

Imperial College of Science, Technology and Medicine
Department of Life Sciences

3D Imaging of the Connectome

Arian Rokkum Jamasb

Submitted in part fulfilment of the requirements for the degree of
Bachelor of Science in Biochemistry of Imperial College London, June 2017

Abstract

Advances in automated light microscopy have enabled the routine acquisition of whole-brain datasets from genetically modified mice that express cell-specific labels. However, analytical tools are now required to simplify the quantification of key anatomical features within these very large datasets. I have developed a flexible methodology that enables automated quantification of cellular distributions following serial two-photon imaging. However, the methodology I have developed could be applied to other whole-brain imaging platforms. This analytical tool was validated using a $\text{Sox14}^{\text{gfp}/+}$ mouse strain that strongly labels an interneuron population in the P20 mouse brain. I have demonstrated that the images acquired were of sufficient resolution to enable identification of individual cell somas and axon fibres. The young mouse brain images we acquired using the $\text{Sox14}^{\text{gfp}/+}$ mouse strain were unsuitable for automated registration to an existing anatomical mouse brain atlas due to significant structural disparities. However, this structural disparity cannot be conclusively attributed to being age-related or an artefact of the preparation procedure in this work. I also demonstrate that this methodology requires iterative parameter refinement to account for cell density variability across hemispheres and in different brain regions, although hemispheric refinement may not be required if preparation artefacts are not required. It is hoped that analytical tools of this type will help generate important information on cell loss during the healthy aging process and following neurodegenerative disorders. More basic information about the mouse connectome could also be facilitated by combining this analysis approach with state-of-the-art genetic labelling tools for synaptic identification and network reconstruction.

Acknowledgements

I would like to express my most humble thanks to my supervisor for this work, Dr Stephen Brickley, for his guidance and patience throughout. I am greatly indebted to Mr Gerald Moore for the most in-depth feedback on my work I think I shall ever be lucky enough to receive. And, of course, the apples.

This work would not have been possible if not for Dr Simon Schultz providing access to the TissueCyte system and a dedicated workstation for data analysis, and Dr Alessio Delogu & Ms Polona Jager from King's College London for providing the Sox14^{gfp/+} mouse.

Once there appeared before King el-Harit, the son of Modad, a wise man who declared he could read destinies in the sand. “Do you make exact calculations?” the king asked him and before the magus could recover from his surprise he added “if you do not know how to calculate exactly, your visions are worth nothing; if you arrive at them through calculation alone, I disbelieve them. In India I learned a proverb that says, ‘Distrust the calculation seven times over, the calculator a hundred times.’”

Malba Tahan, *The Man Who Counted*

Contents

| | |
|---|-----|
| Abstract | i |
| Acknowledgements | iii |
| 1 Introduction | 1 |
| 1.1 Objectives and Motivation | 1 |
| 1.1.1 Objectives | 1 |
| 1.1.2 Motivation | 2 |
| 1.2 Background | 2 |
| 1.2.1 Whole-brain Imaging Technologies | 5 |
| 1.2.2 Fluorescence-based Whole-brain Imaging Technologies | 6 |
| 1.2.3 TissueCyte 1000 Serial Two-Photon Microscope | 10 |
| 2 Materials and Methods | 13 |
| 2.1 Tissue Preparation | 13 |
| 2.1.1 Fixation | 13 |
| 2.1.2 Embedding | 14 |
| 2.2 Image Acquisition | 14 |

| | | |
|----------|--|-----------|
| 2.3 | Image Pre-Processing | 15 |
| 2.4 | Image Registration | 16 |
| 2.4.1 | Automated Registration | 16 |
| 2.4.2 | Manual Registration | 16 |
| 2.5 | Quantification | 17 |
| 2.5.1 | Watershed Algorithm | 17 |
| 2.5.2 | Establishing a Ground-truth Dataset | 19 |
| 3 | Results | 20 |
| 3.1 | Primary GFP Fluorescence Signal Sufficient For Cell Identification | 20 |
| 3.2 | Distribution of Sox14 ⁺ Interneurons | 21 |
| 3.3 | Image Registration | 22 |
| 3.3.1 | Automated Registration | 22 |
| 3.3.2 | Manual Registration | 24 |
| 3.4 | Evaluation of Automated Cell Counts | 25 |
| 3.4.1 | Thresholding & Cell Segmentation | 25 |
| 3.4.2 | Centroid Detection | 28 |
| 3.4.3 | Anatomical Segmentation of Centroids | 29 |
| 4 | Discussion | 30 |
| 4.1 | Summary of Thesis Achievements | 30 |
| 4.2 | Manual Registration | 30 |
| 4.2.1 | Hemispheric Asymmetry of Cell Counting | 31 |
| 4.3 | Algorithm Improvement | 33 |

| | | |
|----------|--|-----------|
| 4.3.1 | Thresholding | 34 |
| 4.3.2 | Overcounting of Clusters | 35 |
| 4.3.3 | Alternative Approaches to Counting Dense Regions | 35 |
| 4.4 | Applications and Further Work | 36 |
| 4.4.1 | Distributed Validation | 36 |
| 4.4.2 | mGRASP Synaptic Counting | 36 |
| 4.4.3 | Rabies Virus Tracing | 37 |
| A | Appendix | 38 |
| A.1 | Coronal Sections Acquired are Off-Axis | 39 |
| B | List of Abbreviations | 41 |
| C | Glossary | 43 |
| | Bibliography | 47 |

List of Figures

| | | |
|-----|--|---|
| 1.1 | <i>Cerebellar Purkinje cell stained with Golgi stain drawn by Ramón y Cajal. Instrumental to the comprehension of the morphological visualisation is the sparse staining of the cells, revealing fine structural detail. Structural inferences are possible due to simplification of image data. a axon, b collateral, c dendrites. Image from Gray's anatomy</i> | 3 |
| 1.2 | <i>Serial Section Analysis flow chart. SSA is a precursor methodology for imaging tissue volumes through manual sectioning, preparation and imaging of tissue sections. Each stage must be repeated for each section. Lack of automation in the procedure introduces sources of heterogeneity between sections. Z registration of images to produce stacks is challenging as slices are not consistently positioned or oriented. Image: Timothy Ragan . . .</i> | 5 |
| 1.3 | <i>Overview of automated Light Microscopy methods for whole-brain imaging. STP is the only method that provides submicron resolution imaging of whole tissue volumes. All methods suffer from background fluorescence arising from nonspecific excitation of bulk sample below the imaging plane. STP, Serial Two-Photon Tomography; fMOST, fluorescence Micro-Optical Sectioning Tomography; LSFM, Light Sheet Fluorescence Microscopy; PMT, Photomultiplier Tube; Image: Osten & Margrie, 2013</i> | 7 |

