Iowa Initiative for Artificial Intelligence

Final Report

Project title:	Gene expression-driven brain parcellation via manifold convolutional neural networks
Principal Investigator:	Thomas Nickl-Jockschat (Associate Professor of Psychiatry, Iowa Neuroscience Institute) Stephen Baek (Assistant Professor of Industrial and Systems Engineering)
Prepared by (IIAI):	Avinash Mudireddy
Other investigators:	Peterson, Zeru J (Research Associate - Psychiatry)
Date:	

Were specific aims fulfilled:	
Readiness for extramural proposal?	Yes
If yes … Planned submission date	Spring/Summer Cycle 22
Funding agency	NIH
Grant mechanism	R01
If no … Why not? What went wrong?	N/A

Brief summary of accomplished results:

Research report:

Aims (provided by PI):

The human brain is the most complex organ known to biology. A crucial factor for its proper development and functioning is gene expression. The formation and maintenance of the enormous cellular diversity and the neuronal circuitry of the brain critically depend on precisely honed gene expression patterns in the respective tissues.

Mapping gene expression in the human brain, however, is incredibly challenging in various regards. First, a significant amount of unspecific variance between donors. Second, body donations to research, in general, and brain donations, in specific, are very scarce. Since only small numbers of brains are available for research, the unspecific variance between individuals becomes a key obstacle, because they do not “average out”, as they would in larger cohorts. Third, given the size of the human brain, specifically the cortex, conventional molecular biological techniques do not allow a continuous mapping of the entire transcriptome across the entire telencephalon.

Gene expression is typically determined only in small biopsies of the cortex. While this is the only feasible approach so far, it leads to large gaps across the brain, for which no data exist. These obstacles, so far, have prevented researchers from answering decisive questions: are there common regional patterns of gene expression in the human brain that are shared across individuals? If so, how do these transcriptomic patterns parcellate the brain and how do they relate to classical – anatomical and

functional – parcellations? Once addressed, these answers would open up an entirely new perspective for neuroscience – how gene expression presumably underlies the structure and function of the human brain. Researchers could capitalize on such a new brain atlas to explore the architecture of the healthy brain, and to develop new insight into what goes awry in neuropsychiatric diseases.

Identifying the transcriptomic architecture of the human cortex by deep learning.

(A) The Allen Human Brain Atlas provides gene expression in ca. 3,700 samples across the human brain. However, inter-individual differences between donors have made it challenging to identify regional transcriptomic patterns.

(B) We will use deep learning to identify regional gene expression patterns that are shared across donors in the cortex. Similar coloring of the sampled MNI-registered tissue loci indicates similar gene expression patterns.

(C) Based on these identified patterns, we will predict gene expression across the cortical surface.

The specific aims of this proposal are to:

(1) develop a usable C++ programming library for manifold CNN, based on Prof. Baek's recent work on Zernike CNN method;
(2) formulate an un-/semi-supervised clustering problem for the gene expression data in the Allen Human Brain Atlas, using the Zernike CNN; and
(3) identify distinctive gene expression patterns for each cluster by visualizing the Zernike CNN

IIAI collaboration:

- Part 1: IIAI consultant will help the team to build an image segmentation network that segments the grey matter from CSF, white matter, and background of a given brain-slice image.
- Part 2: Once the grey-matter is segmented, the next step is to identify separable layers within the grey matter.

Data:

The Allen Human Brain Atlas provides the most comprehensive gene expression atlas so far. It contains expression levels for the entire genome in more than 3,500 tissue samples across six donor brains (Hawrylycz et al., 2012), among which approx. 1,800 samples were taken from the cortical surface.

However, the IIAI team is provided with 2433 image slices of the human brain as input. These slices are random crops of original Allen Human Brain Atlas images (ISH nissls series). Image size is 128*128 pixels.

Using data augmentation techniques, we increased the data size to a total of 24048 images with training, validation, and test split of 21083, 1809, 1156 images respectively.

AI/ML Approach:

Part 1: image segmentation network.

In deep learning, convolutional neural networks (CNNs) are used to analyze vision/image-related tasks. The important applications of CNNs include image and video recognition, recommender systems, image classification, image segmentation, medical image analysis, natural language processing, brain-computer interfaces, and financial time series. In our case, we are interested in Image segmentation.

