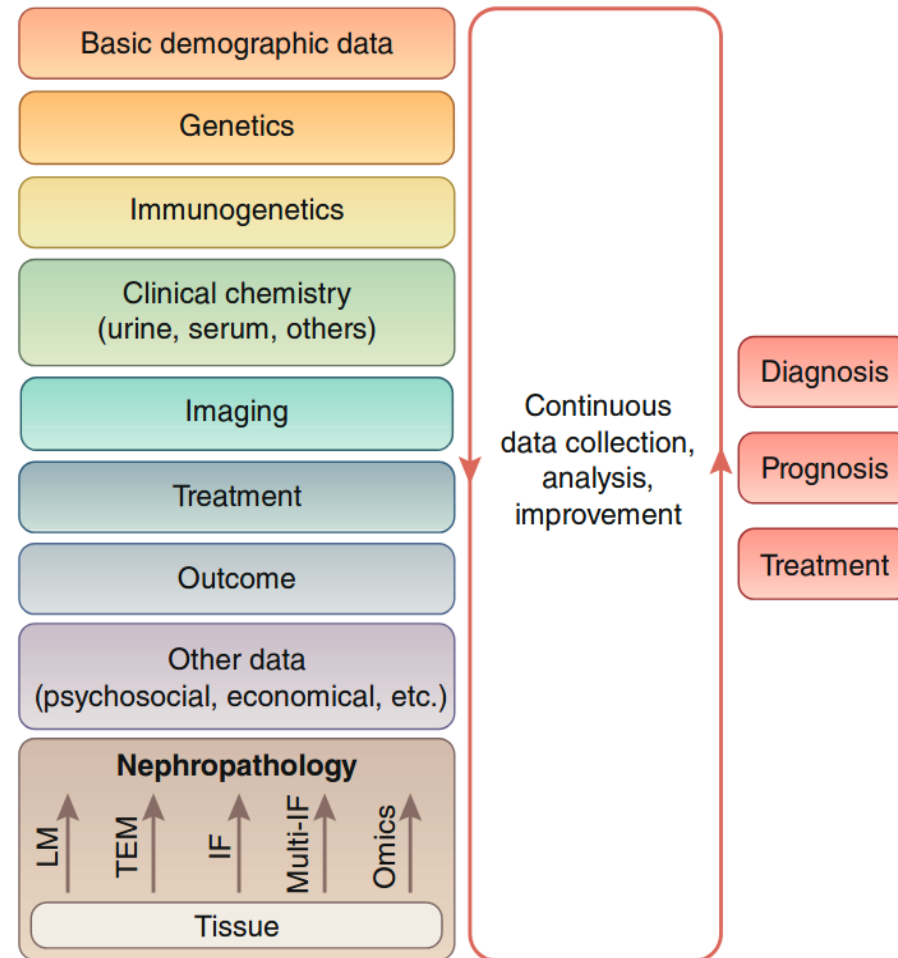


# Machine Learning in Nephropathology: Basic Concepts

Jan U. Becker  
Cologne, Germany

# Nephropathology, Big Data and Precision Medicine



# How to extract extra information from biopsy tissue?

	Reproducibility	Cost	Tissue use	Hypothesis-free
Transcriptomics	++	+++	yes	yes
Proteomics	++	+++	yes	yes
Immunostaining	+++	+	yes	no
ML/computer vision	++++	+?	no	yes

ML/computer vision is the only available reproducible, low cost method for hypothesis-free research without tissue use

# Intelligence

Input



Intelligence

Meaningful output

# Intelligence in nephropathology

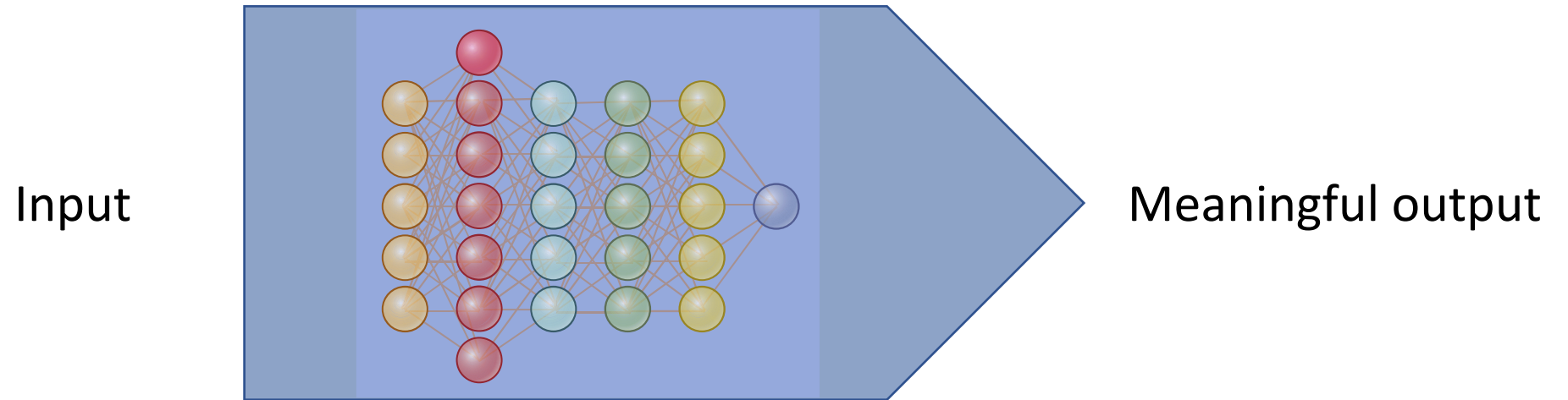
Clinical/laboratory  
Light microscopy  
IF/IHC  
Electron microscopy  
Genetics



Intelligence

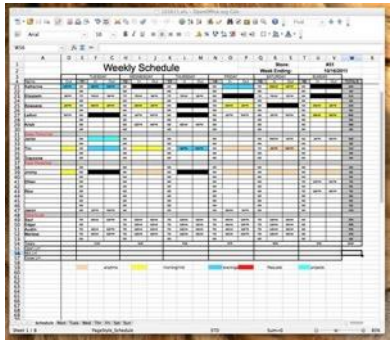
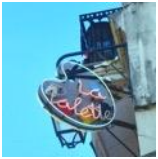
Diagnosis  
Prognosis  
Treatment

# Artificial intelligence/machine learning



# Machine learning

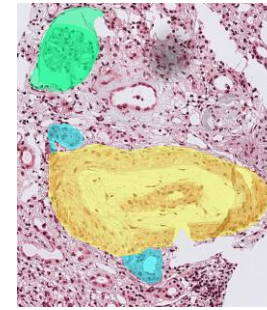
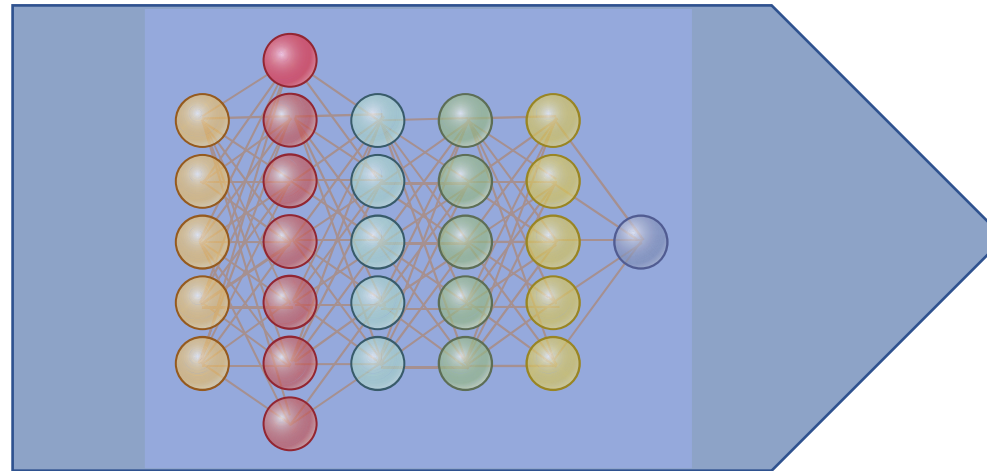
Lorem ipsum dolor sit amet,  
consectetur adipiscing elit,



