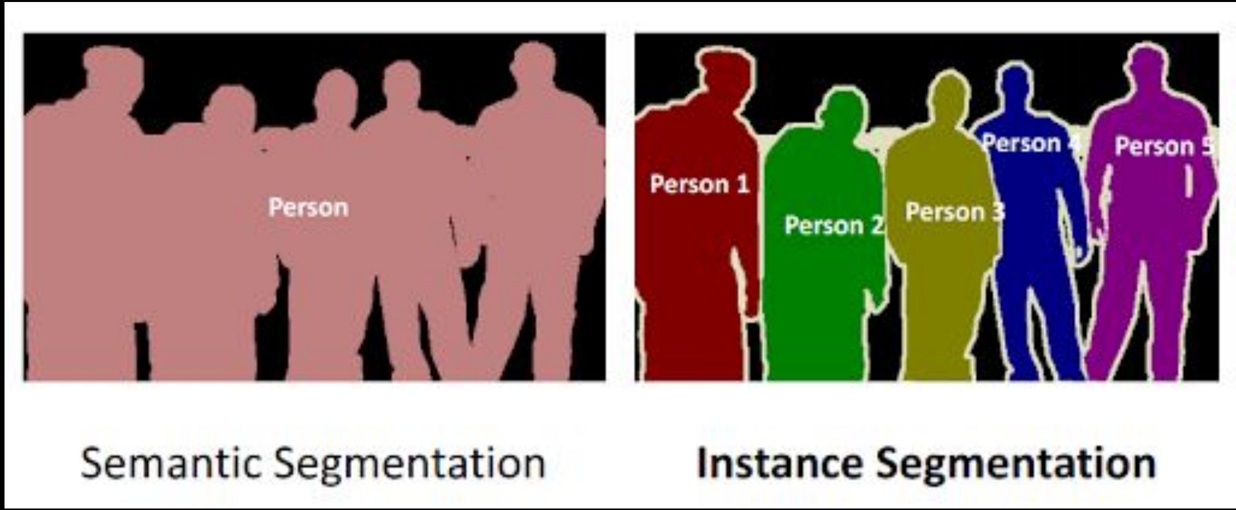
EQUIVARIANT DEEP LEARNING FOR HISTOPATHOLOGY IMAGE SEGMENTATION

Josef Liem 2025

CSCI 1051

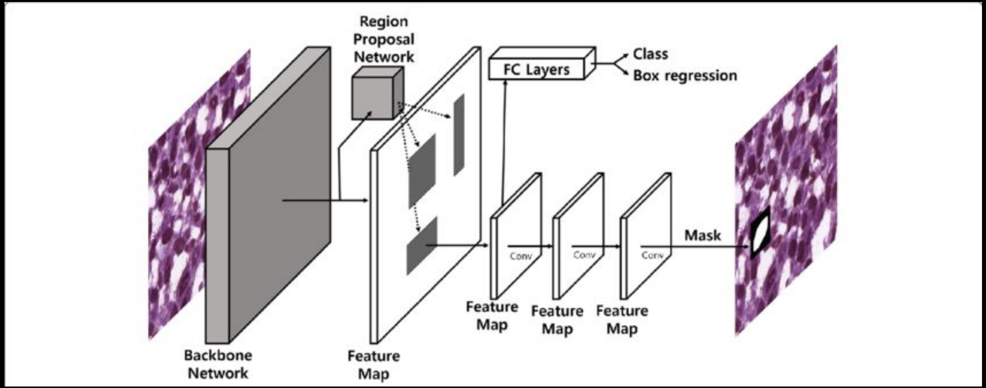
Segmentation



https://www.researchgate.net/figure/Semantic-segmentation-left-and-Instance-segmentation-right-8_fig1_339328277

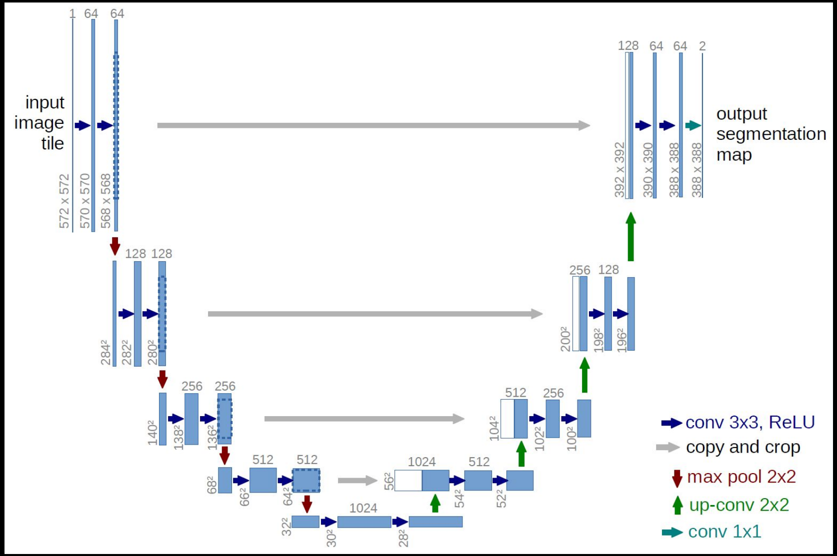
- Pixel-wise classification ≈ Segmentation
- Achievable w/ autoencoder or autoencoder-like architectures

