

Automatic segmentation of cortical layers in histological images of human brain

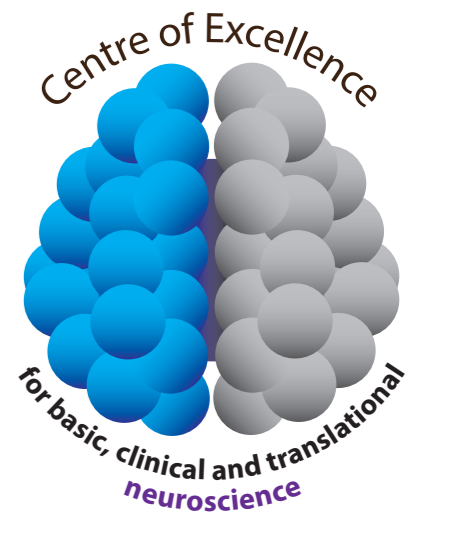
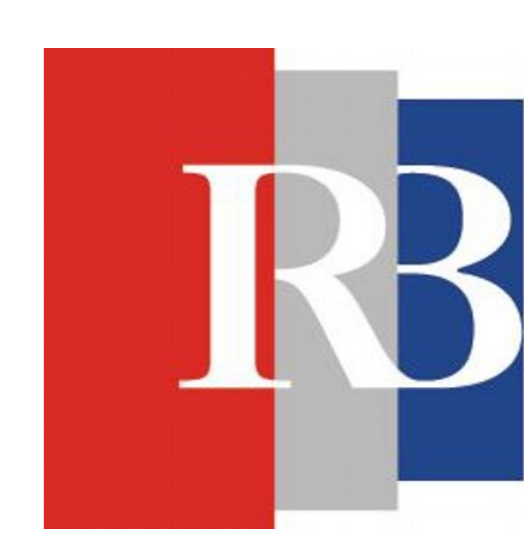
Štajduhar A.¹, Lipić, T.², Judaš M.¹

¹ Croatian Institute for Brain Research, School of Medicine, University of Zagreb

² Ruđer Bošković Institute

astajduhar@hiim.hr

Poster No. Th707



Introduction

Brain structure is defined by organization and distribution of neurons in the tissue. Today, most investigations in brain cytoarchitectonics are not automatized and require significant involvement of the researcher, which is very time consuming and may introduce significant bias. This type of analysis may not be accurate enough or reveal enough information about the structure of the tissue (Meijering, 2016.). Brain tissue forms complex structures whose organization and distribution may be reliably revealed only using systematic computational approach. This research develops novel computational methods for automatic and objective analysis of histological images of human brain suitable for advanced and fast analysis of brain cytoarchitectonics and laminar structure. The methods are developed aiming for a data-driven approach, with as little assumptions about the data as possible.

Materials and Methods

Histological imaging

To examine cellular and structural layout of tissue, we used histological sections of an adult human brain that were stained using NeuN, a common immunohistochemistry method used for labelling neurons in the brain tissue. Samples were obtained from the Zagreb Brain Collection (Judaš, 2011) and scanned using the Hamamatsu Nanozoomer 2.0 scanner (Hamamatsu Photonics, Japan). Images were acquired from a section of 10 μ m thickness at 40x magnification, which corresponds to 0.226 μ m/pixel resolution.

Cortical layer features

Initial research in parcellation of the cortex using neuron density and cell body size presented how the two features, unevenly distributed among the neurons in different layers, reveal laminar structure of the cortex, as shown in Fig. 2. They alone were not enough for precise delineation of the layers, or for more detailed investigations. We develop methods for computing neuron features beyond area and local density that consider more complex aspects of local patterns in the tissue. Two portions of digitized histological preparations of human prefrontal cortex were given to three human experts who manually delineated borders between the layers in the cortex which contained over 50 000 neurons which were automatically segmented from the high-resolution images of the tissue preparation (Štajduhar, 2019). Density-based clustering methods (Fahad, 2014.) identify the radius of 100 – 300 μ m around the neuron for measuring characteristics of local patterns in neuron's neighborhood that is significant for estimation of its features. It usually includes 200-600 nearest neurons. In our investigations, number of neurons was preferred over fixed radius for robustness and computational efficiency.

