from typically "partially labelled" training images

Traditional Tools

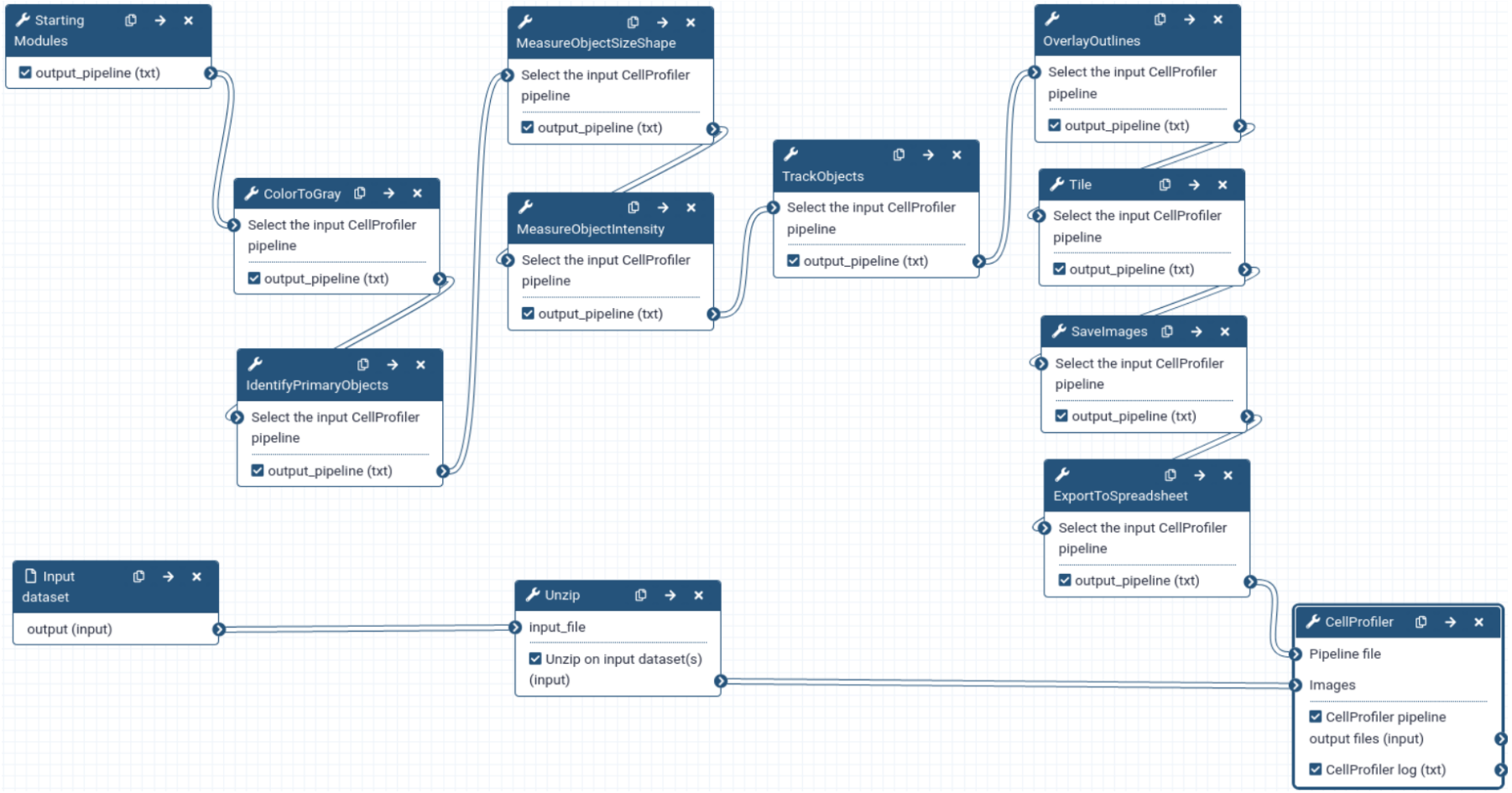


CellProfiler™
cell image analysis software

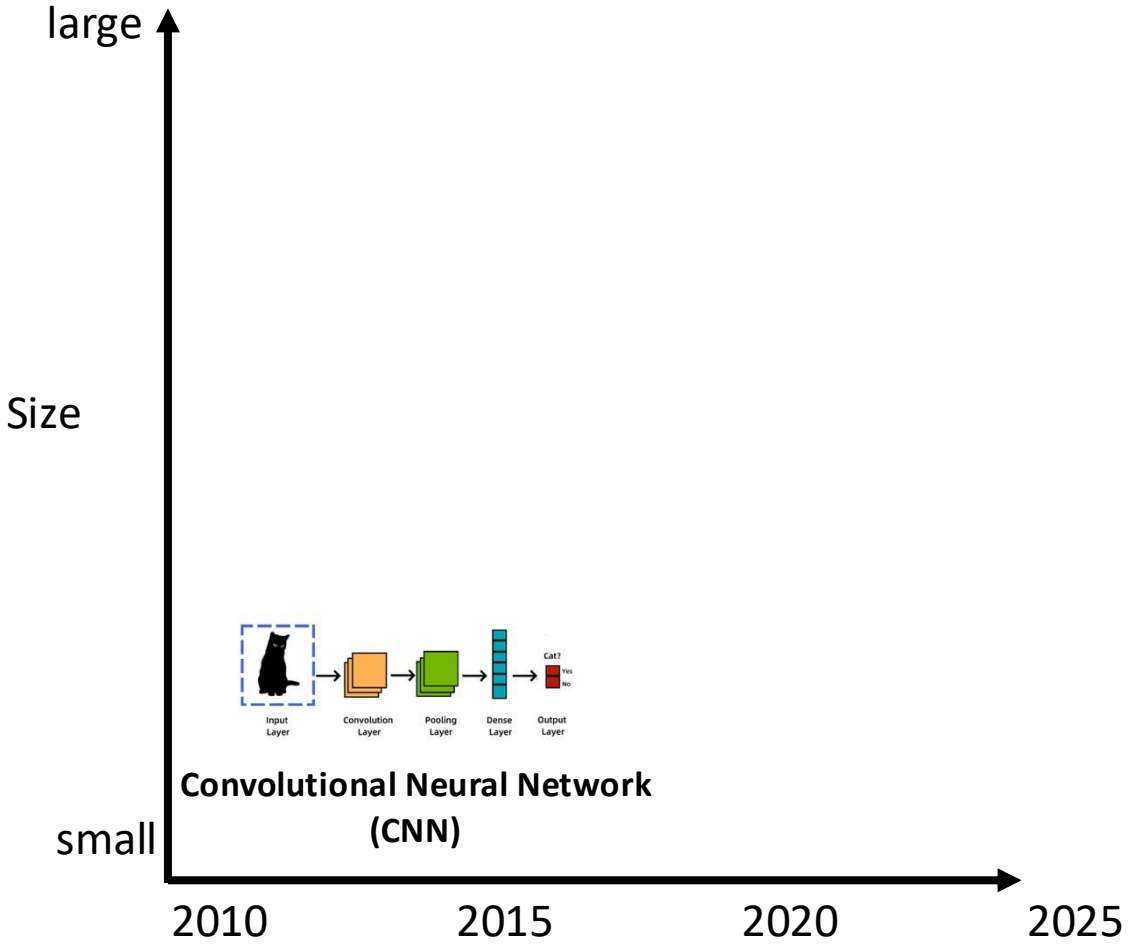
QuPath

ilastik
Object Classification

Workflow

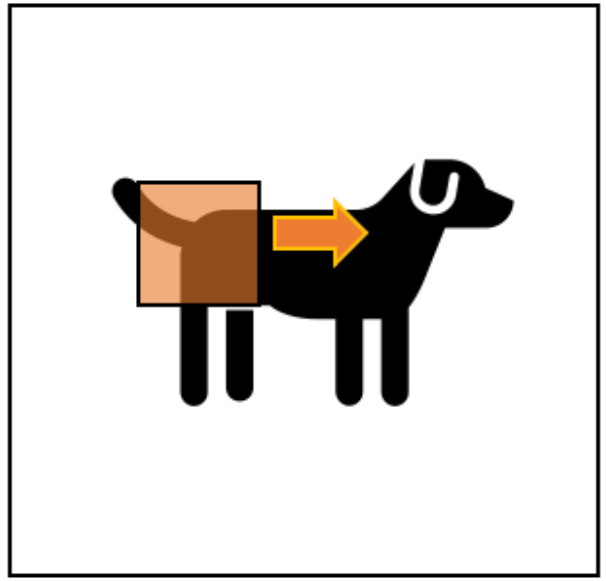

Starting Modules
output_pipeline (txt)
ColorToGray
Select the input CellProfiler pipeline
output_pipeline (txt)
IdentifyPrimaryObjects
Select the input CellProfiler pipeline
output_pipeline (txt)
MeasureObjectSizeShape
Select the input CellProfiler pipeline
output_pipeline (txt)
MeasureObjectIntensity
Select the input CellProfiler pipeline
output_pipeline (txt)
TrackObjects
Select the input CellProfiler pipeline
output_pipeline (txt)
OverlayOutlines
Select the input CellProfiler pipeline
output_pipeline (txt)
Tile
Select the input CellProfiler pipeline
output_pipeline (txt)
SaveImages
Select the input CellProfiler pipeline
output_pipeline (txt)
ExportToSpreadsheet
Select the input CellProfiler pipeline
output_pipeline (txt)
Input dataset
output (input)
Unzip
input_file
Unzip on input dataset(s) (input)
CellProfiler
Pipeline file
Images
CellProfiler pipeline output files (input)
CellProfiler log (txt)

Deep Learning Methods



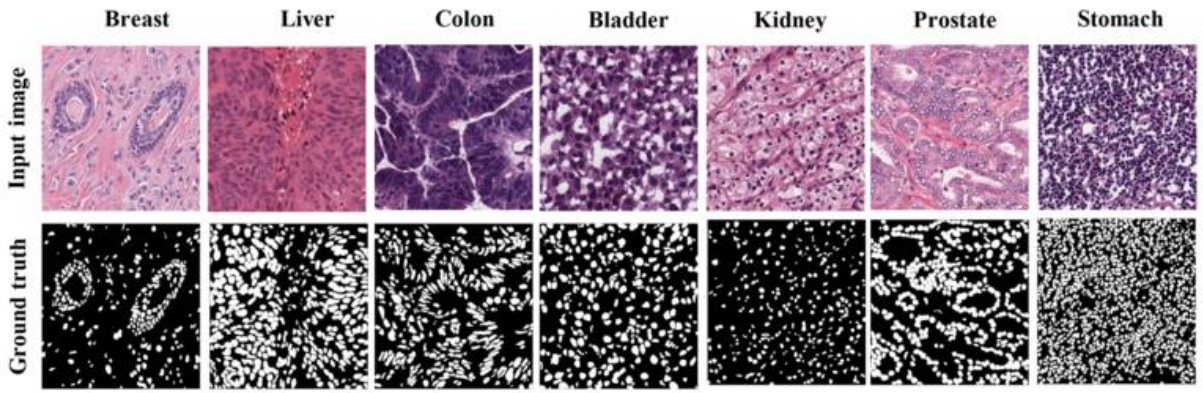
https://www.basic.ai/blog-post/a-comprehensive-guide%3A-what-are-convolutional-neural-networks
https://paperswithcode.com/method/vision-transformer
https://www.v7labs.com/blog/segment-anything-model-sam

CNN

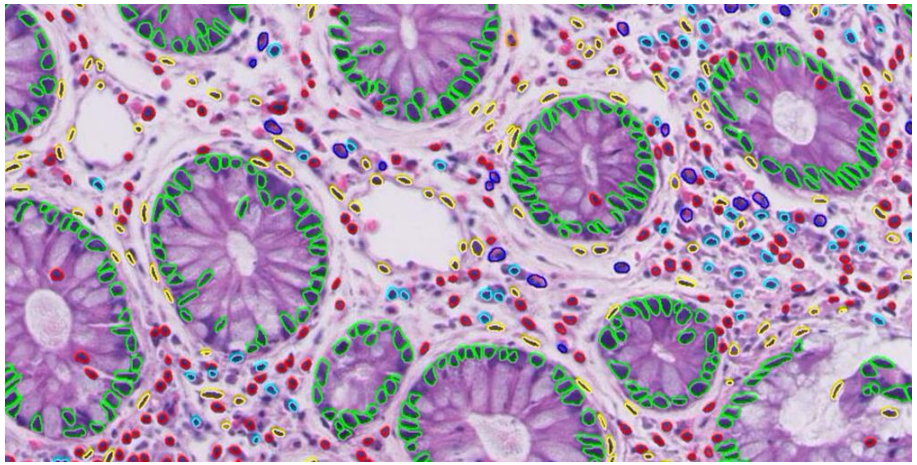


https://medium.com/towards-data-science/recent-developments-and-views-on-computer-vision-x-transformer-ed32a2c72654

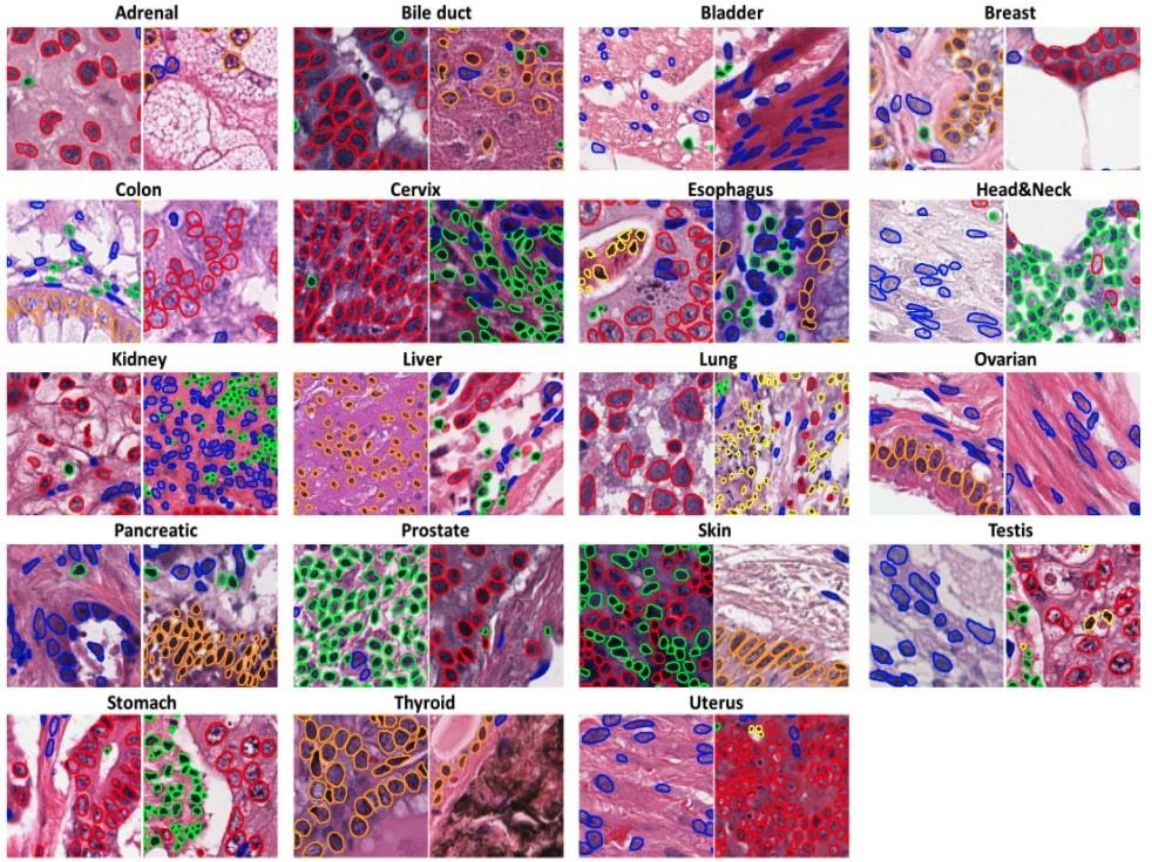
Workflow →Example based: Deep Learning



MonuSeg 2018

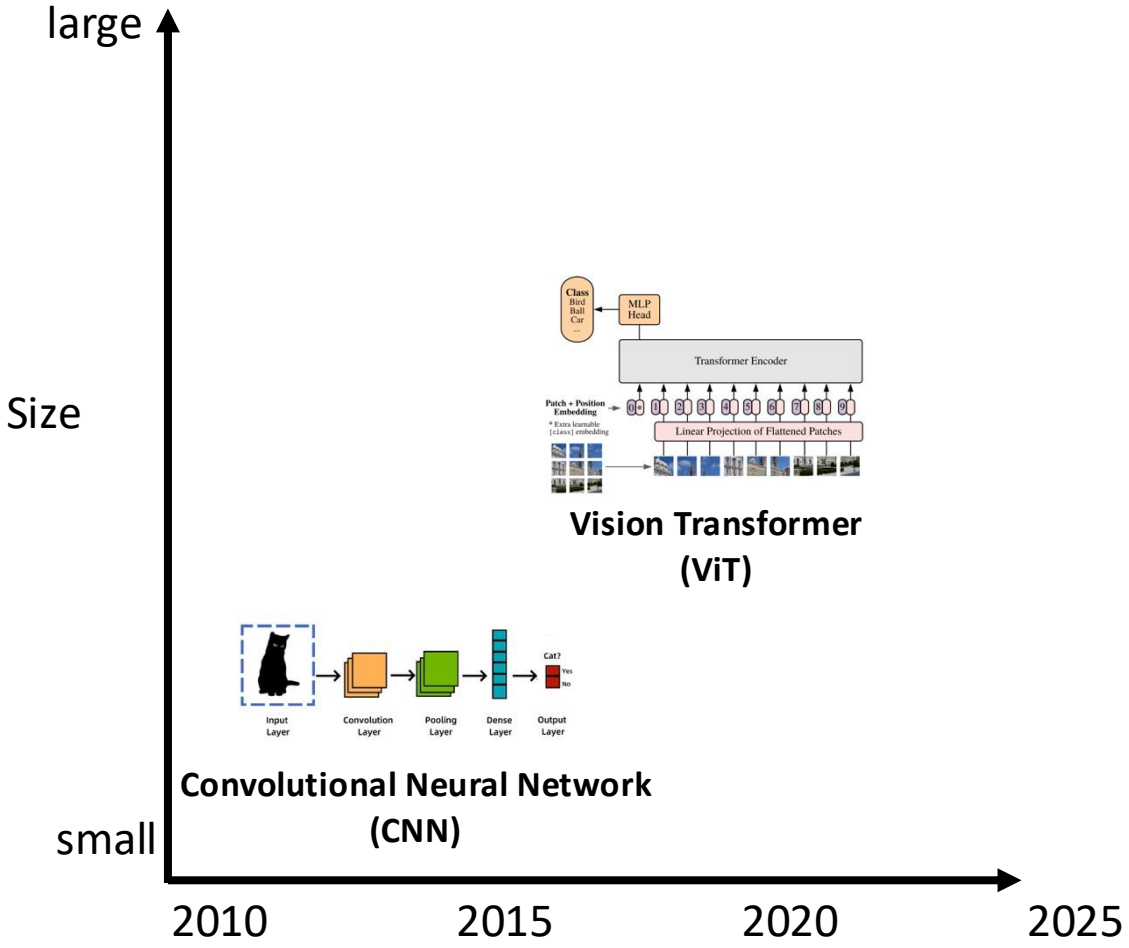


CoNIC 2022



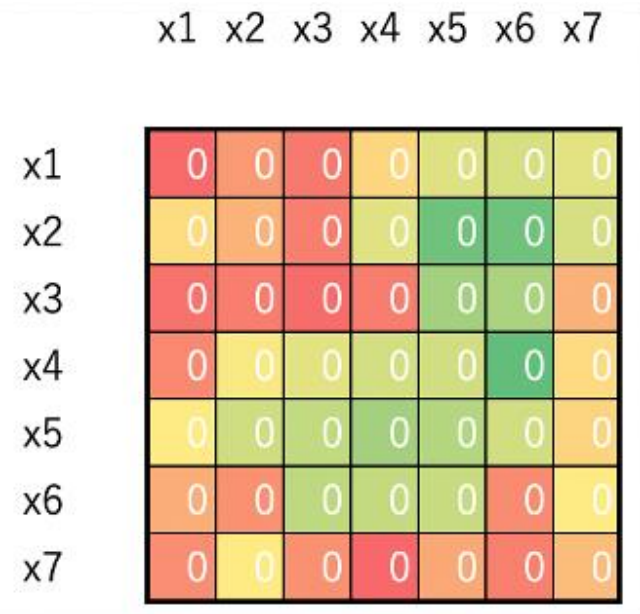
PanNuke 2020

CNN to Transformer



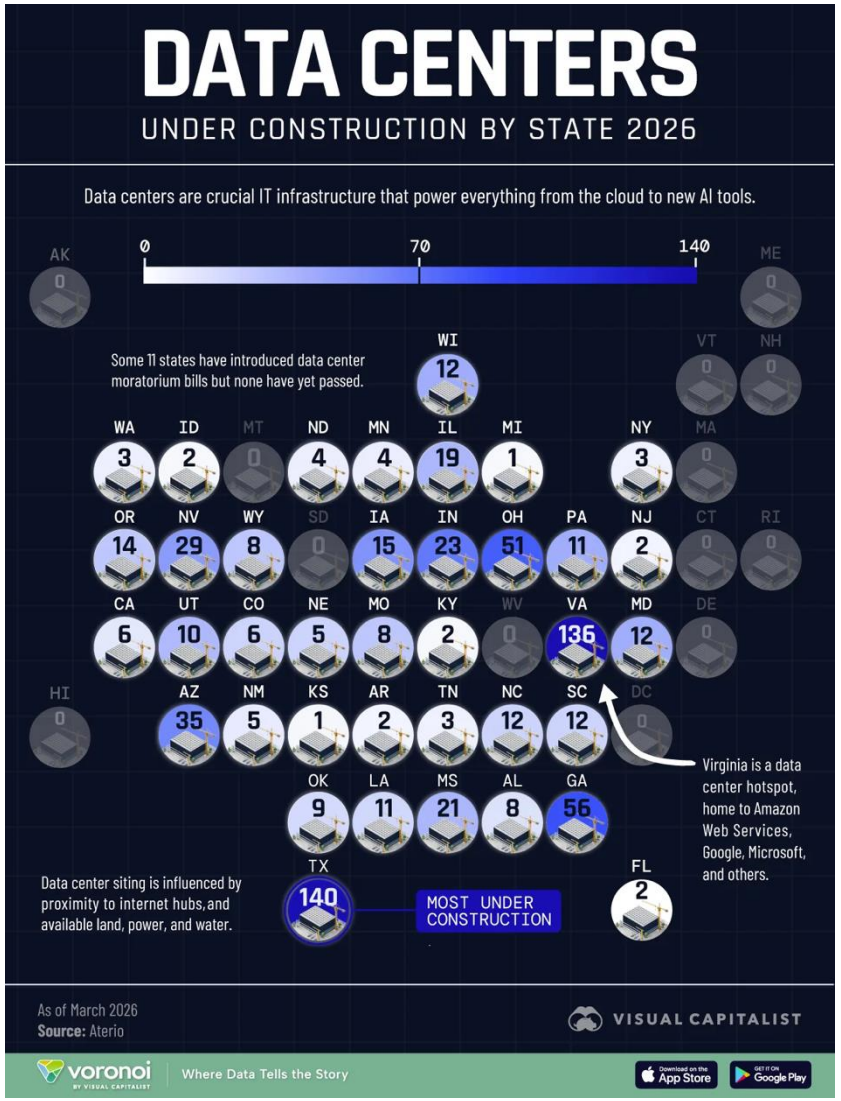
https://www.basic.ai/blog-post/a-comprehensive-guide%3A-what-are-convolutional-neural-networks
https://paperswithcode.com/method/vision-transformer
https://www.v7labs.com/blog/segment-anything-model-sam

Self Attention



https://medium.com/towards-data-science/recent-developments-and-views-on-computer-vision-x-transformer-ed32a2c72654

If we have infinite data and resource?