- 1.4 *Jablonski diagram demonstrating difference between one (**left panel**) and two-photon (**right panel**) excitation. One photon excitation uses a single high energy photon (**blue**) to cause an electronic transition in the fluorophore. Two-photon excitation allows use of two lower energy photons (**red**), each corresponding to roughly half the required energy, to cause the same electronic transition. In both cases, the emission signal (**green**) as the fluorophore returns to the ground state is recorded by photomultiplier tubes to produce an image.* 8
- 1.5 *Overview of TissueCyte imaging setup. Sample is fixed to microscope slide and immersed in PBS to prevent drying. Bath is secured on an x-y-z stage capable of precisely moving the sample between the lens and the microtome.* 10
- 1.6 *Signal path employed in TissueCyte 1000 microscope. **Laser:** excitation generated using Ti:sapphire laser with 400 nm tuning range (680 - 1080 nm), peak power output > 3.5 W at 800 nm. **Objective lens** focusses laser light on sample. Emission signal is routed through lens. **Dichroic mirrors** are used to decompose the signal into three channels on the basis of wavelength prior to collection by PMTs. **Channel 1** collects > 560 nm (autofluorescence signal), **Channel 2** collects 500-560 nm (GFP fluorescence signal), **Channel 3** collects < 500 nm. No filters other than IR filters to block out laser light are installed in the signal collection path. Image: Gerald Moore.* 11
- 1.7 *Microtome sectioning of tissue block results in distortions and aberrations at the surface. TissueCyte instrument is capable of optical sectioning enabling the image acquisition plane to be positioned below surface, preventing histology-level artefacts in imaging. This capability is not unique to the TissueCyte instrument, although two-photon imaging enables greater optical sectioning depth than confocal microscopy and light sheet microscopy does not provide sufficient or consistent resolution throughout the volume. Optical sectioning depth can be improved with tissue clearing.* 12

| | | |
|-----|--|----|
| 2.1 | <i>TissueCyte imaging setup. Agarose-embedded sample is fixed using an adhesive to a magnetic microscope slide and immersed in PBS throughout imaging to prevent drying. Bath is secured on an x-y-z stage capable of precisely moving the sample between the imaging and cutting positions for sectioning by microtome. Image: Gerald Moore</i> | 14 |
| 2.2 | <i>Overview of algorithm implementation employed to count cells. Thresholding and Adjustment: Contrast is enhanced using tophat and bottomhat filtering. Autofluorescence signal is subtracted from GFP fluorescence signal to remove residual background fluorescence. Image is converted from 8-bit grayscale to a binary 1-bit black and white image to enable object detection. Object Detection: Objects are filled to produce solid, continuous objects suitable for watershedding. Morphological opening performed to remove isthmi and protrusions from objects. Objects with fewer than 3 connected pixels removed to further remove noise and fluorescence not corresponding to cell somas. Centroid Quantification Putative centroids recognised using an extended-maxima transform to recognise regional maxima in the image. Watershedding applied to define and enumerate detected centroids. Centroid Segmentation Centroids segmented into anatomical regions of interest using manually registered index images using the colour ontology scheme, and segmented into hemispheres depending on position relative to the image midline.</i> | 17 |
| 3.1 | <i>Third ventricle cell distribution. Left panel: Individual cell somas can be discerned in downscaled JPEG file used for analysis outside regions of saturating fluorescence. Right panel: Colour scheme inverted, enabling visualisation of axon fibres. Scale bar: 100 μm.</i> | 20 |

- 3.2 *3D Reconstruction of whole Sox14^{gfp/+} mouse brain using serial two photon tomography. dLGN shows symmetrical distribution of Sox14^{gfp/+} interneuron population across both hemispheres. DM and surrounding tissues display strong, saturating fluorescence signal. Top row: full brain reconstruction using autofluorescence signal. GFP signal overlaid in green. Middle panel: common reference space used to define orientation. Taken from Allen Brain Atlas. Bottom row: GFP fluorescence visualisation only. dLGN: dorsal lateral geniculate nucleus; DM Dorsomedial nucleus of the hypothalamus.* 21

3.3 *Stereotaxic coordinates (mm) for adult mouse brain. Red line: width indicates 200 μm region where the 10 thalamic sections analysed are derived from. Images analysed are spaced 20 μm apart, the approximate width of a cell soma, to overcome oversampling of cells due to images being taken in 5 μm optical sections. All coordinate measurements relative to bregma. Figure adapted from Franklin and Paxinos 2008.* 22

3.4 *Visualisation of automated registration of coronal sections. Automated registration produces reasonable cortical alignment. Subcortical alignment accuracy is insufficient to allow accurate anatomical segmentation of cellular distribution. Green: autofluorescence intensity signal. Blue: registered anatomical atlas; anatomical regions defined using intensity ontology scheme from Allen reference atlas. Red: Outline of template hippocampus.* 23

- 3.5 Example of a section image manually registered to Paxinos and Franklin's stereotaxic mouse brain co-ordinates. **Top left:** manually registered image with outlines used for manual counting. Alignment of dLGN prioritised to provide uniform dataset for counting. **Top right:** manually registered image with colour ontology scheme overlaid used as an index for anatomical segmentation of cells. **Bottom left:** close up of ventricular misalignment in the hypothalamus. **Bottom right:** close-up of dLGN (right hemisphere) showing regional segmentation. **Blue:** ventricles; **light blue:** granular cell layer of the hippocampus; **red:** dLGN; **yellow:** dorsomedial nucleus of hypothalamus; **green & purple:** arcuate nucleus of hypothalamus (*ArcMP* & *ArcLP*, respectively). 24
- 3.6 Image histogram showing distribution of pixel intensities in a typical section image. **X axis** shows pixel intensity on an 8-bit integer scale (0-255) from black to white. **Y axis** shows the number of pixels $\times 10^6$ in each intensity bin. Section images show a unimodal intensity distribution - typical of fluorescence micrographs due to selective illumination of small objects in a large imaging area. Unimodal distribution prevents use of conventional image segmentation methods that seek to find the optimum threshold between two classes in a bimodal distribution. Instead, a global intensity threshold must be used to segment images. 25
- 3.7 Total cell counts for varying binarization thresholds across 10 thalamic sections for dLGN across both hemispheres (**top panel**), left hemisphere (**bottom left panel**), right hemisphere (**bottom right panel**). **Horizontal line** shows total ground truth counts for each condition. **Dashed line** shows optimal threshold derived from intersection of automated and ground truth counts. **Otsu threshold** yielded a binarization threshold of 0.27. Thresholds selected: 0.65 (left hemisphere), and 0.4 (right hemisphere). 26