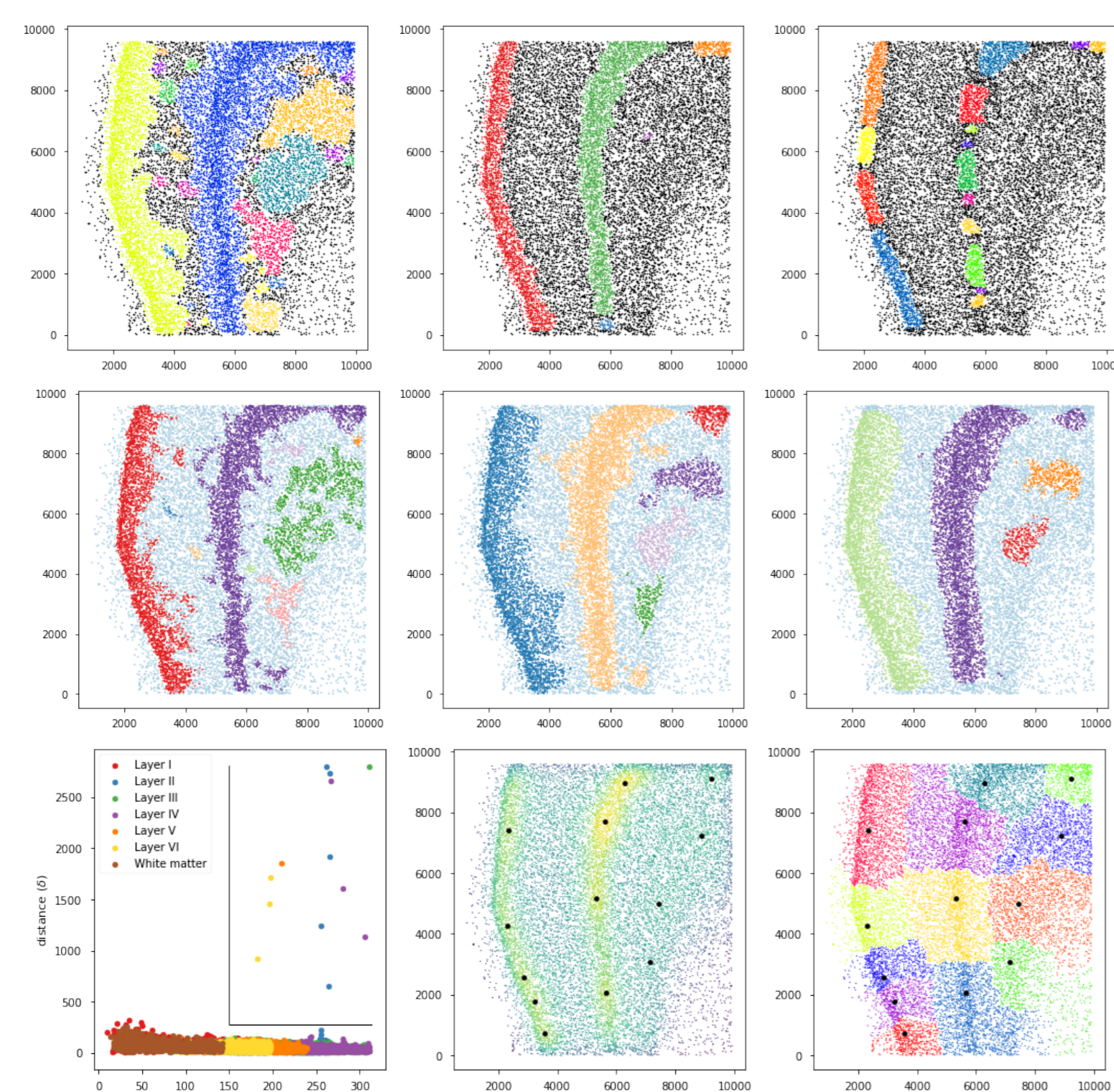


Figure 1: Density-based clustering algorithms cannot create clear distinction between the layers, but give insight into neuron organization. Top to bottom: HDBSCAN, OPTICS, FSDP