Mask-RCNN: *Extension of Fast-RCNN Region Proposal Network*



https://www.researchgate.net/figure/The-overall-network-architecture-of-Mask-R-CNN_fig1_336615317

U-Net: *An Autoencoder-like Architecture (w/ Skip Connections)*



DeepLab V1,2,3+: *Dilated Convolution for Large Scale Context and Spatial Resolution*

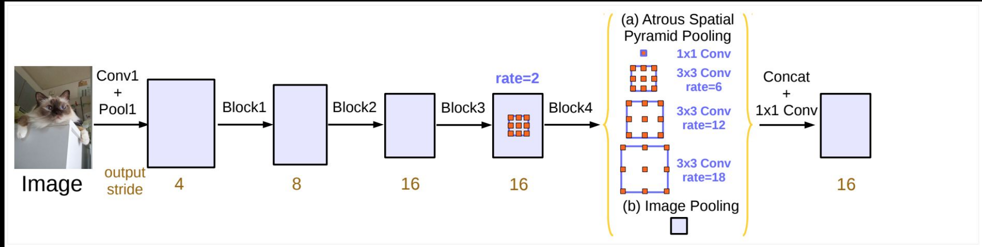


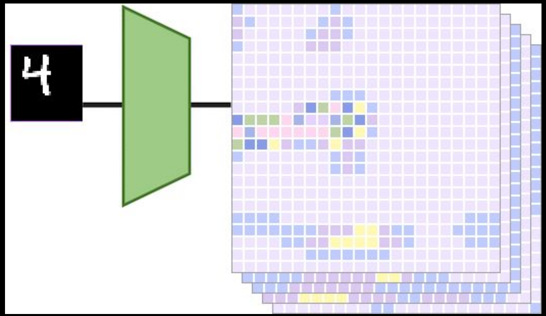
Figure 5. Parallel modules with atrous convolution (ASPP), augmented with image-level features.

https://paperswithcode.com/method/deeplabv3

Equivariance

- We have learned in class (implicitly) that convolution is translationally equivariant:

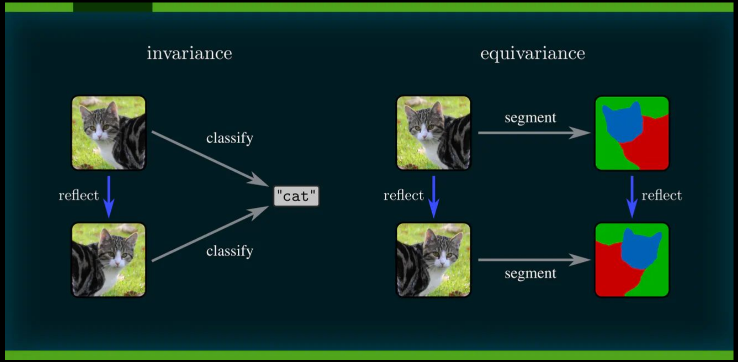
  "shift of position in input image = shift at output of convolution layer" (roughly speaking)
- Not true for other transformations: *rotations/reflections/etc...*



Equivariance: *f(T(x))=T(f(x))for all x*

Invariance: *f(T(x))=f(x)for all x*

(Invariance is just equivariance under the identity)