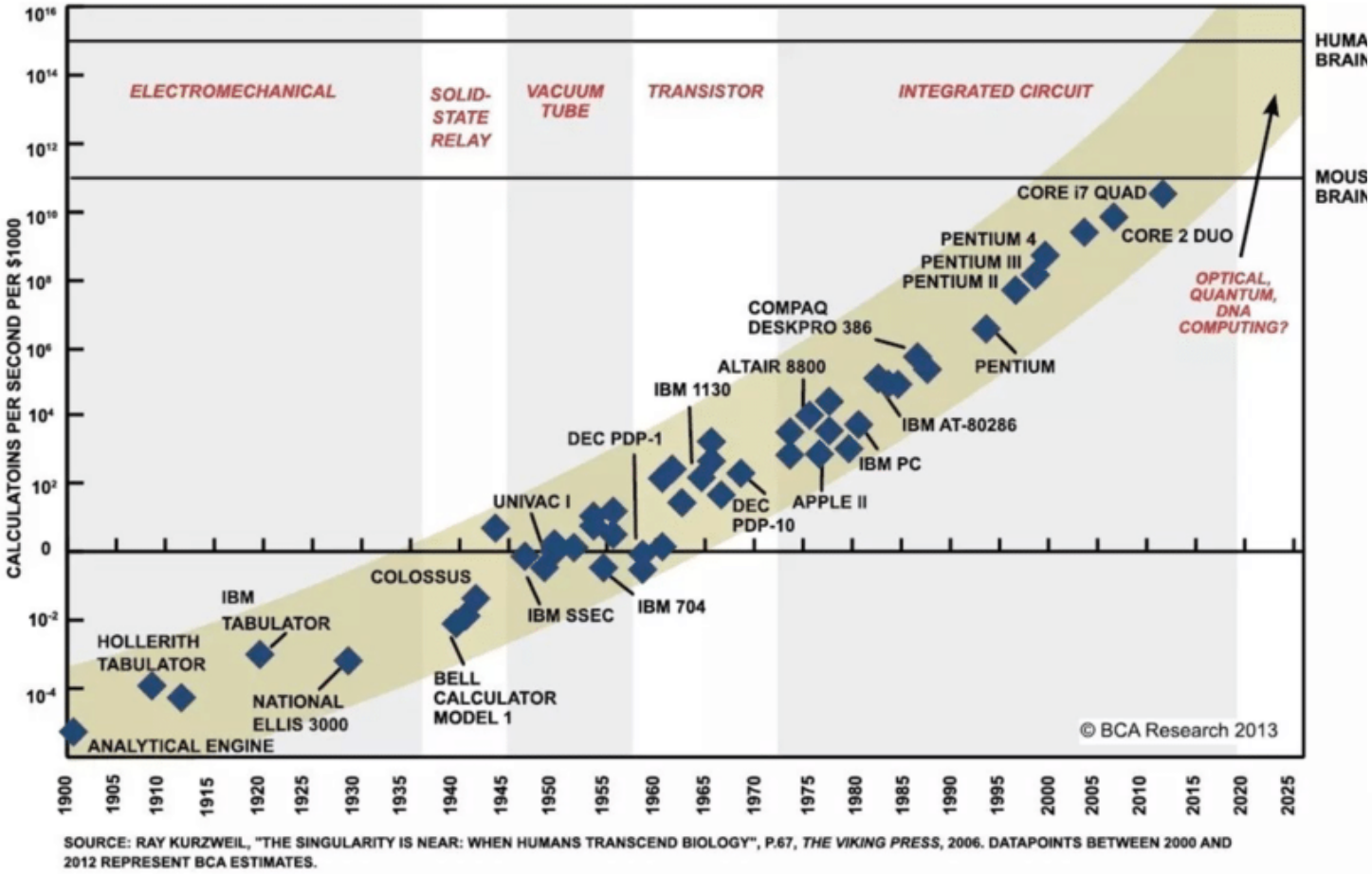




https://www.visualcapitalist.com/mapped-americas-data-center-construction-boom/ https://makeagif.com/gif/inside-a-google-data-center-cm4RCv
https://www.lviassociates.com/en-us/industry-insights/hiring-advice/data-centers-are-forcing-a-new-power-debate

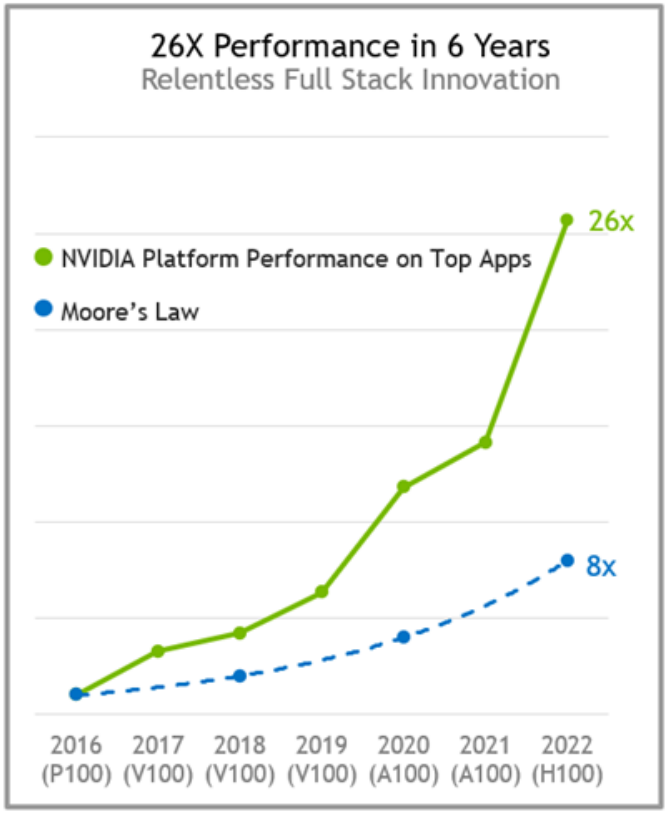
Scaling Law

CPU



SOURCE: RAY KURZWEIL, "THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY", P.67, *THE VIKING PRESS*, 2006. DATAPOINTS BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.

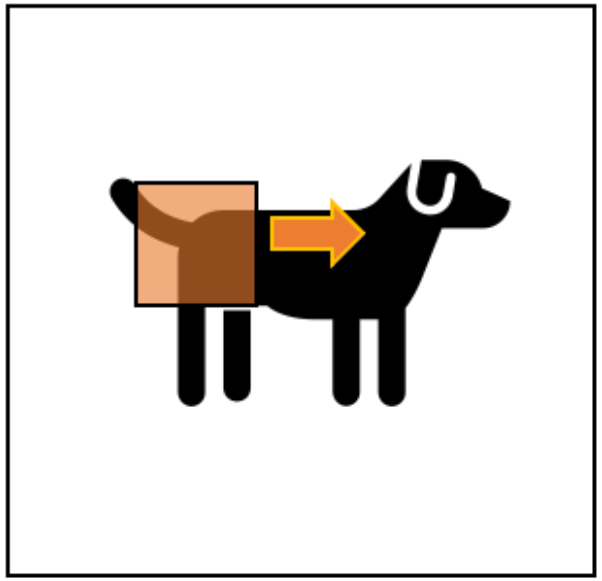
GPU



https://www.nextbigfuture.com/2011/10/kurzweil-discusses-disruptive-technologies-mind-boggling-predictions-and-dangerous-ideas.html

Scale-up

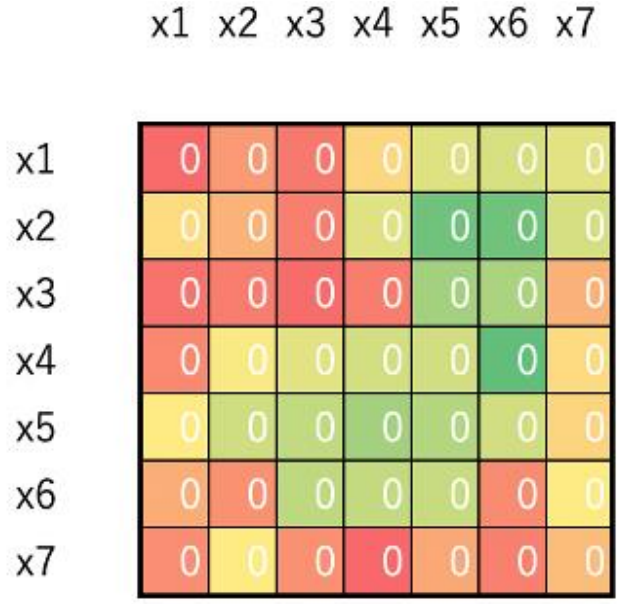
CNN



Why?



Transformer



https://medium.com/towards-data-science/recent-developments-and-views-on-computer-vision-x-transformer-ed32a2c72654

Scale-up

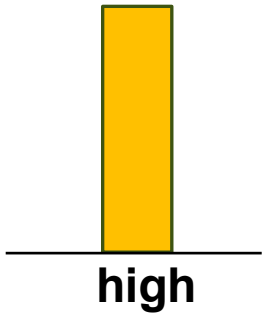


https://www.slamonline.com/the-magazine/every-basketball-movie-and-documentary-where-to-watch-%F0%9F%8E%A5%F0%9F%8D%BF/





Inductive Bias





https://www.wbiw.com/2024/12/18/come-coach-the-next-big-college-basketball-star/ https://medium.com/towards-data-science/recent-developments-and-views-on-computer-vision-x-transformer-ed32a2c72654
https://news.vanderbilt.edu/2021/11/23/the-pioneer-behind-shea-ralph-a-mother-and-daughters-basketball-journey-tells-the-story-of-a-new-era-of-equality/

Scale-up



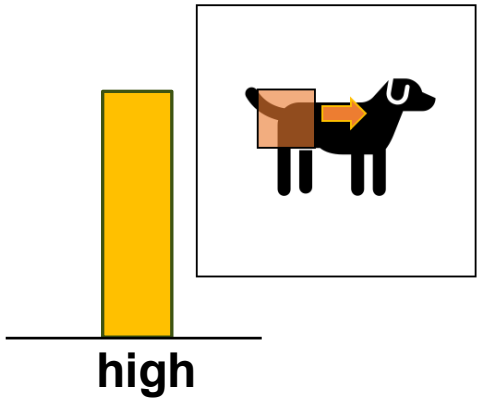
https://www.slamonline.com/the-magazine/every-basketball-movie-and-documentary-where-to-watch-%F0%9F%8E%A5%F0%9F%8D%BF/

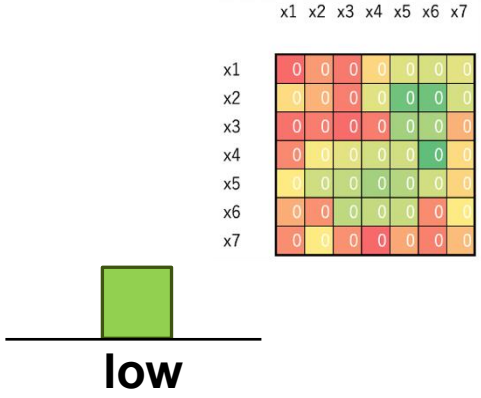
Inductive Bias





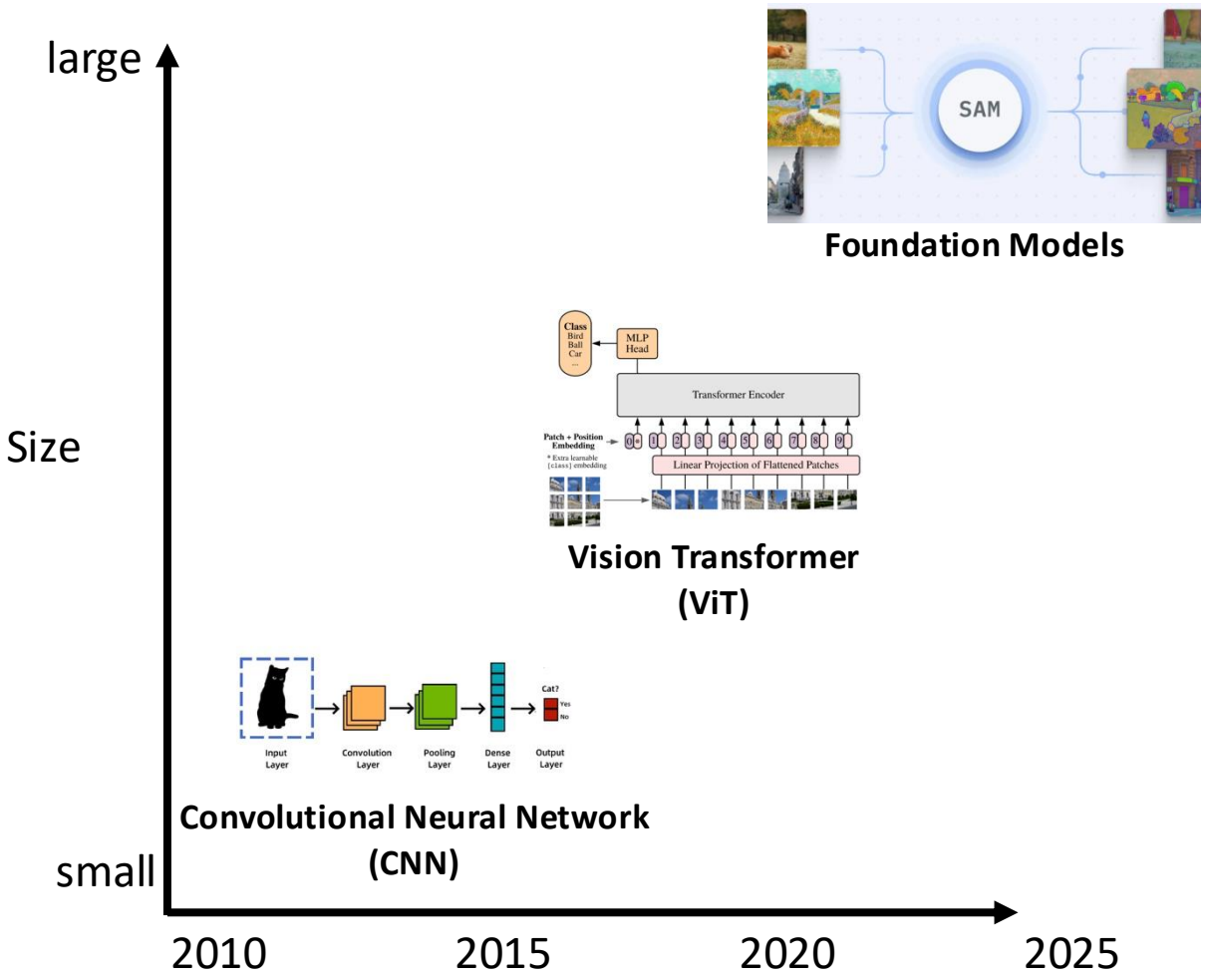
CNN **Transformer**

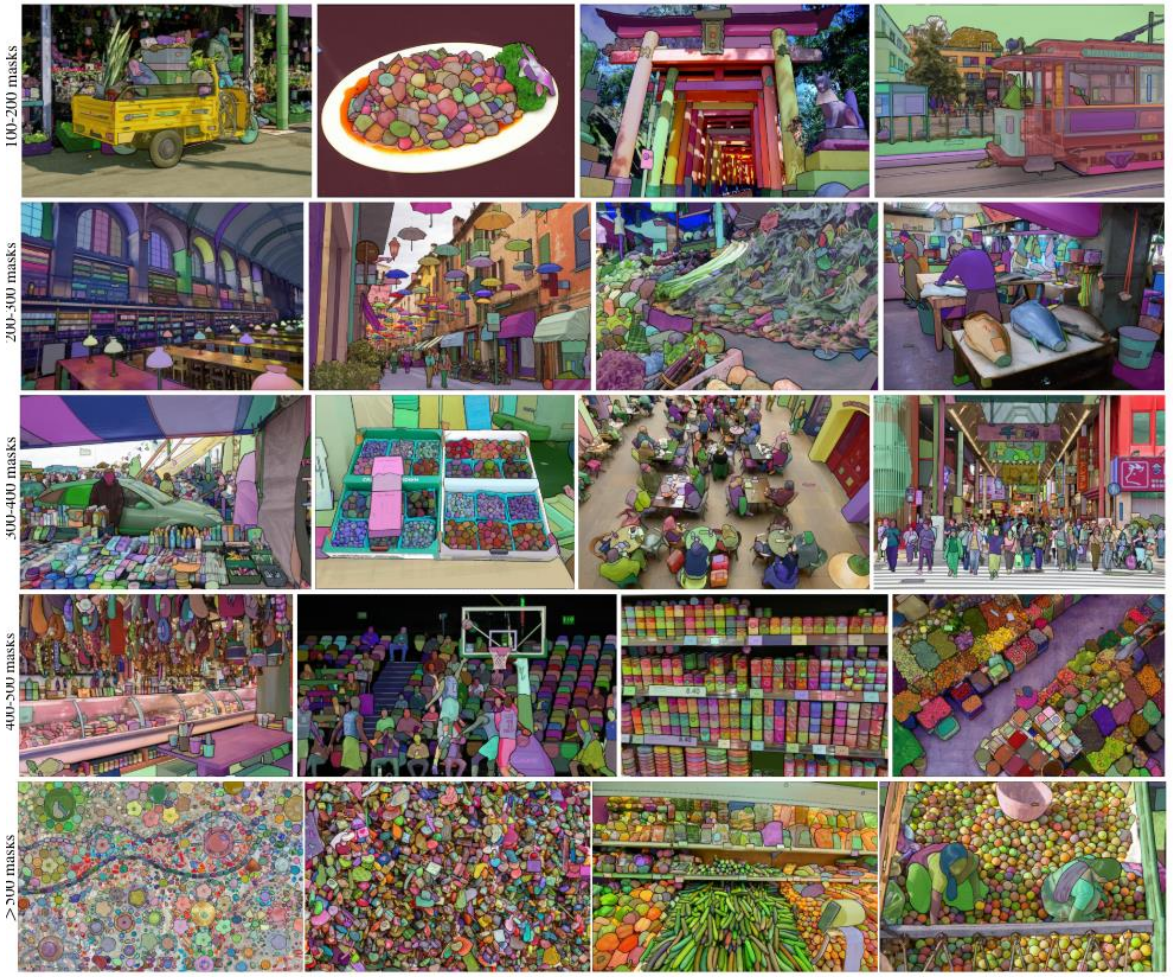




https://www.wbiw.com/2024/12/18/come-coach-the-next-big-college-basketball-star/ https://medium.com/towards-data-science/recent-developments-and-views-on-computer-vision-x-transformer-ed32a2c72654
https://news.vanderbilt.edu/2021/11/23/the-pioneer-behind-shea-ralph-a-mother-and-daughters-basketball-journey-tells-the-story-of-a-new-era-of-equality/

Foundation AI Models





https://www.basic.ai/blog-post/a-comprehensive-guide%3A-what-are-convolutional-neural-networks
https://paperswithcode.com/method/vision-transformer
https://www.v7labs.com/blog/segment-anything-model-sam

https://blog.syntha.ai/p/segment-anything-by-meta-the-future-of-image-segmentation

AI Foundation Models



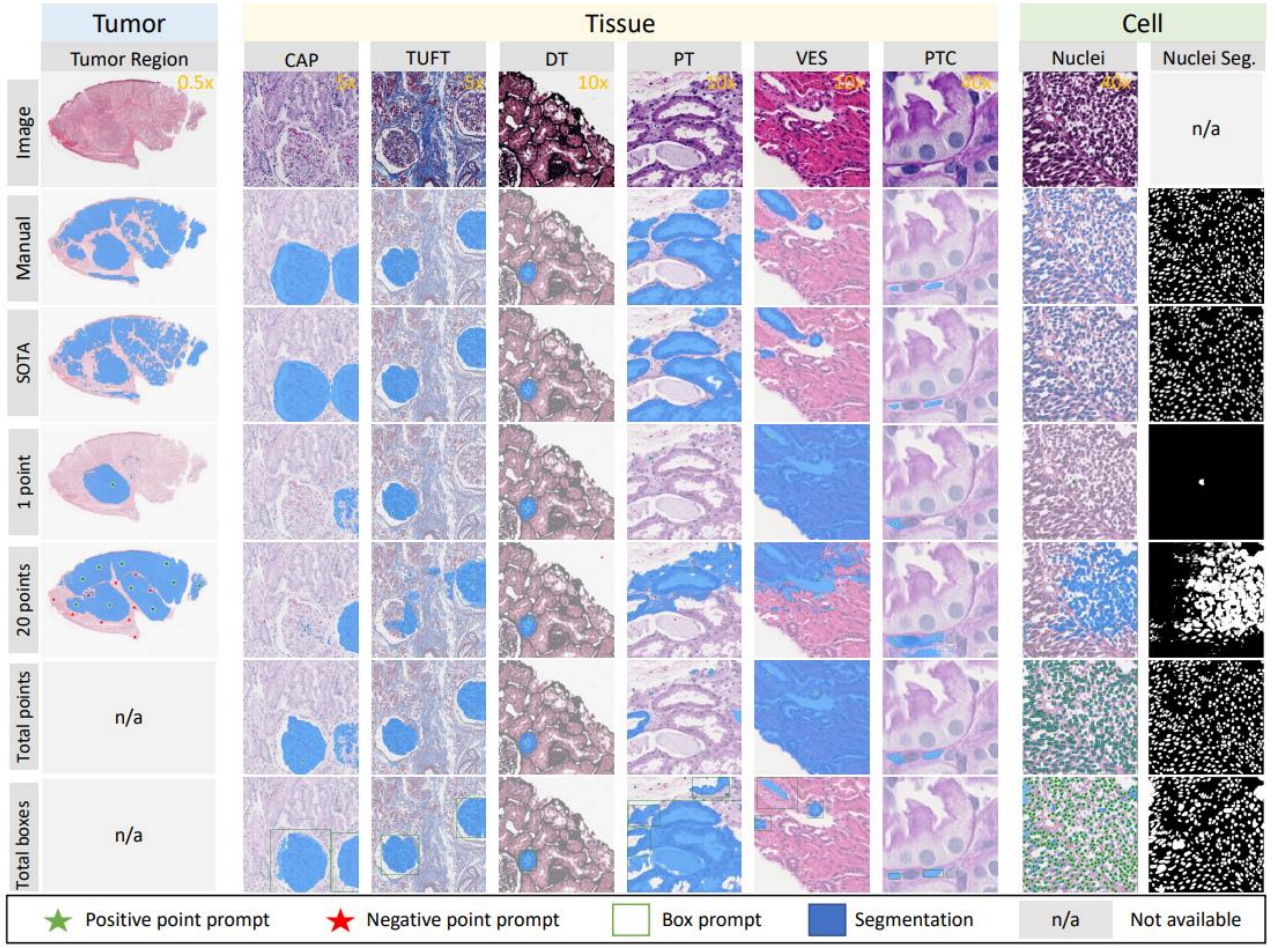


https://www.maginative.com/article/meta-ai-open-sources-its-segment-anything-sa-project/

Foundation AI model for Pathology

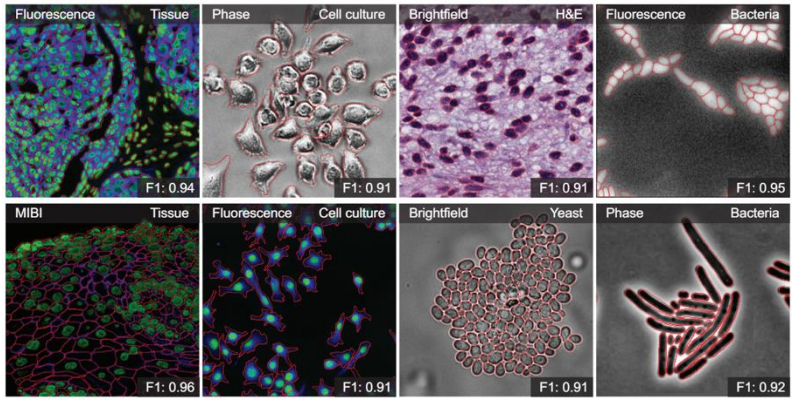




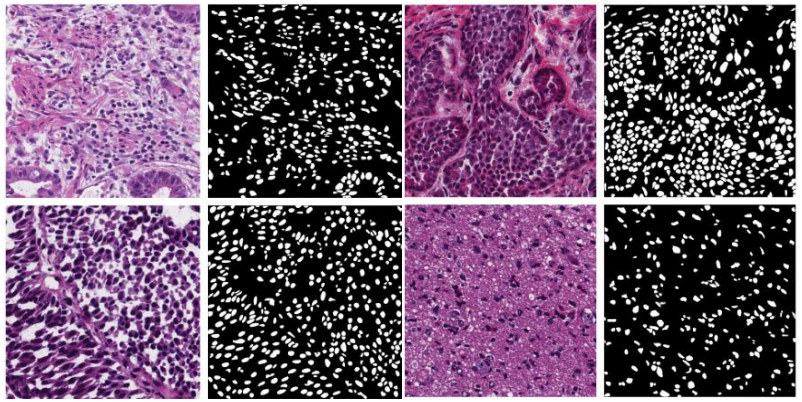


Ruining Deng, …, Yucheng Tang, Yuankai Huo "Segment anything model (sam) for digital pathology: Assess zero-shot segmentation on whole slide imaging." arXiv (2023).

