Investigating Moment Kernel CNNs for Rotation Equivariance in Bioimage Analysis

Faria Tavacoli, Josef T. Liem, Daniel J. Tward

Background

- Equivariance is a special property of functions that specify a predictable relationship between transformations from a given input to an output.
- Convolutional neural networks (CNNs), which are inherently translationally equivariant, are widely used in many natural image contexts for their ability to learn complex patterns.

Challenge

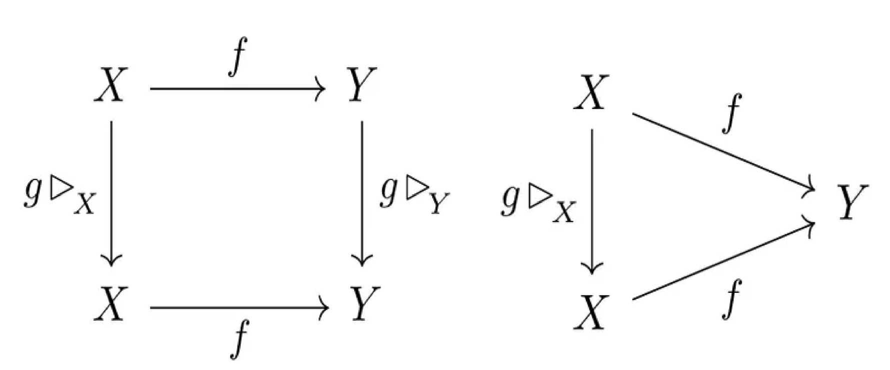
- For bioimaging tasks, there exist additional, more complex symmetries in the data that classical CNN approaches are unable to exploit.
- Existing methods of achieving equivariance to group actions beyond translations are computationally expensive and unscalable to higher dimensions, limiting their adoption.

Approach

- Moment kernels, a novel kind of convolutional kernel equivariant to reflections and rotations were benchmarked on the MedMNIST dataset against classical CNNs and existing equivariant frameworks.
- Moment kernels were used with the Allen Brain Cell Atlas (ABC) to perform classification and segmentation on spatial transcriptomics data.

Conclusion

- Our benchmarking results revealed the moment kernel approach improved accuracy metrics by a few percent over alternatives on most datasets with rotational symmetry.
- Our new approach to analyzing ABC helps in understanding how tissues emerge out of mixtures of cells and may benefit from these symmetries.



$$k(x) = \begin{cases} f_1(|x|) \in \mathbb{R}^{1\times 1} \\ f_2(|x|)x \in \mathbb{R}^{1\times D} \\ f_3(|x|)x^T \in \mathbb{R}^{D\times 1} \\ f_4(|x|)xx^T \in \mathbb{R}^{D\times D} \\ f_5(|x|)id \in \mathbb{R}^{D\times D} \end{cases}$$

Figure 1. Equivariant (left) and invariant (center) mappings (Weiler, 2023). Invariance is simply a special case of equivariance. All 5 moment kernels (right) form a stacked $(D+1)\times(D+1)$ matrix.