In general, the summary of a CNN is given below

- Accepts a volume of size $W_1 \times H_1 \times D_1$
- Requires three hyperparameters:
  - their spatial extent F,
  - the stride S,
- Produces a volume of size $W_2 \times H_2 \times D_2$ where:
  - $W_2 = (W_1 - F)/S + 1$
  - $H_2 = (H_1 - F)/S + 1$
  - $D_2 = D_1$
- Introduces zero parameters since it computes a fixed function of the input
- Note that it is not common to use zero-padding for Pooling layers

Where,

- W_1 x H_1 x D_1 are width, height, and channels of input image respectively.
- F is the number of filters
- S is the stride of the filters over a channel of the image
- W_2 x H_2 x D_2 are width, height, and channels of the transformed image respectively.

*Note: The above image slide is from Fei-Fei Li & Andrej Karpathy & Justin Johnson

In our case, to segment different matter in the brain images, we used a modified version of CNN named U-nets inspired by Olaf Ronneberger et al. "U-Net: Convolutional Networks for Biomedical Image Segmentation." We made few adjustments to the U-net architecture to suit the problem. We used MobilenetV2 as the encoder to the U-net architecture.

The model accepts input images of 128*128*3 dimensions and outputs a segmented image of dimensions 128*128*5. The different channels in the output correspond to the background, CSF, white matter, grey matter, and unlabeled class.

In our problem, we are interested in getting the predictions for grey matter more accurate. In those lines, we have adopted a postprocessing technique for better predictions. The post-processing step includes the following steps:

1. We know that the images(128*128*3) used for training and evaluation are random crops of original Allen Human Brain Atlas images (ISH nissl series). As a post-processing trick, divide the original image which we need to predict, into crops of 128*128*3 dimension images using a 16*16 grid. That is, each original image will be divided into 256 slices irrespective of original image size.
2. Run predictions using the trained model for each of those 256 slices.
3. Now reconstruct the predicted image with the same number of pixels as the original image, using the predictions of the slices. In this case, each pixel in the original image might be a part of

one or more slices as the original image is divided into 256 smaller ones. Hence, for each pixel in the original image, there can be one or more prediction labels as it can be part one or more slices. For such predictions, we adopted a polling mechanism. The most frequent label is considered as the prediction for that pixel. If more than 2 labels have the same count, the grey matter label gets more priority, otherwise, any first occurrence is picked.

We observed huge differences in perditions before and after applying the postprocessing trick.

Part 2: to identify separable layers within the grey matter.

To automatically segment 5 layers (6 borders) from the gray matter regions identified as containing the layer structure, a graph-theoretic approach was used using edge-based cost functions of the dark-to-bright or bright-to-dark transition.

We developed a desktop application with an interface where the user draws a central line on the grey matter portion in an image. Based on this central line, we separate pixels perpendicular to the central line using a threshold-based transition of pixels from dark to bright or bright to dark. All the pixels corresponding to a particular transition belong to the same layer. The thresholds for the transition, number of layers, etc., can be customized in the application.