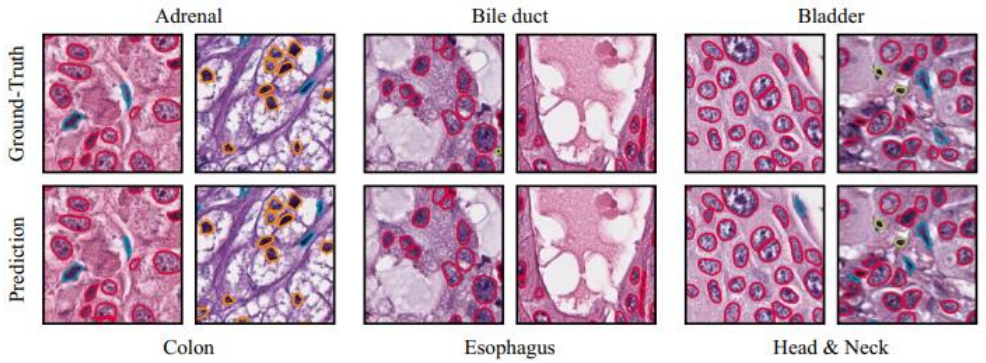
AI Foundation Models



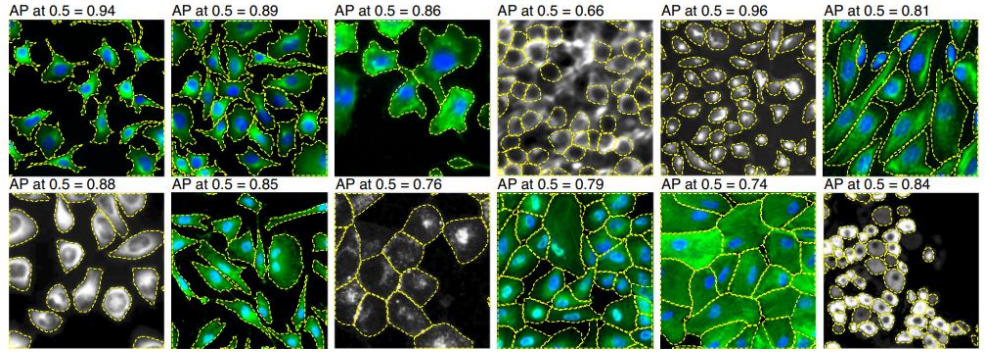
CellSAM (Israel, Uriah, et al.)



Segment Any Cell (Na, Saiyang, et al.)



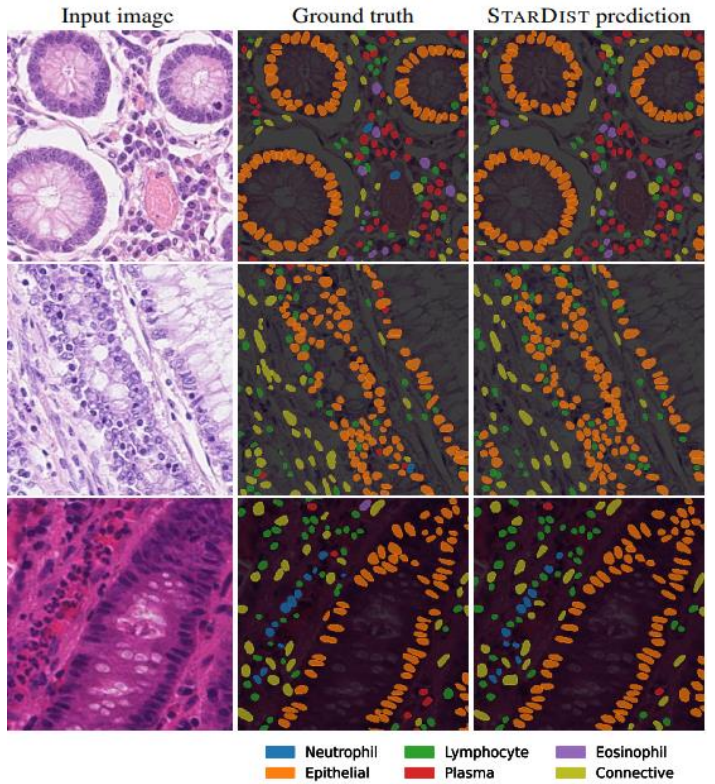
CellViT (Fabian Hörst, et al.)



Cellpose (Carsen Stringer, et al.)

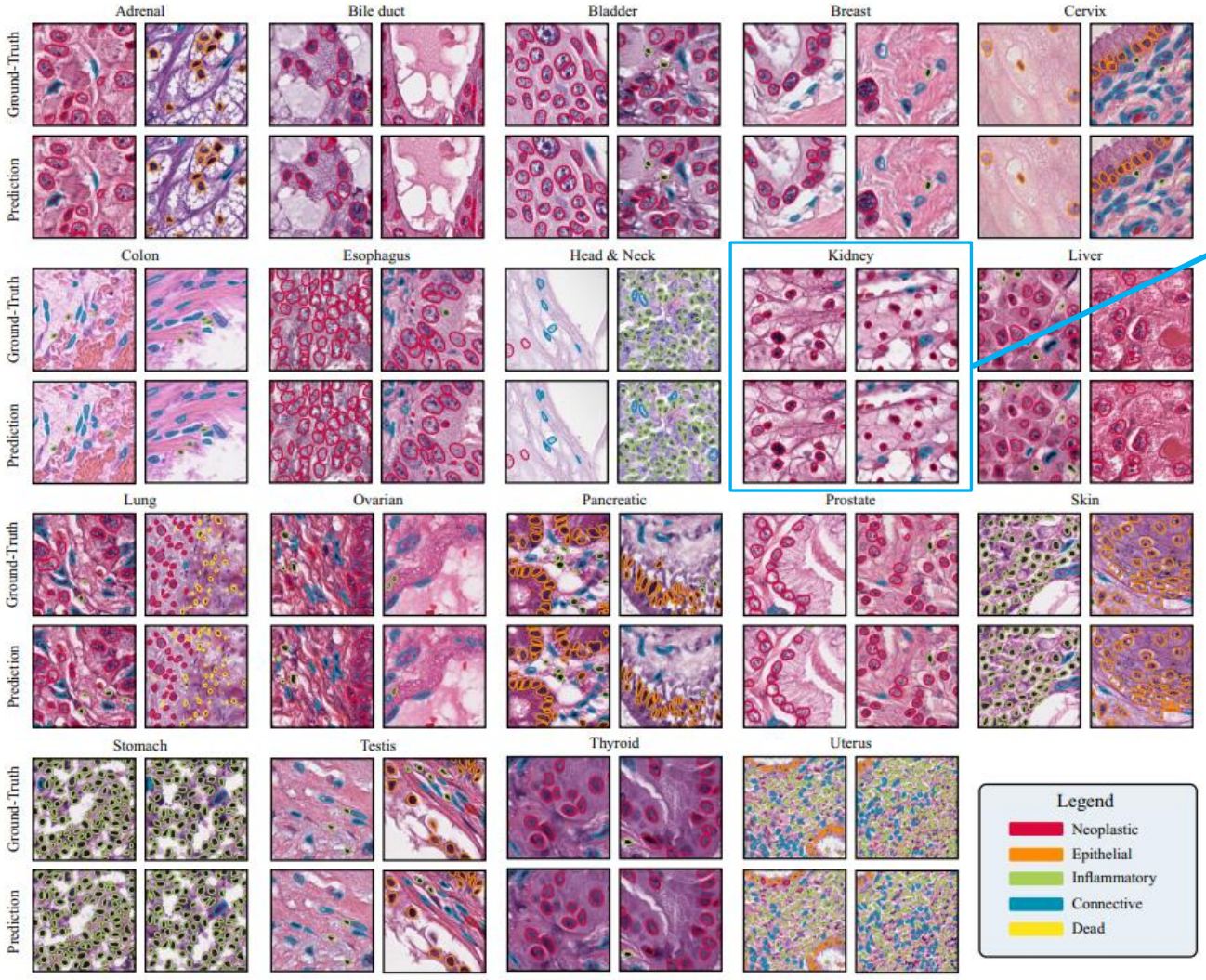
ChatGPT

Gemini

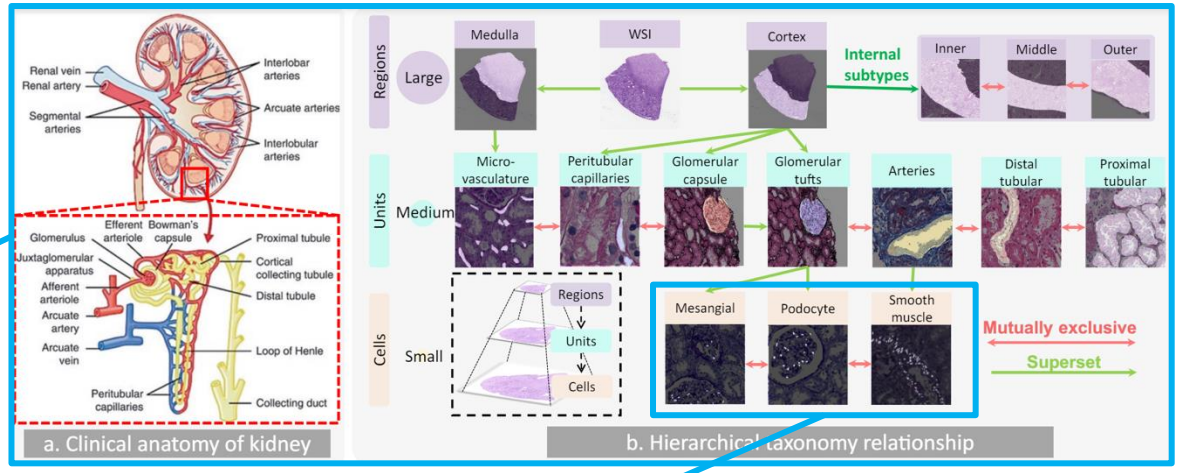


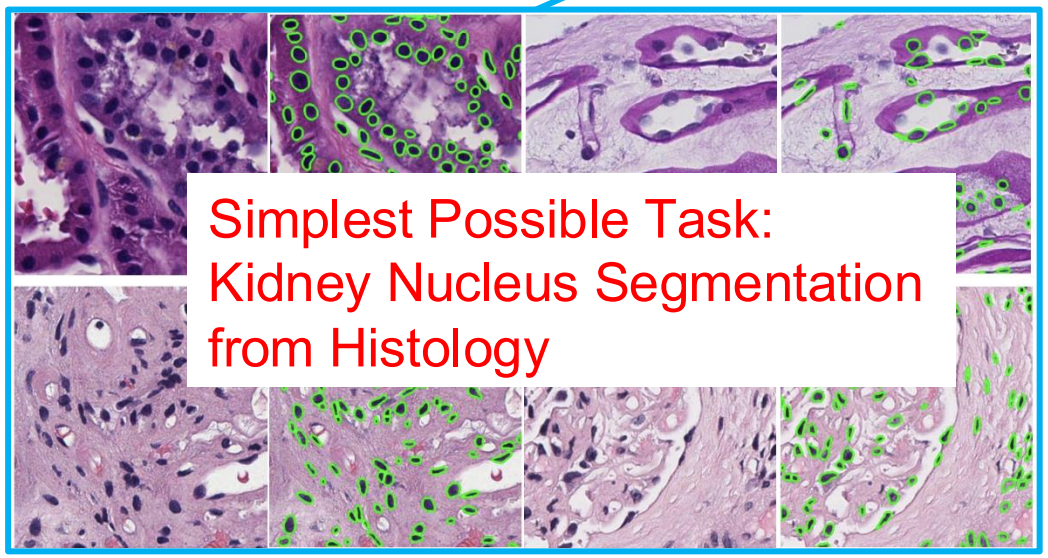
StarDIST (Martin Weigert, et al.)

Let's do a simple test



CellViT (Fabian Hörst, et al. 2023)





How good we are?

communications medicine

Explore content · About the journal · Publish with us

nature › communications medicine › articles › article

Article | Open access | Published: 24 November 2025

Evaluating cell AI foundation models in kidney pathology with human-in-the-loop enrichment

Junlin Guo, Siqi Lu, Can Cui, Ruining Deng, Tianyuan Yao, Zhewen Tao, Yizhe Lin, Marilyn Lionts, Quan Liu, Juming Xiong, Yu Wang, Shilin Zhao, Catie Chang, Mitchell Wilkes, Agnes Fogo, Mengmeng Yin, Haichun Yang & Yuankai Huo

Communications Medicine **5**, Article number: 495 (2025) | Cite this article

2211 Accesses | **1** Citations | Metrics



Junlin Guo



Agnes Fogo

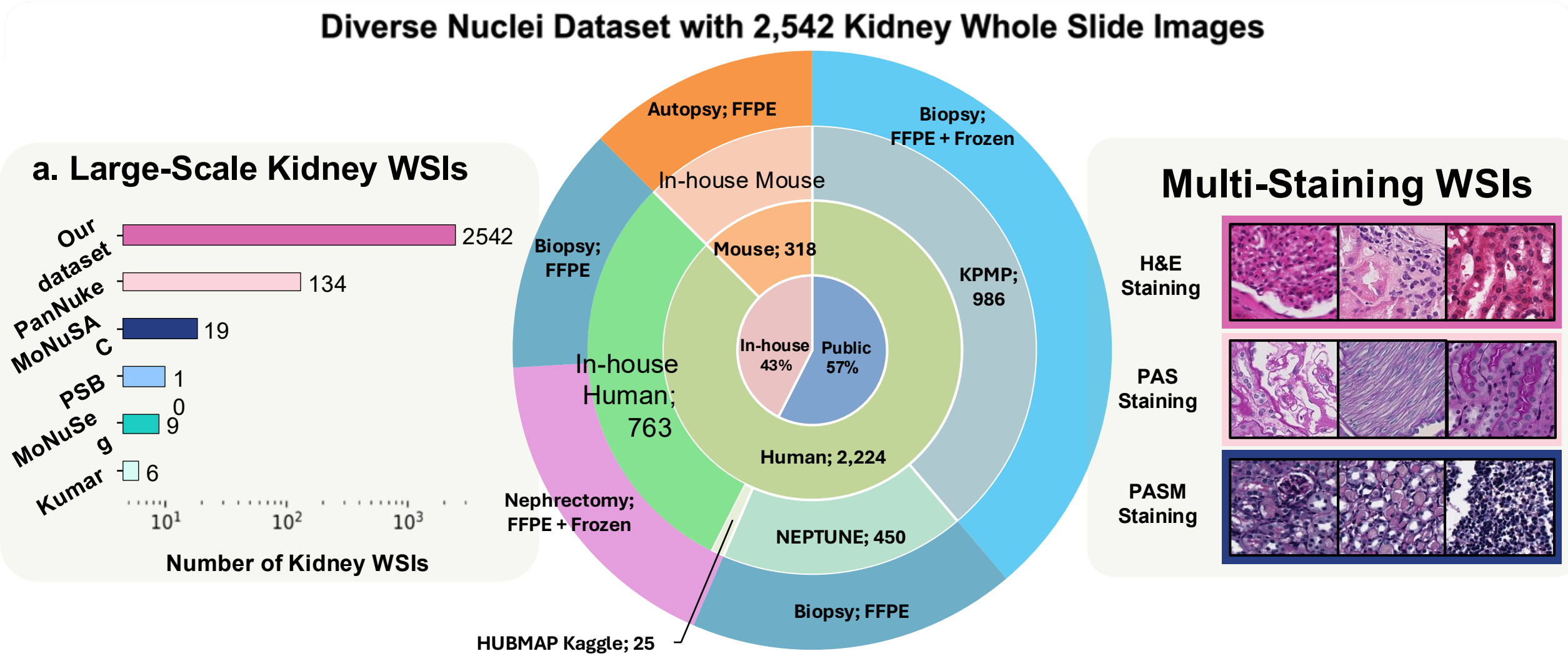


Haichun Yang

Simple, but comprehensive!

Guo et al. Communications Medicine (2025)

Curate the largest dataset



Performance of Each Method

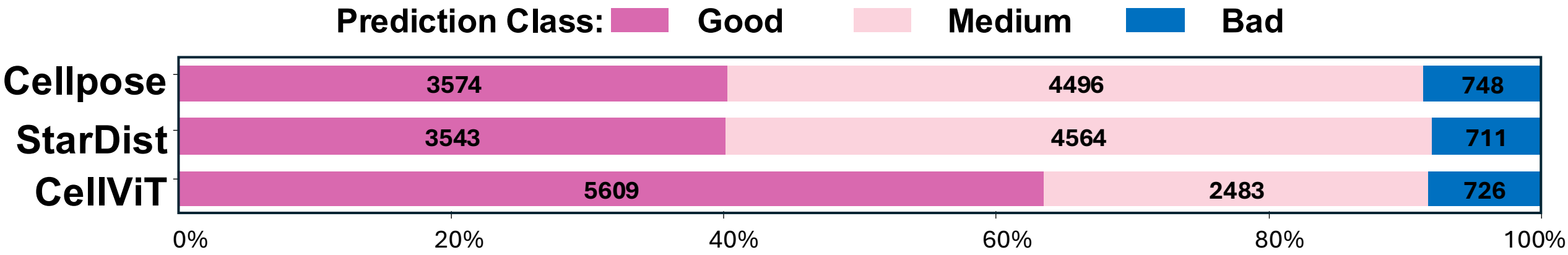
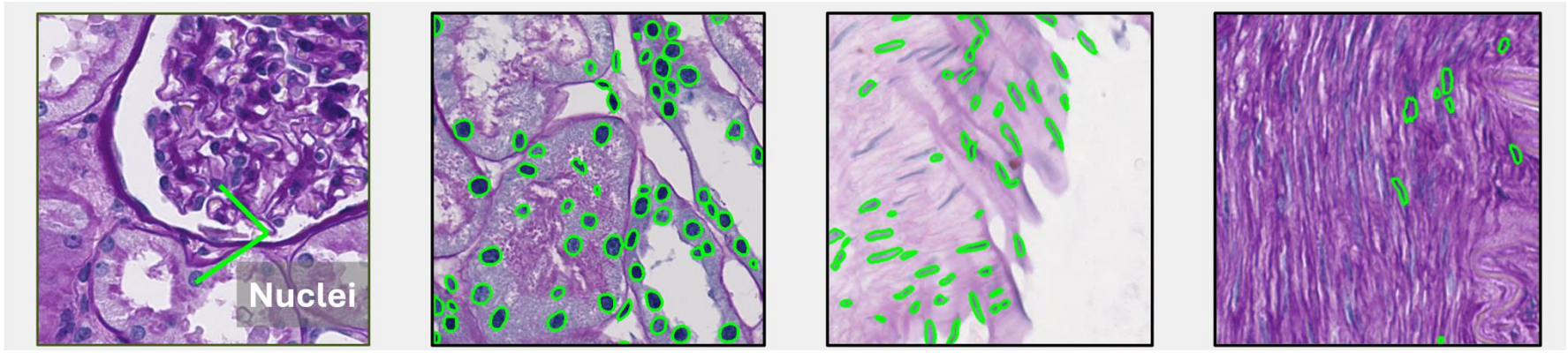


Image Samples

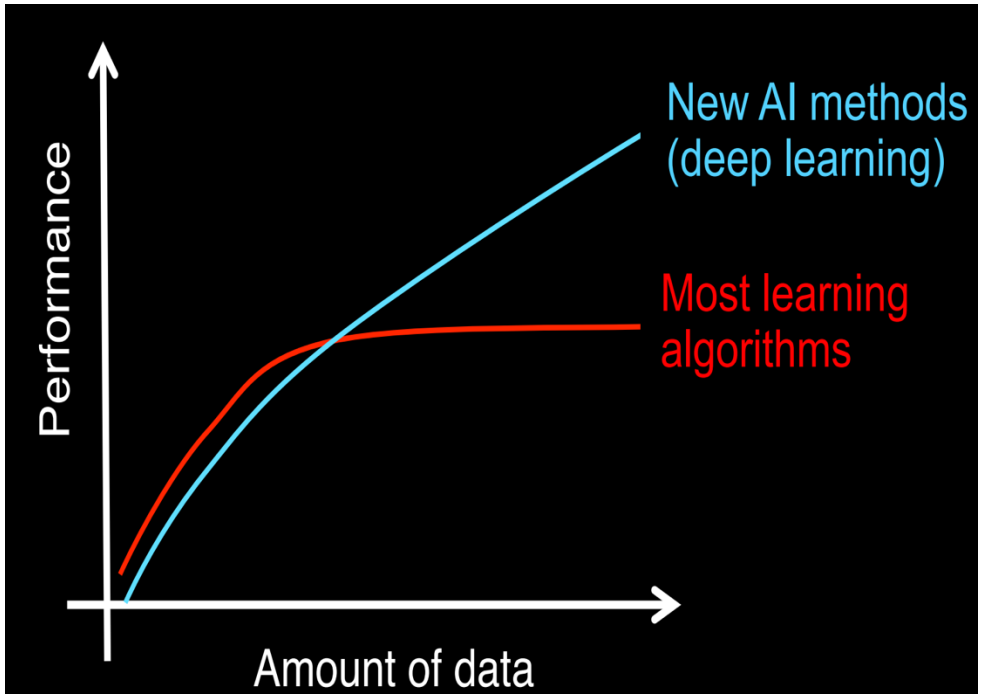


Good
Accuracy ~>90%

Medium
Accuracy ~ 50%-90%

Bad
Accuracy ~< 50%

No Free Lunch – Scale up is not easy for healthcare

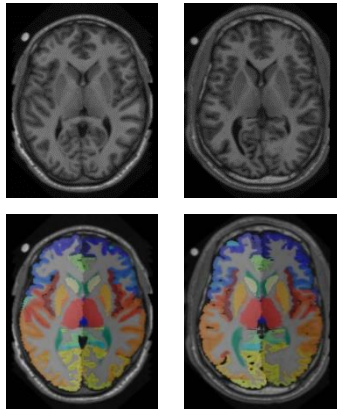


Natural Images





Medical Images





cs229.stanford.edu/materials/CS229-DeepLearning.pdf
https://www.slideshare.net/xavigiro/deep-learning-for-computer-vision-segmentation-upc-2016
https://www.radiologybusiness.com/topics/quality/subspeciality-rates-climb-majority-american-radiologists-remain-generalists

$

$$$$$$$$$$$

Another Direction

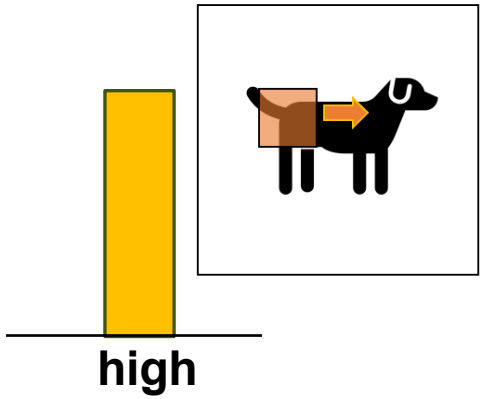


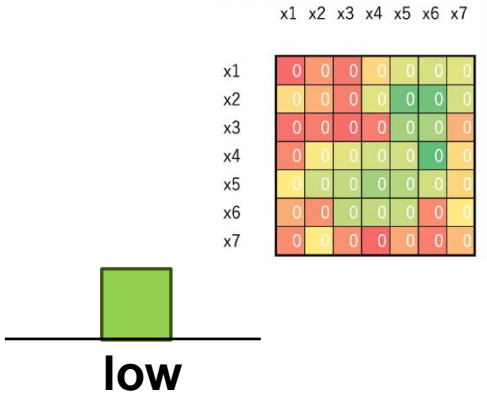


CNN

Transformer

Inductive Bias





https://www.wbiw.com/2024/12/18/come-coach-the-next-big-college-basketball-star/ https://medium.com/towards-data-science/recent-developments-and-views-on-computer-vision-x-transformer-ed32a2c72654
https://news.vanderbilt.edu/2021/11/23/the-pioneer-behind-shea-ralph-a-mother-and-daughters-basketball-journey-tells-the-story-of-a-new-era-of-equality/