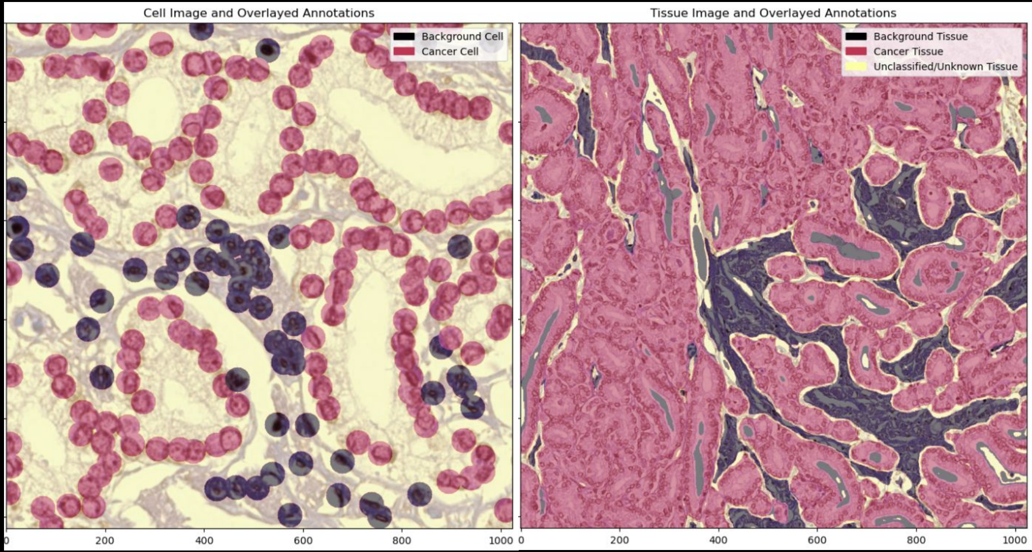


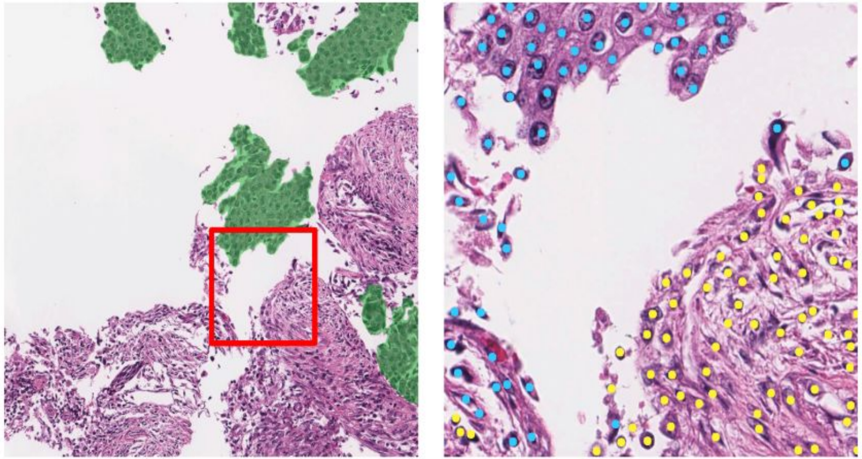
An easy read on equivariance:
https://maurice-weiler.gitlab.io/blog_post/cnn-book_1_equivariant_networks/

Motivation: Medical Imaging

Overlapped Cell On Tissue Dataset for Histopathology MICCAI 2023 Vancouver CANADA

OCELOT





(x_l, y_l^t) (x_s, y_s^c)

Core problem:

How can we use large FOV context to improve smaller FOV classification?

Symmetries are evident on a cellular and tissue level. How can we take advantage of these symmetries?

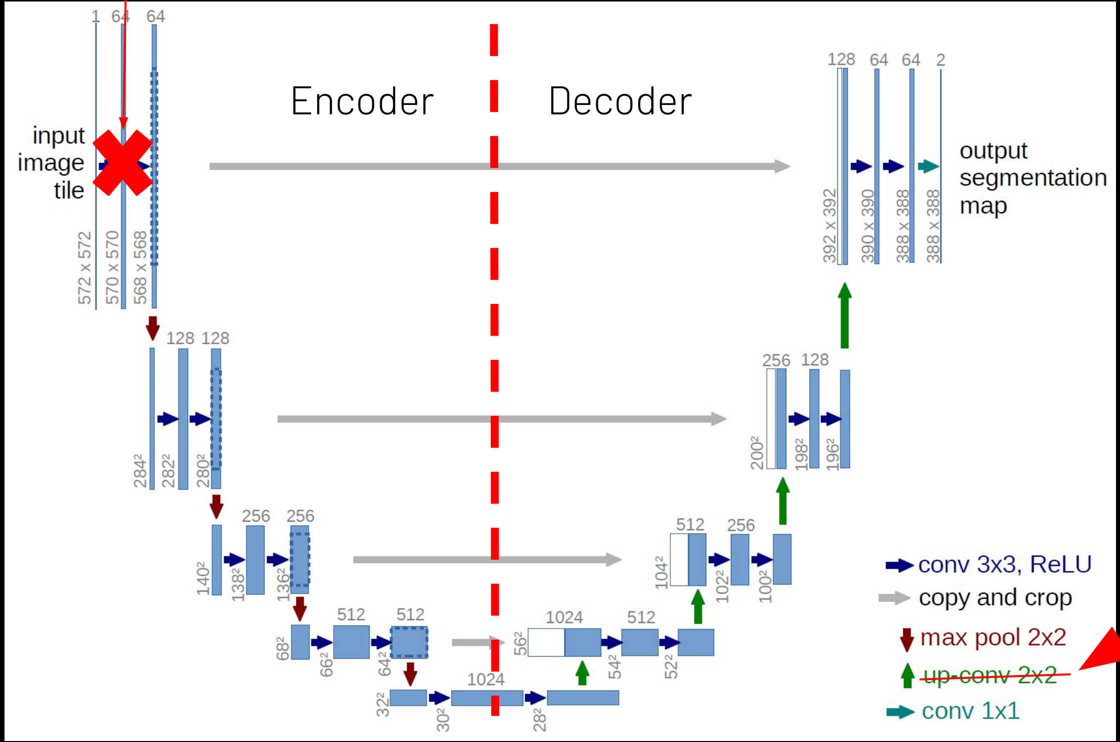
We have limited medical data, and data augmentation will only get us so far... Equivariance?