- 3.8 *Manually defined global intensity thresholding improves counting algorithm accuracy than conventional bimodal thresholding techniques for images with unimodally distributed pixel intensities. Slopechart plots of manual and automated cell counts for dLGN using Otsu and optimal Global intensity threshold. Each pair of connected points corresponds to the counts obtained from one thalamic section image. Top panel: right hemisphere; lower panel: left hemisphere. Black line shows mean for each plot.* 27
- 3.9 *Comparison of centroid detection in dense and sparse regions. Sparse regions: cells are identified by counting algorithm with good accuracy. Small clusters of overcounting can be observed and may be reduced. Dense regions: cells cannot be counted accurately due to saturating fluorescence. Individual cell somas cannot be discerned manually or boundaries defined computationally using standard computer vision toolkits. Development of a fluorescence area-based approximation method or a full 3D rendering of each cell volume from data with greater z-resolution may enable estimation of cell counts for these regions. Colour of point corresponds to anatomical region: Red: detected centroid; blue: dLGN, yellow: ArcMP, green: ArcLP; magenta: DM; cyan: ventricule* 28
- 3.10 *Visualisation of Centroid Segmentation. Left panel: All detected centroids in a section. Red points: detected centroid, blue points: centroid corresponding to dLGN Right panel: All detected centroids corresponding to an anatomical region defined in the colour ontology scheme. Blue: ventricles, red structure: dLGN, magenta: DM, yellow: ArcMP. green: ArcLP.* 29

| | | |
|-----|---|----|
| 4.1 | <i>Incorrect sample preparation results in imaging plane no longer being perpendicular to central axis. Misalignment of imaging and central axes likely result in the need for disparate hemispheric binarization thresholds due to the underlying cellular populations providing unequal fluorescent backgrounds. Registration of young mouse brain may be possible with true coronal sections aligned with the central axis. Alignment problem can be overcome by reslicing of data on a tilt, however this is limited by z-resolution to enable accurate interpolation of images between layers.</i> | 32 |
| A.1 | <i>Visualisation of delayed structure emergence in the olfactory bulb (anterior region of the brain). 150 µm in the z-axis between first and last image</i> | 39 |
| A.2 | <i>Visualisation of delayed cortical emergence in the hindbrain (posterior region of the brain). 150 µm in the z-axis between emergence of structure in each hemisphere</i> | 40 |

Chapter 1

Introduction

1.1 Objectives and Motivation

1.1.1 Objectives

The objective of this work is to develop a method by which accurate automated counts of fluorescently-labelled target cells can be rapidly produced from *ex vivo* whole-brain two-photon imaging. Currently, there is no suitable anatomical atlas for automated segmentation of a young mouse brain. In the process, we seek to manually segment only the subcortical region of a young, P20, mouse brain. We will count Sox14⁺ thalamic interneurons, transgenically labelled with a Sox14-GFP fusion protein, as we have previously observed a strong primary fluorescence signal signal. Primarily, we focus on a sparse interneuron population located in the dorsal lateral geniculate nucleus (dLGN) and a dense population located in the dorsomedial nucleus (DM) of the hypothalamus. However, these cells are not the specific focus of this investigation as the goal is to produce a generalised image analysis pipeline for quantifying cellular distribution across anatomical regions in the brain.

1.1.2 Motivation

Neurodegenerative disorders are characterized by a reduction of cell numbers in the brain leading to functional decline. The nature of this loss, as a result of different mechanisms of neurodegeneration, will differ in the cell types and regions affected, and relative intensities of loss. It has suggested that while the number of synapses decline with age, age-related neuron loss been shown to be modest and may in fact not be the main driver of age-related cognitive decline (Gómez-Isla et al. 1996). Despite this, it is feasible to propose that certain neuronal and non-neuronal cellular populations in the brain will decline with age. It has, however, been shown neurons undergo morphological changes with age - such as a reduction in the number of dendritic spines on pyramidal cells of the prefrontal cortex - that may correlate with cognitive decline (Duan et al. 2003).

Developing a counting method can allow unique insight to the differential cellular population dynamics and parameters underlying neural cell loss in mouse models of neurodegeneration. Furthermore, a flexible counting method not restricted in application to cells would enable enumeration of morphological features of neurons, providing insight into the mechanisms of age-related cognitive decline.

1.2 Background

Ramón y Cajal's double impregnation adaption of Golgi's stain allowed him a glimpse into the structural architecture of the brain. The key to the success of the stain was that it only labelled a few neurons, reducing the complexity of the image thereby allowing cellular morphology to be discerned (Figure 1.1). Advances in imaging methods now enable us to acquire whole-brain microscopy datasets, birthing a host of large research initiatives. Amongst these initiatives are the Allen Brain Atlas for gene expression (Sunkin et al. 2013), the ongoing Mouse Brain Architecture Project for mesoscopic connectivity (Bohland et al. 2009) and the FlyEM project, which has recently acquired a whole-brain electron microscopy (EM) volume of an adult female fly at synaptic resolution, hoping to

produce a completed connectome of *Drosophila Melanogaster* (Zheng et al. 2017).