Another Direction



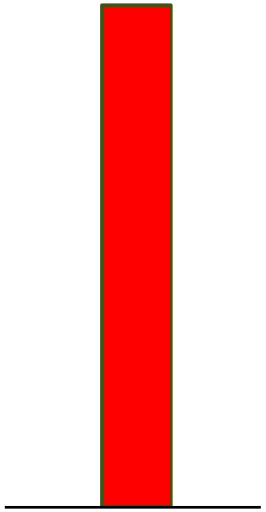




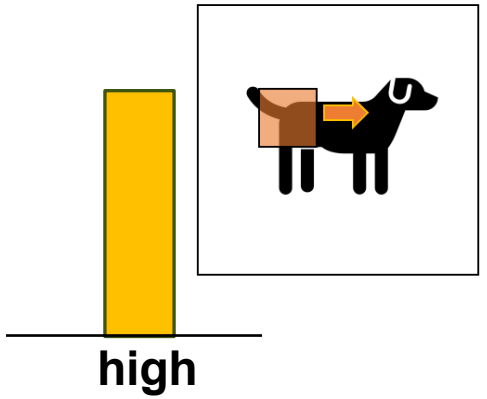
CNN

Transformer

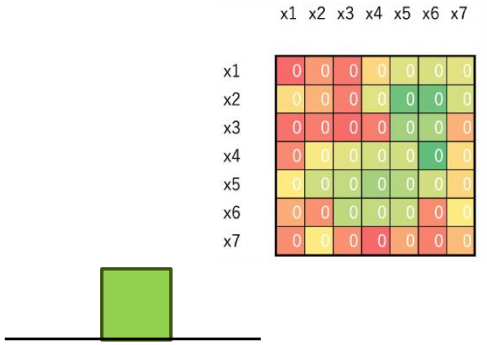
Inductive Bias



high



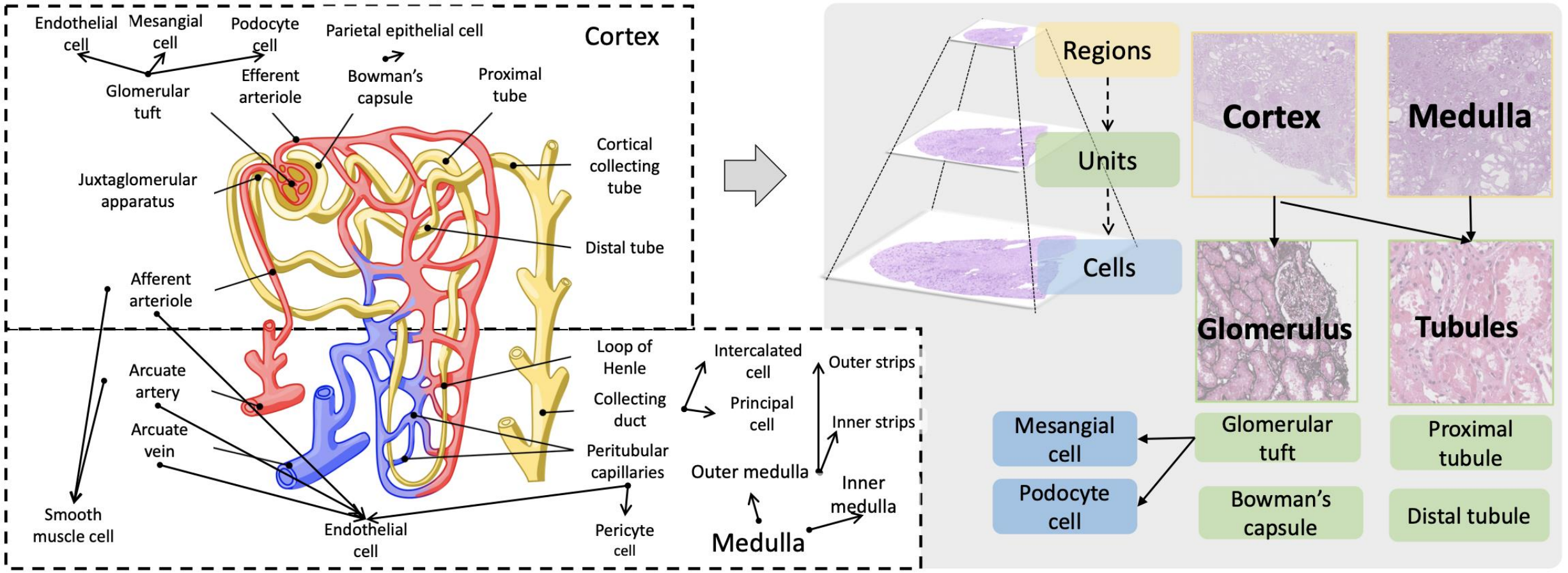
high



low

https://www.wbiw.com/2024/12/18/come-coach-the-next-big-college-basketball-star/ https://medium.com/towards-data-science/recent-developments-and-views-on-computer-vision-x-transformer-ed32a2c72654
https://news.vanderbilt.edu/2021/11/23/the-pioneer-behind-shea-ralph-a-mother-and-daughters-basketball-journey-tells-the-story-of-a-new-era-of-equality/

Knowledge-infused Learning

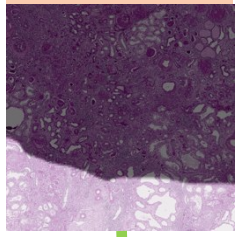


Ruining Deng, Quan Liu, Can Cui, Tianyuan Yao, Jialin Yue, Juming Xiong, Lining Yu, Yifei Wu, Mengmeng Yin, Yu Wang, Shilin Zhao, Yucheng Tang, Haichun Yang, Yuankai Huo. "Prpseg: Universal proposition learning for panoramic renal pathology segmentation." *CVPR*, 2024.

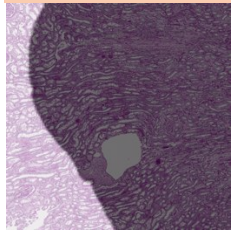
Panoramic Pathology Image Analysis

Regions

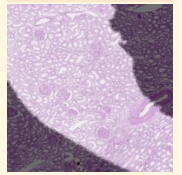
Medulla



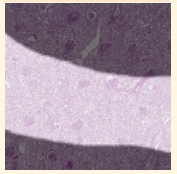
Cortex



Inner



Middle

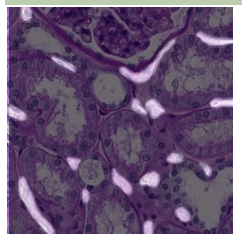


Outer

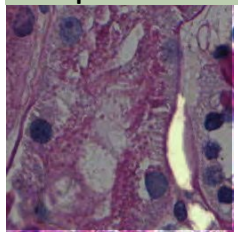


Units

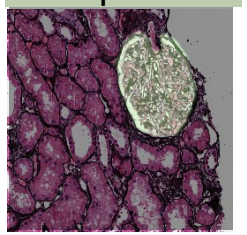
Micro-vasculature



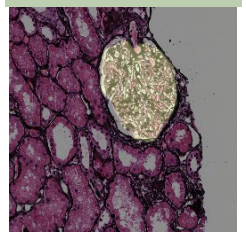
Peritubular capillaries



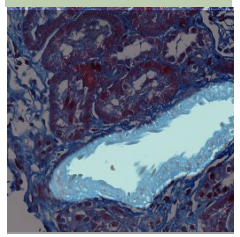
Glomerular capsule



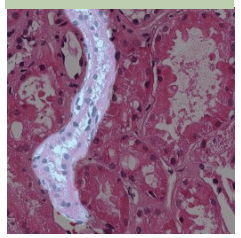
Glomerular tufts



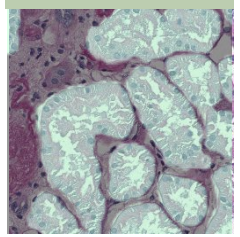
Arteries



Distal tubular

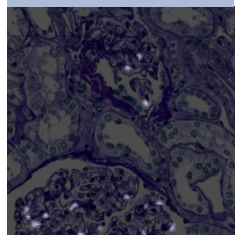


Proximal tubular

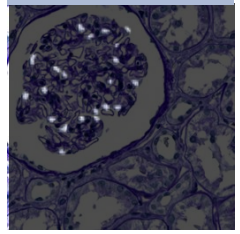


Cells

Mesangial



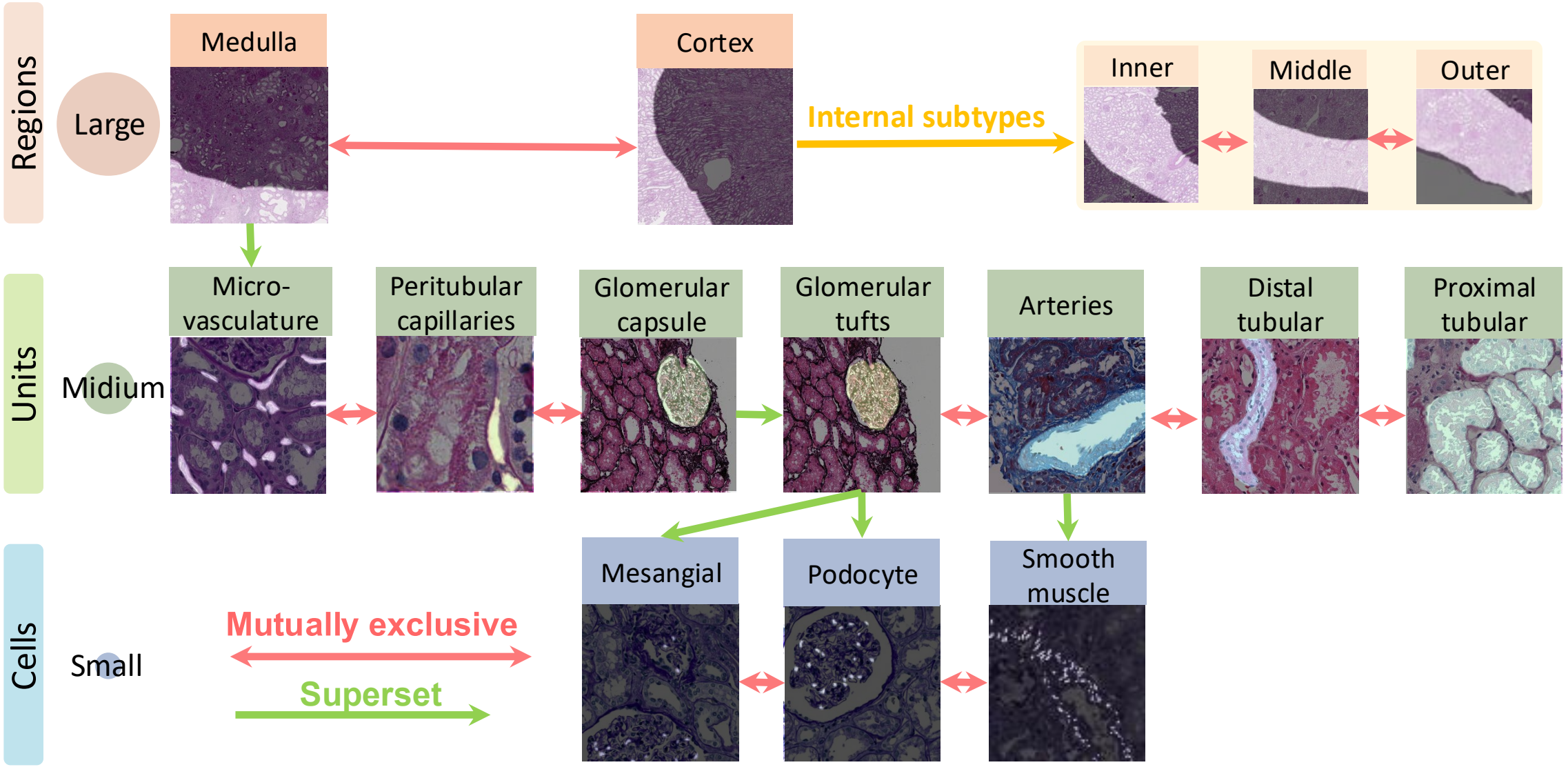
Podocyte



Smooth muscle

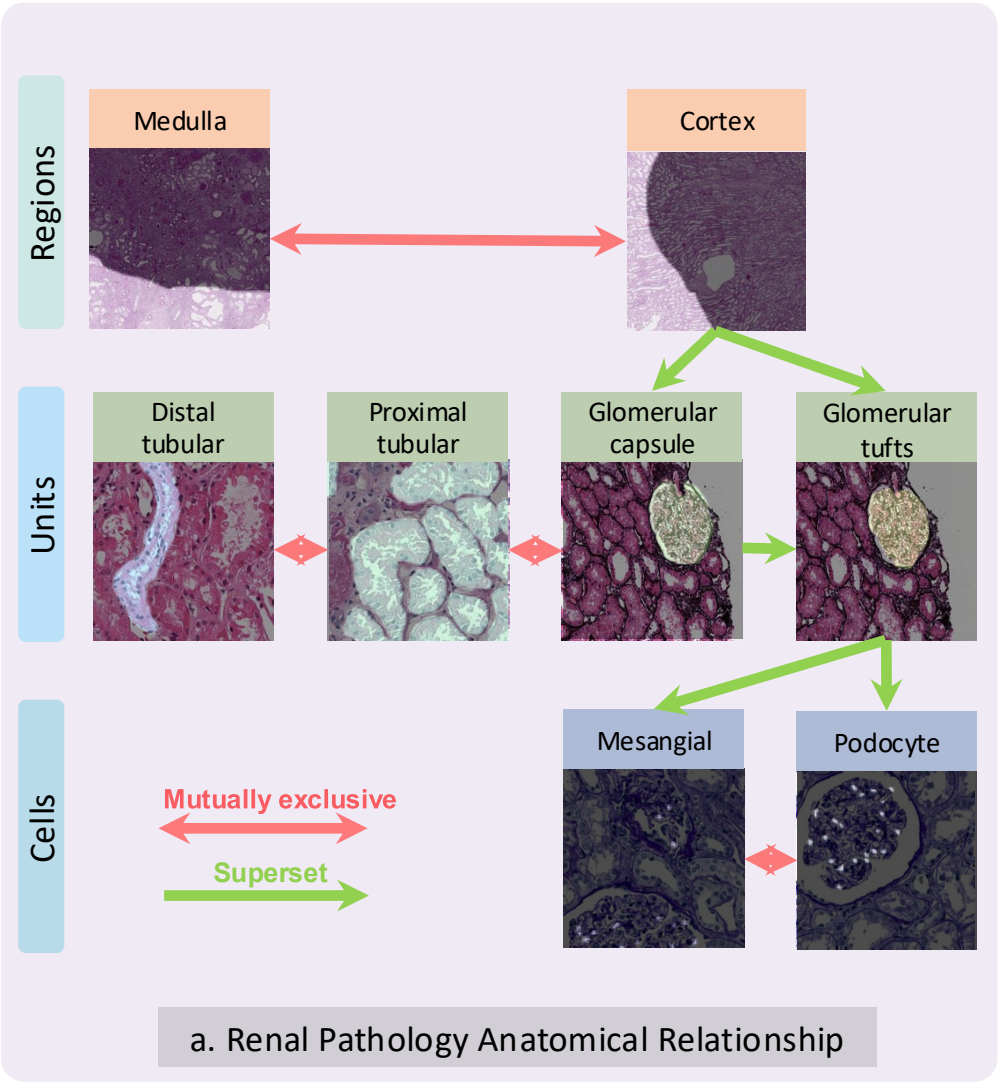


Hierarchical taxonomy relationship



PrPSeg

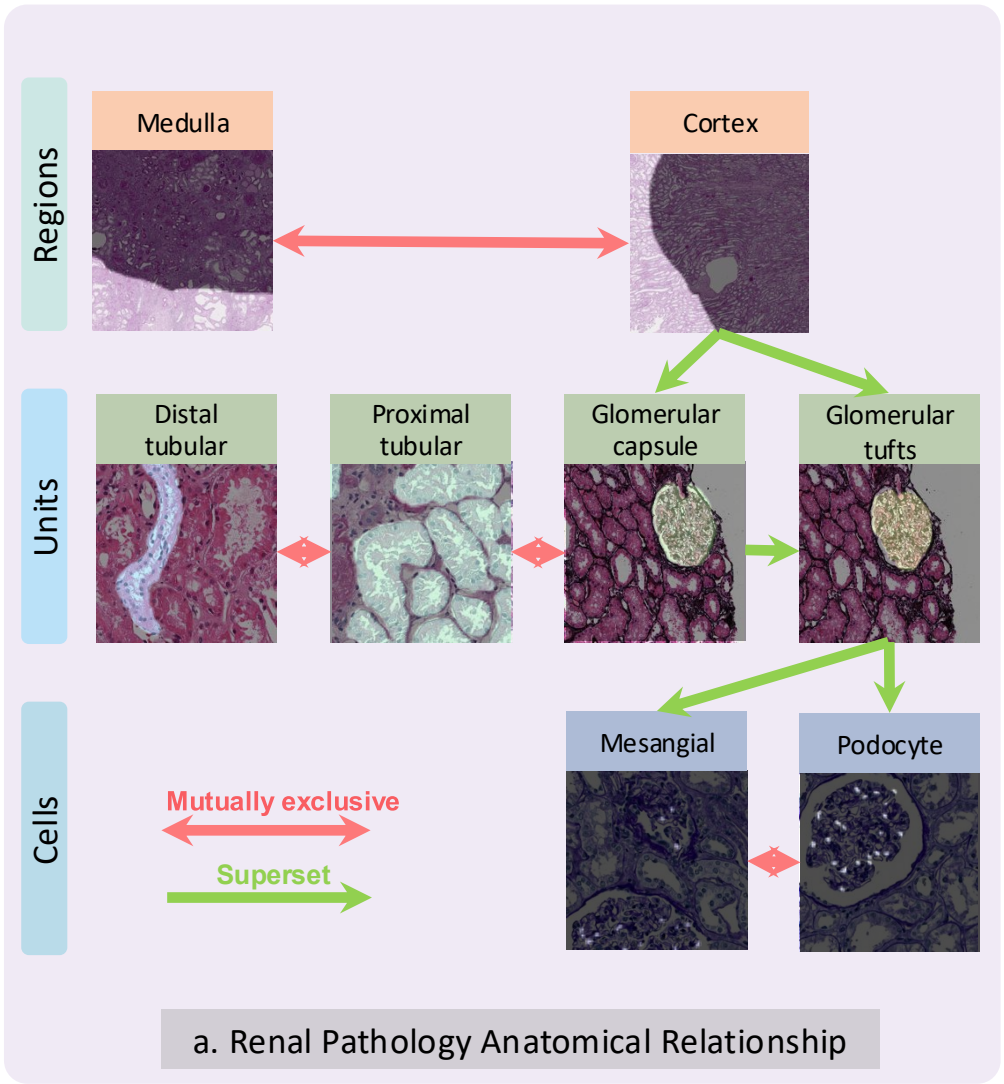
Ruining Deng, … , Yuankai Huo "PrPSeg: Universal proposition learning for panoramic renal pathology segmentation." *CVPR*. 2024.



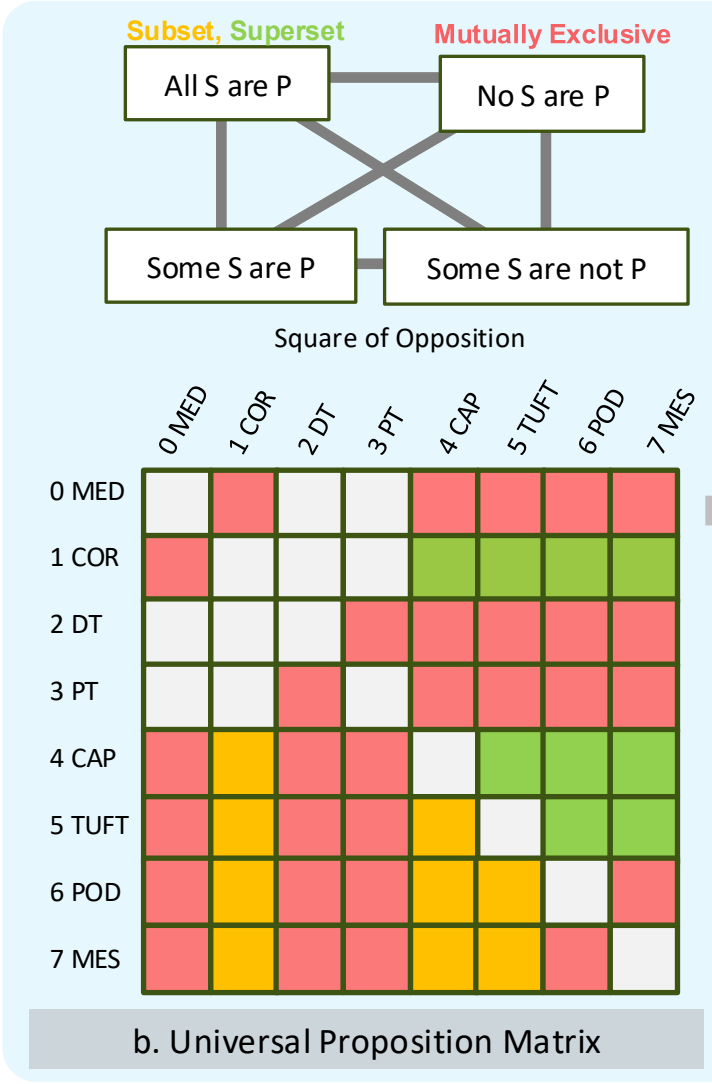
a. Renal Pathology Anatomical Relationship

PrPSeg

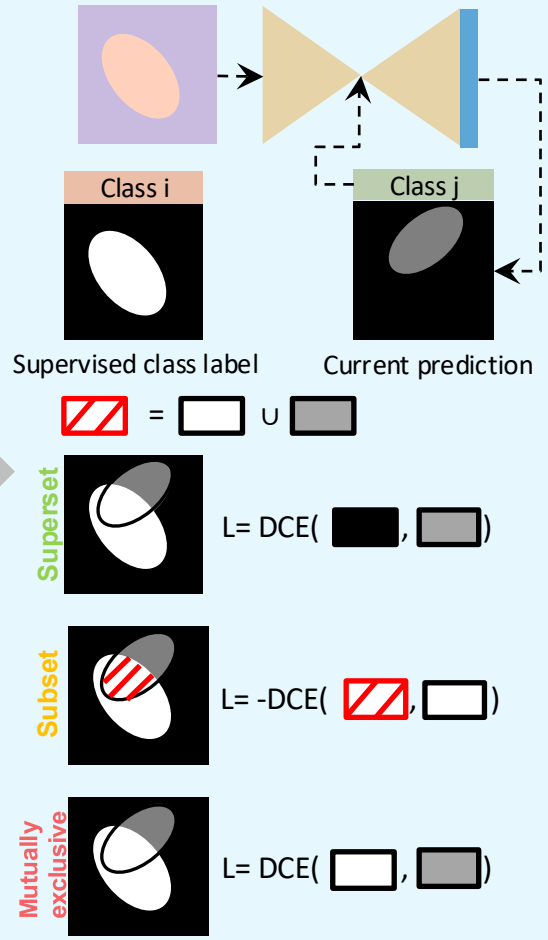
Ruining Deng, … , Yuankai Huo "PrPSeg : Universal proposition learning for panoramic renal pathology segmentation." *CVPR*. 2024.