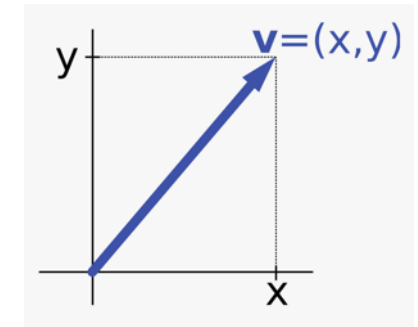
Weekly Schedule						
Day	Mon	Tue	Wed	Thu	Fri	Sat
Class	Math	Science	History	Art	Music	Physical Education
Homework	Math	Science	History	Art	Music	Physical Education
Projects	Math	Science	History	Art	Music	Physical Education
Tests	Math	Science	History	Art	Music	Physical Education



Input

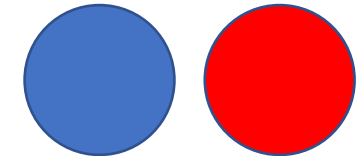


Segmentation



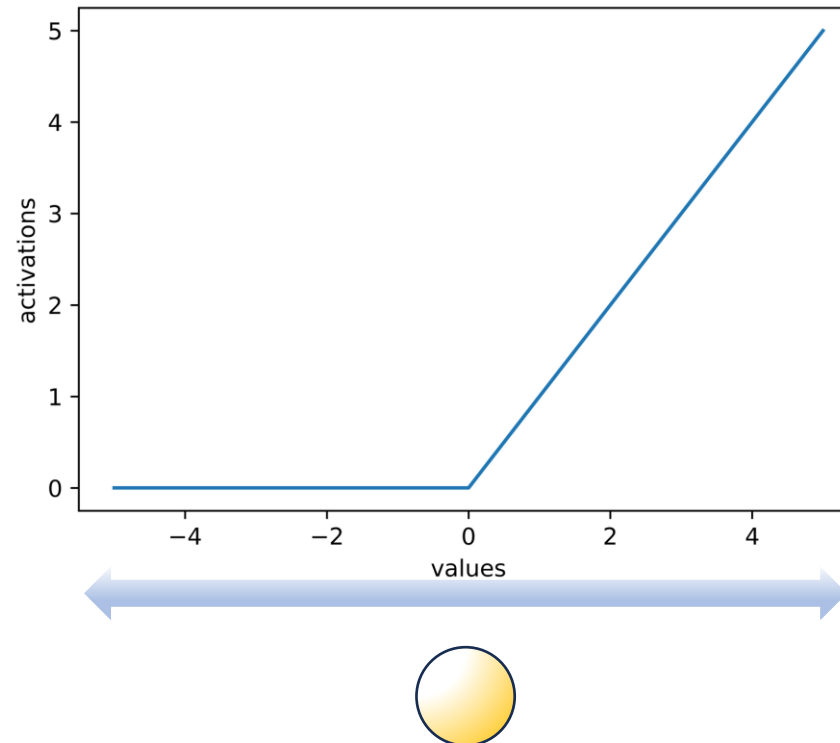
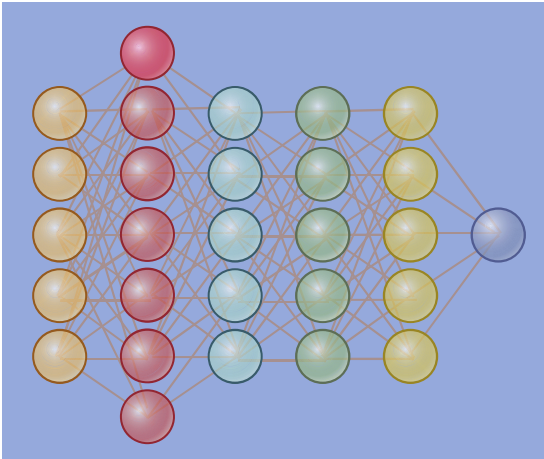
Regression

Meaningful output



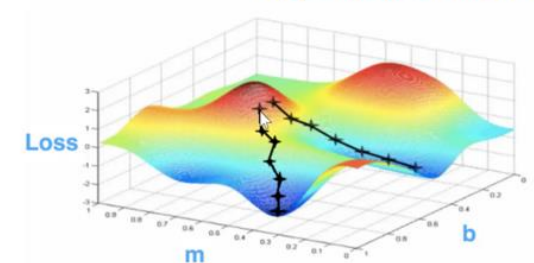
Classification (binary, multi)

# How to train a Neural Network



## Gradient Descent

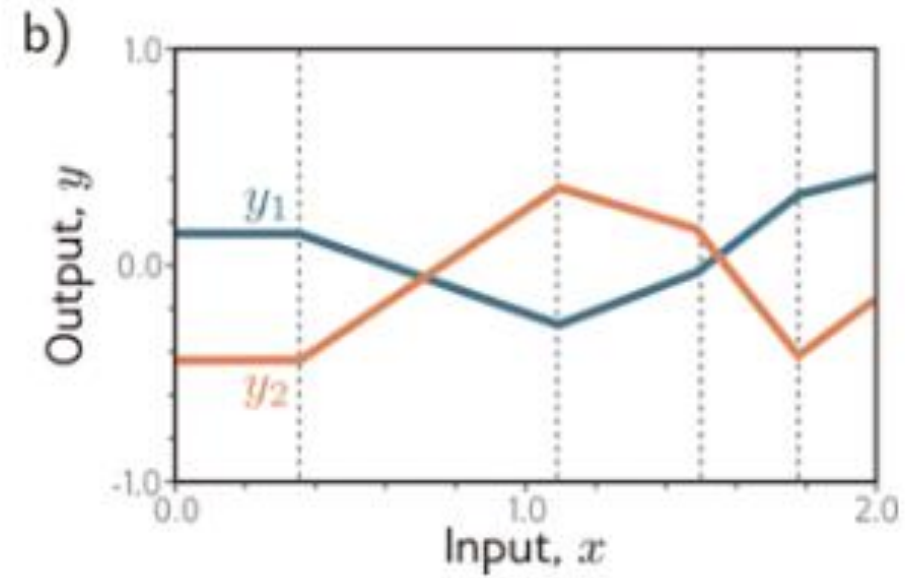
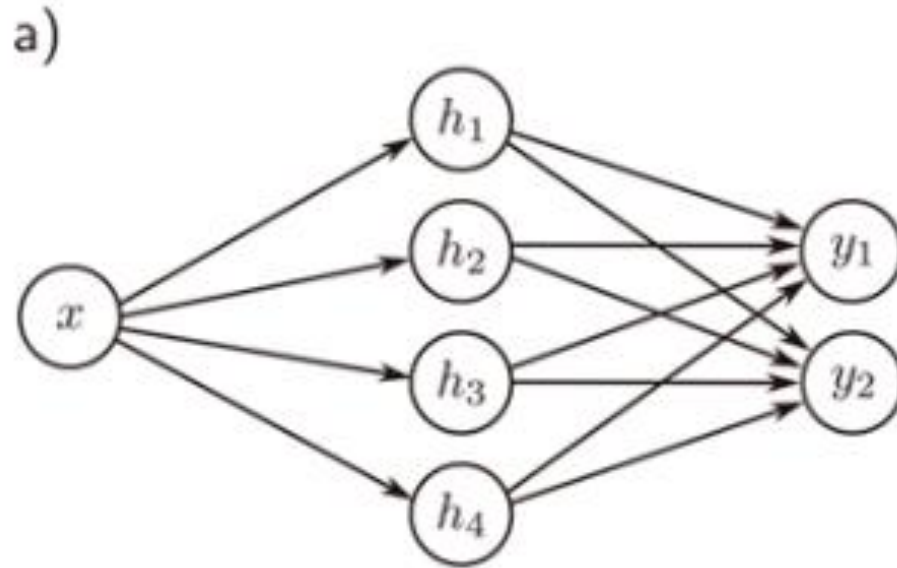
$f(x) = \text{nonlinear function of } x$