Dataset size per organ and data subset

Organs	# Slides			# Patch Pairs		
	Train	Val	Test	Train	Val	Test
Bladder	35	14	14	82	29	26
Endometrium	38	13	13	86	29	25
Head-and-neck	13	5	6	27	9	10
Kidney	47	15	18	122	41	41
Prostate	25	12	10	47	17	16
Stomach	15	6	5	36	12	12
Total	173	65	66	400	137	130

Proposed solution: Use equivariant convolution to improve segmentation of histopathology slides, potentially in place of heavy data augmentation. Compare against the non-equivariant equivalent with data augmentation.

Replace w/ single conv of scalar to scalar equivariance



To map from a scalar to a scalar:

$$K(x) = f(|x|)$$

To map from a scalar to a vector:

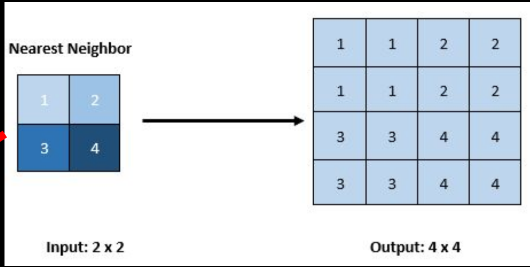
$$K(x) = f(|x|) \cdot x$$

To map from a vector to a scalar:

$$K(x) = f(|x|) \cdot x^T$$

To map from a vector to a vector:

$$K(x) = f_1(|x|) \cdot I + f_2(|x|) \cdot xx^T$$

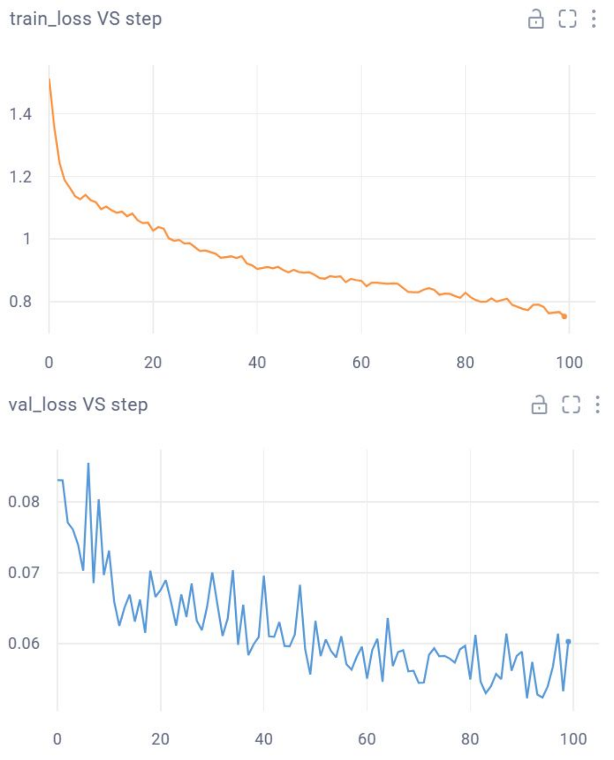


Nearest Neighbor Upsampling

Dice Cross Entropy Loss

Equivariant Loss

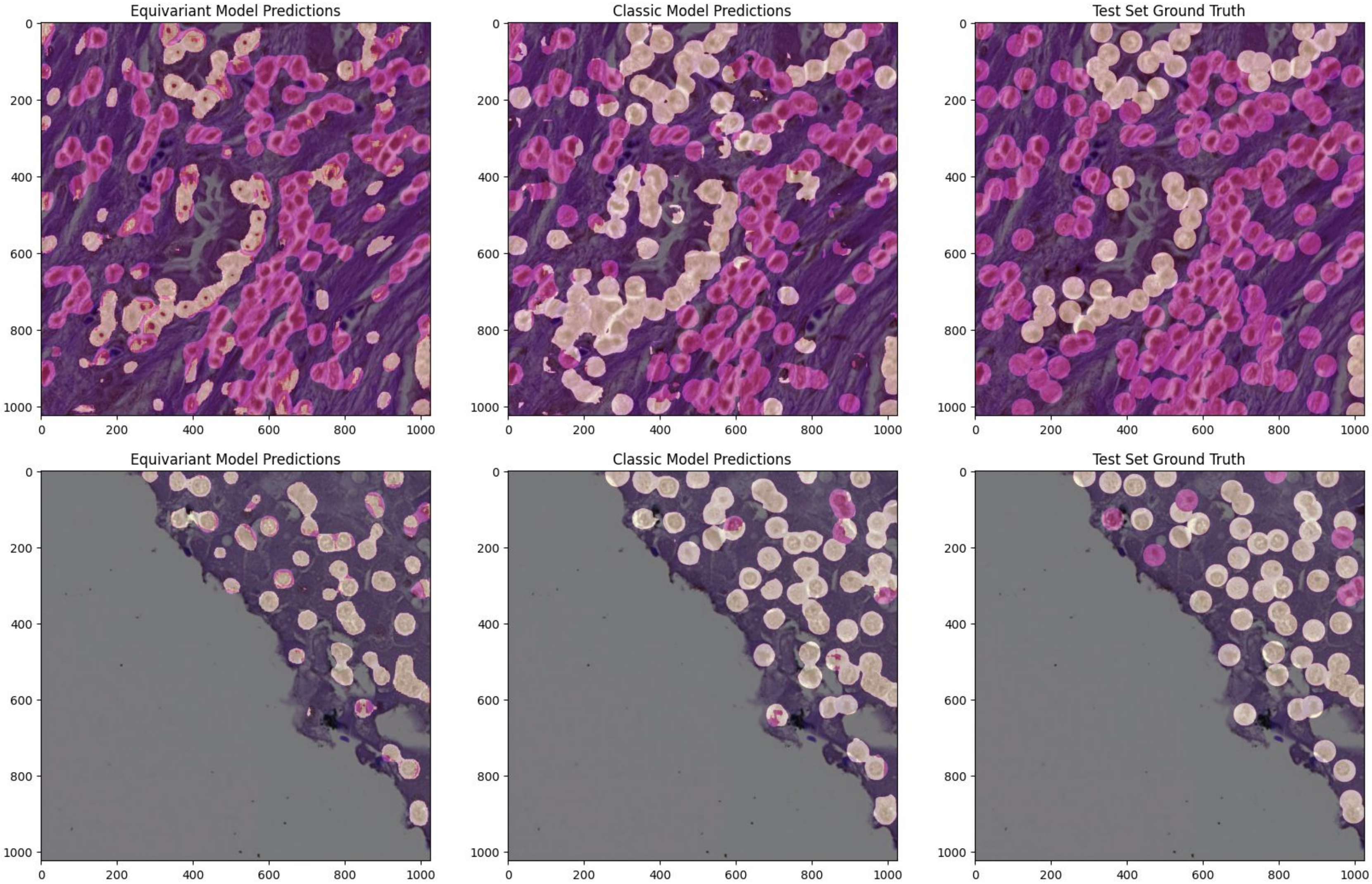
Non-Equivariant Loss

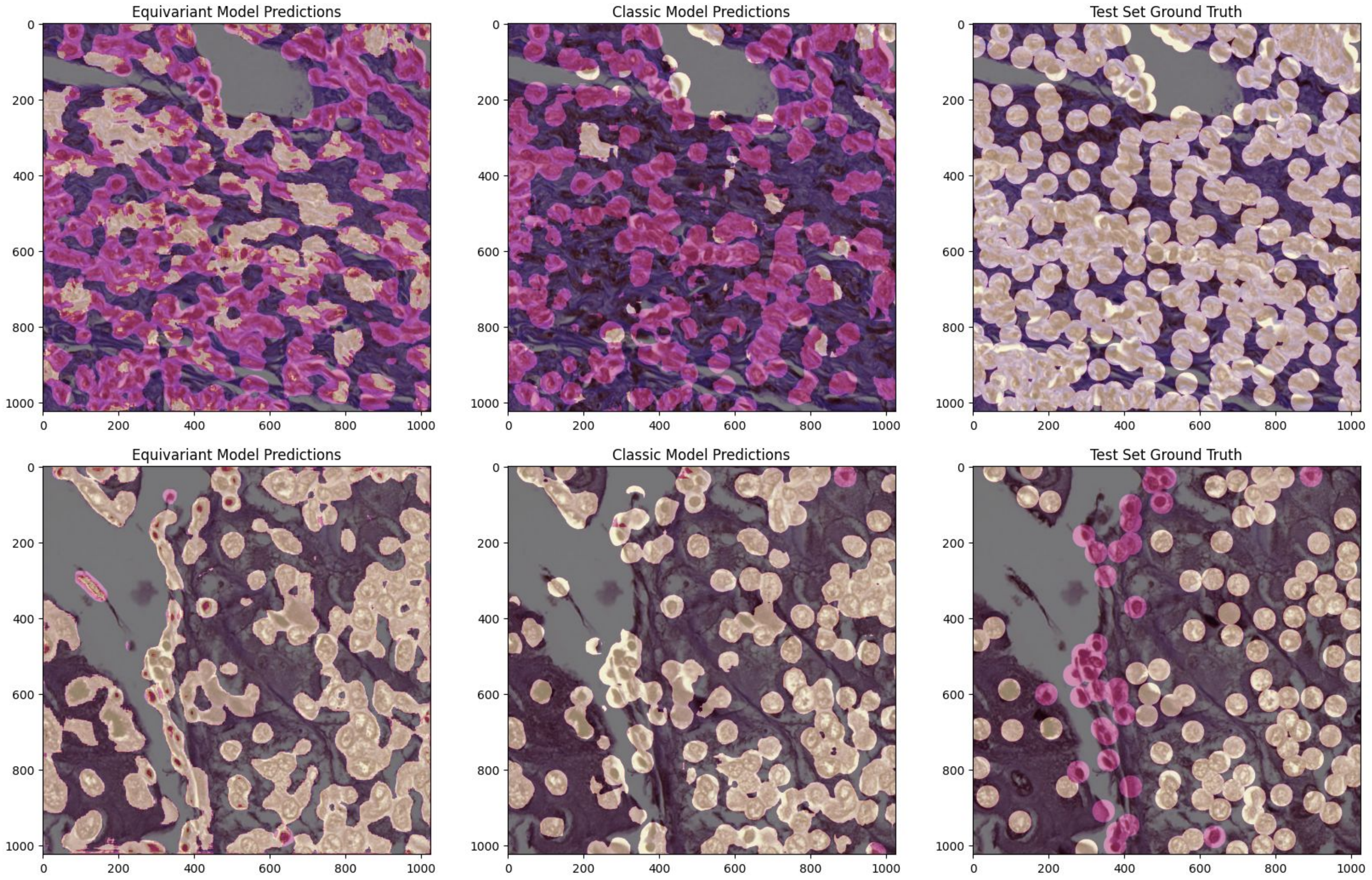