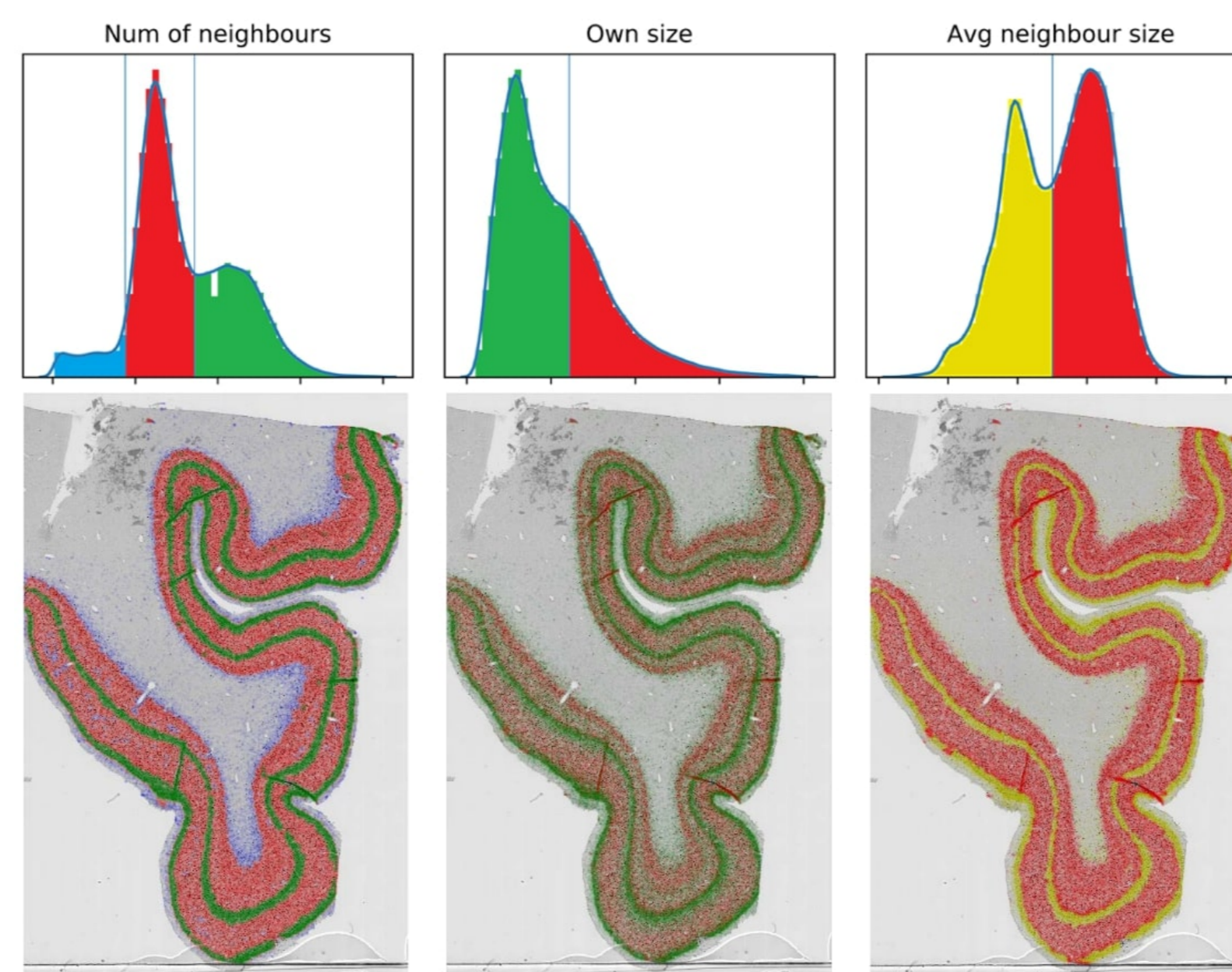


Figure 2: Cell features exhibit multimodal distributions. Thresholds that separate the distributions were obtained using minimization of intraclass variance. Left: Three types of neurons were distinguished by cell density in their surrounding area: very sparse (blue), sparse (red) and dense (green). Middle: Larger (red) and smaller (green) neurons. Right: Average size of neighboring neurons is also a feature that contributes to layer segmentation.

Results

More than 300 features of neurons are developed in several feature classes. The most successful are those derived from density and area of the neurons by computing various statistics and performing oriented measurement taken in different directions from a neuron. Methods based on tree ensembles (Prokhorenkova, 2018.) are used to predict neurons' layer within the cortex in a classical machine learning pipeline. Through performance measurements we show that the machine learning model's output agrees with manual labels in the same amount as they agree with each other, thus creating results indistinguishable from that of human experts. The model is able to generalize well on a new dataset containing over 500 000 neurons and has produced visually correct parcellation of cortical layers on whole histological sections. In cortical structure analysis using machine learning methods, an important aspect is the interpretability of results and visualization of features that contributed to making a prediction either on a whole model level or for a single instance of the data. Importance of developed features on both global model level and individual neuron prediction level is estimated using SHAP analysis (Lundberg, 2018.).