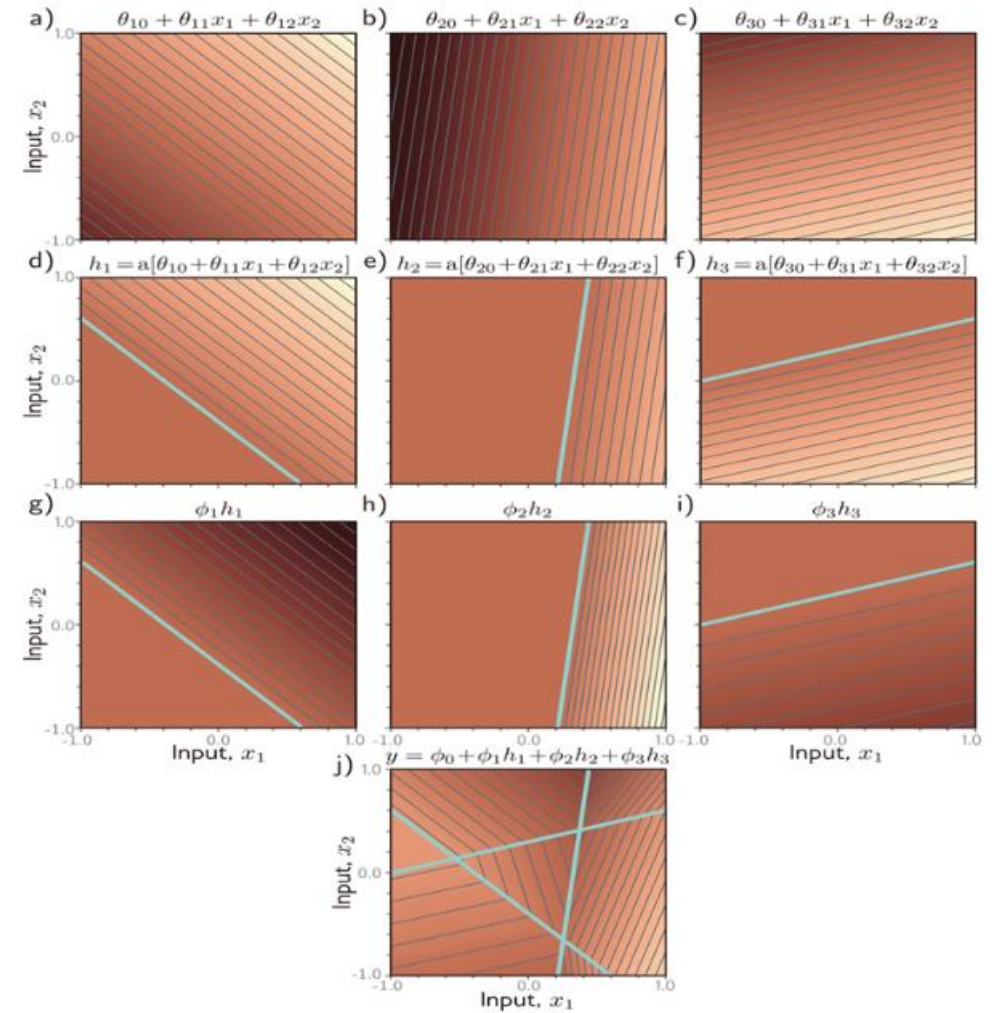
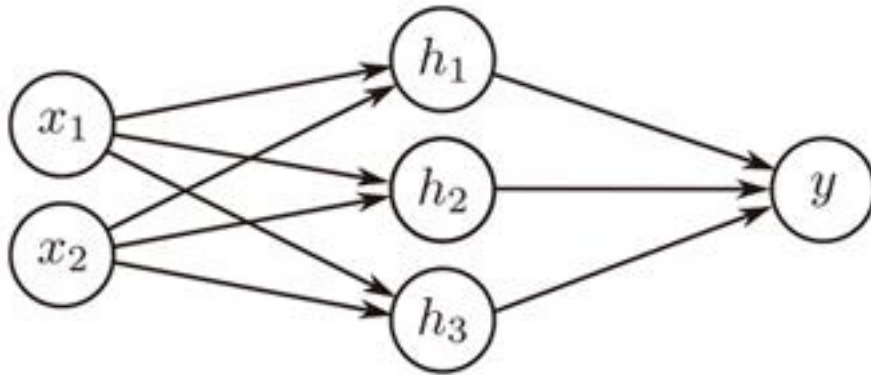
Minimise Loss Function  
(Empirical Values vs. Model Predictions)



