# **Computational Neuroanatomy (Project 5)**

**From the proposal:** "*In stark contrast to brainstem structures with clearly delineated cytoarchitectural borders ... the sensorimotor integration and premotor circuitry exist in the broadly expansive reticular formation. This region lacks clear cytoarchitectural boundaries. As such, those studying brainstem circuits are disadvantaged in their ability to reproducibly locate labeled neurons within this region. To surmount this challenge, we used computer vision and expert training data for the automated detection of brain textures. ... Our methodology permits data from different brains to be superimposed on a common reference atlas, which includes the mean and variation in the location of landmarks*."

**Aim 1: Creation of the Trainable Texture-based Digital Atlas.**

This Aim is on hold until the conceptual issues on Aim 3 are resolved.

**Aim 2: Construction of a fluorescence-based atlas to include spectrally distinct specific cell markers.**

***Automatic detection of marked cells.*** One of the more labor intensive tasks faced by anatomists in our group is identifying individual fluorescently labeled neurons. For example, this was the dominant sink of time for our anatomical study of substantial nigral pars reticular projections (*McElvain et al, Neuron 2021*).

We use machine learning to create a marked cell detector, as depicted in **Figure P5.1**. Both human and computer detection are based on two channels, a channel for the fluorescent marker of projected cells and a reference channel, typically the Nissl-like counterstain Neurotrace. Positive detections are based on the presence of a bright cell signal in the first channel that overlies a signal in the second channel. The detector combines a large number of features that quantify the shapes of the blobs using XGBoost.

A screenshot of a video game

Description automatically generated

**Figure P5.1 Data to illustrate automated detection of cells along with quality control.** The same section is imaged in two channels. The left panel shows the Neurotrace counterstain and detector outputs (dots); the right panel images the GFP fluorescently labeled premotor neurons and detector outputs (dots). Dot size and color corresponds to machine-only vs human QC evaluation of the machine detections. Yellow dots represent confident machine-detections, large cyan dots represent unconfident detections. Small dots correspond to human feedback, white corresponds to positive feedback from the human (that is, the machine detection is correct), purple corresponds to negative feedback (incorrect machine detection). Finally premotor neurons that are missed by the two classes of machine detections (confident and unconfident) are also marked, as shown by the smallest yellow dots.

Importantly the large dots with a purple concentric dot correspond to errors in machine detection. The false positive or false negative detections shown within the blue brackets are replaced by detections in slightly different locations. As the majority of machine-placed circles contain no purple dot, the human and confident machine detections are in close agreement in these images.

Our approach is incremental, rather than a wholesale replacement of the human anatomist by an AI, we use the computer to detect the easily marked cells, and solicit human feedback on the hard ones. Using the concept of boosting margins, we partition the detections into confident, or easy, detections vs. unconfident or hard examples. This scheme is successful if the number of marked cells that it misses, and the number of confident detections that are wrong are both small.

The initial training set was collected by manually locating the marked cells in one out of every four sections, for 20 sections total. This took 20 work hours and generated 2805 human detections.

After training a detector on this data, we applied it to all sections, including the training sections. **Table P5.1** summarized the initial results. The main conclusion is that the confident predictions by the detector are very reliable, only 50 out of the more than 12,000 confident detections were marked by the human as false detection. Detecting that number of marked cells would have taken 85 hours of time. Instead, the human only needs to label the 1420 unconfident predictions, i.e., 10 hours of time (actually less, as visual scanning of the section was unnecessary) . A significant problem is the number of marked cells that were missed (299); reducing this number is a priority.

**Table P5.1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Total** | **Training detections** | **Human +** | **Human -** | **No feedback** |
| Computer confident | 12,224 | 1,599 | 0 | 50 | 12,220 |
| Computer unconfident | 1,420 | 111 | 402 | 247 | 634 |
| No detection |  | 121 | 299 |  |  |

**Aim 3. Improving the computer vision used to detect landmarks.**

In Year 4 we have further improved our structure detectors and made the detections “explainable”. To test the accuracy of those detectors we collected manual annotations of landmark structures from 10 brains (**Table P5.2**).

Table

Description automatically generated

**Table P5.2**: **Structures in the mouse brainstem.** **n** is the number of structures detected (left and right combined). **Mean** and **std** correspond to the mean and std of the n confidence scores. The rows are sorted by decreasing values of the mean confidence. Not surprisingly, 6N and Amb are the two structures easiest to detect.