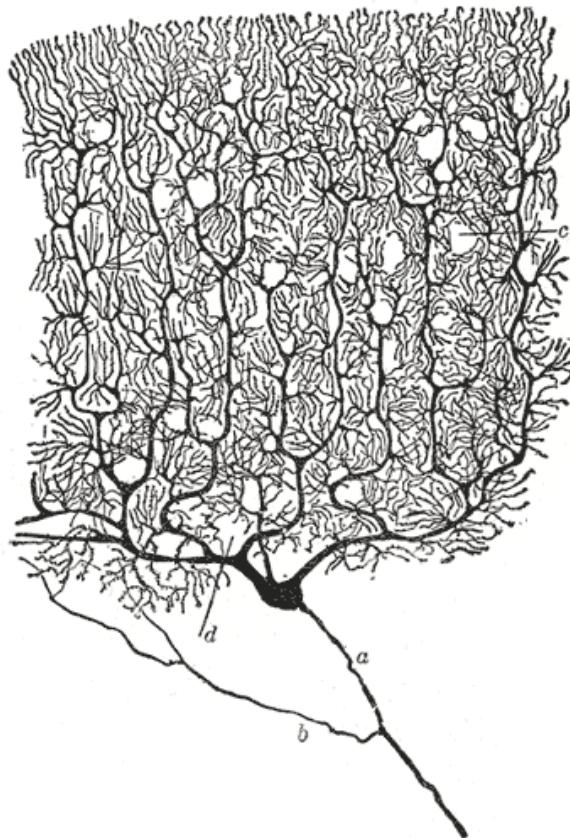


Figure 1.1: *Cerebellar Purkinje cell stained with Golgi stain drawn by Ramón y Cajal. Instrumental to the comprehension of the morphological visualisation is the sparse staining of the cells, revealing fine structural detail. Structural inferences are possible due to simplification of image data.* **a** axon, **b** collateral, **c** dendrites. Image from Gray's anatomy

This focus on systematic generation of whole-brain datasets requires high-throughput instrumentation and an analytical toolbox to, like Cajal, simplify the data. In a way we make same compromise as Cajal in this work: though we can now resolve much greater structural detail across much larger tracts tissue, imaging is restricted to a subset of cells. From observing the morphologies of limited numbers of cells, Cajal was able to pose a great number of hypotheses of the inner workings of the brain. However, a full digital reconstruction of the whole tissue volume is still rich with data beyond cellular morphology that can be mined and extracted from the images.

High throughput acquisition of such datasets have bolstered connectomics, an *omics* sub-field of neuroscience seeking to construct a complete map of brain connectivity. Previously,

whole-brain analyses were the preserve of large initiatives but the advent of automated microscopy tools has brought the possibility into the university laboratory. To fully exploit the research output of these initiatives, flexible methods that complement whole-brain imaging techniques, simplify the data, and optimise for throughput are required. It is only in doing so that adequate sample sizes of whole-brain analyses can be acquired to generate sufficient statistical power to support structural and histological inferences regarding an organ characterised by complexity and variability.

Early work in microscopy treated the raw images as pictorial data analysed only visually, either by the experimenter or recorded using a *camera lucida* or photographic plates. Digital image acquisition has transformed the value of microscopy data; images are represented as numerical arrays where integer values correspond to pixel intensities (gray-scale images use an 8-bit integer scheme). This fundamentally changes what can be achieved with microscopy, enabling computational analysis of image data, thereby rendering microscopy a truly quantitative technique.

Currently, there are many groups working on whole-brain analyses using different methods of image acquisition. Tools with a strong focus on cross-platform compatibility are required to maximise sharing and comparison of datasets. The simplest and, arguably, most universal metric obtainable is cell counts. However, raw whole-brain counts of cells are of little use without their distribution in an anatomical context. Thus, we seek to produce a generalised counting methodology that will output the distribution of a fluorescently-labelled cell-type within each anatomical region of the mouse brain and is sufficiently flexible for deployment on datasets acquired on different platform and quantifying morphological features.

1.2.1 Whole-brain Imaging Technologies

For this work, we used a TissueCyt^e 1000 Serial Two-photon Tomographic (STP) microscope (Ragan et al. 2012). For image acquisition, this is not the only available technology but it is the most suited to our purposes (Figure 1.3). Current approaches can be categorised into two groups: block-face imaging, where the sample undergoes serial sectioning and imaging, and Light Sheet Fluorescence Microscopy (LSFM) on cleared tissue samples, where the imaging plane is illuminated laterally.

Despite differences in acquisition procedure, the imaging methods all experience the same problems of anatomical segmentation for young mouse brains, and quantification discussed in §2.3-2.5. However, the counting method developed should be applicable to datasets acquired using other instruments with some platform-specific parameter optimization.

Serial Section Analysis

Serial sectioning and imaging of tissues has been used in biological research for many years at a variety of scales. The methodology (Figure 1.2) consists of entirely manual sectioning of a sample using a microtome, mounting the sections onto slides, staining, and subsequently imaging. The acquired images then need to be aligned in the z -dimension before stacks can be produced, often hampered by heterogeneity in slide preparation.

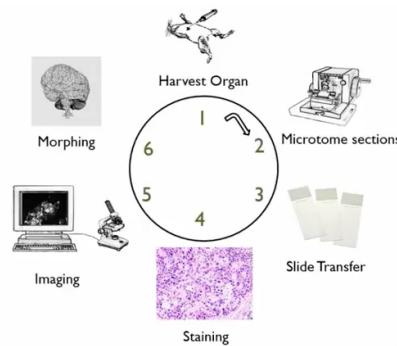


Figure 1.2: *Serial Section Analysis flow chart.* SSA is a precursor methodology for imaging tissue volumes through manual sectioning, preparation and imaging of tissue sections. Each stage must be repeated for each section. Lack of automation in the procedure introduces sources of heterogeneity between sections. Z registration of images to produce stacks is challenging as slices are not consistently positioned or oriented. Image: Timothy Ragan

The technique requires modernisation in order to obtain a comprehensive whole-brain dataset. Microtome sectioning of tissue prior to imaging introduces irreducible histological deformation artefacts. As a manual technique, it is labour-intensive, limiting the throughput and rendering the technique non-scalable to imaging large tissue volumes.

Electron Microscopy

EM allows imaging of biological tissue with nanometre resolution through using a focussed beam of electrons to overcome the diffraction limit typically associated with light microscopy. Compared to the other methods outlined, EM gives unrivaled imaging resolution; only EM enables complete 3D reconstruction of neural tissue at nm resolution. This, however, comes at a significant time cost in both data acquisition and analysis.

It remains unclear whether such resolution is required to further understanding of circuit and network function. The recently acquired adult *Drosophila* EM volume produced a 106 terabyte dataset over several years; tracing neurons through the volume for *in silico* reconstruction will be ongoing for many years into the future. Maximising analytical throughput is central to this work, making EM wholly unsuitable for this application.