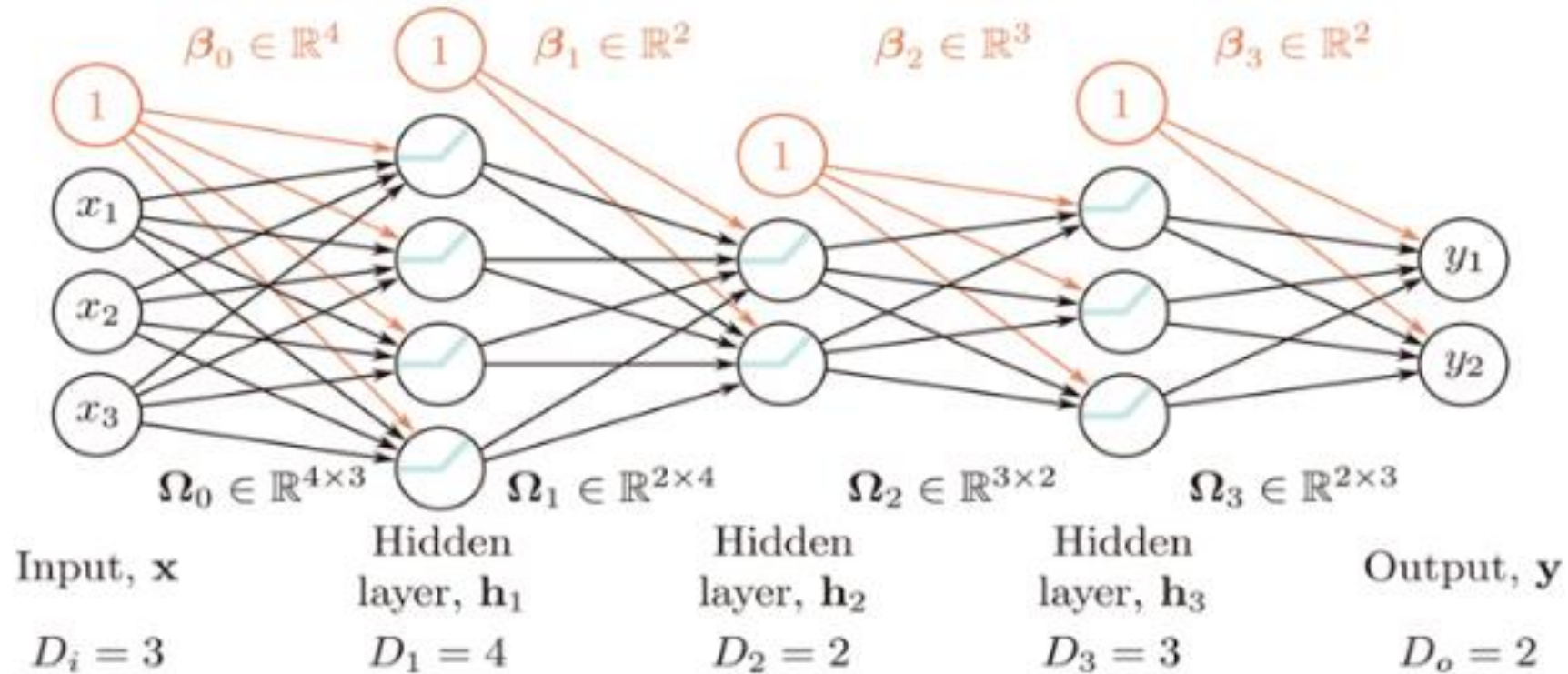
# Shallow Neural Networks



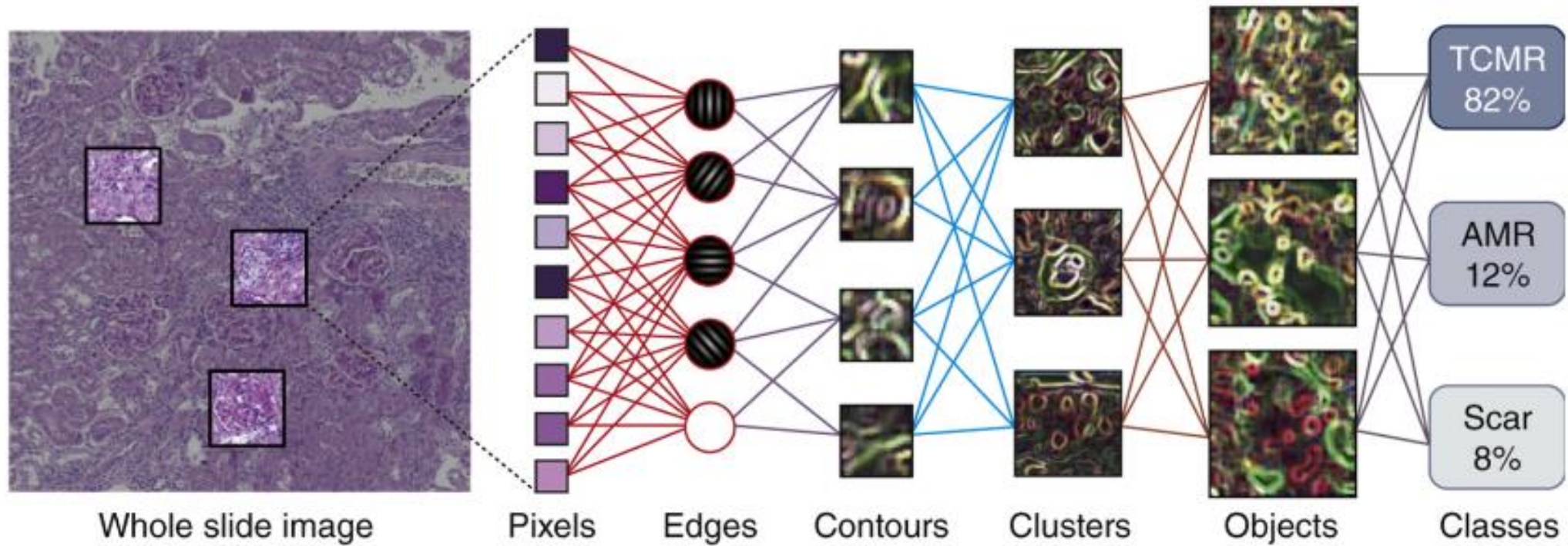
# Shallow Neural Networks



# Deep Neural Networks



# Convolutional Neural Network



# Machine Learning

- Supervised  
Expert labels (ground truth) for all instances in a dataset
- Weakly supervised  
e.g. only biopsy label (ground truth diagnosis, prognosis, treatment response) for entire dataset
- Unsupervised  
No labels at all (foundation models)

# Supervised Learning

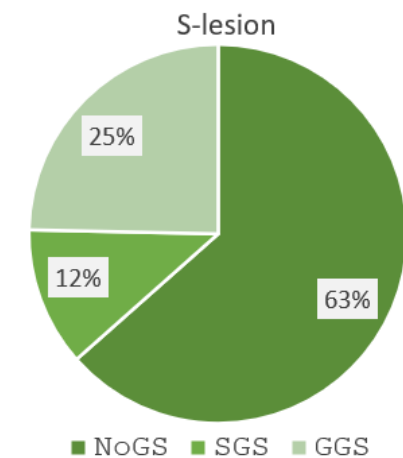
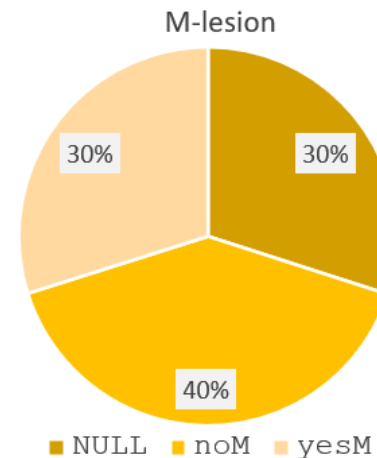
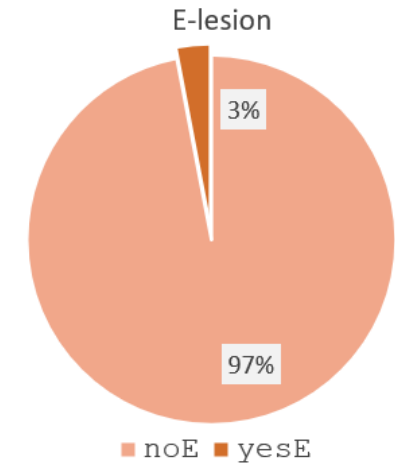
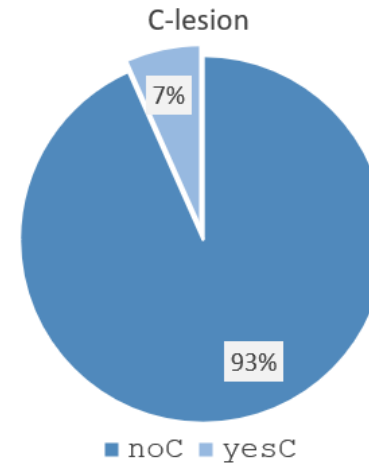


# ML reproduction of Oxford classification for M, E, S, C: MESCCnn

- Only glomerular components (M, E, S, C)
- Supervised approach
- Dataset of n=6206 expert-labeled glomerular images (split n=4298 training, n=1908 validation), 3 institutions (Bari, Köln, Szeged)
- Sample size augmentation with auto augment Cubuk et al. arXiv 2019
- Several recent ConvNets and a Vision Transformer