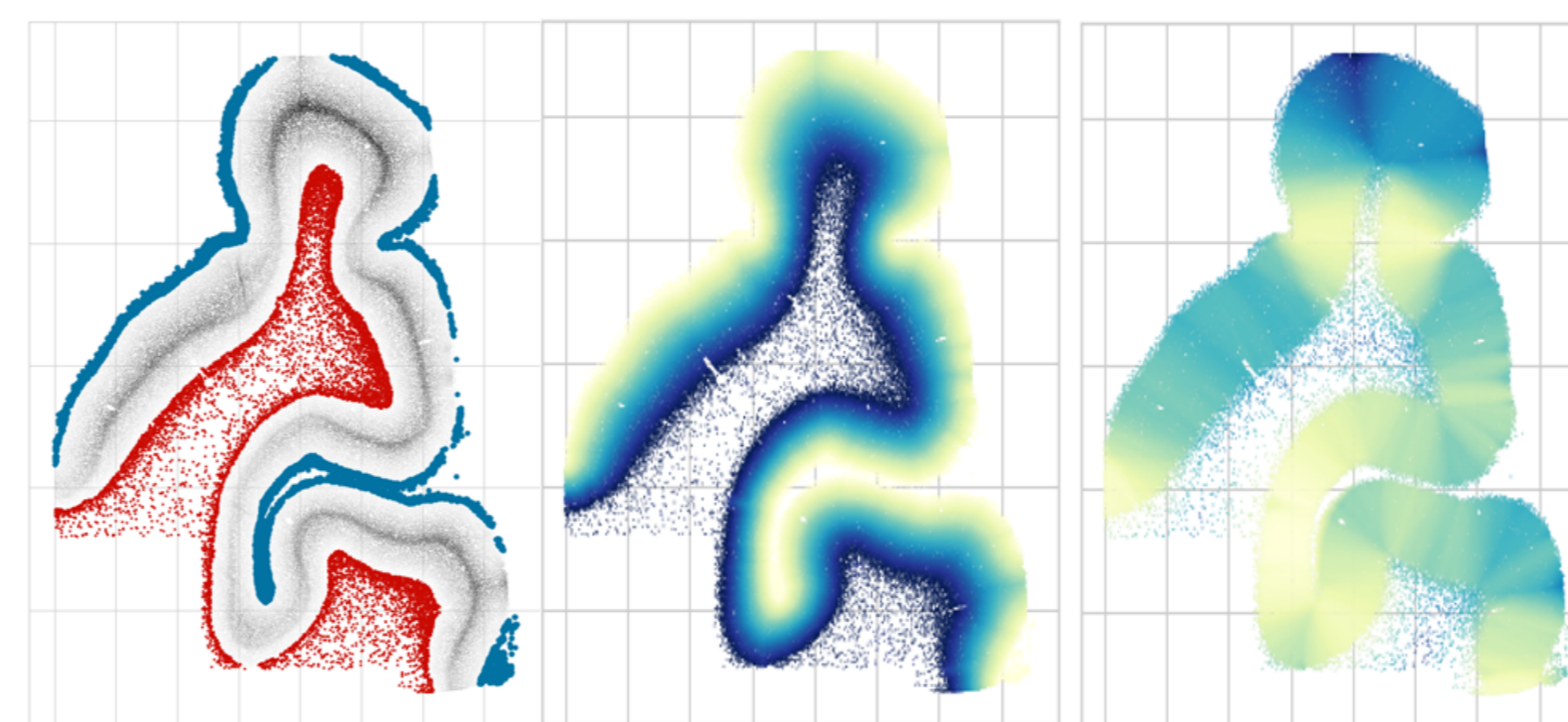


Figure 3: Developed high-level features based on local density and convex hull radius are used to obtain tissue features without drawing profiles and sampling perpendicular to the cortex. Left: sparse areas are separated into layer I and white matter. Middle: cortical depth. Right: cortical thickness.