1.2.2 Fluorescence-based Whole-brain Imaging Technologies

The key advantage of fluorescence microscopy is selective excitation and visualisation of fluorescent objects. Using genetic-labelling tools available for mice, fluorescence microscopy affords the ability to simplify imaging by discriminating between cell-types of interest and background. Fluorescence microscopy requires refinement to enable acquisition of whole-brain data sets. The spatial scale of brain tissue (cm), strongly optically scattering tissue and photodamage arising from prolonged imaging sessions must be addressed.

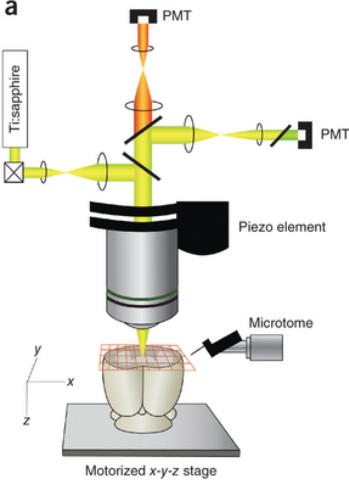
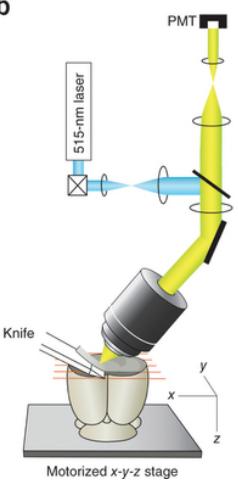
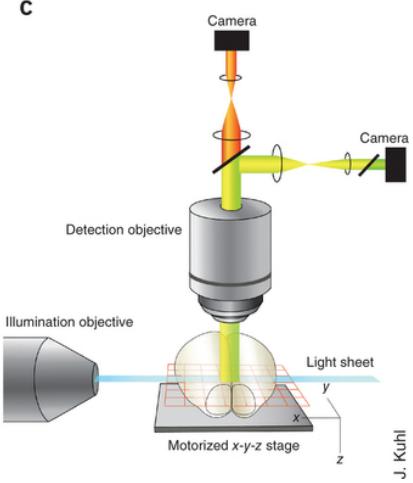
| STP | fMOST | LSFM |
|---|---|---|
|  |  |  J. Kuhl |
| Two-Photon | Confocal Line Scan | Light Sheet |
| Resolution | 1 μm | Deteriorates with depth |
| Submicron | | |
| Sectioning Method | | |
| Microtome | Diamond Knife | Optical Sectioning |
| Section Thickness | | |
| 50 μm mechanical, 1 μm optical | 1 μm | 5 μm |
| Maximum Optical Sectioning Depth | | |
| 600-800 μm | 200 μm | Whole brain |
| Image Tiling | | |
| Mosaic | Line scanning | Mosaic |
| Signal Collection | | |
| PMT | PMT | Camera |
| Sample Preparation | | |
| Embedding procedure maximises preservation of fluorescence signal | Sample requires dehydration and plastic embedding | Requires tissue clearing (attenuates fluorescence signal) |
| Advantages | | |
| Sections can be collected for post-imaging analysis | Making knife opaque to fluorescence signal could remove background fluorescence | Sample intact post-imaging |

Figure 1.3: *Overview of automated Light Microscopy methods for whole-brain imaging. STP is the only method that provides submicron resolution imaging of whole tissue volumes. All methods suffer from background fluorescence arising from nonspecific excitation of bulk sample below the imaging plane.* **STP**, Serial Two-Photon Tomography; **fMOST**, fluorescence Micro-Optical Sectioning Tomography; **LSFM**, Light Sheet Fluorescence Microscopy; **PMT**, Photomultiplier Tube; Image: Osten & Margrie, 2013

Two-Photon Microscopy

Two-photon excitation allows the a number of imaging limitations associated with whole-brain samples to be overcome. Two photons - each corresponding to approximately half the required energy - are used to produce a higher energy electronic transition in a fluorophore, and the resulting emission signal is recorded to produce the image (Figure 1.4). This requires using pulsed, high power laser light to initiate large numbers of photon-photon interactions. As a consequence of using lower energy photons, longer wavelengths are used. This limits photobleaching and photodamage of the sample, vital for prolonged imaging sessions and provides greatly increased spatial and axial resolution. Longer excitation wavelengths in the infrared range afford greater optical penetration of the tissue, enabling imaging deeper below the surface as this corresponds to the optical window of the tissue - the range of wavelengths in which tissue absorbs relatively little light.