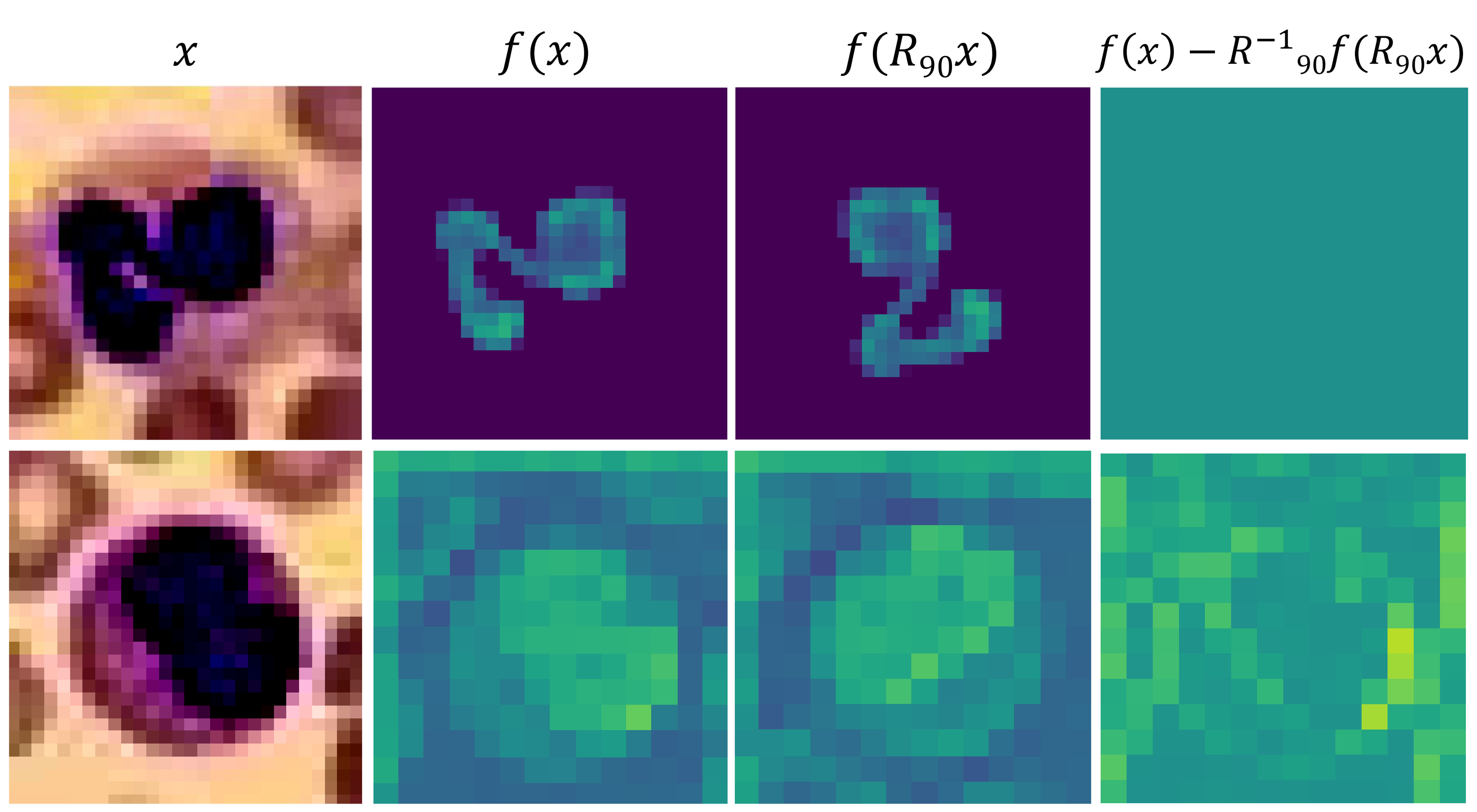


Figure 2. Input BloodMNIST image, and visualization of equivariance to 90-degree rotations of randomly sampled feature maps at the first layer of a CNN implemented with moment kernels under the trivial irreducible representation (top row), and a non-equivariant CNN model (bottom row). Under the definition of equivariance, there should be no difference between directly passing the input through the model and applying a transformation to the input, followed by applying the inverse transformation to the model's output (right column).