a. Renal Pathology Anatomical Relationship

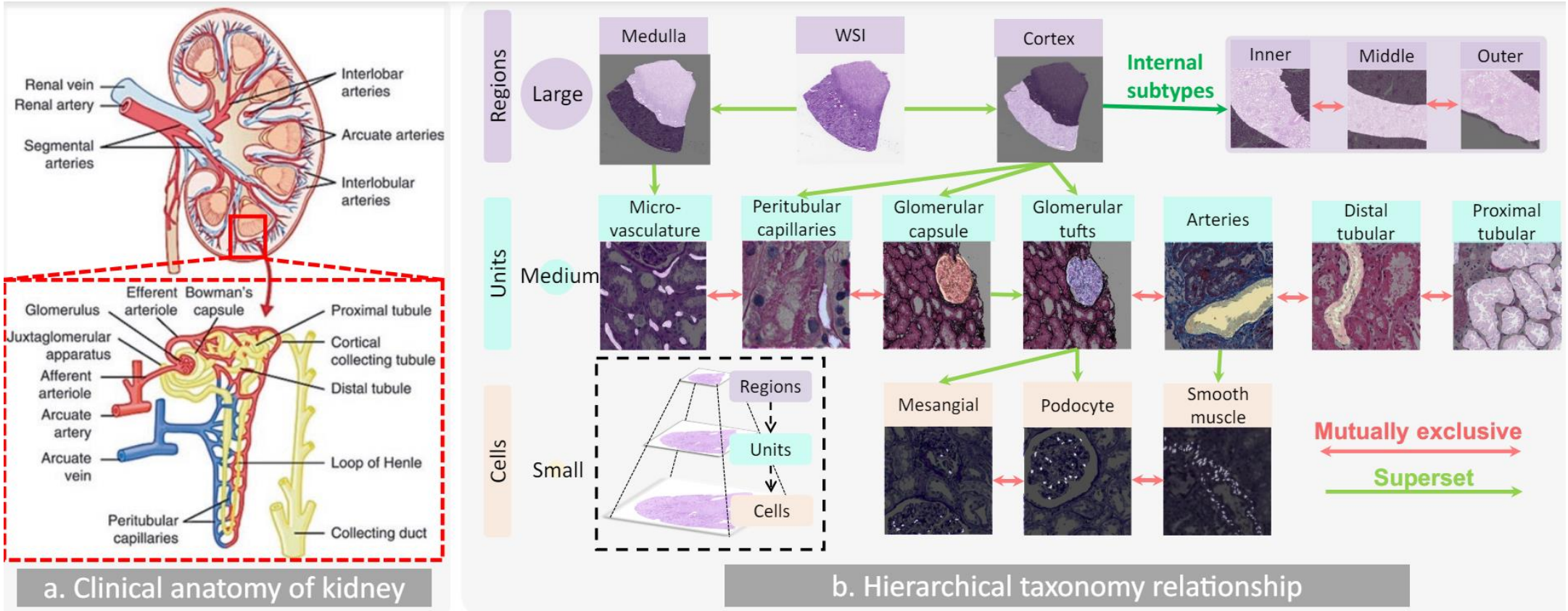


b. Universal Proposition Matrix



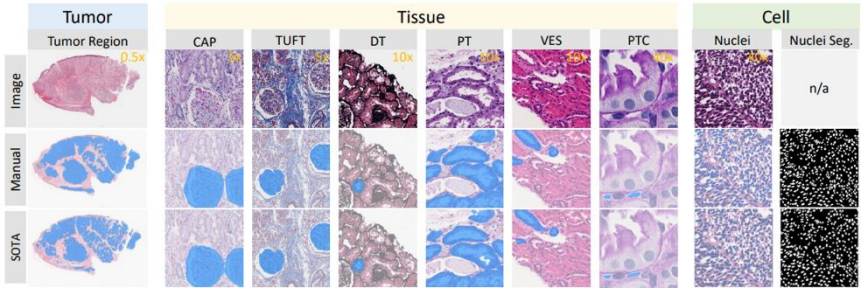
c. Anatomy Loss

HATs model

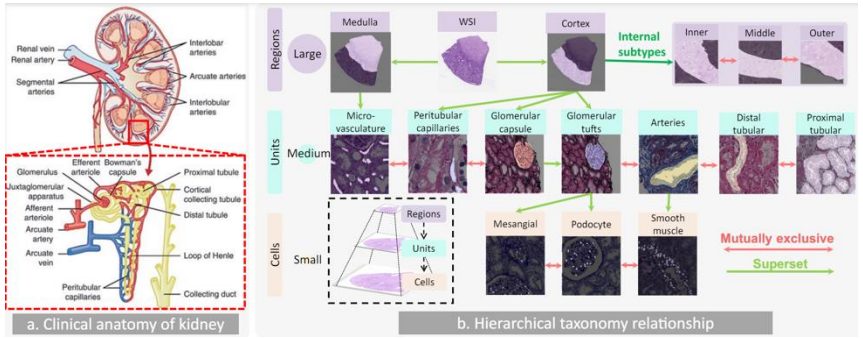


Deng, Ruining, …, Yuankai Huo "Hats: Hierarchical adaptive taxonomy segmentation for panoramic pathology image analysis." MICCAI 2024

We have tools to segment the objects

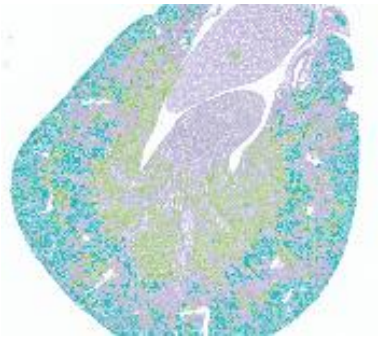


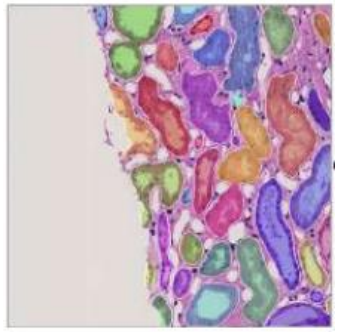
Foundation models

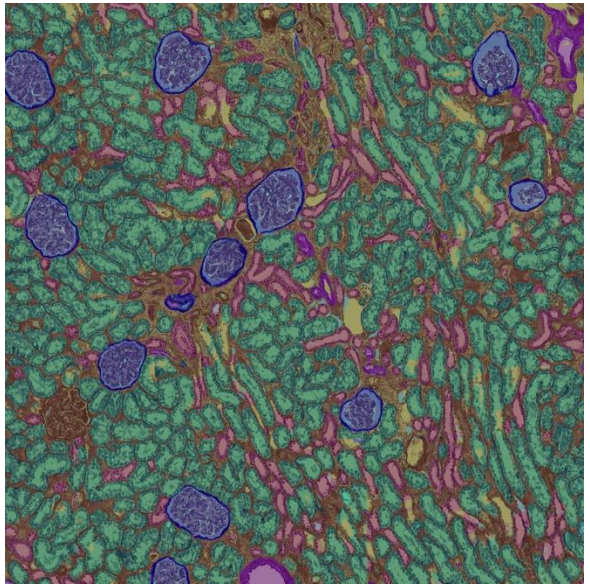


Expert models



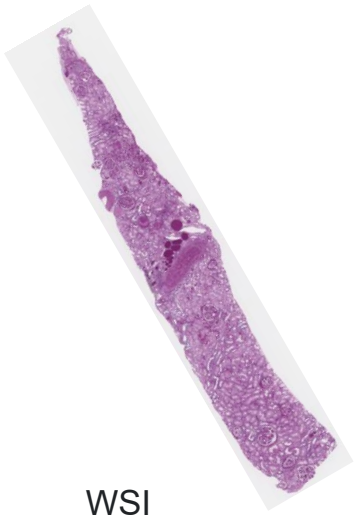




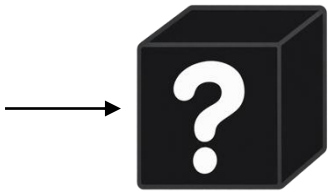


Segment N Objects M Types

Kidney Pathology

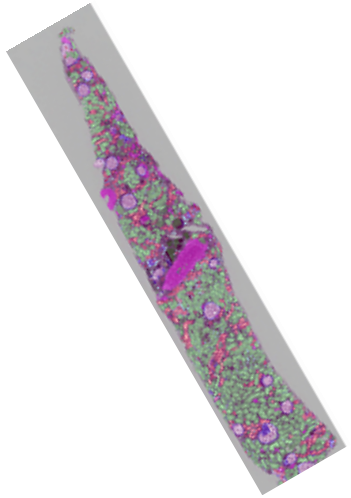


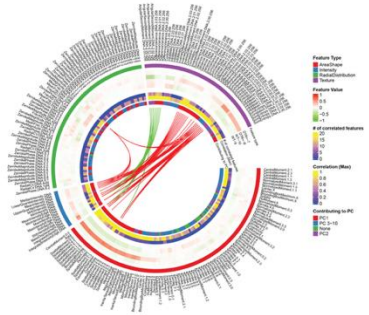
WSI



Still Black Box Model

Segmentation AI





Evidence

Pathomics



Analyze

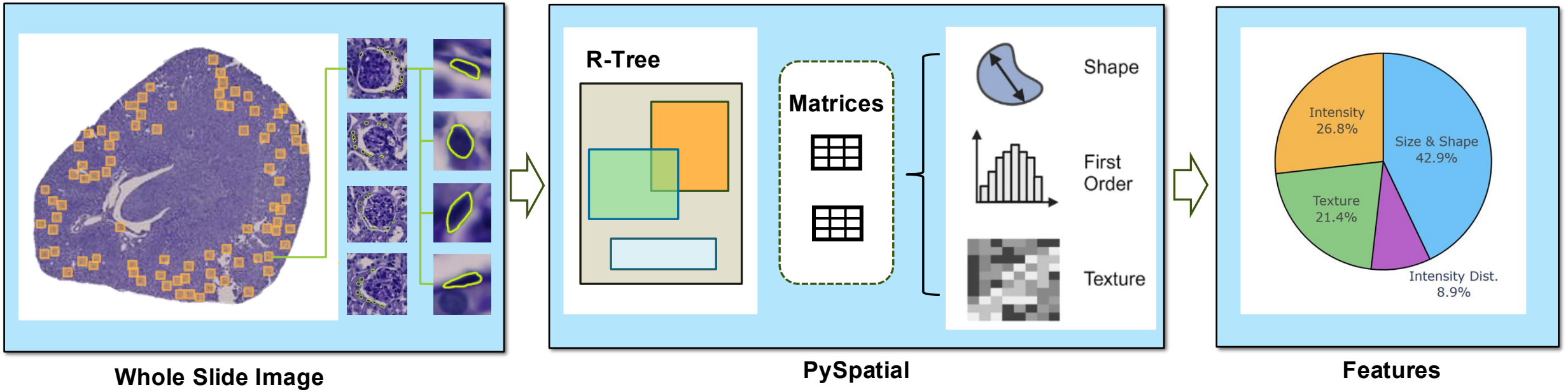
Challenge 1

Challenge 2

Challenge 3

Challenge 2: Interpretability

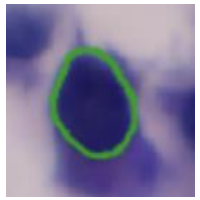
We also have tools to extract pathomics features automatically.



Whole Slide Image **PySpatial** **Features**

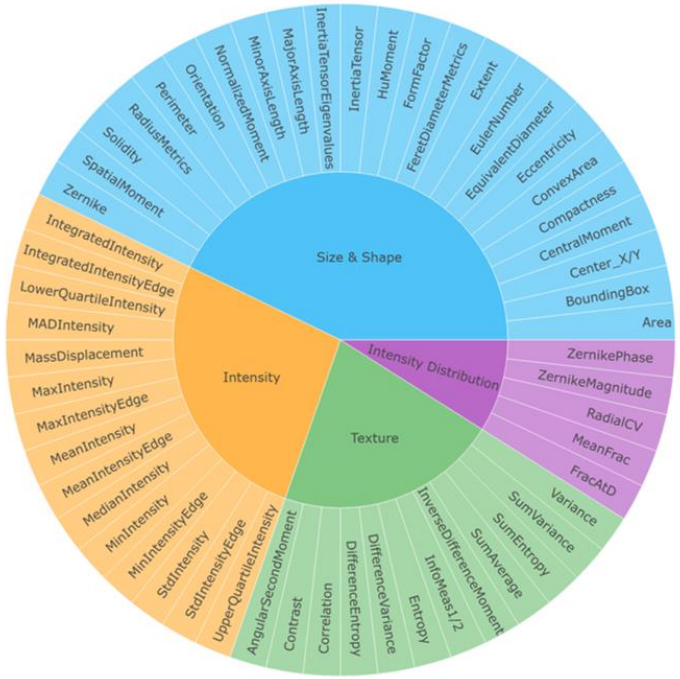
Yuechen Yang, …, Haichun Yang, Yuankai Huo. "PySpatial: A High-Speed Whole Slide Image Pathomics Toolkit." *IS&T International Symposium on Electronic Imaging* 2025

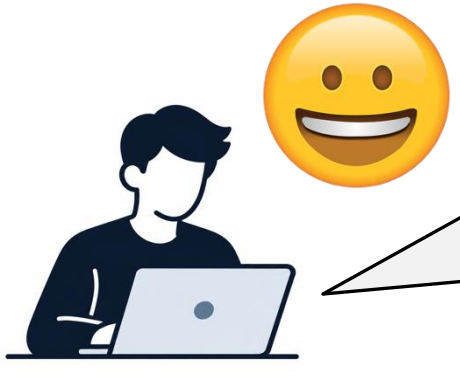
Example of Pathomics



PySpatial

247-dimensional feature vector for each object



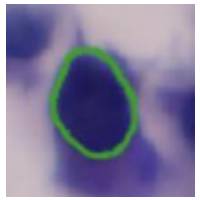


I convert every object into mathematically interpretable features! Let do some large scale analysis!

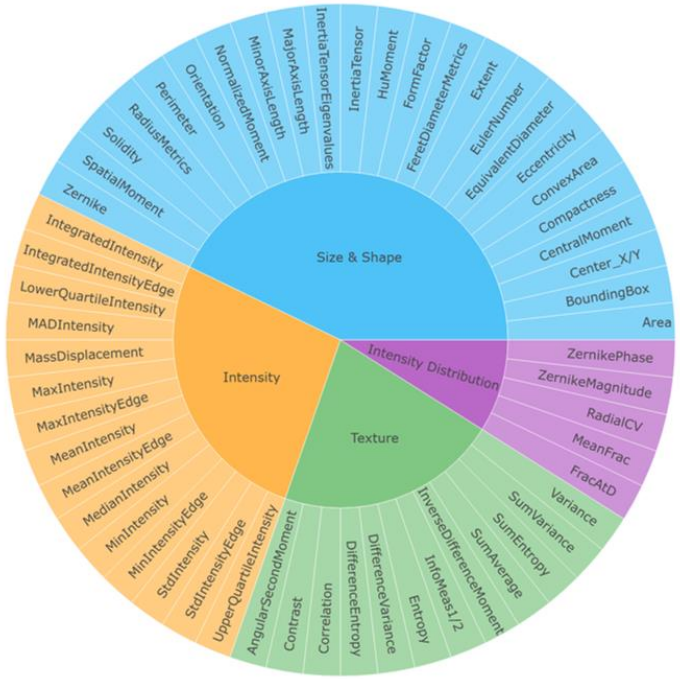
Wait …
What are those?



Example of Pathomics



247-dimensional feature vector for each object



mathematically interpretable Human interpretable ?

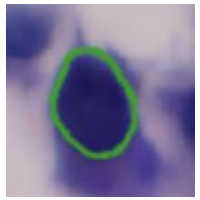
1. Size & Shape (Blue)
- Area
- BoundingBox
- Center_XY
- Compactness
- ConvexArea
- ConvexPerimeter
- Eccentricity
- EquivalentDiameter
- Extent
- FeretDiameter
- FilledArea
- MajorAxisLength
- MinorAxisLength
- Orientation
- Perimeter
- Roundness
- Solidity
- Zernike
- SurfaceArea
- FormFactor
- AspectRatio
- Elongation

2. Intensity (Orange)
- Uniformity
- InterquartileRange
- Energy
- Entropy
- Kurtosis
- Maximum
- Mean
- MeanAbsoluteDeviation
- Median
- Minimum
- Range
- RootMeanSquared
- Skewness
- StandardDeviation
- Variance
- Percentile
- MeanDisplacement
- MaxIntensity
- MinIntensity
- RobustMeanAbsoluteDeviation

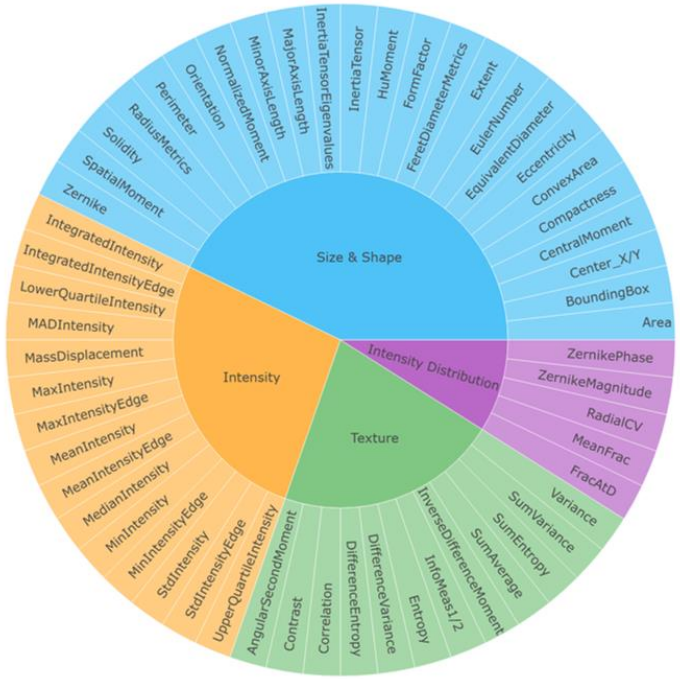
3. Texture (Green)
- Autocorrelation
- Contrast
- Correlation
- DifferenceEntropy
- DifferenceVariance
- Dissimilarity
- Entropy
- Homogeneity
- InformationalMeasureOfCorrelation
- InverseDifferenceMoment
- InverseDifferenceMomentNormalized
- InverseDifferenceNormalized
- MaximumProbability
- SumAverage
- SumEntropy
- SumVariance
- ClusterShade
- ClusterProminence

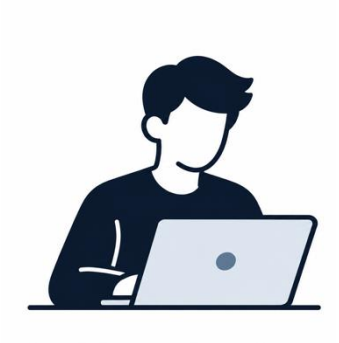
4. Intrinsic Distribution (Purple)
- ZernikePhase
- ZernikeMagnitude
- RadialCV
- MeanFrac
- FracAtD

Example of Pathomics



247-dimensional feature vector for each object





Is there anyway to let people understand the pathomics features directly!