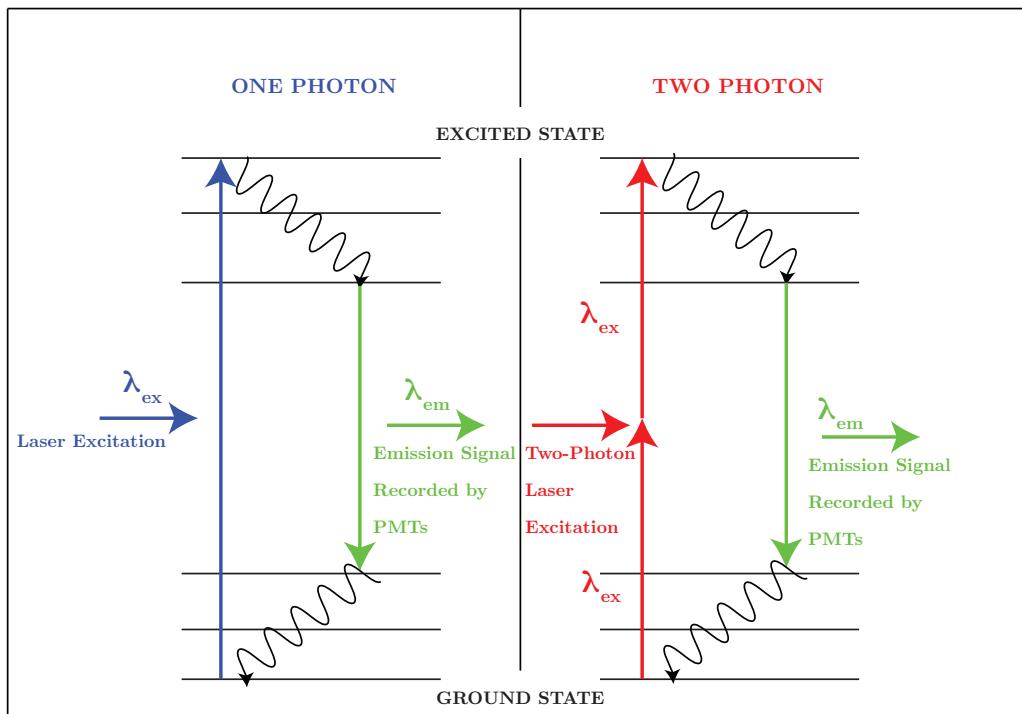


Figure 1.4: *Jablonski diagram demonstrating difference between one (**left panel**) and two-photon (**right panel**) excitation. One photon excitation uses a single high energy photon (**blue**) to cause an electronic transition in the fluorophore. Two-photon excitation allows use of two lower energy photons (**red**), each corresponding to roughly half the required energy, to cause the same electronic transition. In both cases, the emission signal (**green**) as the fluorophore returns to the ground state is recorded by photomultiplier tubes to produce an image.*

Fluorescence Micro-Optical Sectioning Tomography

fMOST utilises a diamond knife to section the sample into ribbons typically $450\ \mu\text{m}$ in width. These ribbons are imaged using a confocal microscope concurrently to sectioning, greatly increasing the speed of data acquisition. This platform has the potential to reduce issues of background fluorescence through making the ventral surface of the knife blade opaque to the emission wavelengths. However, the voxel resolution of this technique is limited to $1\ \mu\text{m}$ in its current iteration.

Light Sheet Fluorescence Microscopy

LSFM enables rapid imaging of cleared tissue samples through selective horizontal plane illumination. Imaging is again performed in a mosaic tile fashion. Due to strong optical scattering from brain tissue, the tissue must be cleared. Briefly, the sample is stabilized in a hydrogel matrix and lipids, responsible for a large portion of the optical scattering, removed electrophoretically. Clearing enables greater permeability of the tissue to antibodies can also be used with the other imaging platforms such as the TissueCyte to increase optical sectioning capabilities.

1.2.3 TissueCyte 1000 Serial Two-Photon Microscope

The TissueCyt^e 1000 (Ragan et al. 2012) is a high speed multiphoton microscope with an attached vibrating microtome capable of acquiring whole-organ image stacks through serial sectioning and imaging (Figure 1.6). In many ways, the TissueCyt^e 1000 resembles a robotic implementation of SSA. However, there are a number of key features that suit it to our purposes.

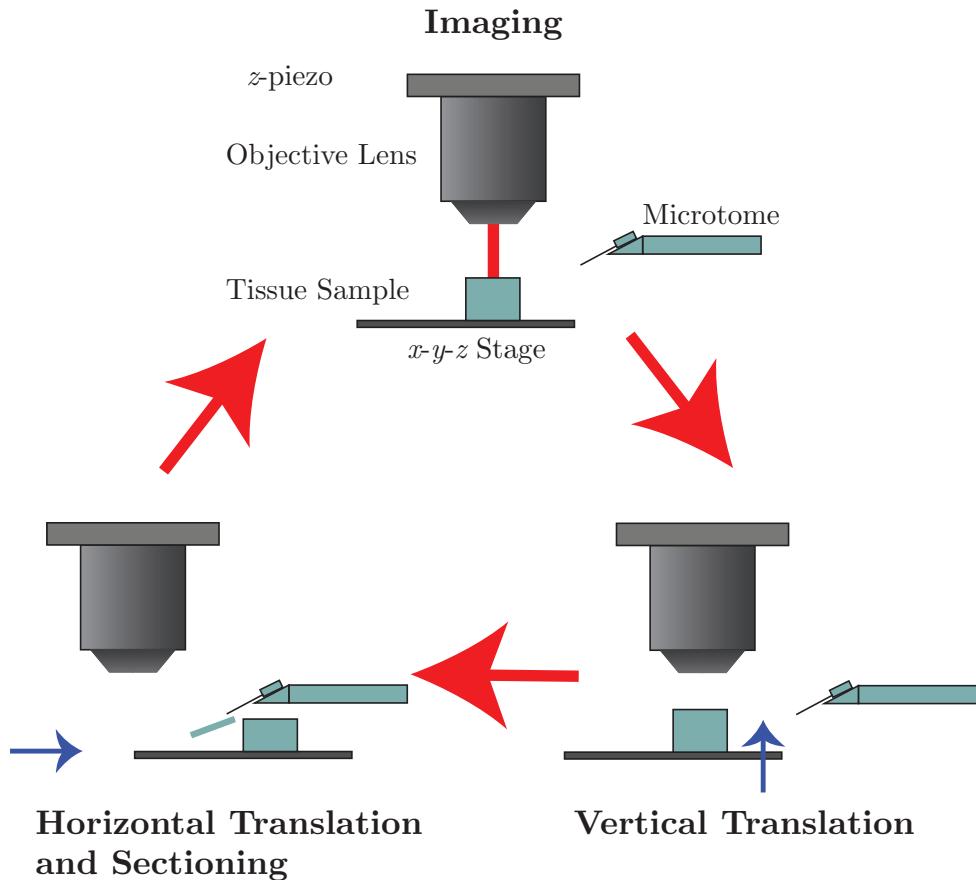


Figure 1.5: *Overview of TissueCyt^e imaging setup. Sample is fixed to microscope slide and immersed in PBS to prevent drying. Bath is secured on an x-y-z stage capable of precisely moving the sample between the lens and the microtome.*

Crucially, it is capable of acquiring images with submicron resolution, enabling visualisation of individual cell somas and axon fibres. At this resolution, it is even capable of resolving individual dendritic spines and, potentially, synapses. Due to the automation of the instrument, this platform provides significant improvements in experimental reproducibility and throughput. Automated sectioning provides immediate z -registration of the

images in production of image stacks. Furthermore, the instrument is capable of multi-spectral image acquisition and is compatible with a host of both virally and transgenically expressed fluorophores, enabling specific visualisation of many cellular subpopulations.

For excitation, the TissueCyte system uses a Ti:Sapphire Laser with a 400 nm tuning range between 680 nm and 1080nm, and a peak power output >3.5 W at 800 nm. For imaging, the two-photon laser light is focussed into the sample and a galvanometer mirror is used to scan across the imaging tile. The emission signal is collected on three PMTs, separated by wavelength using dichroic mirrors. This is repeated for each image tile in a section until the entire section has been imaged. The sample is then sectioned using the microtome and the procedure repeated for the next section.