The user features include

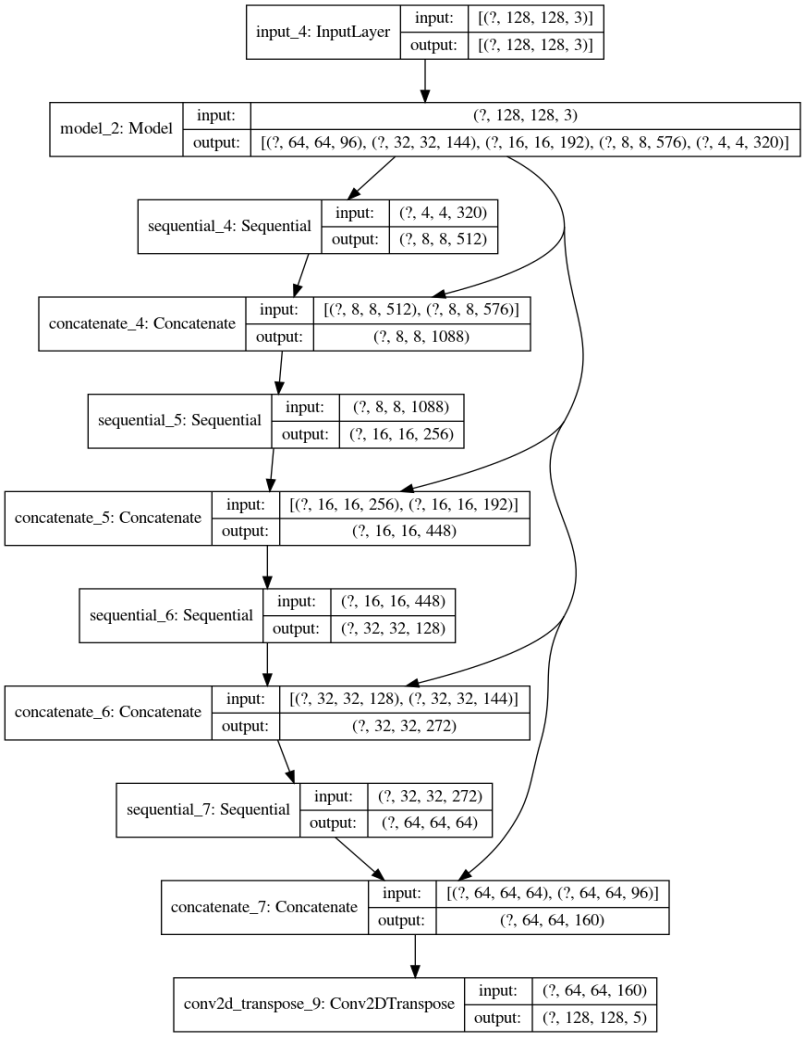
- loading an image and its prediction from part 1
- User to draw a central line
- Algorithm automatically detects the layers depending on the customizable thresholds provided as an input. The application provides to save the detected layers in the form of an XML file and to reload the images and layers
- User can delete the layers and modify the layers

Experimental methods, validation approach:

As described in the “Data” section, we used a total of 24048 image slices with training, validation, and test split of 21083, 1809, 1156 images respectively.

Model description:

The model contains the following layers:



We trained the model for 100 epochs with a batch size of 32 images. We chose Adam optimizer with a learning rate of 0.0001 with a decay of $2e^{-4}$ and a clip value of 0.5. We chose SparseCategoricalCrossentropy as our loss function to optimize. Our evaluation metric is accuracy.

Results:

<u>Part 1: image segmentation network.</u>

After many iterations, our best model has the following metrics:

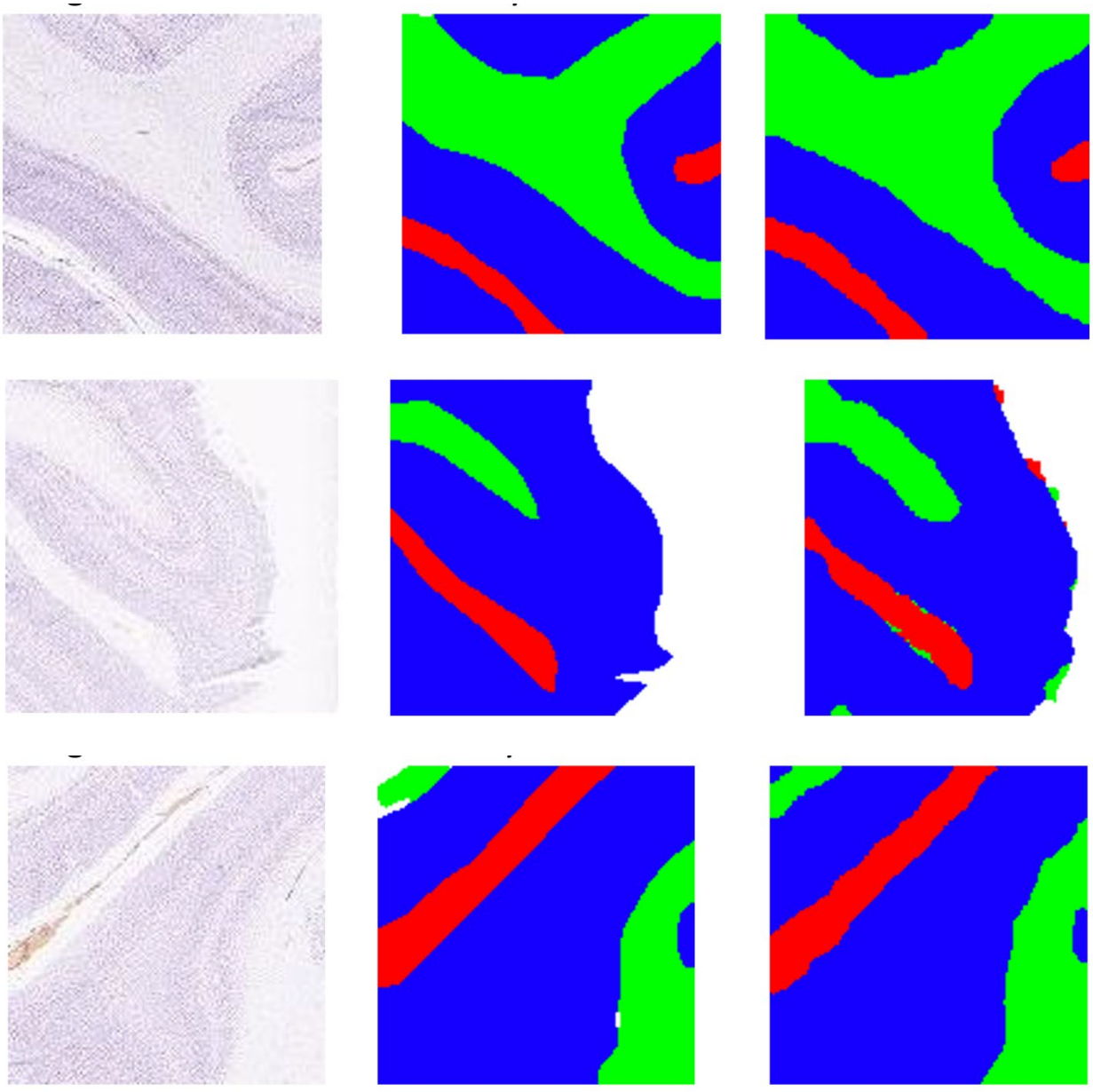
Training loss: 0.2230 Training accuracy: 98.87 %

Validation loss 0.5611 Validation accuracy: 88.89 %

Here are some sample results following the color labels given below:

- Blue = gray matter
- Green = white matter,
- Red = cerebrospinal fluid, and
- White = background

Original image | Manually-labelled mask | Predicted mask

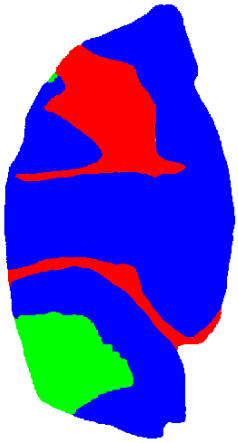


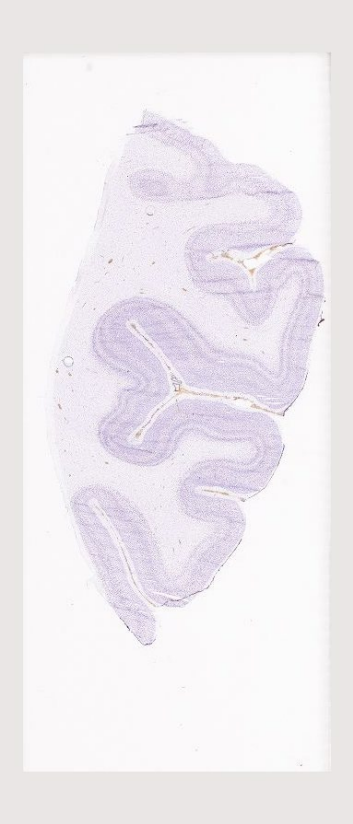
In the above images, the purple/magenta color corresponds to grey matter. Other colors correspond to other matter in the image, but we are not very concerned about the accuracies in the prediction of other matter.