$$\frac{2 * |X \cap Y|}{|X| + |Y|}$$

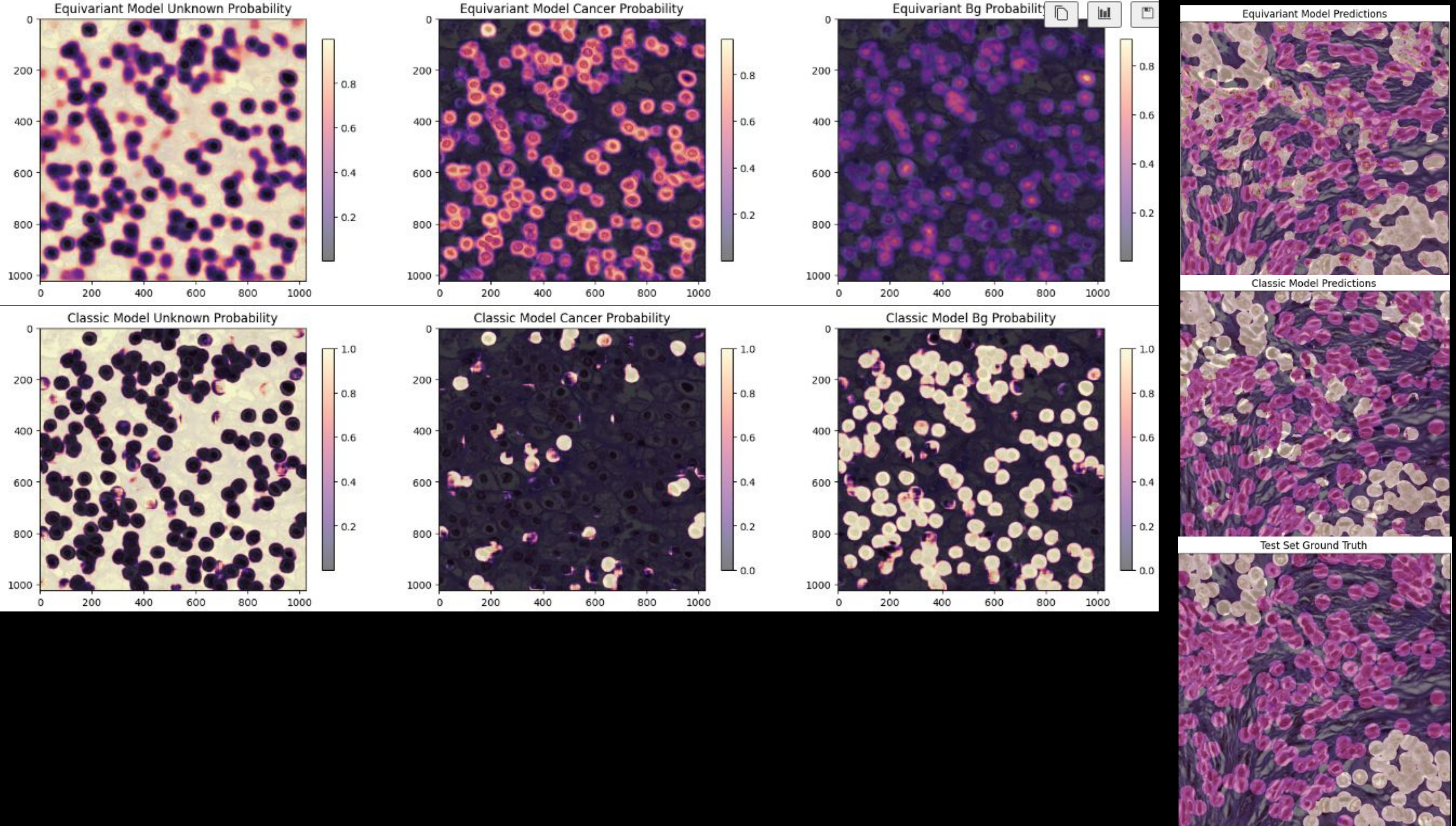
```
Equivariant Model Loss: 1.1681814392407734
Classic Model Loss: 2.2941786273131295
```

```
Equivariant Model Dice Score: 0.6273064613342285
Classic Model Dice Score: 0.6840800642967224
```