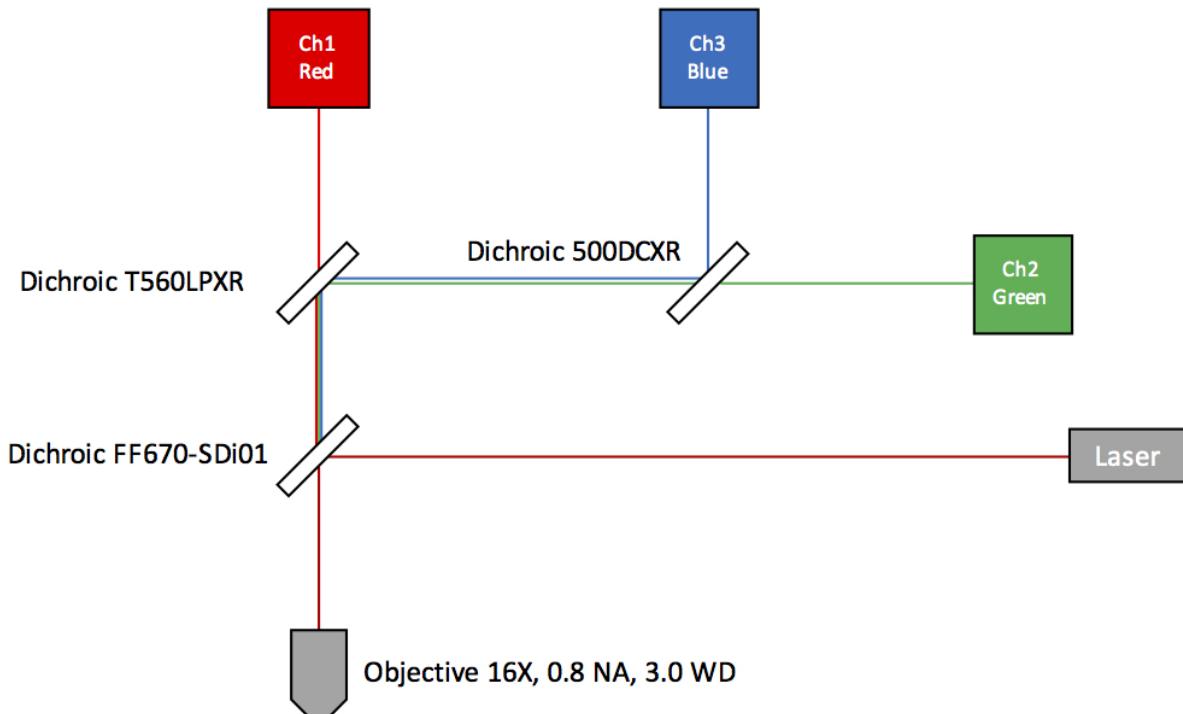


Figure 1.6: *Signal path employed in TissueCyte 1000 microscope. **Laser:** excitation generated using Ti:sapphire laser with 400 nm tuning range (680 - 1080 nm), peak power output > 3.5 W at 800 nm. **Objective lens** focusses laser light on sample. Emission signal is routed through lens. **Dichroic mirrors** are used to decompose the signal into three channels on the basis of wavelength prior to collection by PMTs. **Channel 1** collects > 560 nm (autofluorescence signal), **Channel 2** collects 500-560 nm (GFP fluorescence signal), **Channel 3** collects < 500 nm. No filters other than IR filters to block out laser light are installed in the signal collection path. Image: Gerald Moore.*

The TissueCyte system does not produce deformation artefacts typically associated with microtome sectioning. Instead of imaging the surface of the sample, the optical sectioning capabilities of the instrument enable it to focus on an imaging plane beneath the distorted surface. In this way, the tissue is imaged intact prior to sectioning (Figure 1.7). This capability is not unique to TissueCyte but is greatly enhanced by use of multi-photon imaging.

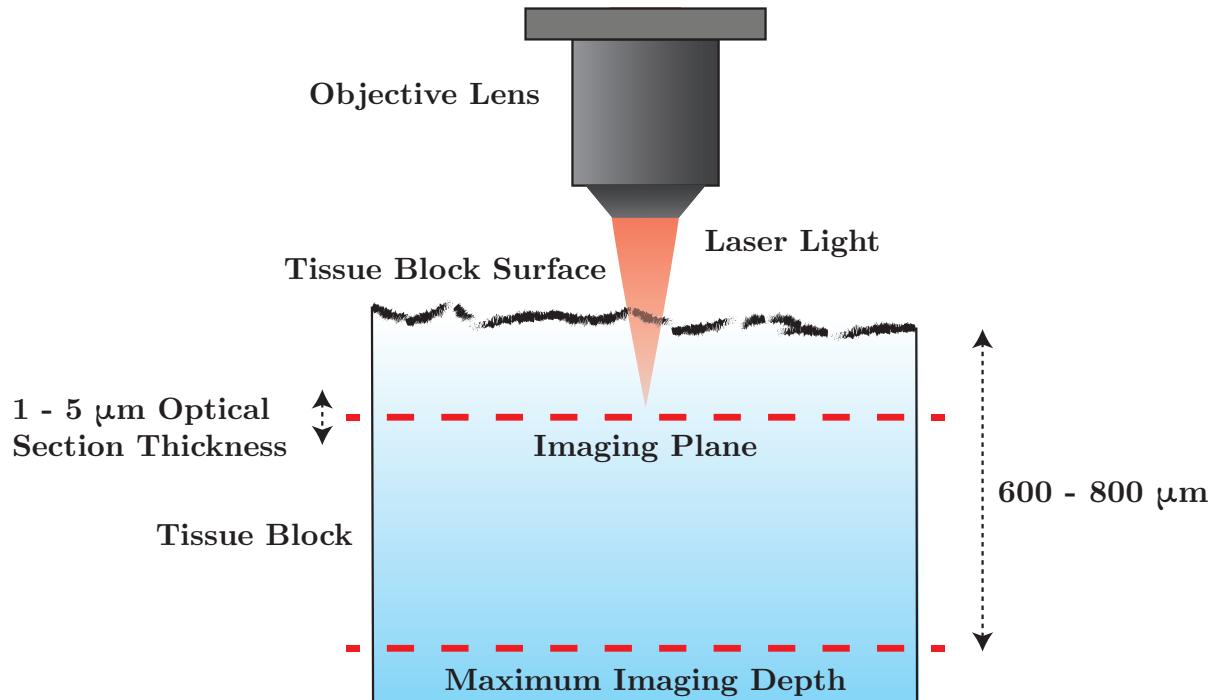


Figure 1.7: *Microtome sectioning of tissue block results in distortions and aberrations at the surface. TissueCyte instrument is capable of optical sectioning enabling the image acquisition plane to be positioned below surface, preventing histology-level artefacts in imaging. This capability is not unique to the TissueCyte instrument, although two-photon imaging enables greater optical sectioning depth than confocal microscopy and light sheet microscopy does not provide sufficient or consistent resolution throughout the volume. Optical sectioning depth can be improved with tissue clearing.*