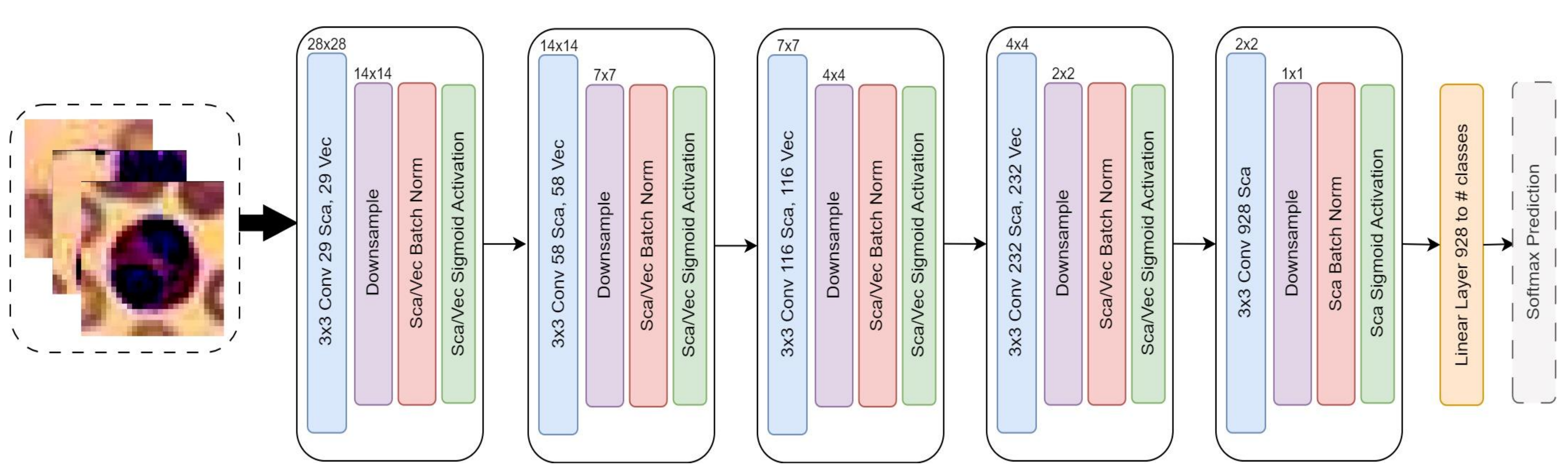


Figure 3. Architecture of the trivial irreducible representation model using moment kernels on the BloodMNIST dataset. Models under this representation achieve equivariance via scalar and vector mappings.

Model	# Channels in First Layer	# Parameters
Trivial Irreducible Moment Kernel	58 (29 scalars + 29 vectors)	1565621
Trivial Moment Kernel	55	1554527
Trivial Irreducible ECNN	62 (31 scalars + 31 vectors)	1562221
Trivial ECNN	67	1540404
Regular ECNN	29	1500090
CNN	32	1574151

Table 1. The number of channels were specified such that models had approximately the same number of parameters as the default CNN. Each model consists of 5 layers which progressively double the number of channels. Equivariant CNN (ECNN) models were implemented using the ESCNN library released by Qualcomm-University of Amsterdam lab (QUVA).