Detection is done in two steps: Rough alignment followed by structure detection (**Figure P5.3**). The structure detector uses the locations estimated by the rough alignment as starting points for detecting the structure.

* **Rough Alignment** is based on grey-level matching at a resolution of 20 micrometers per pixel (**Figure P5.3**). For this operation we use public code from itk.org. The advantage of this approach is that it uses large grey level features, such as high contrast boundaries and does not require any training. The disadvantages are:
  + - It requires a reference brain that is aligned to the atlas and imaged in the same way as the target brain.
* Limited accuracy.

Chart, scatter chart

Description automatically generated

**Figure P5.2**: Comparing the error of the rough alignment (vertical axis) to the error of the detection (horizontal axis).

Chart, scatter chart

Description automatically generated

**Figure P5.3**: Detection error relative to human (horizontal axis) vs confidence (vertical axis).

* **Structure Detection** uses a separate trained detector for each structure. This detector is cell-based and was described in the Year 3 progress report.

Two important features were added to the detector in Year 4.

1. **Confidence Score:** Some structures are easier to detect than others. We define a confidence score which is based on the auto-correlation of the detection score (**Figure P5.3**). When the confidence score is high the detection is more likely to be at the correct location. When performing alignment we use the rough detections and the high confidence detections first, and then perform corrections using the less confident detections.
2. **Explanation for detections:** Detectors are complex adaptive algorithms. The generated detections are 3D locations (COMs) in the sectioned brain. The confidence score identifies the structures for which the computer’s detection is confident. When these detections are vetted by an anatomist it is very helpful to know **why** the detection is confident. (this is part of a general trend called “Explainable AI”).

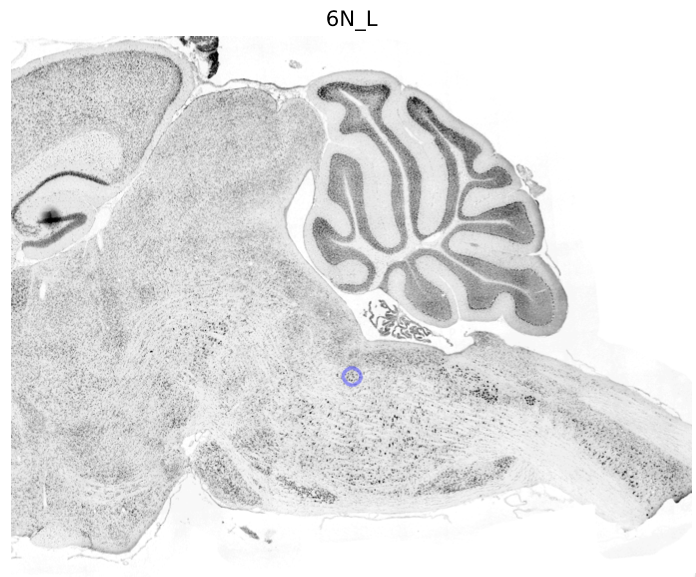
We identify the cells whose features provide the principle evidence for a landmark (inside) versus surrounding regions (outside). This identification makes use of the cell-based aspect of the detector, and that the computational process of Boosting tends to select a sparse representation of all possible features.

To quantify the accuracy of our detector we collected manual structure centers (COMs) for 12 landmarks in 10 brains.

We then compared the automatic detections, after rough alignment and after rough alignment + detection, with the human detection. The figures demonstrate several facts: (1) The detections are significantly better than the rough alignment. (2) Most of the detections have error smaller than 200 micrometers. (3) most detections whose confidence score is higher than 3,000 have error smaller than 100 micrometers.

**Plans for Year 5:**

1. Continue improving structure detectors and identify more structures that can be detected with high confidence. With respect to the data of **Figure P5.4,** the top brain is stained darker than the bottom one, which is likely to confuse a method that is not cell based.
2. Continue improving the marked cell detector and apply it to other tasks.

A picture containing text, envelope, fabric

Description automatically generatedA picture containing text, fabric

Description automatically generated

**Figure P5.4: Explaining detection.** Visual explanations for the detection of 6N\_L in two brains. The cells come from three sections, 100 micrometers apart. The red marked cells are evidence of the inside of the structure, the green cells, which are much smaller, are evidence of the outside. Note that section z=5 has very few cells inside the structure. The detector overcomes this by combining information from all sections. The top brain is stained darker than the bottom one, which is likely to confuse a method that is not cell based.