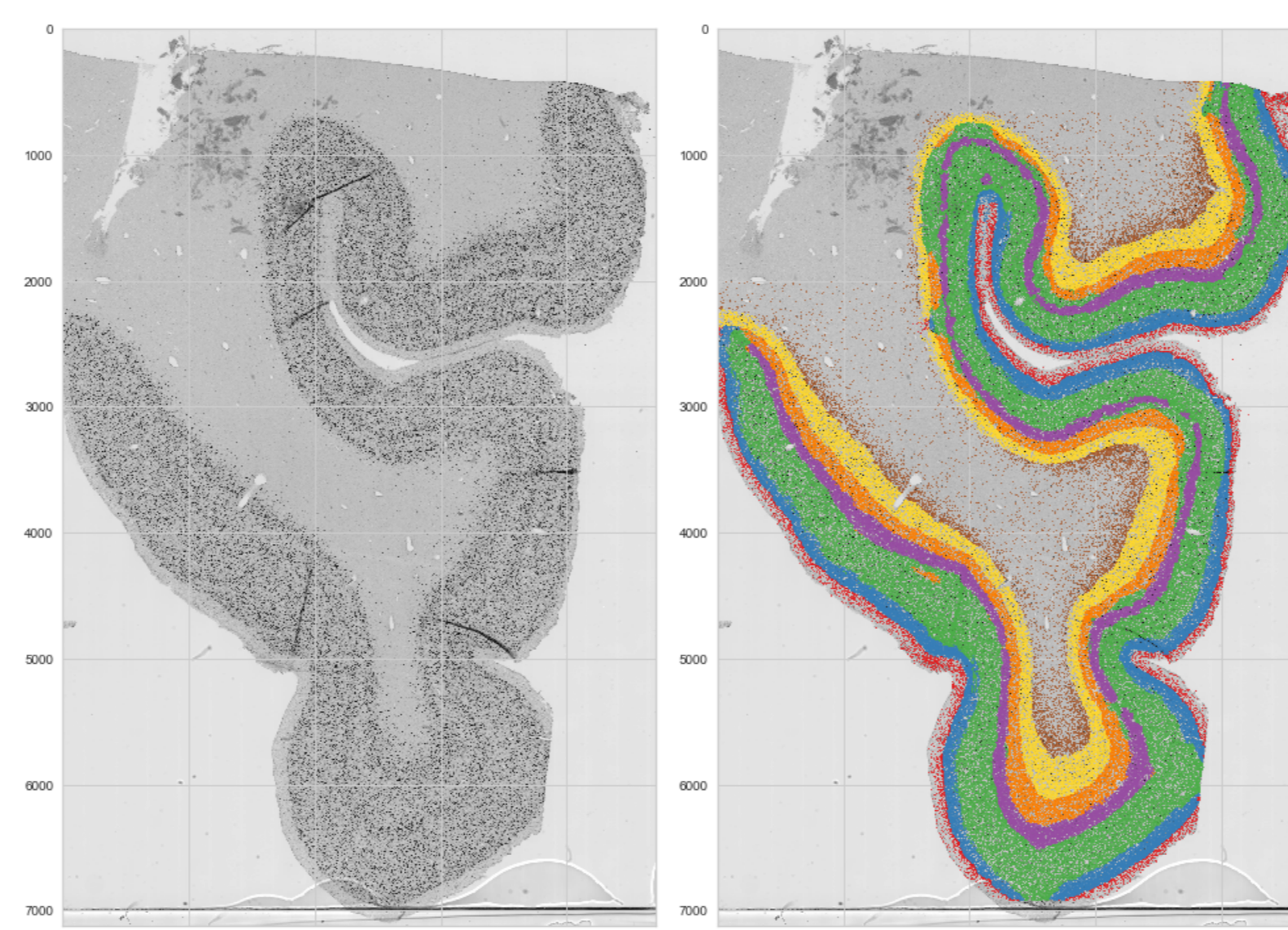


Figure 4: Final classification of cortical layers using CATBoost, a tree-ensemble method.

Fig. 5 shows individual feature attributions at model level, and Fig. 6 shows which neuron features of a single neuron of layer VI contributed to increase from the

base SHAP value and making the prediction. The image also shows the importance of features that decreased the output value for prediction of the same neuron as a white matter neuron. Features of the highest level have largest importance, as they integrate low level features with variations in wider area around neurons. Features based on oriented measurements also have considerable feature importance, identifying neurons on the layer boundaries.