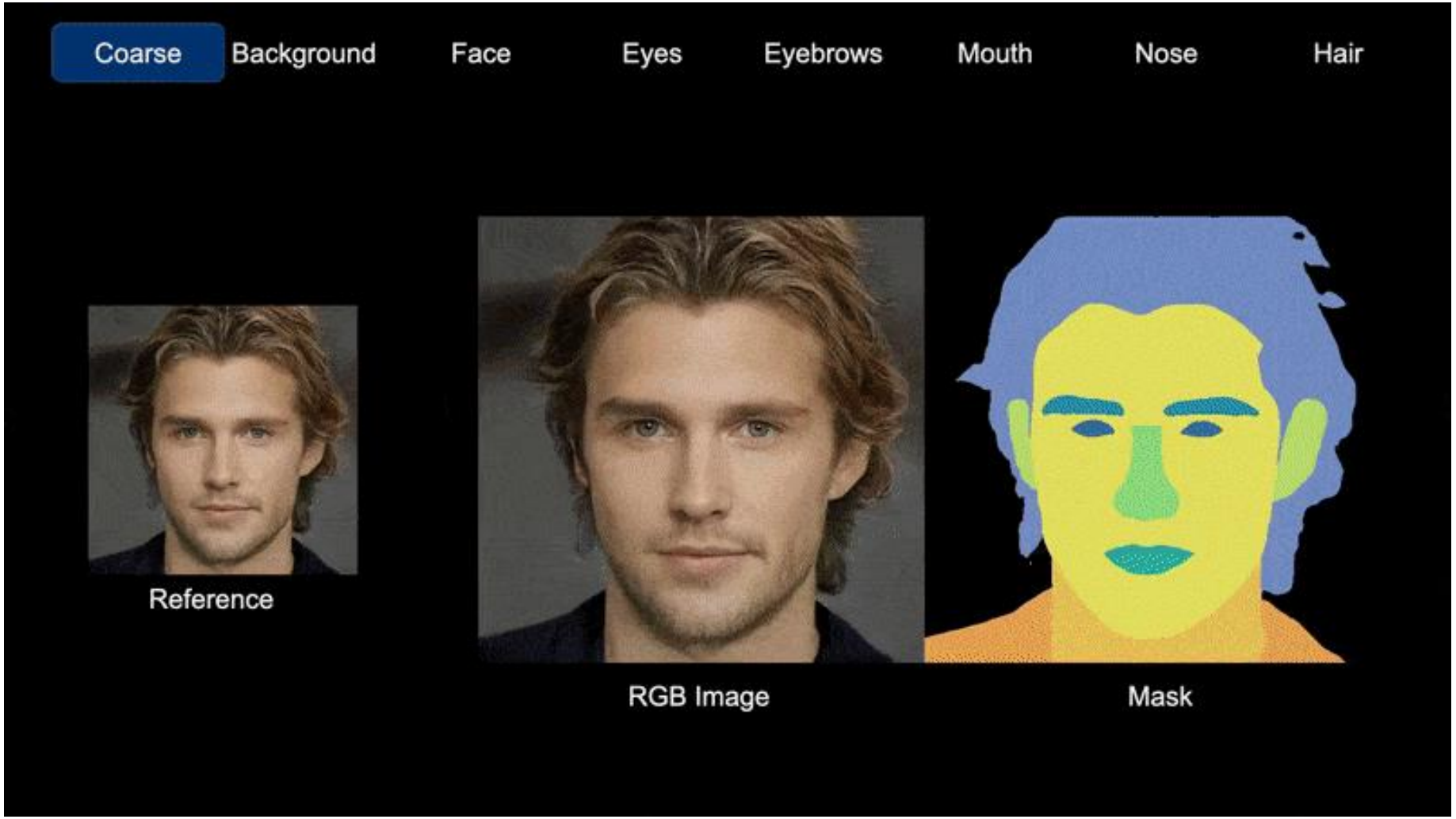


Generative AI



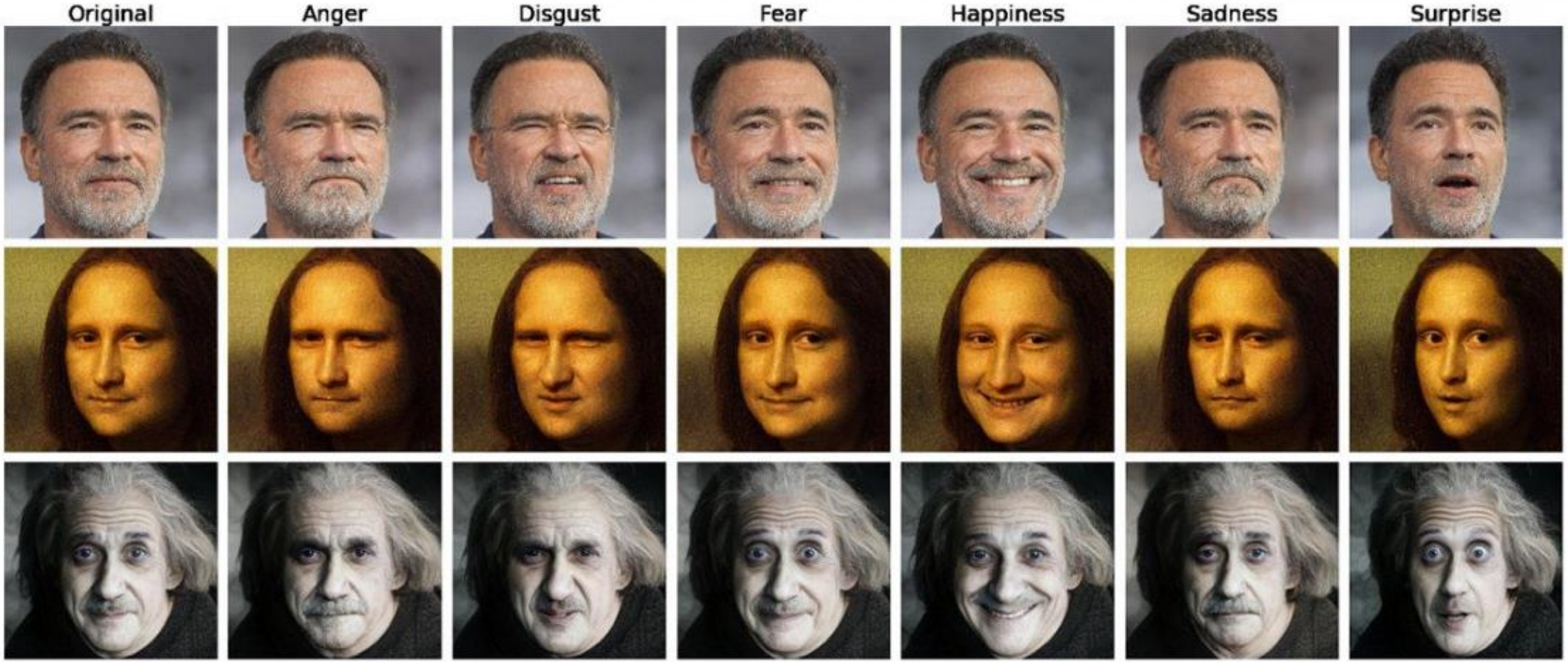
Karras et al. CVPR. 2019

Conditional Generative AI



https://semanticstylegan.github.io/

Conditional Generative AI



https://arxiv.org/pdf/2205.06102.pdf

Interpretation

Upprunalegt

Ótti

Undrun

Reiði

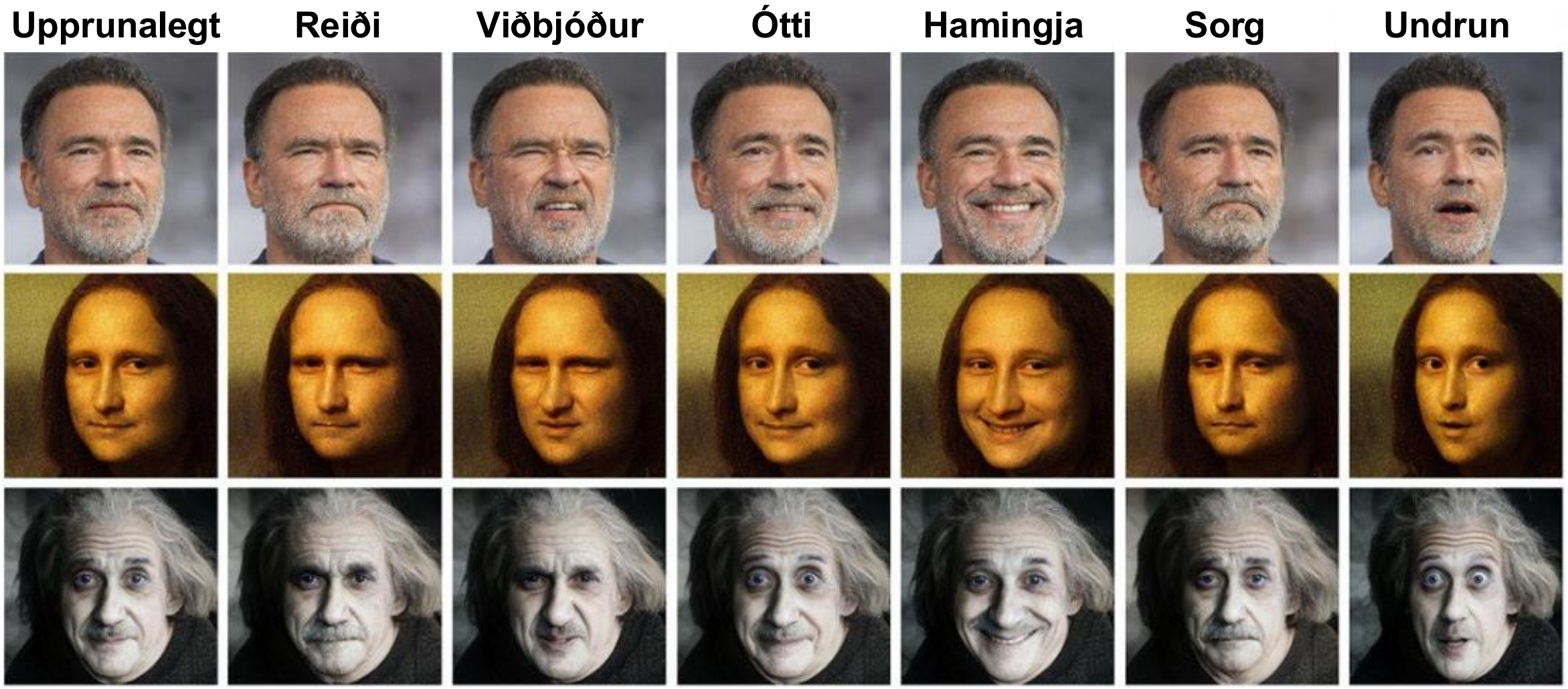
Hamingja

Viðbjóður

Sorg

Icelandic

Conditional Generative AI

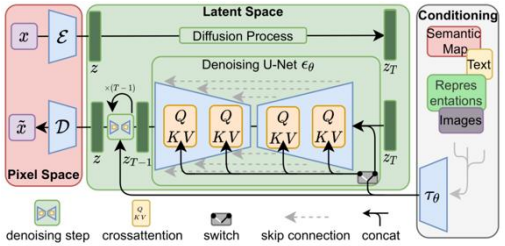


https://arxiv.org/pdf/2205.06102.pdf

Example







Latent Diffusion Models
Rombach et al. 2021

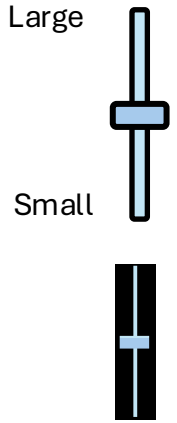


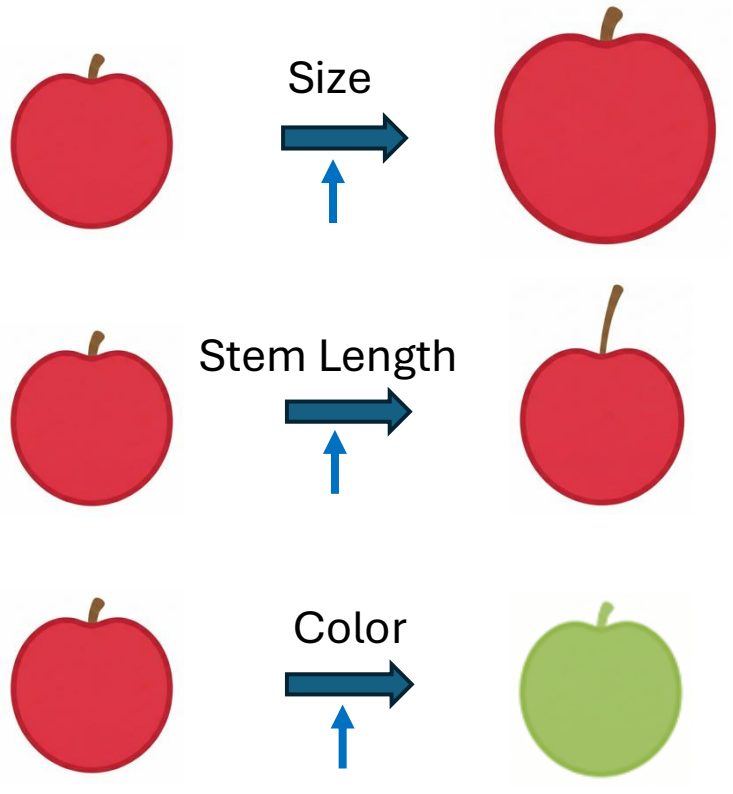
Feature 1

Feature 2

Feature 3

...

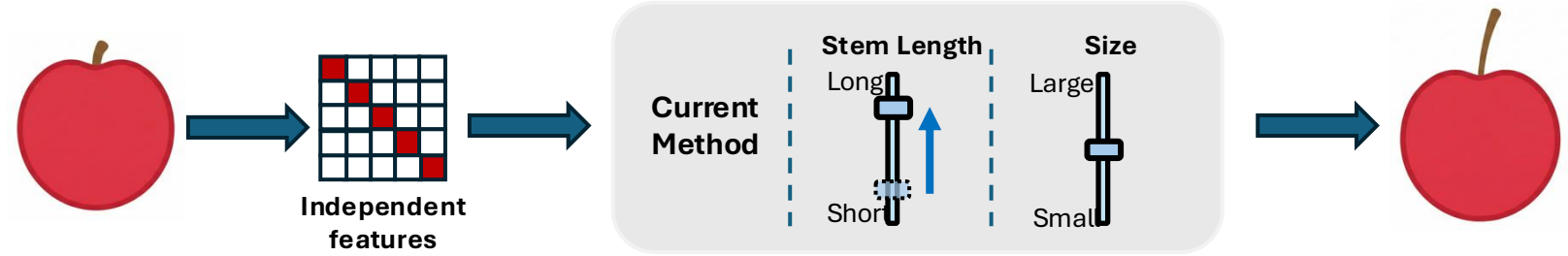




Yang, Yuechen … Yuankai Huo. “Explainable Pathomics Feature Visualization via Correlation-aware Conditional Feature Editing.” *Machine Learning in Medical Imaging* (MIDL 2026):.

Possible for Pathomics?

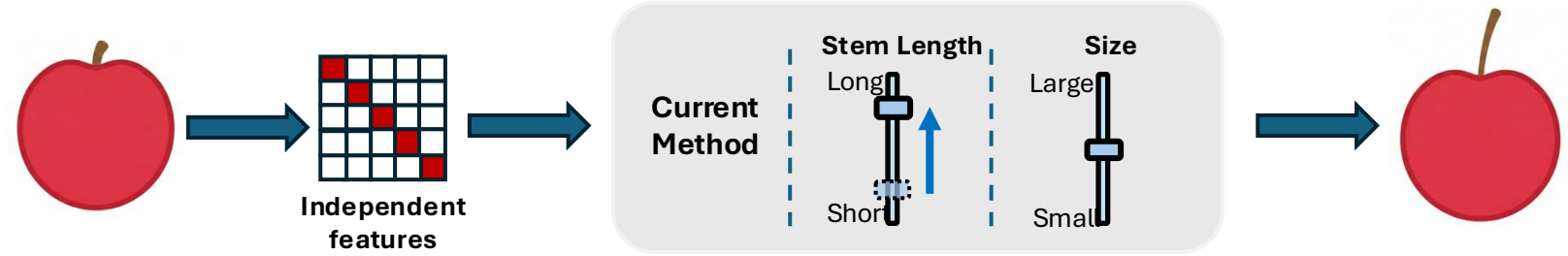
(a) Ideally: An editing-based explanation AI could independently visualize features using conditional diffusion.



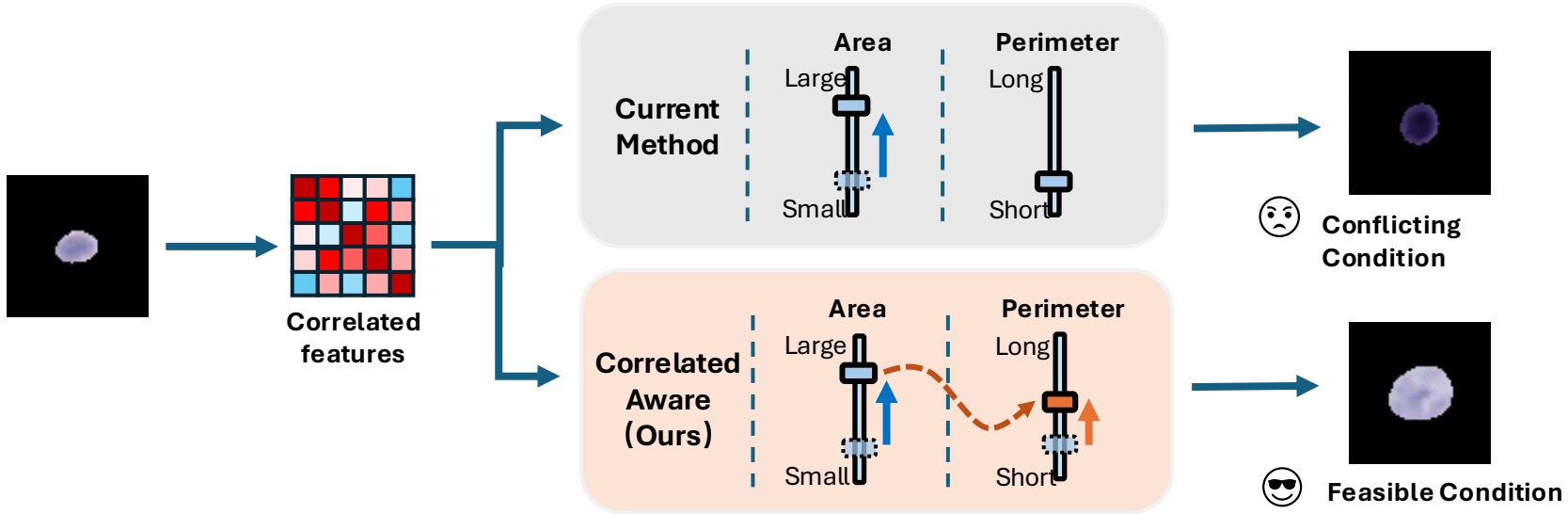
Yang, Yuechen … Yuankai Huo. “Explainable Pathomics Feature Visualization via Correlation-aware Conditional Feature Editing.” *Machine Learning in Medical Imaging* (MIDL 2026):.

Explainable Pathomics

(a) Ideally: An editing-based explanation AI could independently visualize features using conditional diffusion.

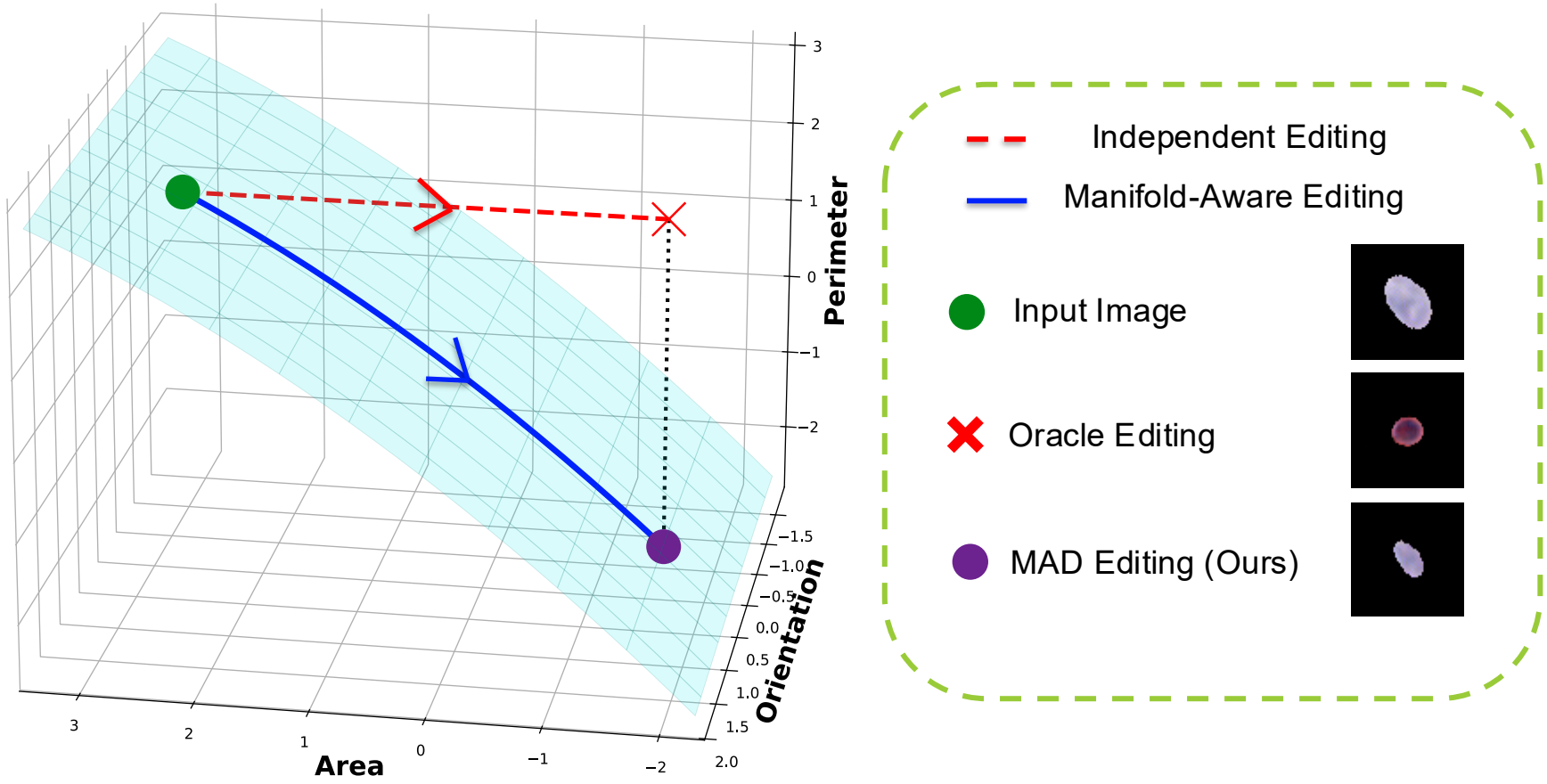


(b) In reality: Explaining pathomics features with such an AI is challenging because the features are often correlated, so editing a single feature may introduce conflicts.



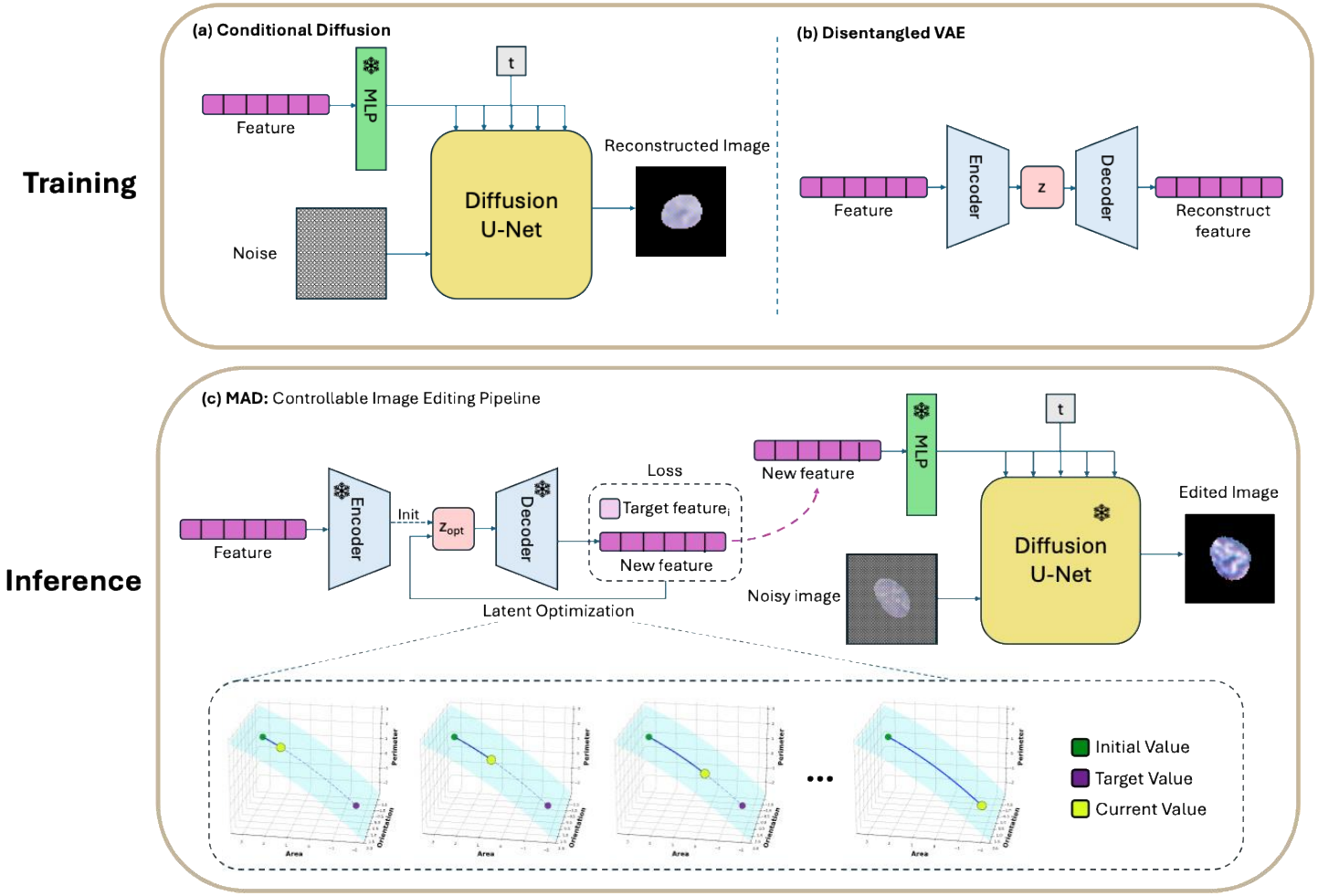
Yang, Yuechen … Yuankai Huo. "Explainable Pathomics Feature Visualization via Correlation-aware Conditional Feature Editing." *Machine Learning in Medical Imaging* (MIDL 2026):.

Explainable Pathomics



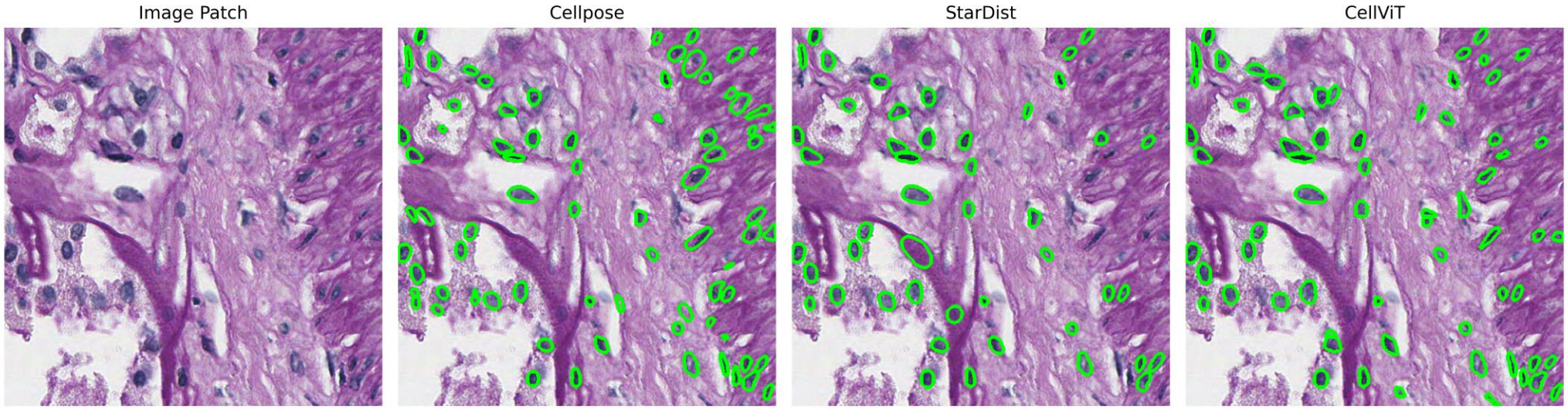
Yang, Yuechen … Yuankai Huo. “Explainable Pathomics Feature Visualization via Correlation-aware Conditional Feature Editing.” *Machine Learning in Medical Imaging* (MIDL 2026):.

Explainable Pathomics



Yang, Yuechen … Yuankai Huo. "Explainable Pathomics Feature Visualization via Correlation-aware Conditional Feature Editing." *Machine Learning in Medical Imaging* (MIDL 2026):.