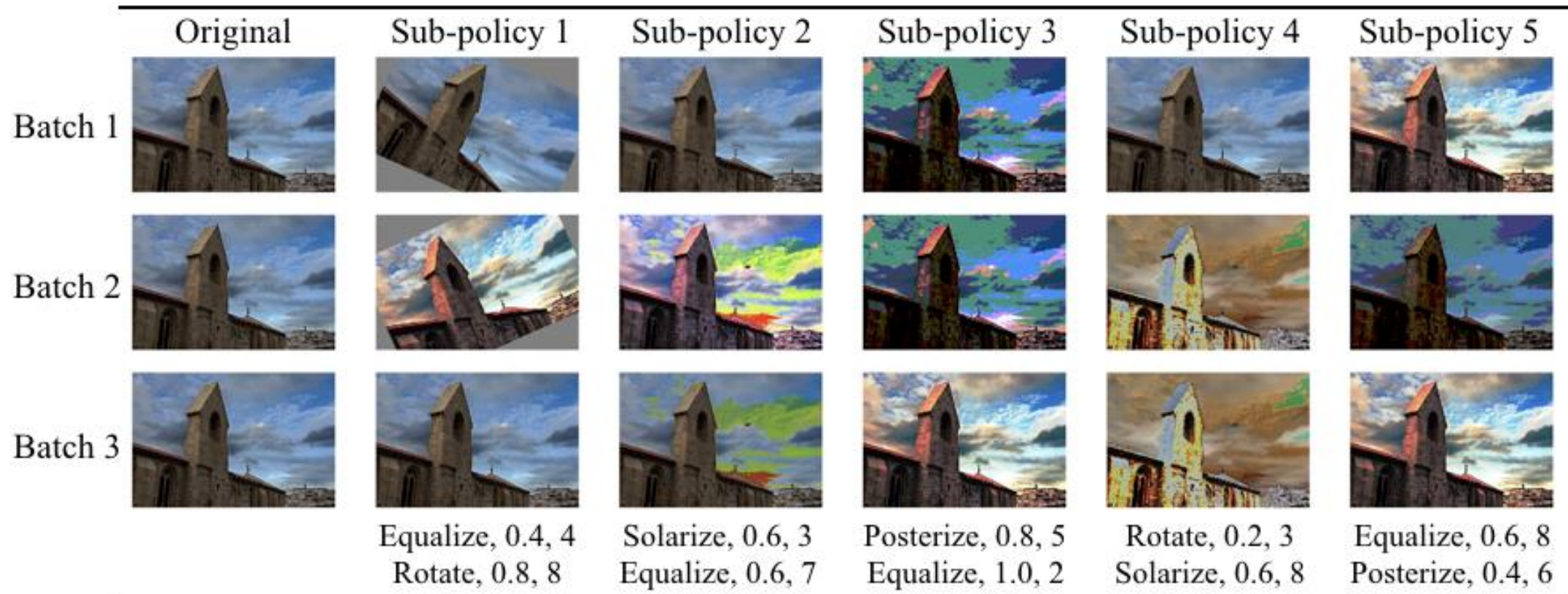
# Dataset for MESCnn

- Mesangioproliferation (M\_score)
  - No = 2494
  - Yes = 1855
  - Null = 1860
- Endocapillary hypercellularity (E\_score)
  - No = 6022
  - Yes = 187
- Glomerulosclerosis (S\_score)
  - No = 3941
  - Segmental = 737
  - Global = 1531
- Active crescent (C\_score)
  - No = 5795
  - Yes = 414





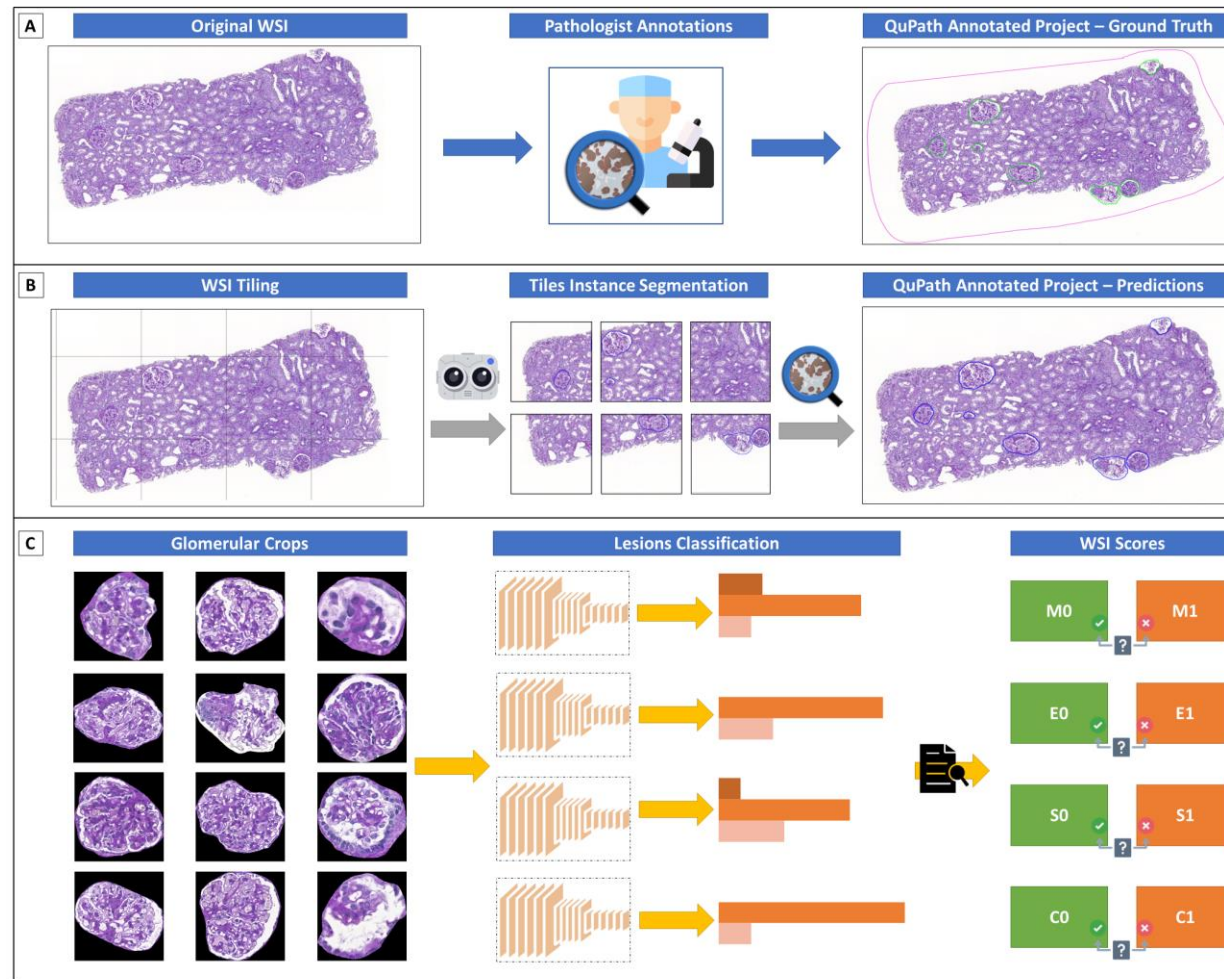
# Sample Size Augmentation



# Inter-rater intraclass-correlation coefficient

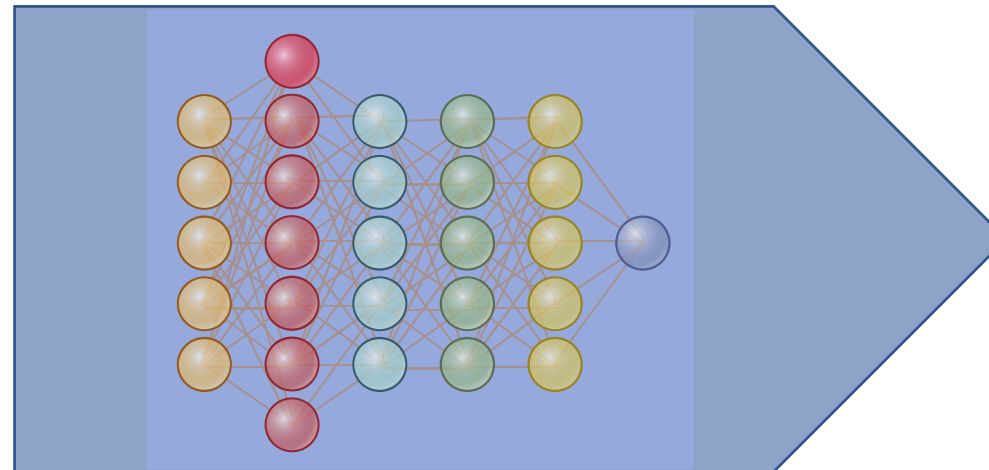
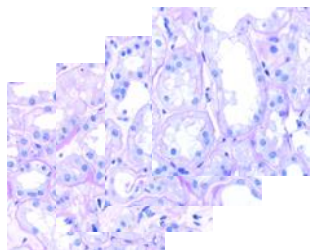
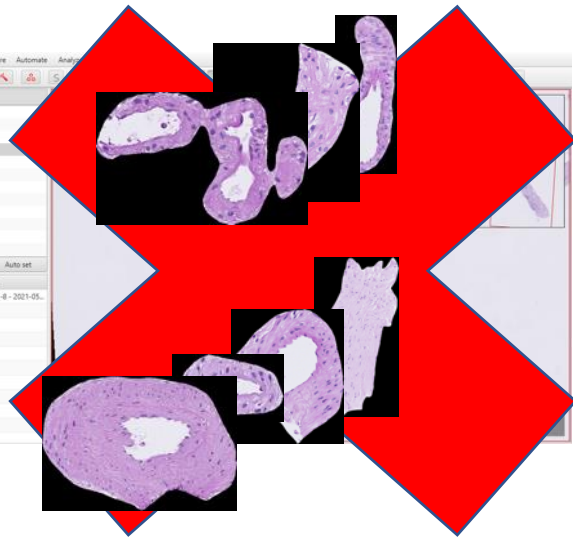
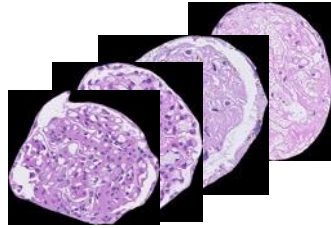
Glomerular expert label vs.	ICC(3,1)			
	M	E	S	C
efficientnetv2-l	0.781	0.790	0.617	0.648
efficientnetv2-m	0.891	0.742	0.779	0.712
efficientnetv2-s	0.836	0.774	0.713	0.569
densenet161	0.769	0.746	0.648	0.614
densenet121	0.869	0.801	0.853	0.624
resnet50	0.768	0.819	0.697	0.484
resnet34	0.632	0.689	0.738	0.585
mobilenetv2	0.779	0.814	0.706	0.644
squeezenet	0.845	0.000	0.720	0.000
PretrainedViTB32	0.589	0.716	0.322	0.158
PretrainedViTL32	0.399	0.336	0.576	0.000