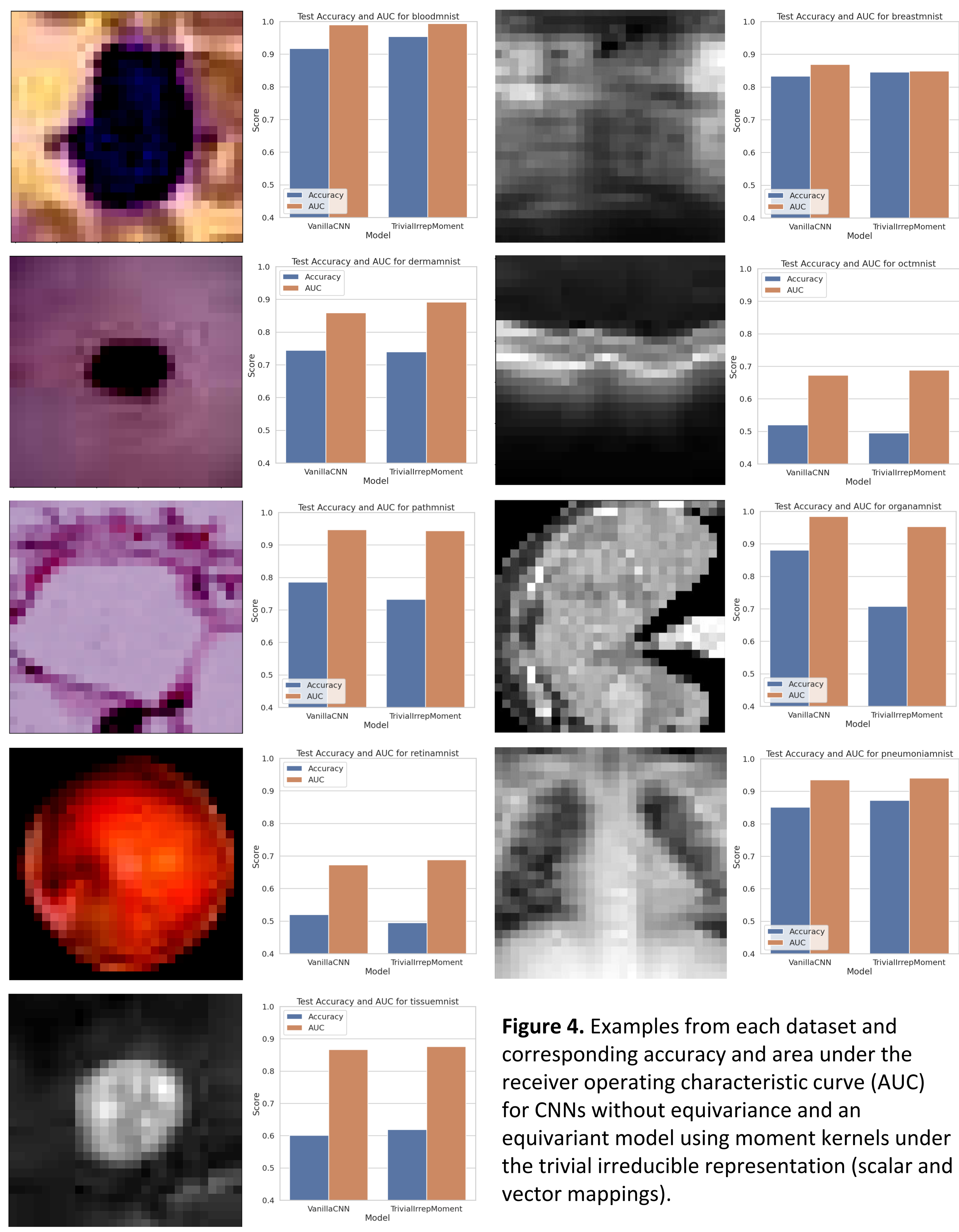


Figure 4. Examples from each dataset and corresponding accuracy and area under the receiver operating characteristic curve (AUC) for CNNs without equivariance and an equivariant model using moment kernels under the trivial irreducible representation (scalar and vector mappings).