Equivariant Model Predictions
Classic Model Predictions
Test Set Ground Truth
Equivariant Model Predictions
Classic Model Predictions
Test Set Ground Truth
0
200
400
600
800
1000

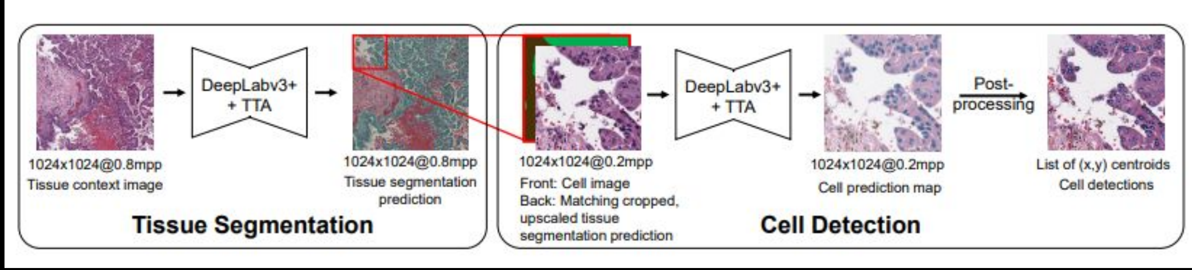


Equivariant Model Predictions
Classic Model Predictions
Test Set Ground Truth
Equivariant Model Predictions
Classic Model Predictions
Test Set Ground Truth
0
200
400
600
800
1000



Equivariant Model Unknown Probability
Equivariant Model Cancer Probability
Equivariant Bg Probability
Classic Model Unknown Probability
Classic Model Cancer Probability
Classic Model Bg Probability
Equivariant Model Predictions
Classic Model Predictions
Test Set Ground Truth

Takeaways



SoftCTM

- Many building blocks, but how do we put them together?
- Potential to expand on tissue/cell patch pair problem
- Consider a combination of data augmentation AND equivariance?
- Good data his hard to come by in certain domains

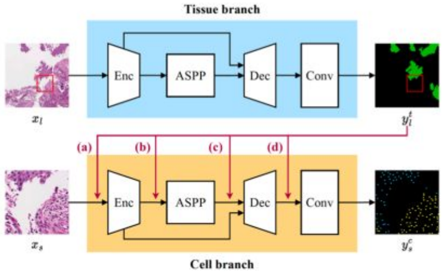
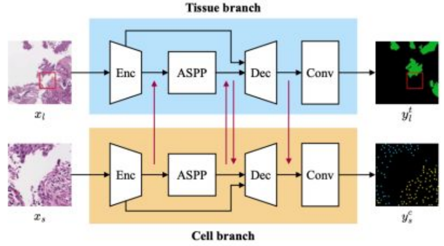


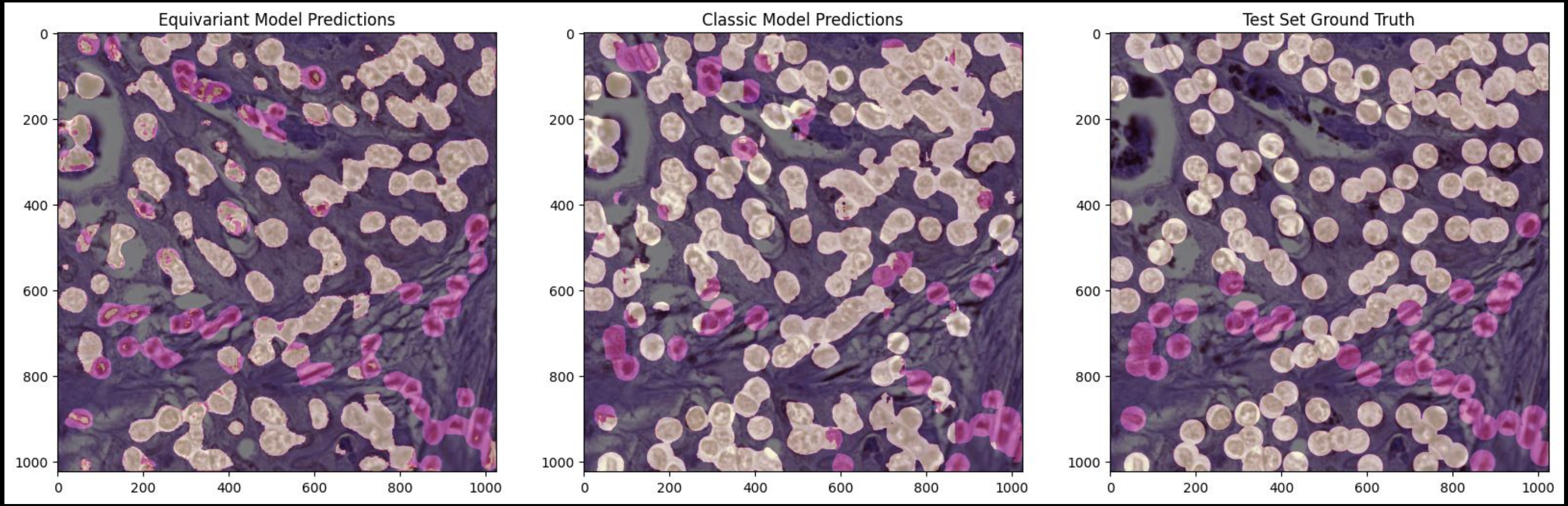
Figure 2.3.: Tissue prediction injection models. (Figure taken from Ryu et al. [2023]).