# MESCnn tool permits quality control by nephropathologists



# Weakly Supervised Learning

# ML as Multiple-Instance Learning



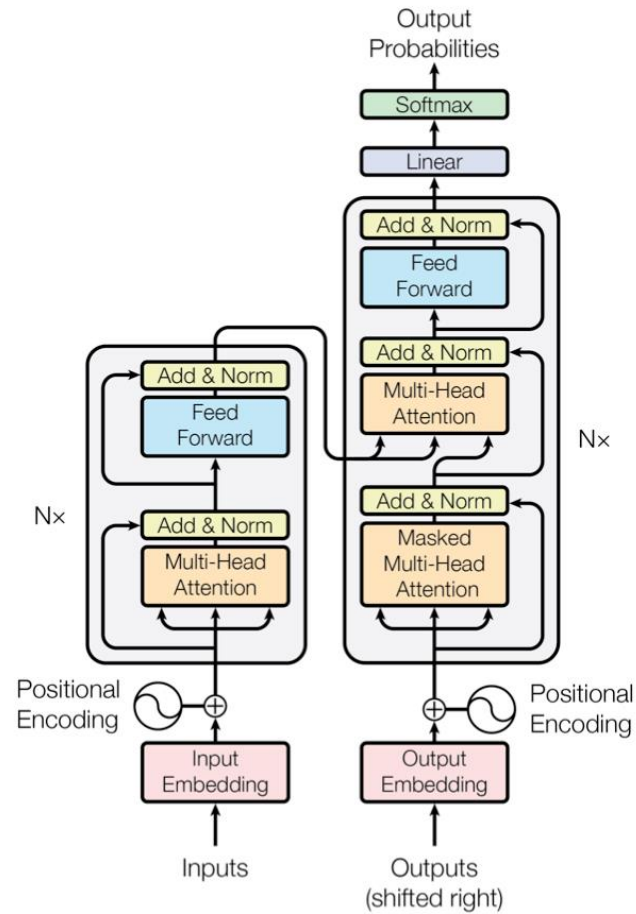
Meaningful output

# AMR Diagnosis with ML

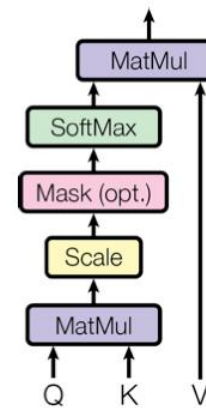
- Semi-supervised approach (biopsy label, binary „AMR“ vs. „noAMR“)
- N=89 biopsies, n=51 AMR, n=38 noAMR, only Köln
- Only PAS-WSIs
- N=1655 glomerular transections, expert-annotated



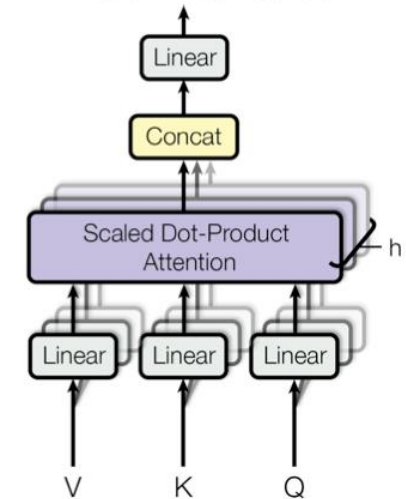
# Self Attention (Transformers)



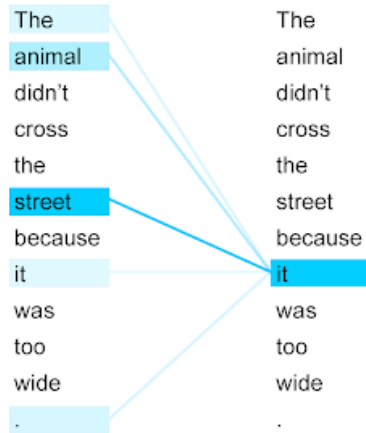
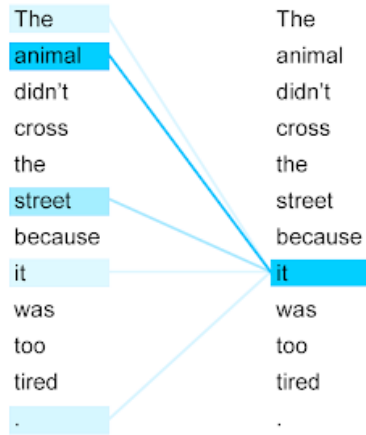
Scaled Dot-Product Attention