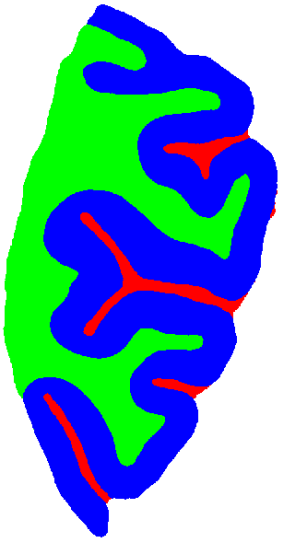
Note that, the accuracy number presented earlier, does include the error in predicting other matter too.

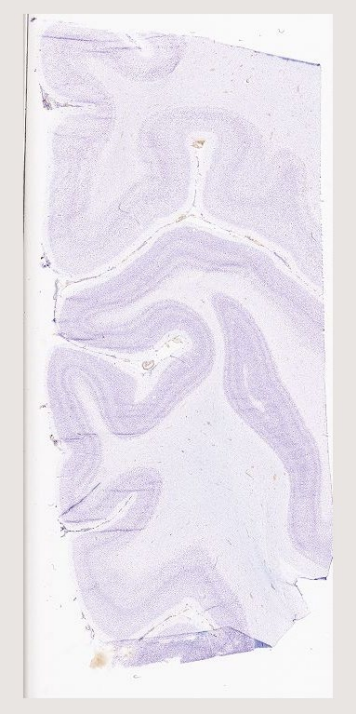
However, we see that in few predictions, the grey matter is not accurately predicted. To have better predictions, we used the post-processing trick as explained in the “AI/ML approach” section.

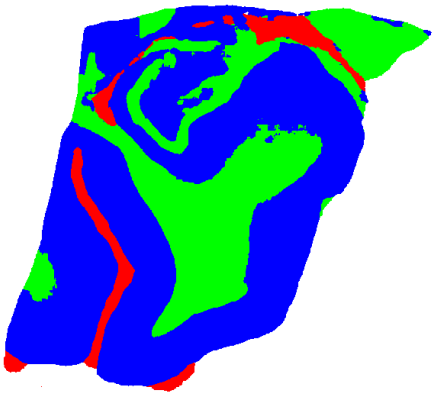
Here are few examples of predictions after the Post-processing trick. Note that the blue color in the below images corresponds to grey matter.