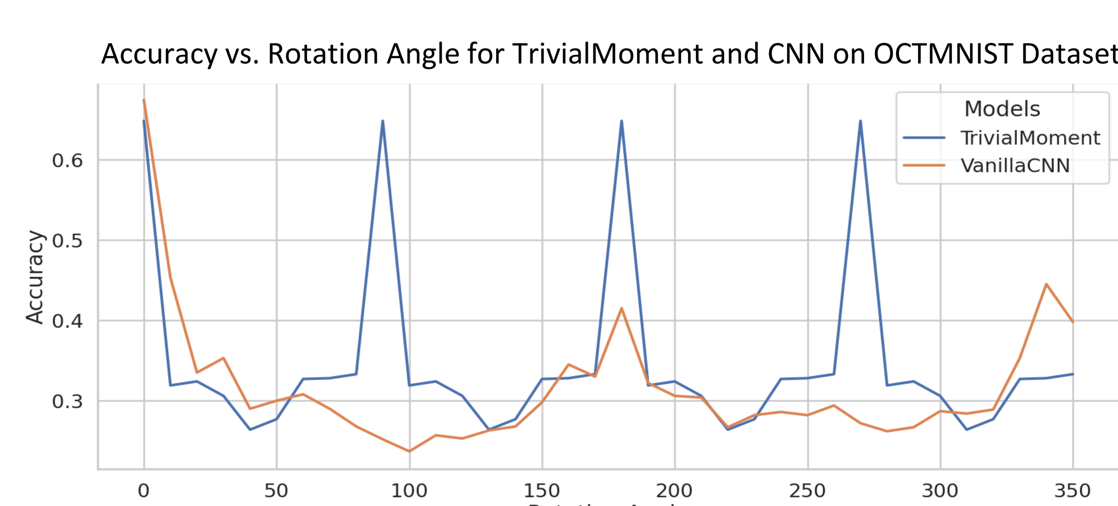


Figure 5. Accuracy of a D_4 equivariant model using the trivial representation (scalar mappings only) vs. a non equivariant CNN at different rotations of images in the OCTMNIST dataset.

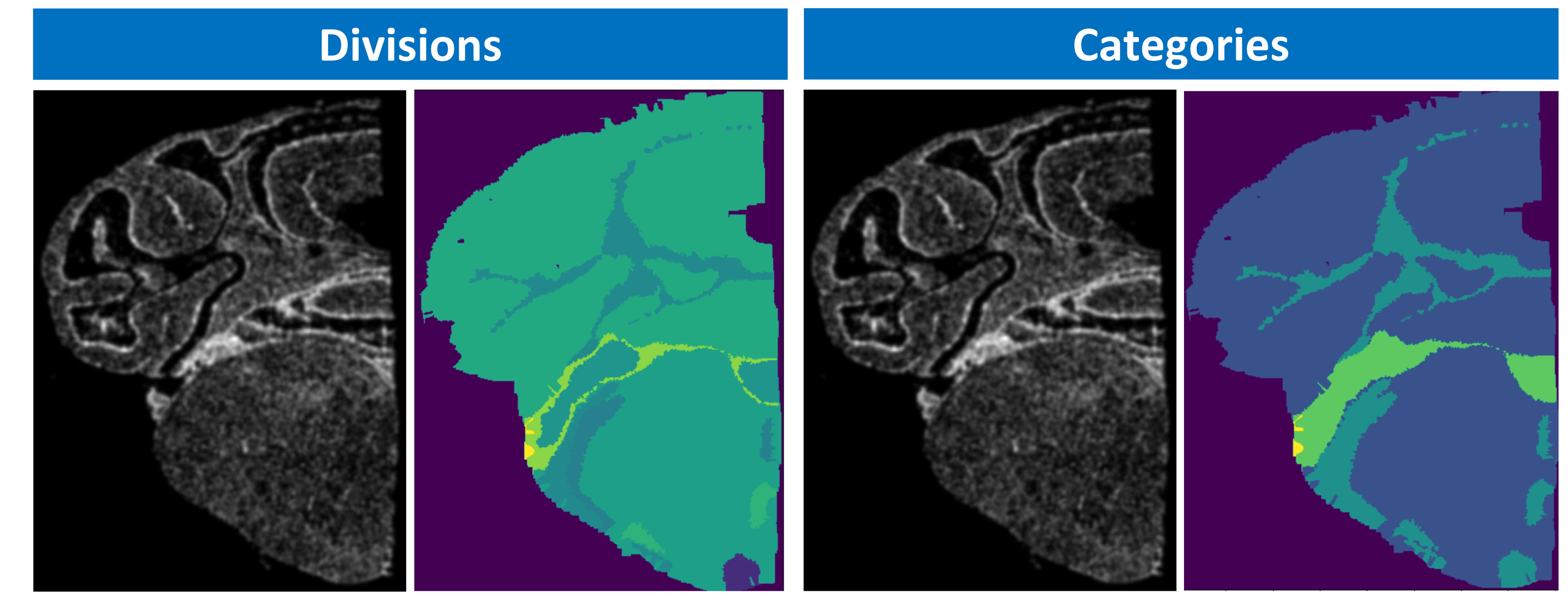


Figure 6. Illustration of cells in the ABC atlas, assigned to different tissue types at different levels of granularity.