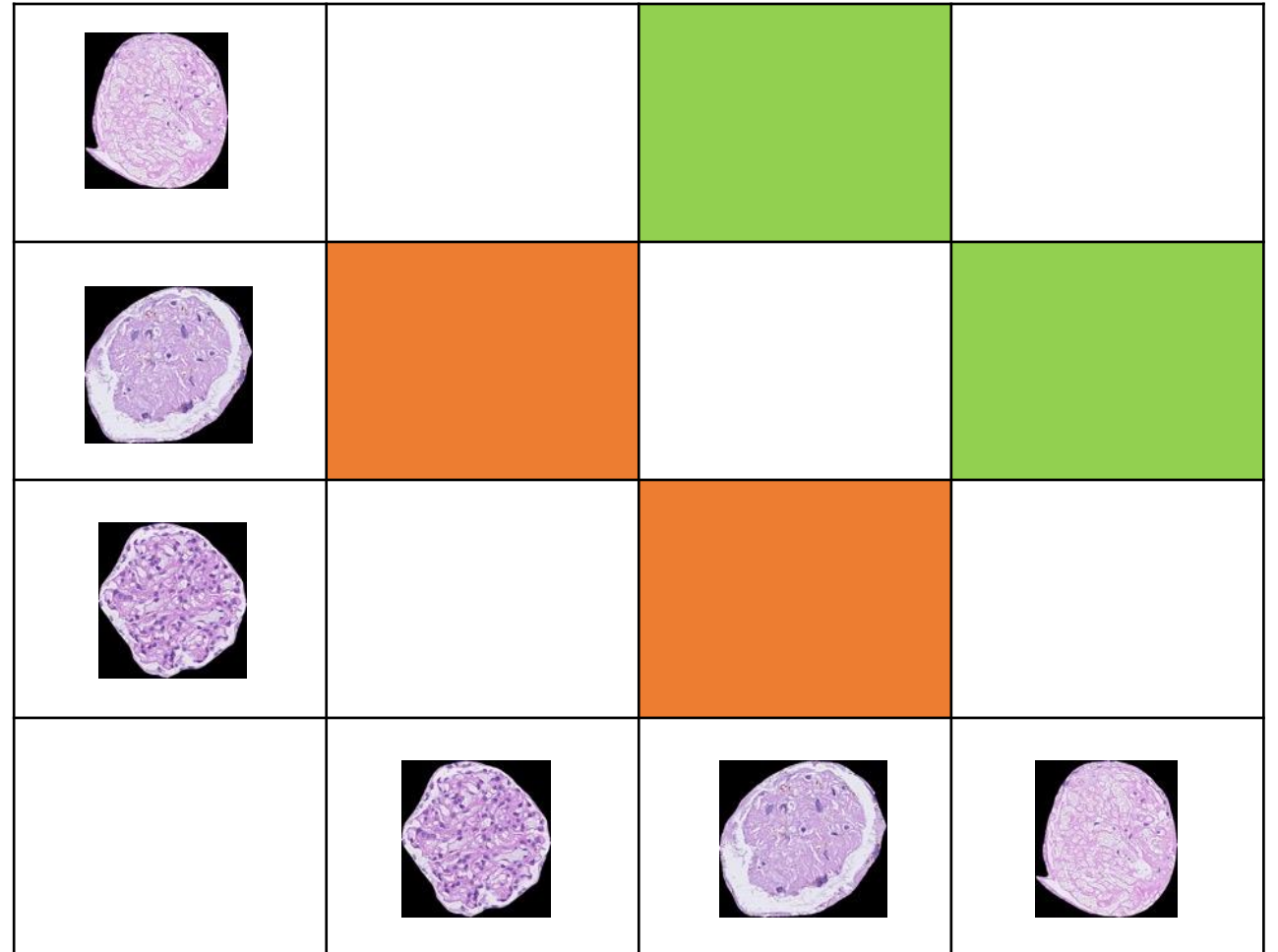
Multi-Head Attention



# Self-Attention (Transformer Networks)

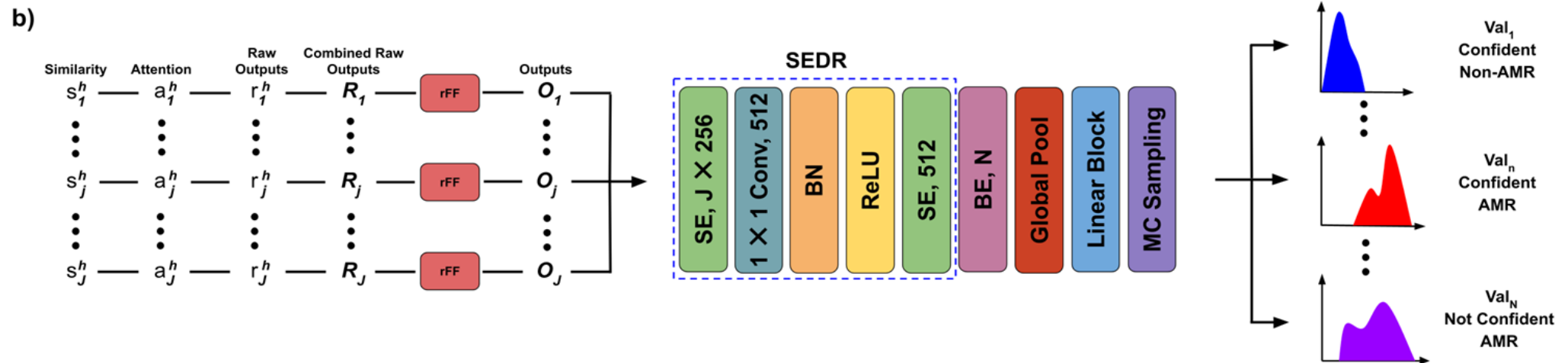
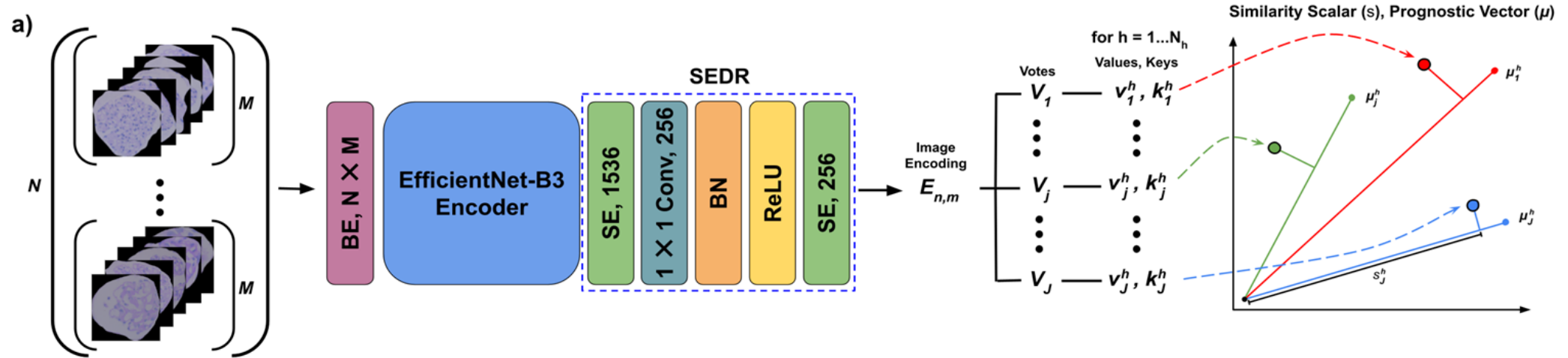