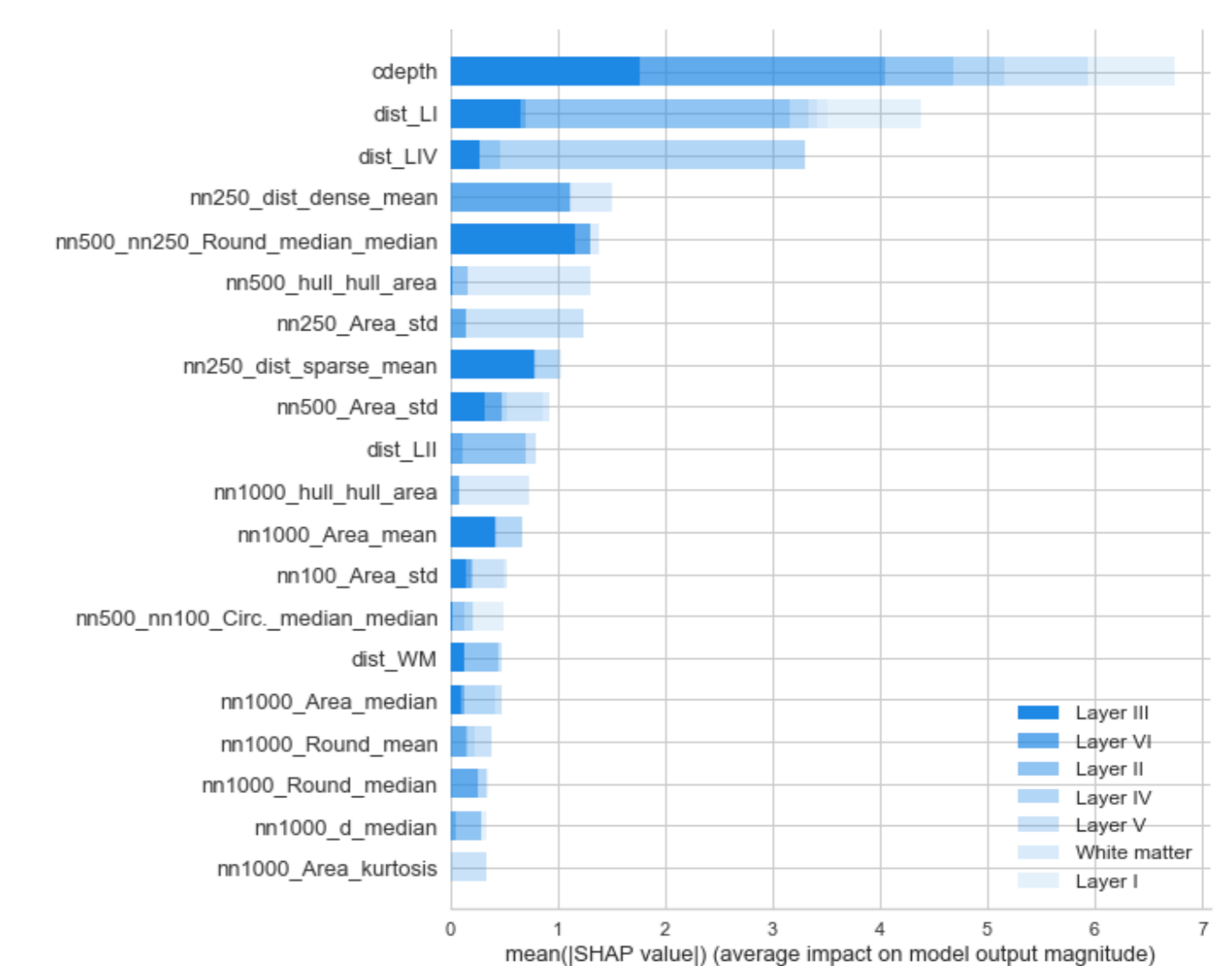


Figure 5: Model-level neuron feature attributions as SHAP values

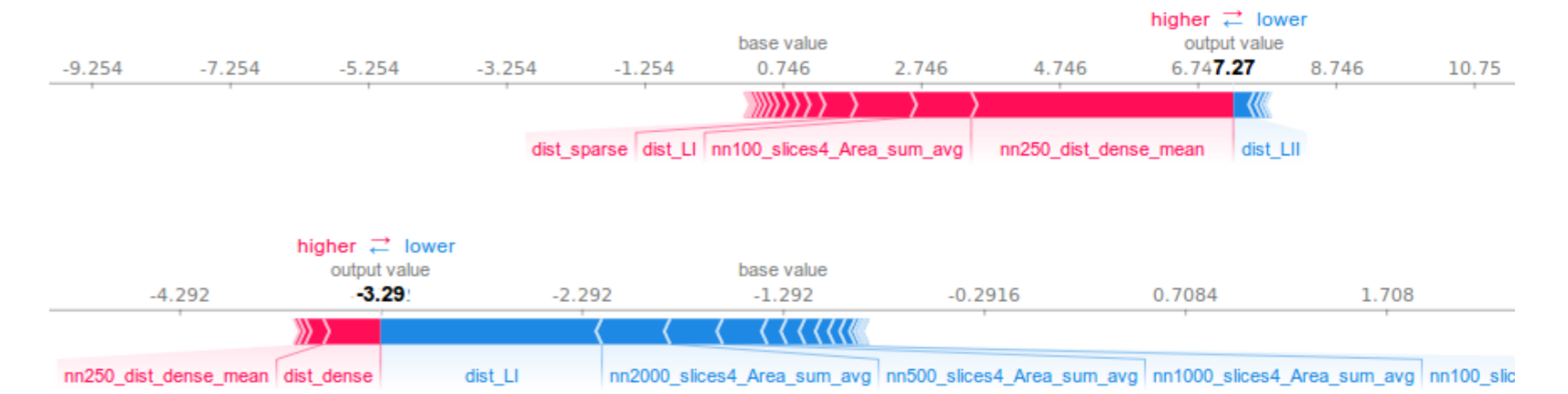


Figure 6: Contribution of different features for making a prediction of a neuron's layer. Top: neuron features of a single neuron of layer VI that contributed to increase from the base SHAP value and making the prediction. Bottom: the importance of features that decreased the output value for prediction of the same neuron as a white matter neuron.

Conclusion

We show that through development of complex neuron features it is possible to learn the distribution of neurons within the cortex and descriptors of cortical layers. Through analysis and comparison of experts' performance we address the question of correctness and existence of true baseline in such tasks. A machine learning model was able to create precise segmentation of cortical layers. Quantitative analysis of neuronal distribution and exploration of underlying principles of variation that facilitate parcellation of cortical layers result in novel framework and methodology for cellular-level investigations of brain tissue. Future work will develop end-to-end procedures based on graph neural network for automatic learning of neuron-level representations of neuron neighborhood.

Acknowledgements

This poster was supported by the European Union through the European Regional Development Fund, Operational Programme Competitiveness and Cohesion, grant agreement No. KK.01.1.1.01.0007, CoRE - Neuro and the Canada First Research Excellence Fund, awarded to McGill University for the Healthy Brains for Healthy Lives initiative. Special thanks to Dora Sedmak and Goran Sedmak from Croatian Institute for Brain Research and Jennifer Novek from Montreal Neurological Institute for providing manual segmentations and valuable discussions.