Cell-Tissue feature sharing model

To introduce greater diversity and flexibility in the flow of cell-tissue information, the OCELOT team developed an additional set of models for cell-tissue feature sharing. These models enable information exchange between the cell detection model and the tissue segmentation model at various feature representation levels in each network. This approach allows for fluid and dynamic information sharing between the two models. The resulting best model is illustrated in Figure 2.4.



THANK YOU



Equivariant Model Predictions
Classic Model Predictions
Test Set Ground Truth
0
200
400
600
800
1000

- To map from a scalar to a scalar:

$$K(x) = f(|x|)$$

- To map from a scalar to a vector:

$$K(x) = f(|x|) \cdot x$$

- To map from a vector to a scalar:

$$K(x) = f(|x|) \cdot x^T$$

- To map from a vector to a vector:

$$K(x) = f_1(|x|) \cdot I + f_2(|x|) \cdot xx^T$$

Example: *a mixed representation (scalar/vector) equivariant convolution over three scalar channels (an image)*

- Let $I_i(x)$ be the i-th of 3 scalar channels in the input image (where $i = 1$ means red, 2 means green, and 3 means blue).
- Let $J_s^j(x)$ be the j-th of 2 scalar channels in the output image.
- Let $J_v^j(x)$ be the j-th of 2 vector channels in the output image.

The equation with kernels (f_{ij}) is:

$$J_s^1(x) = \int f_{11}(|x - x'|) I_1(x')\, dx' + \int f_{12}(|x - x'|) I_2(x')\, dx' + \int f_{13}(|x - x'|) I_3(x')\, dx'$$

$$J_s^2(x) = \int f_{21}(|x - x'|) I_1(x')\, dx' + \int f_{22}(|x - x'|) I_2(x')\, dx' + \int f_{23}(|x - x'|) I_3(x')\, dx'$$

For the vector part, the kernels are (g_{ij}):

$$J_v^1(x) = \int g_{11}(|x - x'|)(x - x') I_1(x')\, dx' + \int g_{12}(|x - x'|)(x - x') I_2(x')\, dx' + \int g_{13}(|x - x'|)(x - x').$$

Note that ($J_v^1(x)$) has two components (x and y):

$$J_v^2(x) = \int g_{21}(|x - x'|)(x - x') I_1(x')\, dx' + \int g_{22}(|x - x'|)(x - x') I_2(x')\, dx' + \int g_{23}(|x - x'|)(x - x').$$

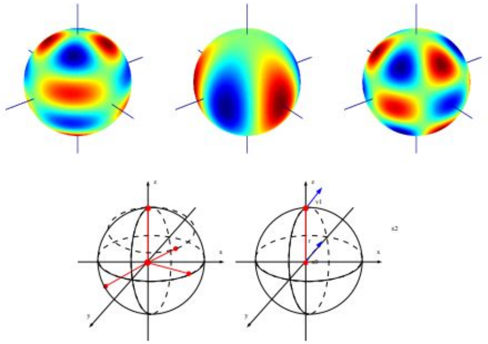
By Prof. Daniel J. Tward, UCLA
Departments of Computational Medicine and Neurology
Ahmanson-Lovelace Brain Mapping Center

There are many ways (representations) for achieving equivariance. Some representations allow for equivariance to more group actions, but sometimes at a high cost (training & compute)...

- `e2cnn.group.Group.irrep()`,
- `e2cnn.group.Group.regular_representation()`,
- `e2cnn.group.Group.quotient_representation()`,
- `e2cnn.group.Group.induced_representation()`,
- `e2cnn.group.Group.restrict_representation()`,
- `e2cnn.group.directsum()`,
- `e2cnn.group.change_basis()`

e2cnn
0.2.3

On Invariance, Equivariance, Correlation and Convolution of Spherical Harmonic Representations for Scalar and Vectorial Data.



Janis Keuper
keuper@imla.ai[1,2]

[1]Institute for Machine Learning and Analytics (IMLA), Offenburg University
[2]CC-HPC, Fraunhofer ITWM, Kaiserslautern

July 10, 2023

POLAR TRANSFORMER NETWORKS

Carlos Esteves, Christine Allen-Blanchette, Xiaowei Zhou, Kostas Daniilidis
GRASP Laboratory, University of Pennsylvania
{machc, allec, xiaowz, kostas}@seas.upenn.edu

Neural Fourier Transform: A General Approach to Equivariant Representation Learning

Masanori Koyama, Kenji Fukumizu, Kohei Hayashi, Takeru Miyato

*I won't be able to tell you how some of these other representations work... Please don't ask...