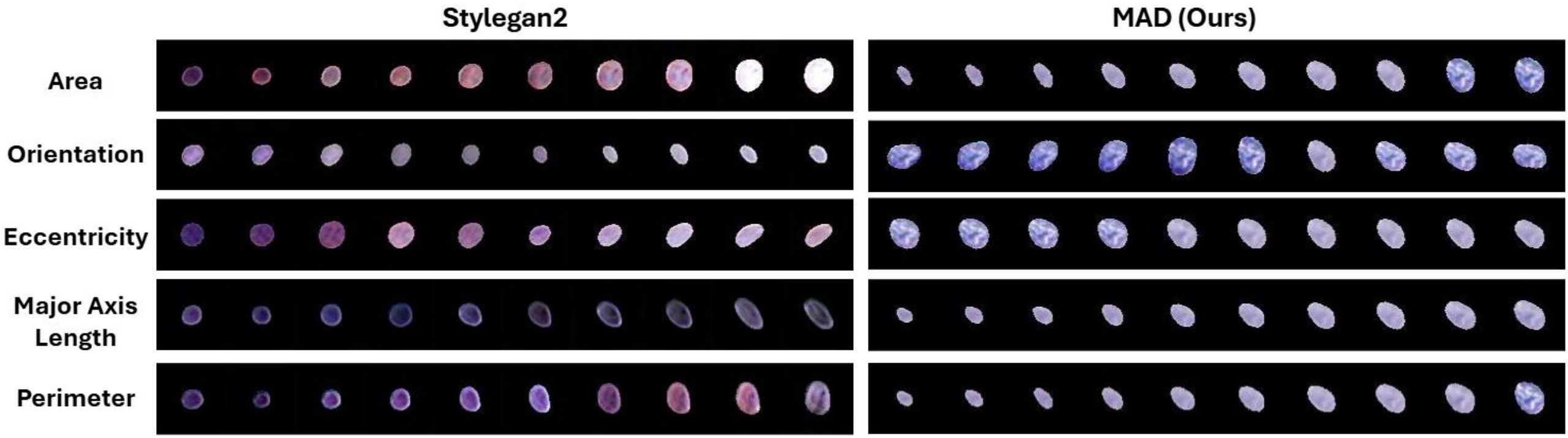
Explainable Pathomics



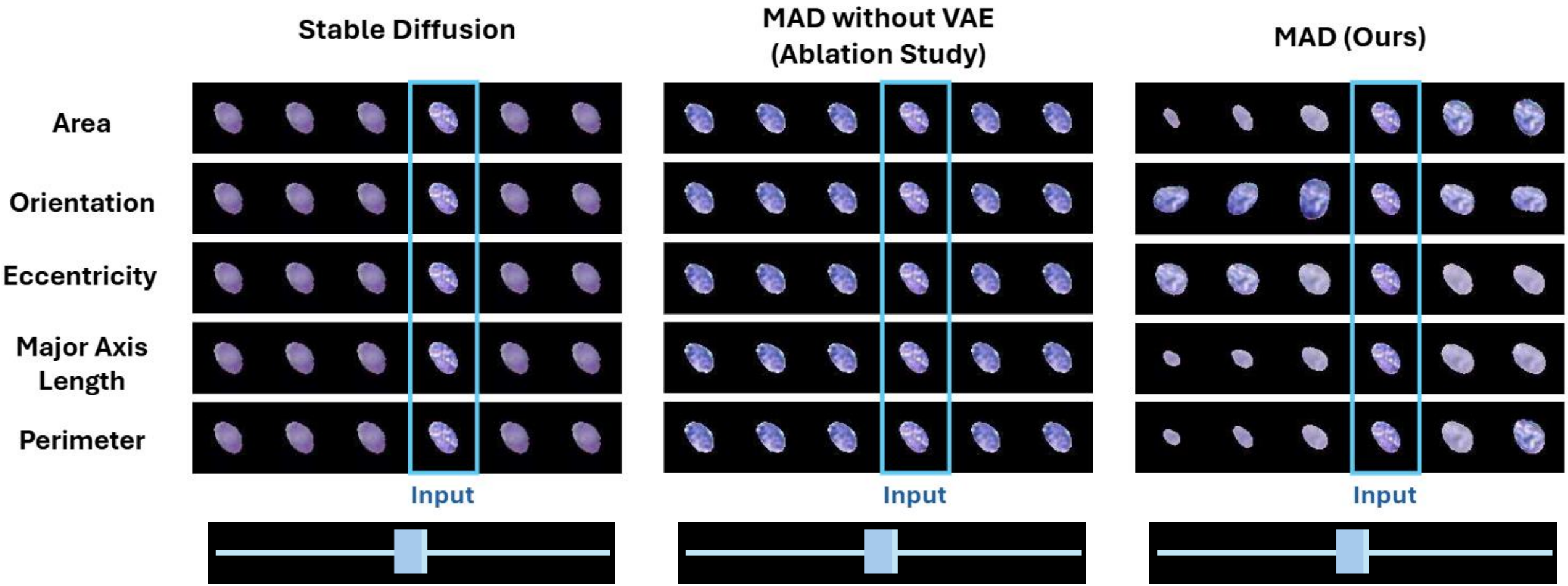
- 1,556 whole-slide images (WSIs) of kidney tissue at 40x magnification
- 5 random 512x512 patches extracted per WSI
- Each nucleus cropped and resized to 64x64 patches
- 93,643 nucleus-centered patches
- 75-dimensional feature vector per nucleus, extracted using Pyspatial

Yang, Yuechen … Yuankai Huo. “Explainable Pathomics Feature Visualization via Correlation-aware Conditional Feature Editing.” *Machine Learning in Medical Imaging* (MIDL 2026):.

Explainable Pathomics



Explainable Pathomics





Size 0%
Perimeter 0%
Intensity 0%
Orientation 0°



Size 0%
Perimeter 0%
Intensity 0%
Orientation 0°



Size 0%
Perimeter 0%
Intensity 0%
Orientation 0°



Size 0%
Perimeter 0%
Intensity 0%
Orientation 0°



Size 0%
Perimeter 0%
Intensity 0%
Orientation 0°



Size 0%
Perimeter 0%
Intensity 0%
Orientation 0°



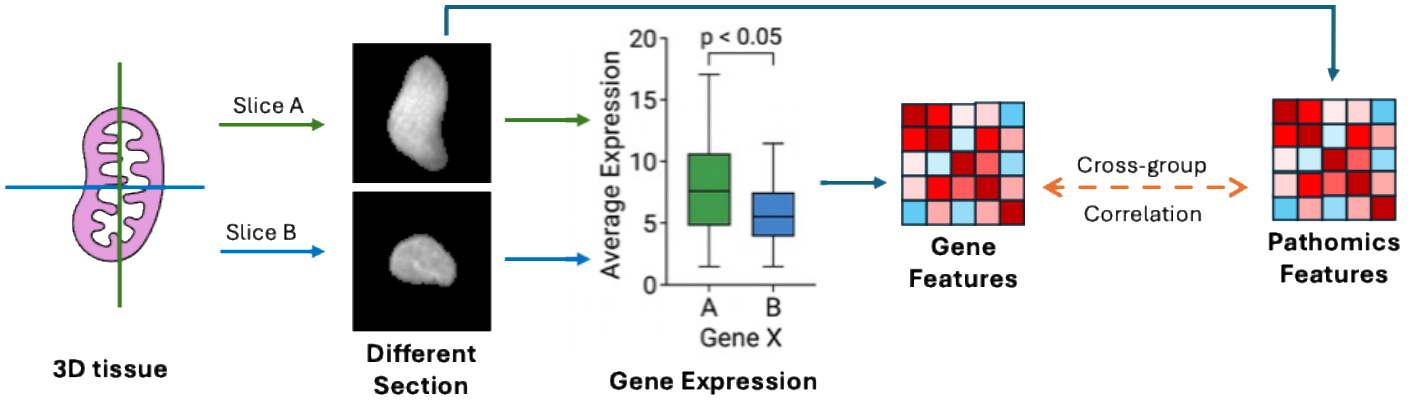
Size 0%
Perimeter 0%
Intensity 0%
Orientation 0°



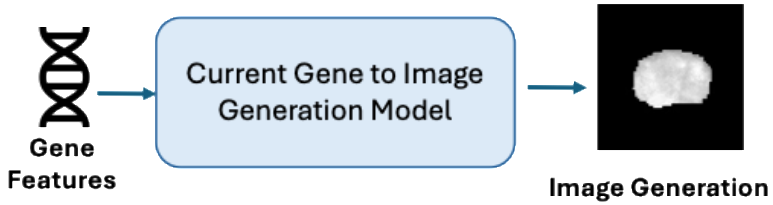
Size 0%
Perimeter 0%
Intensity 0%
Orientation 0°

Associate with Spatial Omics

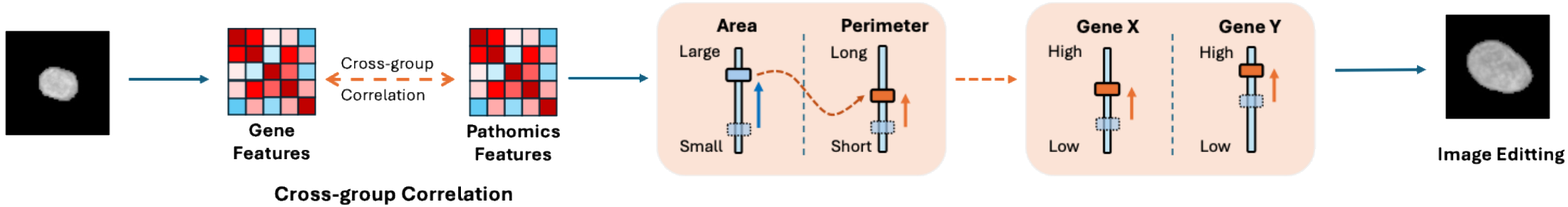
(a) Problem: The orientation of tissue sectioning alters the 2D cross-sectional composition, which acts as a confounding factor in the correlation between pathomic features and measured gene expression levels.



(b) Current: Most gene to image model is unconditional generation.

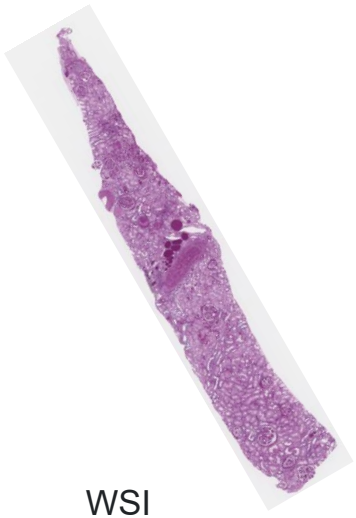


(c) Our Method: By integrating pathomics data, the proposed model achieves condition-driven image editing.

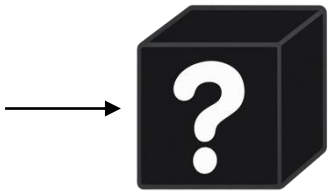


Yang, Yuechen … Haichun Yang, Yuankai Huo. “From Cross-Section to Cross-Omics: Conditional Cell Nucleus Image Editing in Spatial Transcriptomics” MICCAI MISO Workshop 2026.

Kidney Pathology

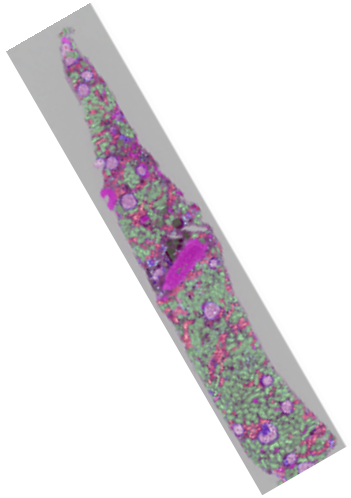


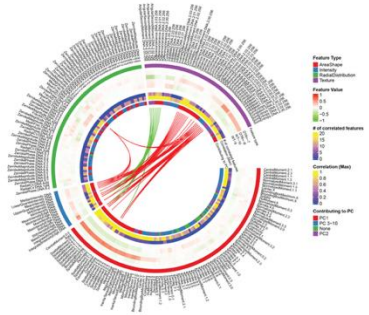
WSI



Still Black Box Model

Segmentation AI





Evidence

Pathomics



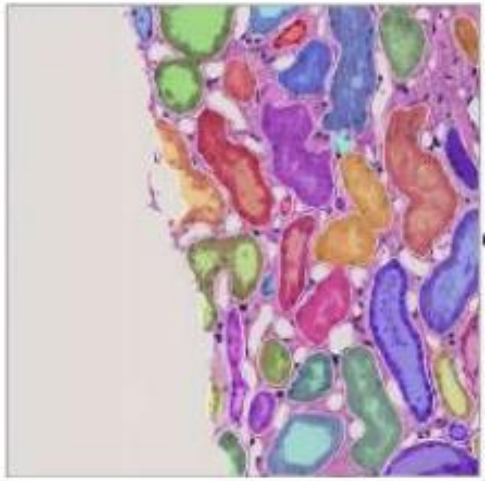
Analyze

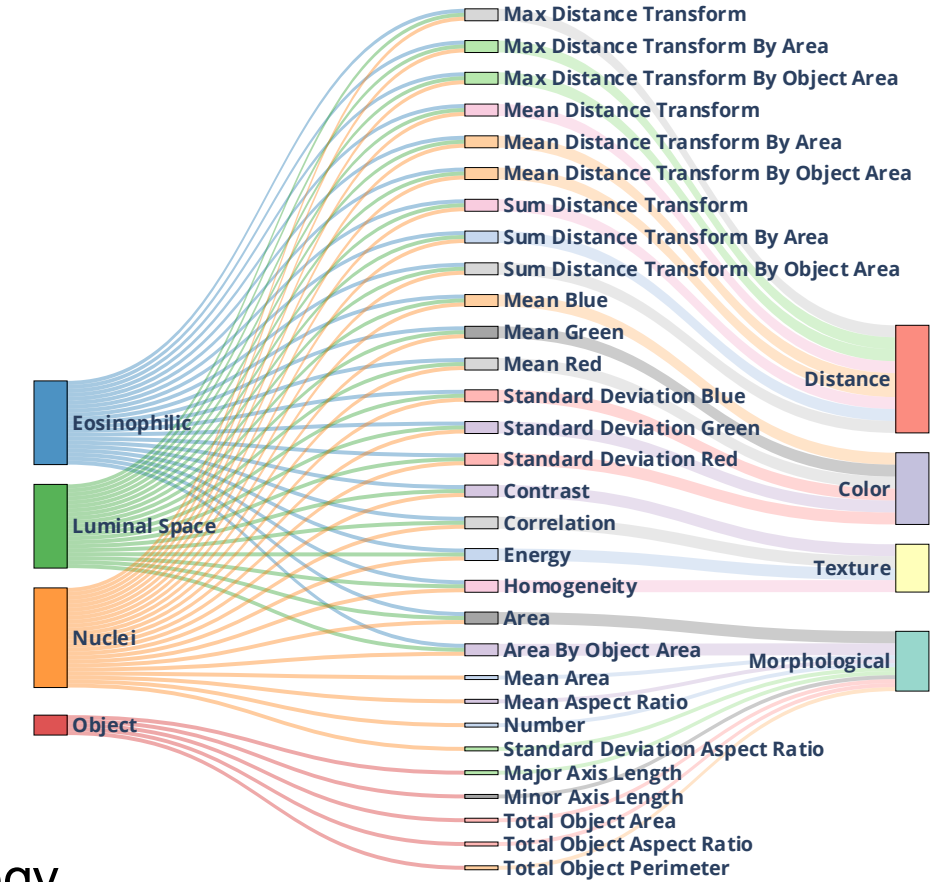
Challenge 1

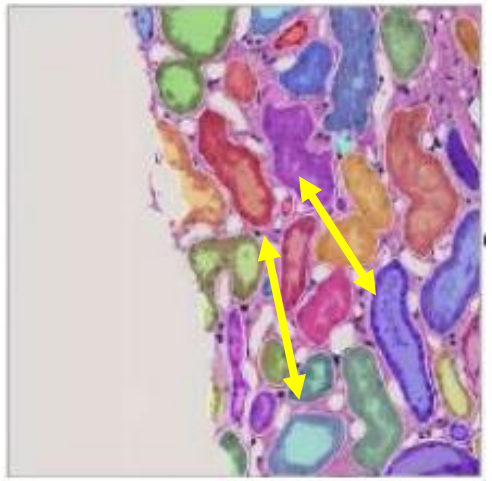
Challenge 2

Challenge 3

Challenge 3 How to Analyze







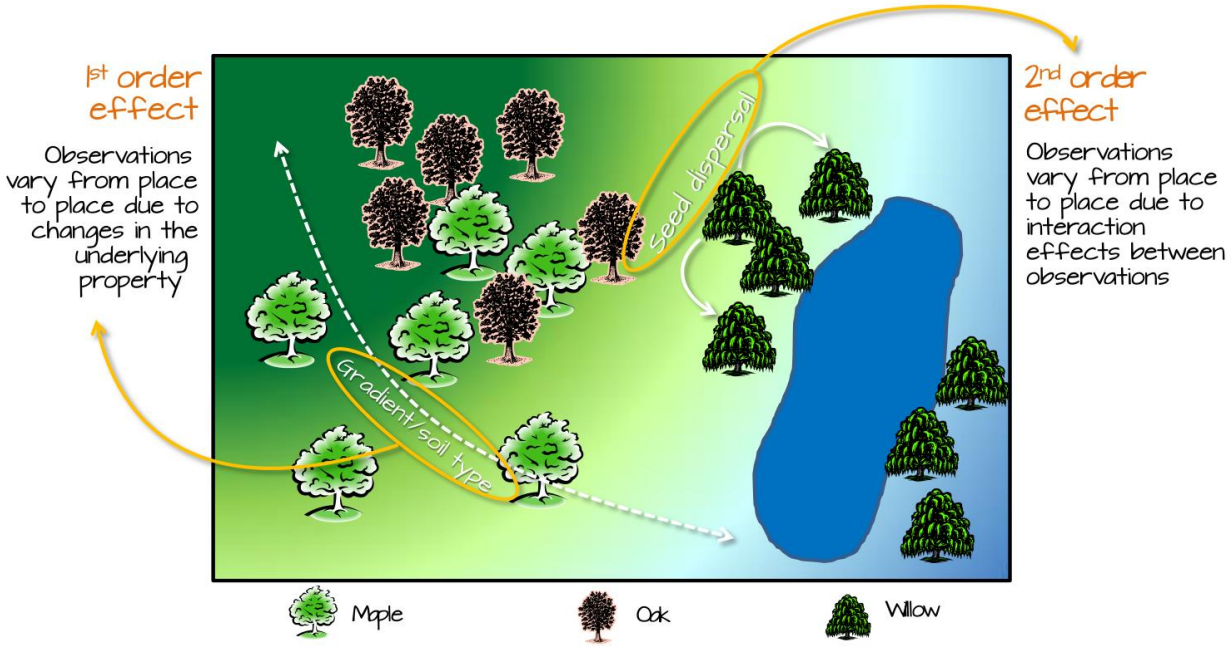
Morphology

Spatial Distribution?

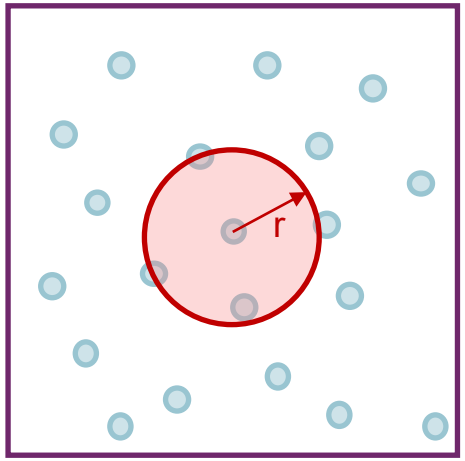
Jiayuan Chen … Yuankai Huo. “Spatial Pathomics Toolkit for Quantitative Analysis of Podocyte Nuclei with Histology and Spatial Transcriptomics Data in Renal Pathology” SPIE Medical Imaging 2025.

Geographic Information System (GIS)

Spatial Analysis

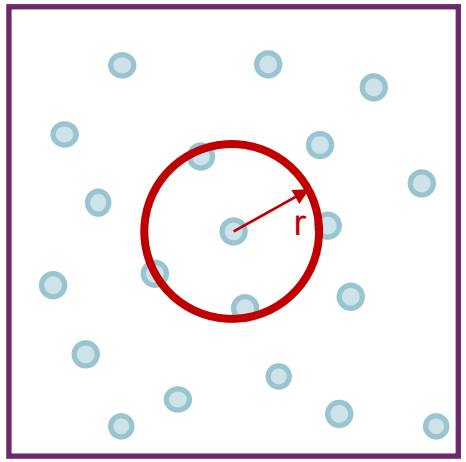


L-function



Shows whether points are clustered, randomly distributed, or dispersed

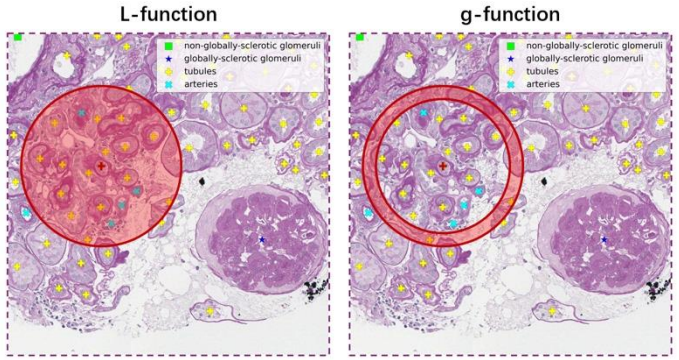
g-function

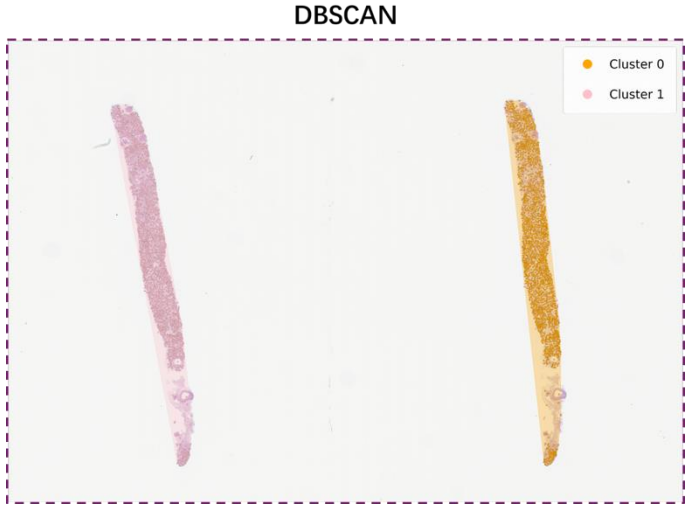


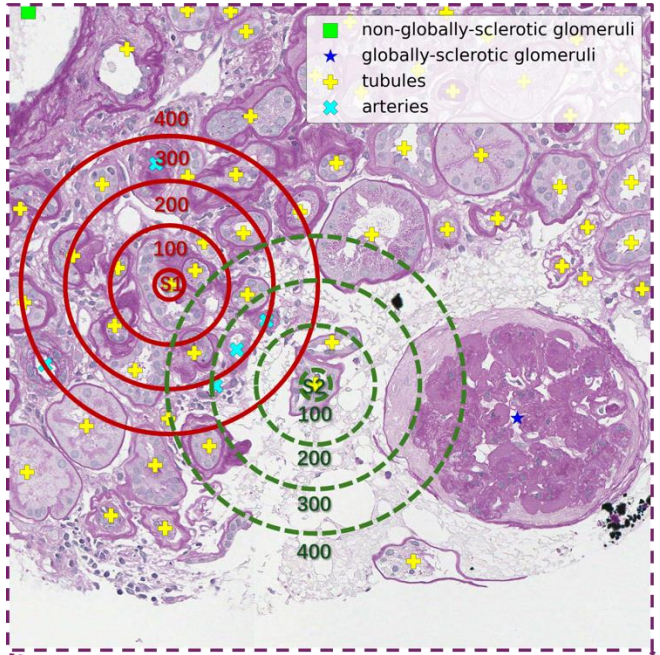
Shows how close points are to one another

First-order effects occur when environmental or socioeconomic factors cause values to vary across locations. **Second-order** effects occur when observations influence one another, creating spatial patterns such as clustering, dispersion, or autocorrelation.