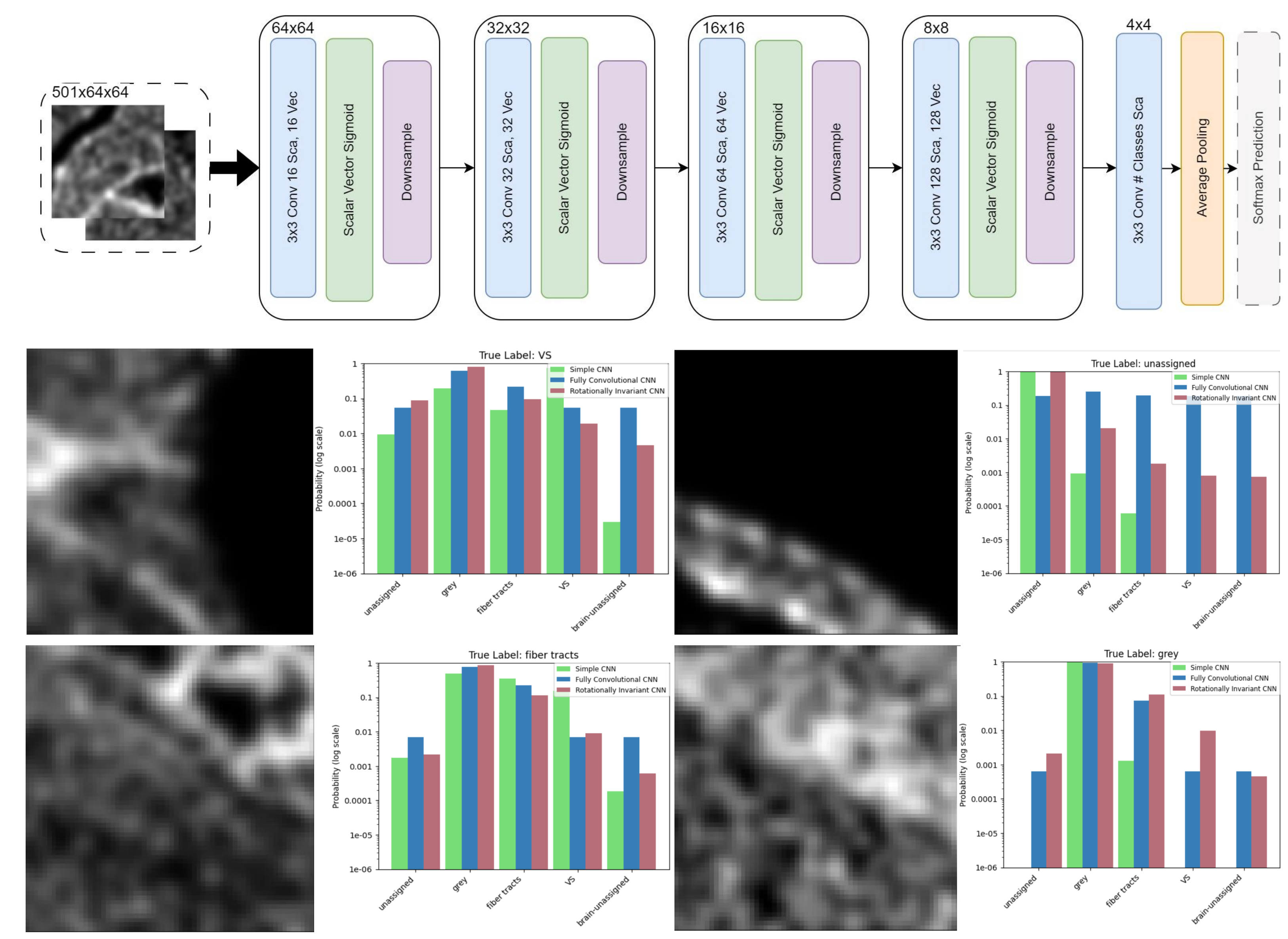


Figure 7. Classification neural network architecture (top) and several examples of input images and classification probabilities (bottom) at the level of categories.

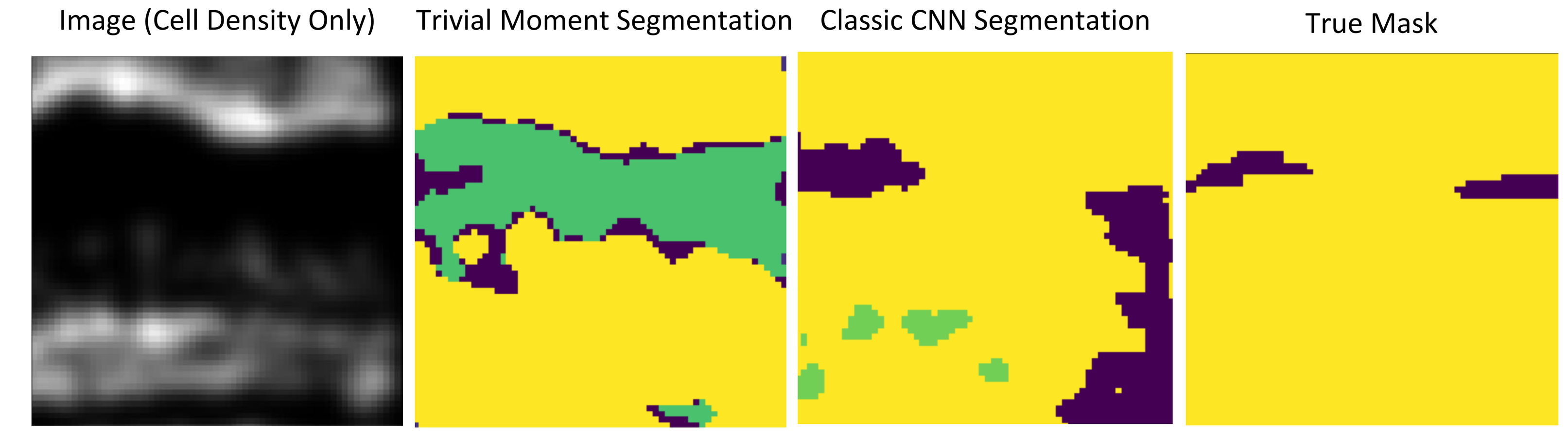


Figure 8. Segmentation of ABC dataset ROIs at the level of divisions with a UNet-like architecture using moment kernels with the trivial representation (scalar mappings) and using non-equivariant convolution. The original image (left) is displayed with only cell density – the 0th of 501 channels.

Acknowledgements

We would like to extend our sincere thanks to Bryson Gray, Gary Zhou, Siqi Fang, Yeun Kim, and members of the Brain Mapping Center for their guidance. This project was also possible thanks to the Bruins in Genomics (B.I.G.) Summer Program and UCLA's department of Computational Medicine.

References

Cohen, Taco, and Max Welling. "Group equivariant convolutional networks." International conference on machine learning. PMLR, 2016.

Yao, Zizhen, et al. "A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain." *Nature* 624.7991 (2023): 317-332.

Yang, Jiancheng, et al. "Medmnist v2-a large-scale lightweight benchmark for 2d and 3d biomedical image classification." Scientific Data 10.1 (2023): 41.