Googleblog



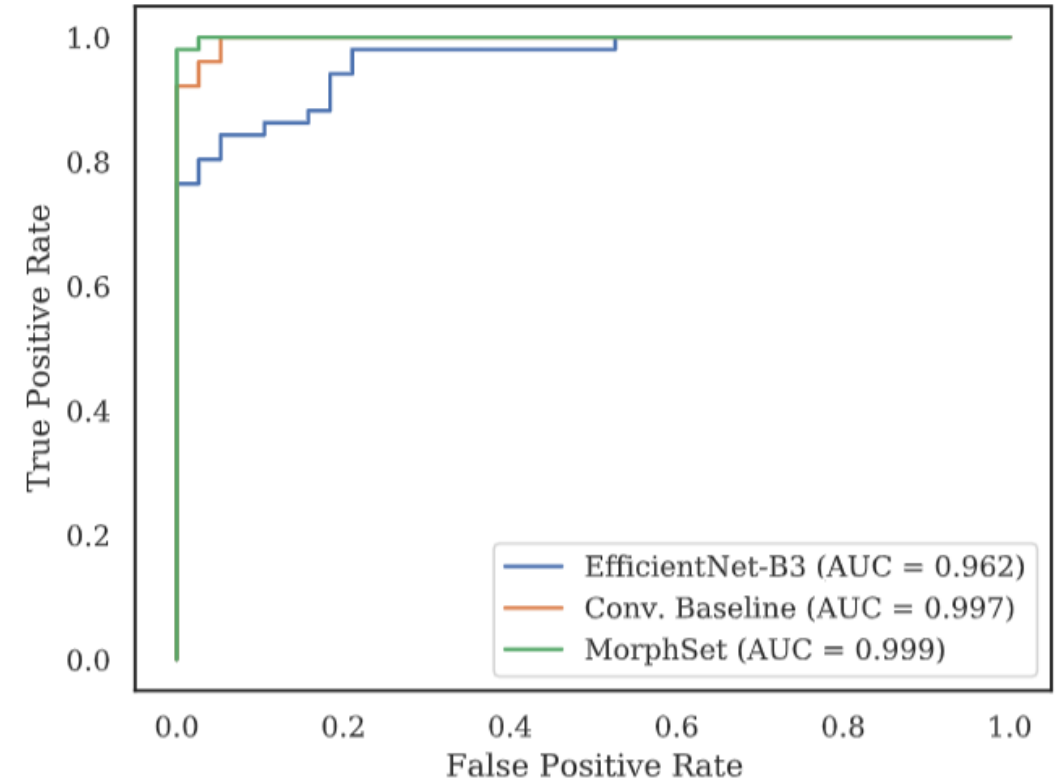


# MorphSet Architecture



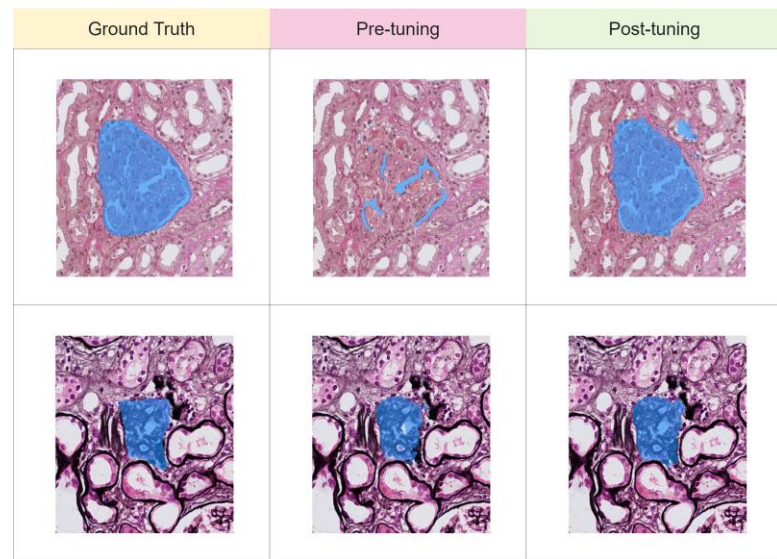
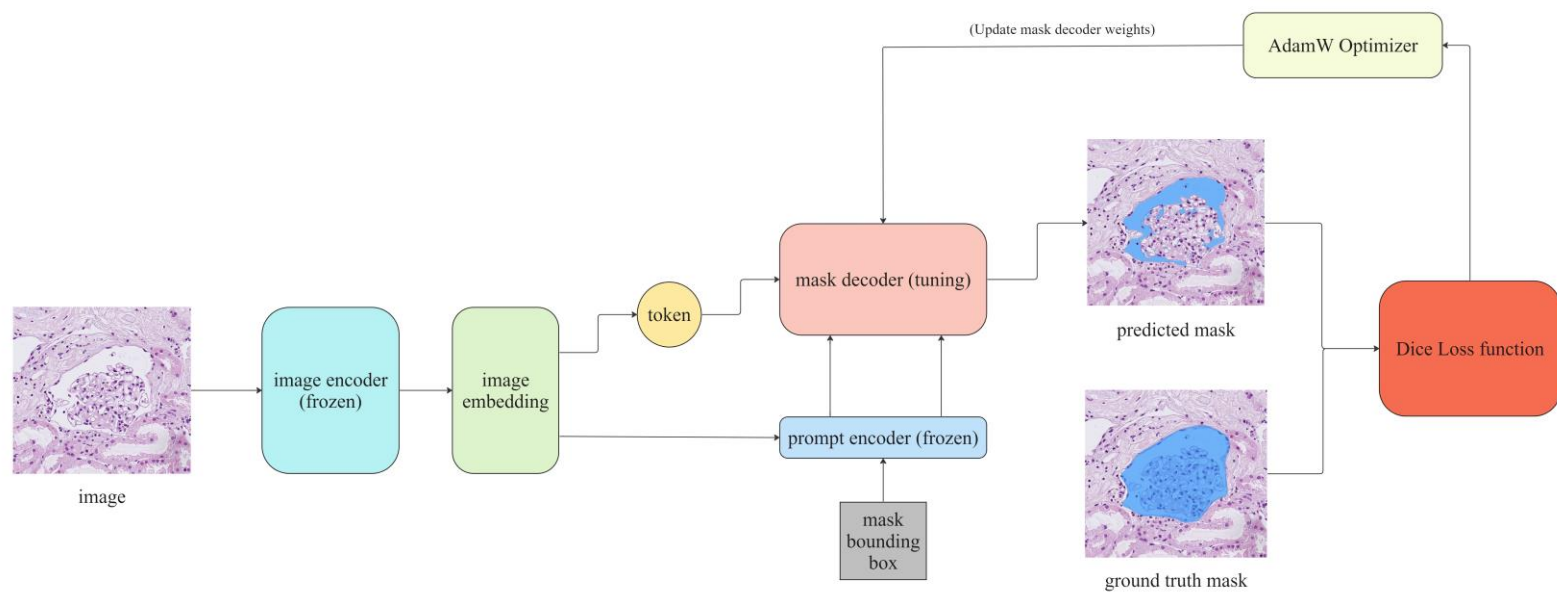
# AMR diagnosis with ML

- EfficientNet-B3-Encoder trained with fully-supervised expert consensus label
- Semi-supervised ConvNet model and MorphSet as least as good as fully-supervised model



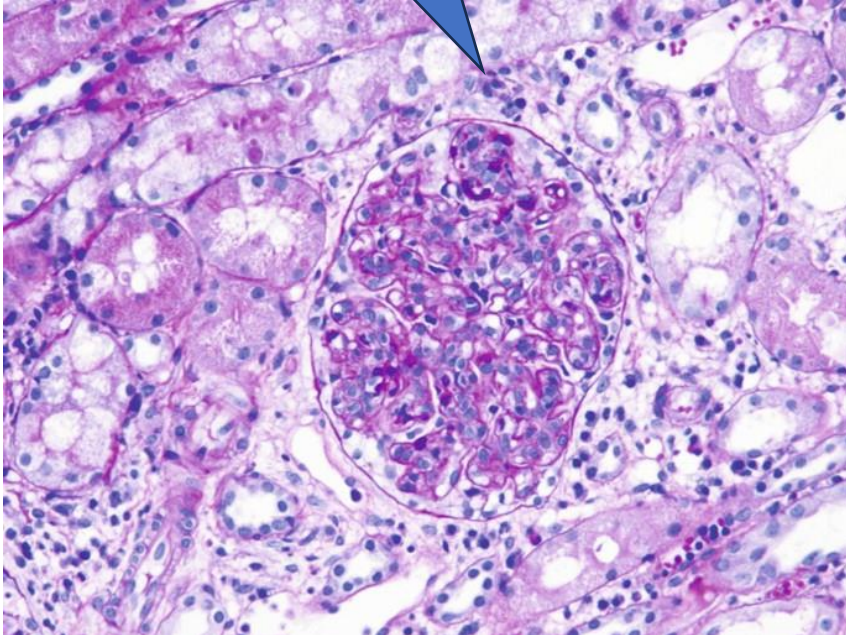
# Unsupervised learning (Foundation Models)

# Foundation Models: Tuning SAM (Segment Anything)



# Tuning Large Language-Vision Models: GlomChat

Create questions and  
answers to this image  
prompt



1. What is the predominant histological feature observed in this image?

Answer: Endocapillary hypercellularity within the glomerulus.

2. Is there evidence of inflammatory cell infiltration in the glomeruli?

Answer: Yes, there appears to be an increase in the number of cells within the capillary loops, which may include inflammatory cells.

3. How would you describe the distribution of hypercellularity? Is it focal or diffuse?

Answer: The hypercellularity appears to be diffuse within the glomerulus observed.

4. What is the condition of the glomerular basement membrane (GBM)?  
(Multiple choice)

A) Thickened

B) Normal

C) Ruptured

D) Cannot be determined from this image

Answer: D) Cannot be determined from this image, as special staining and higher magnification would be required to assess the GBM.

Thank you!