Bibliography

- Economou, C. and Koskinas, G.N., 1925. *Die cytoarchitektonik der hirnrinde des erwachsenen menschen*. J. Springer.
- Judaš, M. et al. (2011) *The Zagreb Collection of human brains: a unique, versatile, but underexploited resource for the neuroscience community*. *Zagreb Collection, Annals of the New York Academy of Sciences*, 1225(S1), pp. E105–E130. doi: 10.1111/j.1749-6632.2011.05993.x.
- Deroulers, C. et al. (2013) *Analyzing huge pathology images with open source software*, *Diagnostic Pathology*, 8(1). doi: 10.1186/1746-1596-8-92.
- Fahad, A., Alshatri, N., Tari, Z., Alamri, A., Khalil, I., Zomaya, A. Y., Fofouo, S., Bouras, A., (2014) 'A survey of clustering algorithms for big data: Taxonomy and empirical analysis', *IEEE transactions on emerging topics in computing*, Vol. 2, No. 3, pp. 267–279.
- Štajduhar, A., Džaja, D., Judaš, M., and Lončarić, S., (2019) *Automatic Detection of Neurons in NeuN-stained Histological Images of Human Brain*, *Physica A: Statistical Mechanics and its Applications*, Vol. 519, pp. 237–246.
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V. and Gulin, A., (2018) *CatBoost: unbiased boosting with categorical features*, *Advances in Neural Information Processing Systems*, pp. 6637–6647.
- Lundberg, S.M., Erion, G.G. and Lee, S.I., (2018) *Consistent Individualized Feature Attribution for Tree Ensembles*, arXiv preprint arXiv:1802.03888.