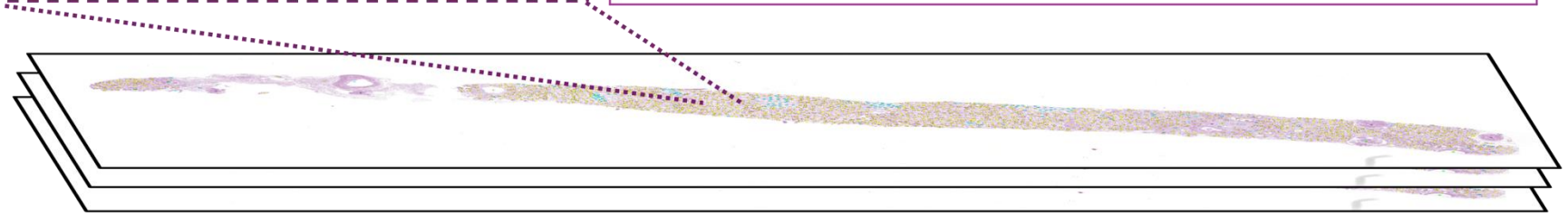
https://mgimond.github.io/Spatial/





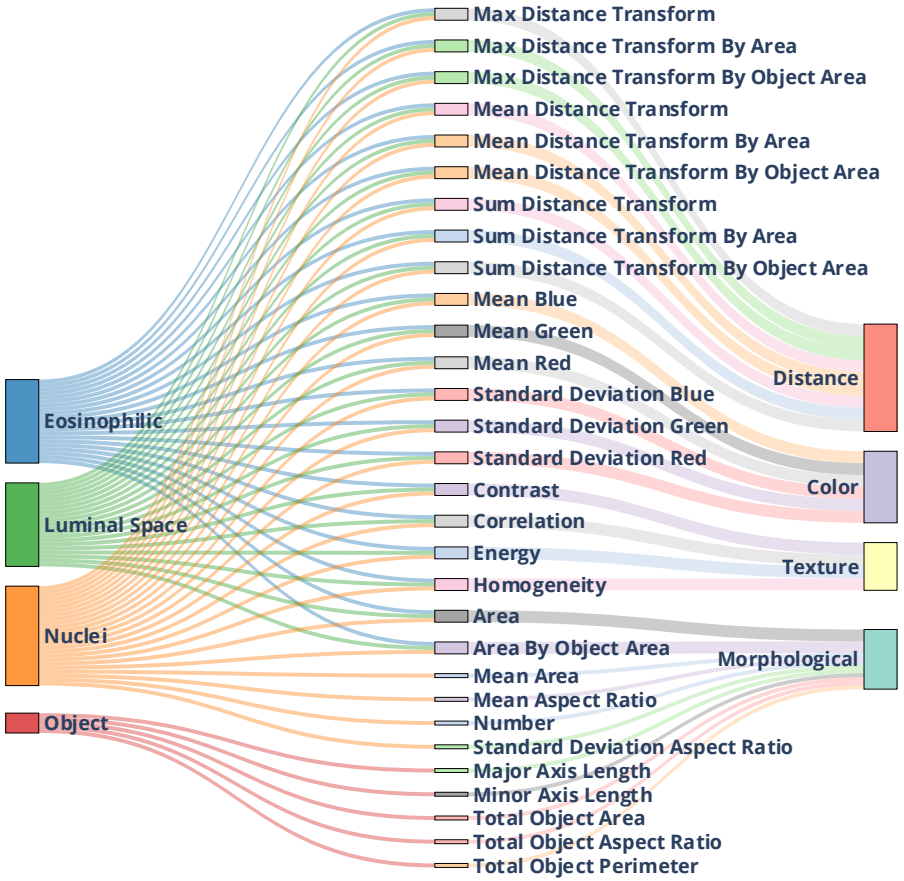


Distance (pixel)	Events From S1	Events From S2	Events From Si	K
100	0	0	···	0.012
200	3	1	···	0.067
300	9	1	···	0.153
400	14	4	···	0.269

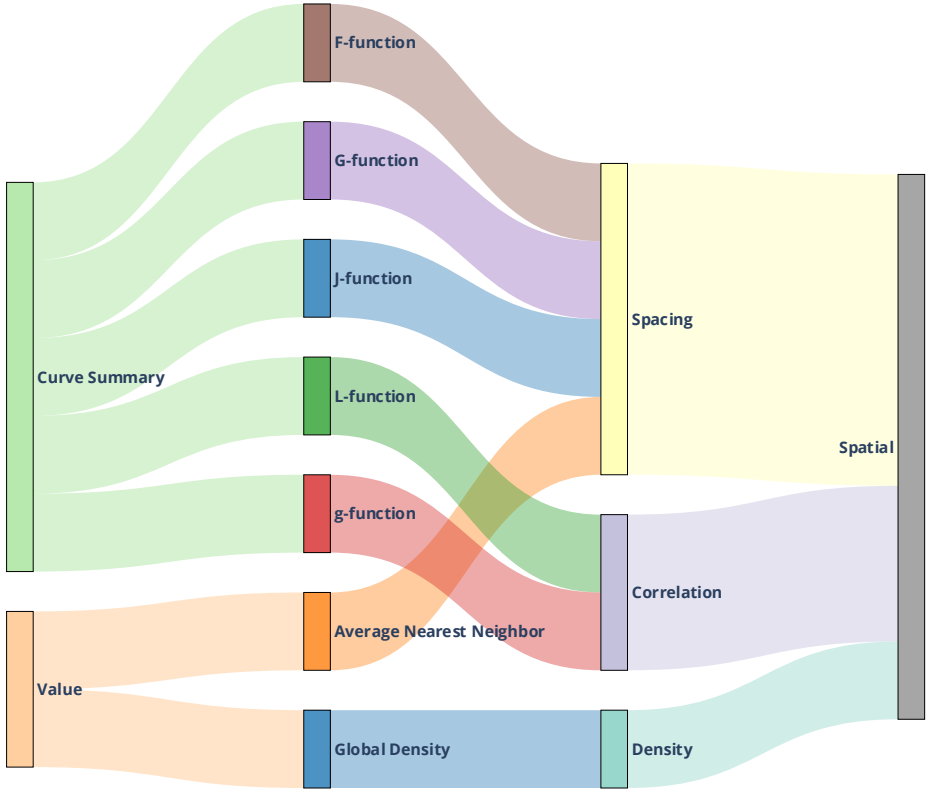


Yang, Yuechen … Jovan Tanevski, Laura Barisoni, Avi Z. Rosenberg, Yuankai Huo. “HistoWAS: A Pathomics Framework for Large-Scale Feature-Wide Association Studies of Tissue Topology and Patient Outcomes” Electronic Imaging 2026.

Many Objects

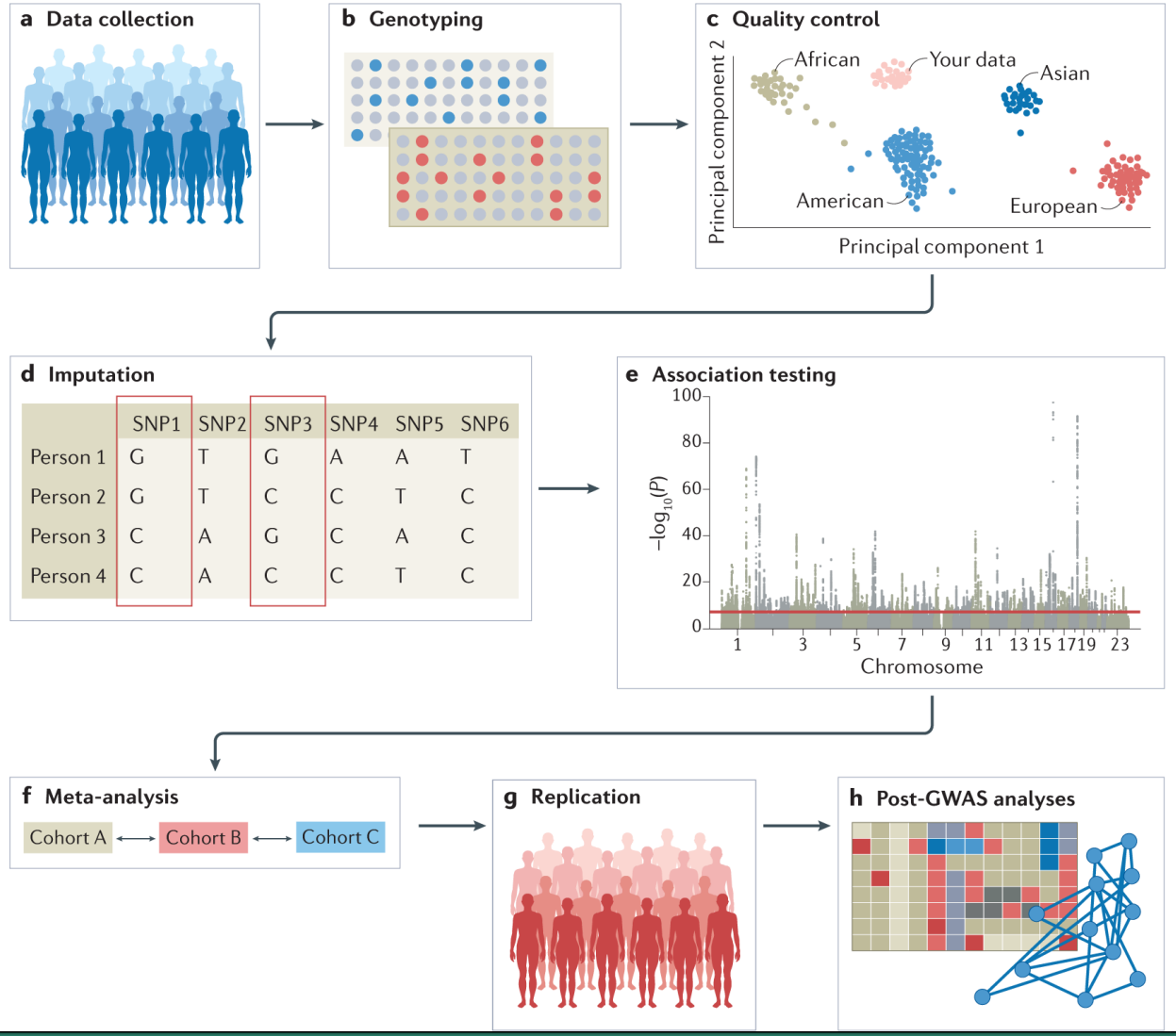


Morphological Pathomics Features



Spatial Pathomics Features

Associate with Clinical Features?



Primer | Published: 26 August 2021

Genome-wide association studies

Emil Uffelmann, Qin Qin Huang, Nchangwi Syntia Munung, Jantina de Vries, Yukinori Okada, Alicia R. Martin, Hilary C. Martin, Tuuli Lappalainen & Danielle Posthuma

Nature Reviews Methods Primers **1**, Article number: 59 (2021) | Cite this article

Save article

519k Accesses | **2038** Citations | **553** Altmetric | Metrics

KPMP – Descriptor Score

Descriptor Score：Serving as the Phenotype

- Clinical descriptor score assessed by pathologists

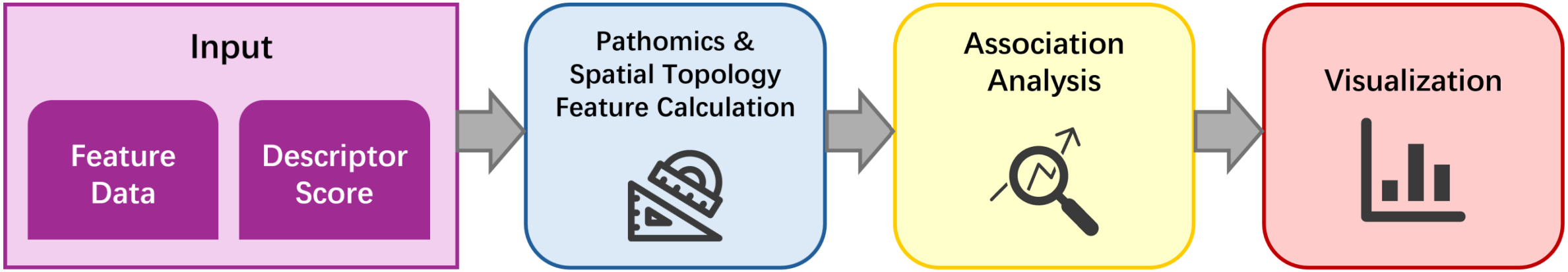
206 Number of Descriptor Scores
385 Number of Whole slide images

	Score 1	Score 2	…	Score N
Biopsy 1	XXX	XXX	…	XXX
Biopsy 2	XXX	XXX	…	XXX
Biopsy 3	XXX	XXX	…	XXX
…	…	…	…	XXX
Biopsy N	XXX	XXX	…	XXX

Number of biopsy: 385

Yang, Yuechen … Jovan Tanevski, Laura Barisoni, Avi Z. Rosenberg, Yuankai Huo. “HistoWAS: A Pathomics Framework for Large-Scale Feature-Wide Association Studies of Tissue Topology and Patient Outcomes” Electronic Imaging 2026.

HistoWAS



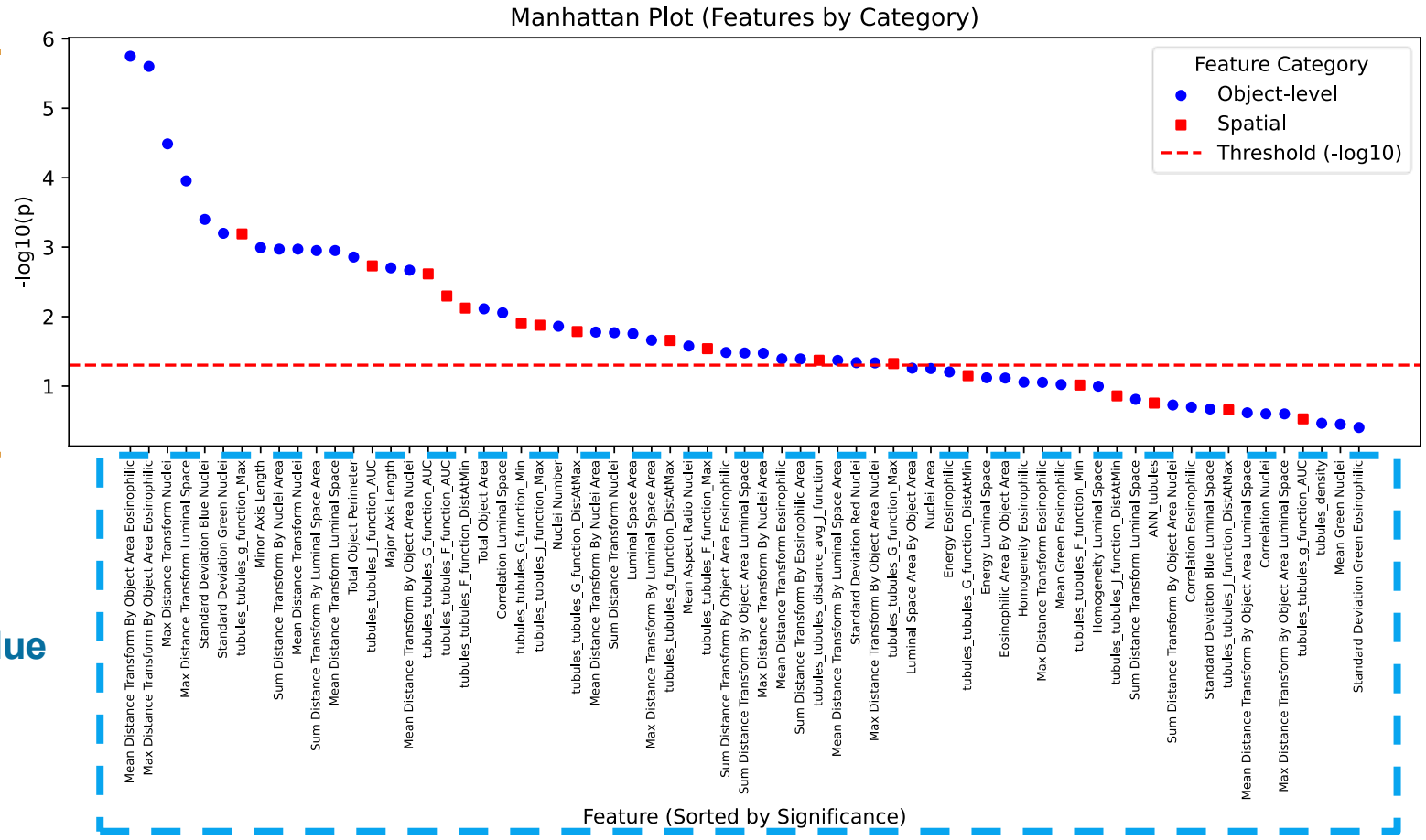
Yang, Yuechen … Jovan Tanevski, Laura Barisoni, Avi Z. Rosenberg, Yuankai Huo. “HistoWAS: A Pathomics Framework for Large-Scale Feature-Wide Association Studies of Tissue Topology and Patient Outcomes” Electronic Imaging 2026.

Manhattan plot

Goal: Identify the statistically significant features

Interstitial Fibrosis Score as an example.

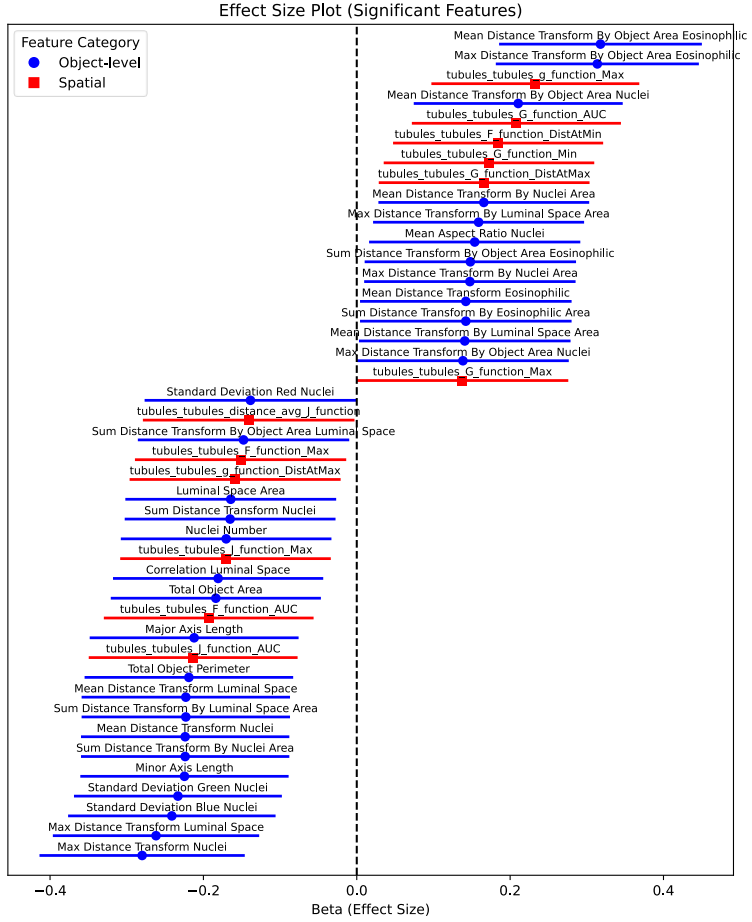
$-\log_{10}(p)$



Features: ranked by their p-value

Effectsize Plot

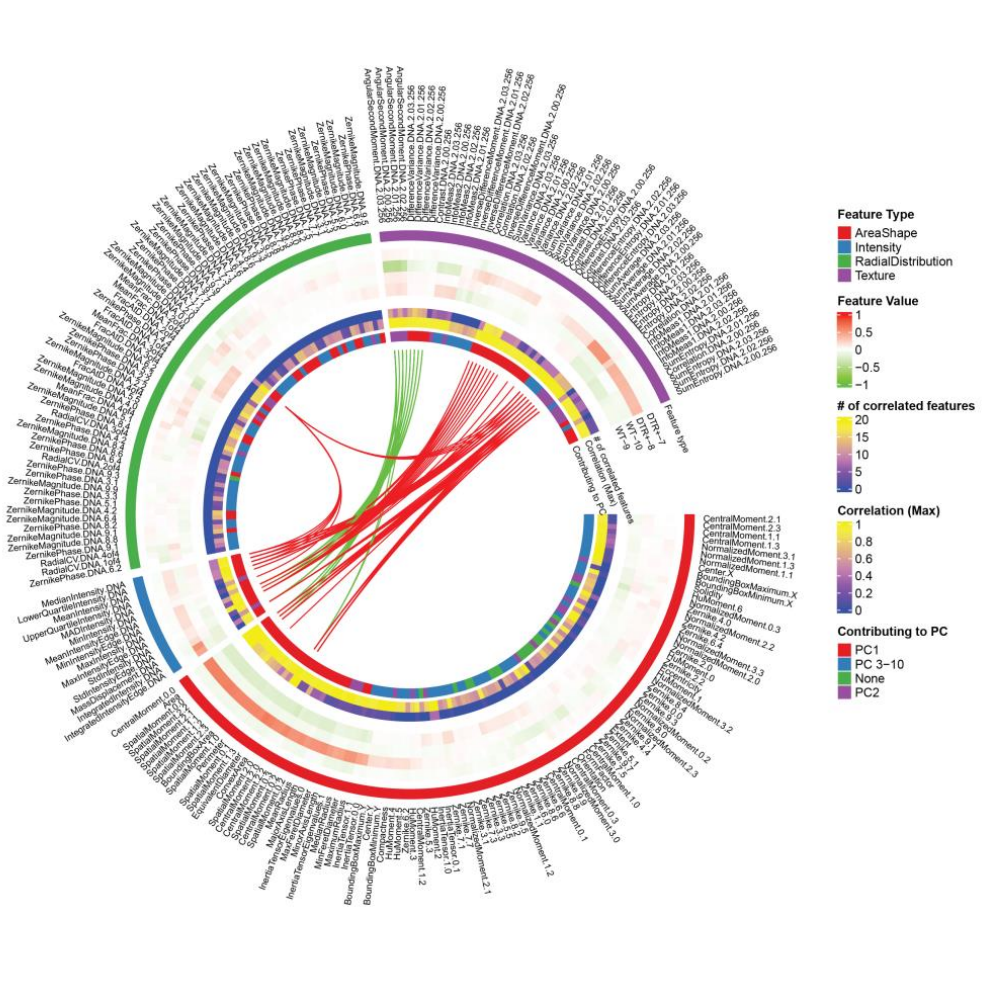
Goal: To visualize the effect direction and magnitude for those significant features.

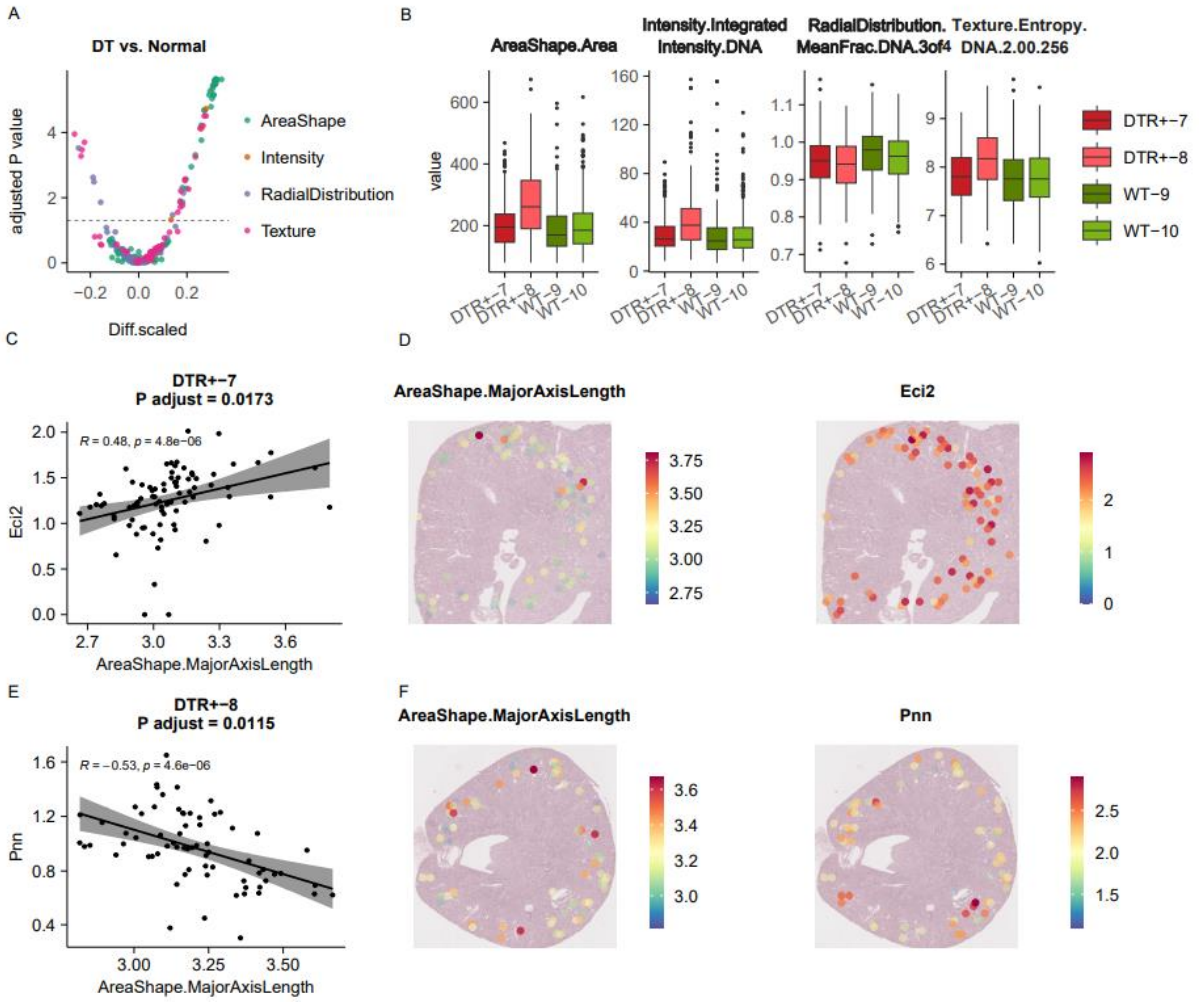


Negative effect — **Positive effect**

Stronger ← → **Stronger**

Spatial Pathomics





Jiayuan Chen … Yuankai Huo. “Spatial Pathomics Toolkit for Quantitative Analysis of Podocyte Nuclei with Histology and Spatial Transcriptomics Data in Renal Pathology” SPIE Medical Imaging 2025.

Conclusions

1. **Evidence-based explainable AI** offers a promising pathway toward enhancing clinical confidence and supporting reliable decision-making in AI-driven medical image analysis.
2. **Comprehensive spatial pathomics** could serve as a powerful learning framework for developing evidence-based explainable AI, enabling more interpretable and biologically grounded predictions.





Thank you!



We appreciate the support from the funding agencies!

Acknowledgment:

- NIH R01 DK135597
- NSF 2434229
- DoD HT9425-23-1-0003
- NEH HAA-293452-23
- NSF 2419793
- NIH P30CA068485-29S1
- ACR FCRI
- ONR N00014-21-1-2468
- KPMP Young Investigator Travel Grant
- KPMP Glue Grant
- Vanderbilt Seed Success Grant
- Vanderbilt Discovery Grant
- Vanderbilt-Liverpool Collaboration Grant
- VISE Seed Grant
- NIH R01 EB033385
- NIH R01 DK132338
- NIH R01 MH125931
- NVIDIA Hardware funding
- The Leona M. and Harry B. Helmsley Foundation