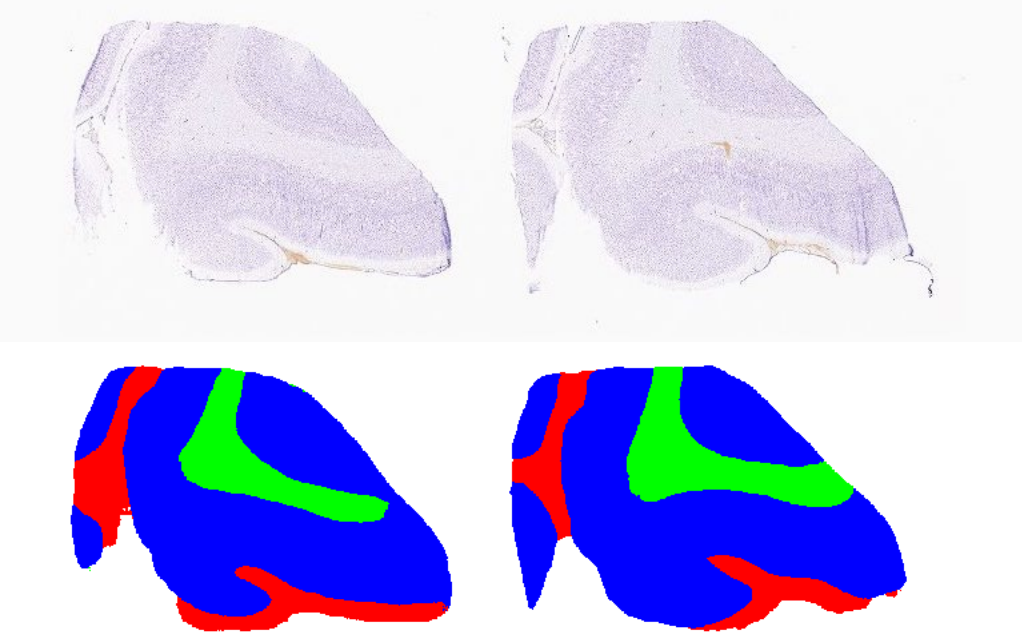






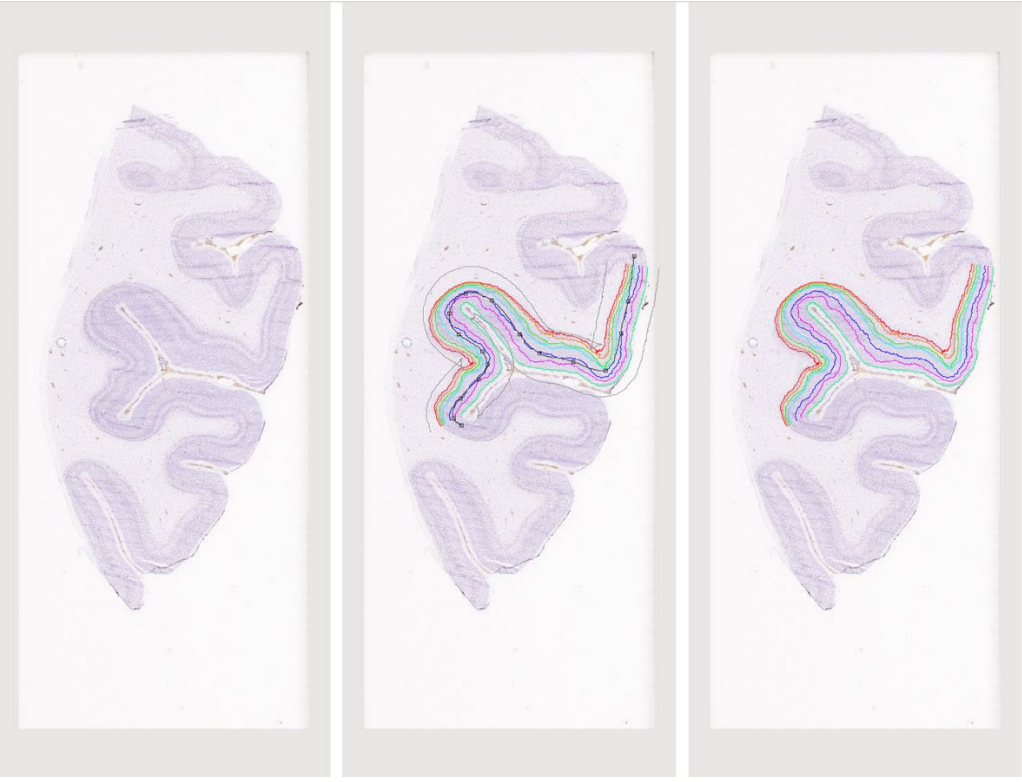


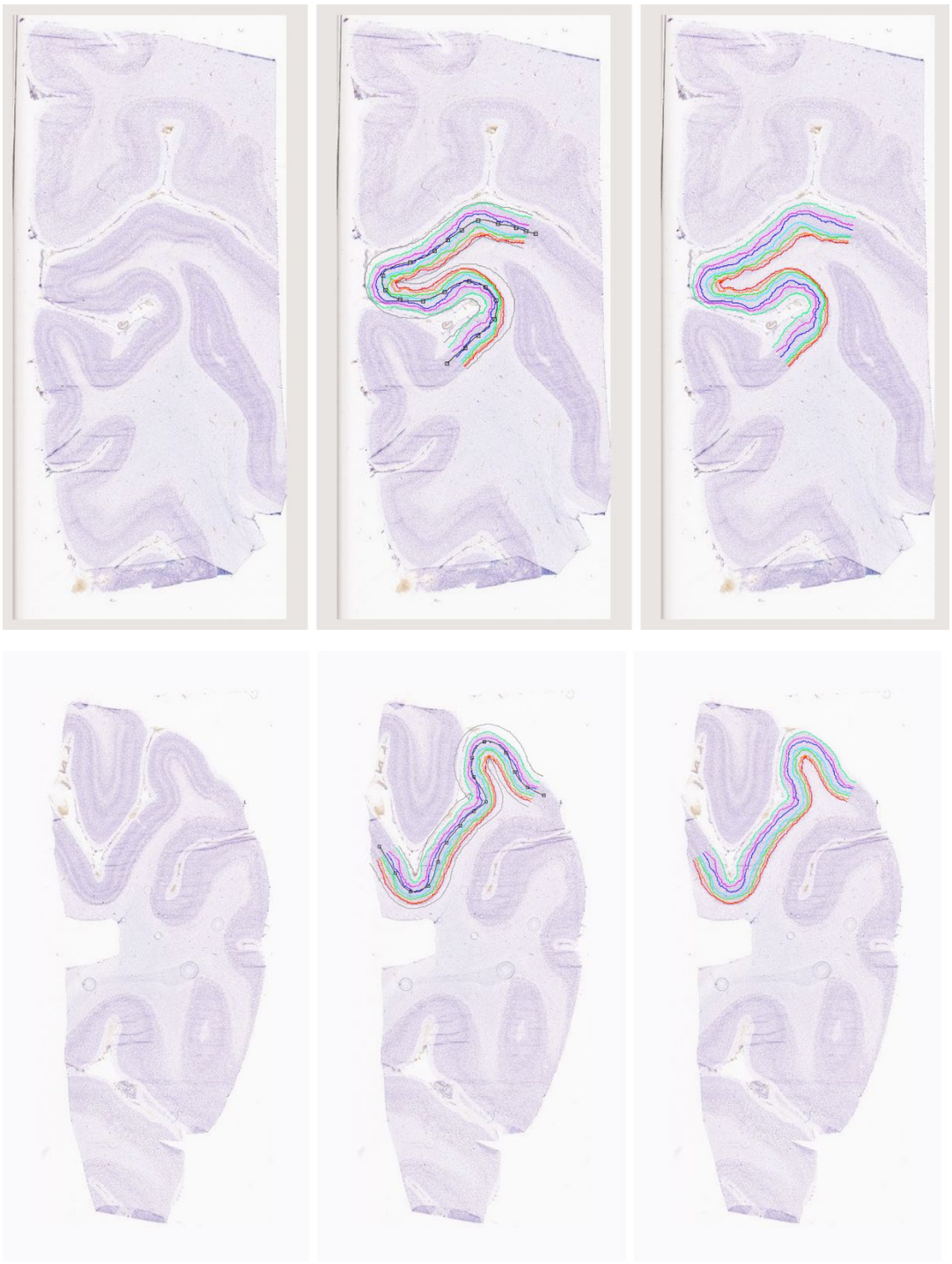




Part 2: to identify separable layers within the grey matter.

As explained earlier, here are examples of the outputs of the desktop application used to identify separable layers within the grey matter.





The blue line denotes the central line drawn by the user. The layers parallel to the blue line are separated by the algorithm. The area covered by the gray line in the middle image is the area selected for the analysis by the algorithm. The right image represents the final output.

Ideas/aims for future extramural project:

The successful implementation of this project would open up two entirely new perspectives for neuroscience. First, magnetic resonance imaging (MRI) studies have been a major avenue in the last

decades to study the structure and function of the brain. A parcellation of the brain based on gene expression would allow a comparison of these patterns to the results of MRI studies. This would enable researchers to actually search for the molecular underpinnings of brain structure and function. Second, vulnerability to neuropsychiatric disorders emerges from the interplay of hundreds of genetic risk variants. While the genetic architecture of these diseases has been well delineated, polygenic interactions still are not understood at all. Our new atlas will provide us with a completely novel tool that can identify susceptibility genes that are likely to interact with each other, as genes that are expressed in the same brain regions are the most likely candidates for such an interaction. This will propel the understanding of the mechanistic underpinnings of neuropsychiatric disorders. These two major advances will provide the foundation for R01 grant applications at the National Institute of Health (NIH), in particular the National Institute of Mental Health (NIMH), the National Institute of Neurological Disorders and Stroke (NINDS), and the National Institute of Child Health and Human Development (NICHD).

Publications resulting from project:

- Nickl-Jockschat et al. “The transcriptomic architecture of the human cerebral cortex” *In preparation.*
- Peterson et al. “Layer-specific gene expression patterns in the human temporal and occipital cortex” *In preparation*