Chapter 2

Materials and Methods

2.1 Tissue Preparation

Experiments were performed on a single heterozygous Sox14^{gfp/+} P20 mouse (Delogu Lab, King's College London). Cells expressing the Sox14 transcription factor are tagged with GFP in the transgenic line, thereby labelling interneurons in the thalamus and other subcortical regions. The preparation procedure outlined below serves to retain the structural integrity of the tissue sample throughout imaging and maximise preservation of the primary GFP fluorescence signal.

2.1.1 Fixation

The animal was sacrificed and subjected to a standard transcardial perfusion tissue fixation procedure (Gage et al. 2012) using 4% paraformaldehyde (PFA, Sigma). Transcardial perfusion allows greater and more uniform penetration of the tissue than typical immersion fixation through utilisation of the animal's vasculature. The brain is excised and stored in 4% PFA at 4 °C for 24 hours. Subsequently, the sample is washed three times and stored in phosphate-buffered saline (PBS, pH 7.4) at 4 °C.

2.1.2 Embedding

To facilitate positioning of the sample, a portion of the cerebellum is removed in the dorsal-ventral plane to form a flat surface on which the sample can stand. The sample is placed in a cubic mould (VWR) with the anterior surface facing up, and embedded in 4.5% oxidised agarose to ensure homogenous sectioning and prevent tissue deformation resulting from the microtome (Sallee and Russell 1993). Briefly, agarose is oxidised by addition of 10 mM sodium periodate (NaIO_4 , Sigma) solution over 2 hours at room temperature, washing thrice using a vacuum pump and resuspension in PBS to bring the final concentration to 4.5%. Covalent cross-linking between brain surface and agarose was induced by equilibrating in excess 0.5-1% sodium borohydrate (NaBH_4 , Sigma) in 0.05 M sodium borate buffer (pH 9.0-9.5) overnight at 4 °C. Cross-linking in this way ensures the sample remains securely embedded within the agarose block during sectioning.

2.2 Image Acquisition

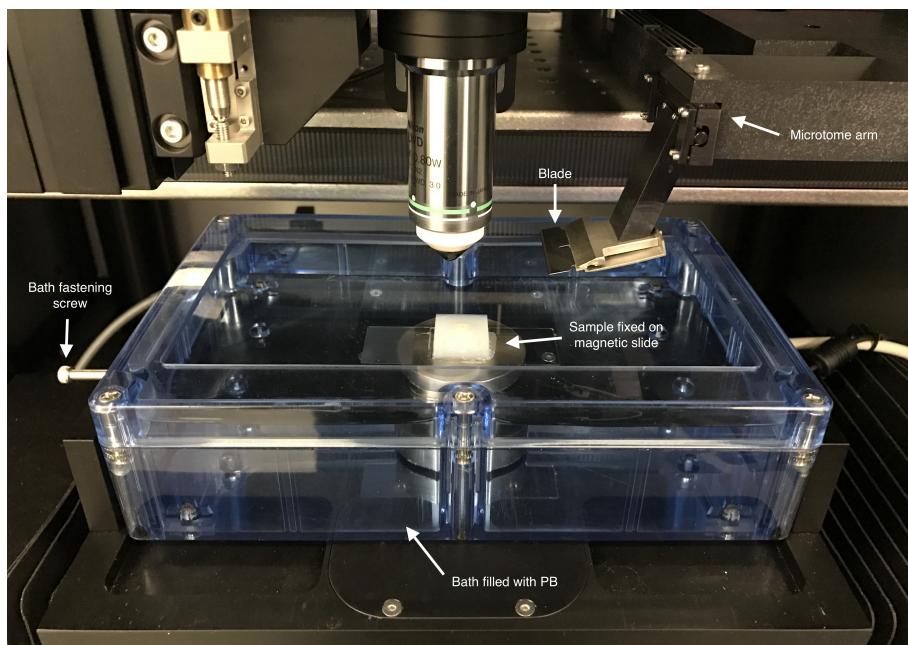


Figure 2.1: *TissueCyte imaging setup. Agarose-embedded sample is fixed using an adhesive to a magnetic microscope slide and immersed in PBS throughout imaging to prevent drying. Bath is secured on an x-y-z stage capable of precisely moving the sample between the imaging and cutting positions for sectioning by microtome. Image: Gerald Moore*

The sample is fixed to a microscope slide, which is immersed in a chamber containing PBS for imaging with a TissueCyte 1000 serial two-photon tomographic microscope (TissueVision) (Figure 2.1). The cutting position and sample dimensions are input manually. Images were acquired using an excitation wavelength of 920 nm produced by a Ti:Sapphire Laser source (Chameleon Ultra II, Coherent). Imaging was performed using 50 μm vibratome sectioning and 5 μm optical sectioning. Autofluorescence signal (<560 nm) is collected on channel 1 and is used for template registration; the GFP fluorescence signal (500-560 nm) is collected on channel 2 and is used for cell quantification.

2.3 Image Pre-Processing

The TissueCyte system outputs high-resolution image tiles, indexed by section and tile position. These are processed and stitched to form an image representation of the entire cross-sectional surface of the brain for each section. Stitching is computationally intensive and is performed asynchronously, as the images are acquired. Due to the terabyte size of the data sets acquired, images are stored on a network attached storage drive (NAS) from which they are pulled in order to perform processing operations. Image tiles are cropped by 1.8% pixels from each edge to remove illumination artefacts, and rotated 90° for correct orientation. Images are then corrected for differences in illumination intensity across a tile by constructing an average illumination tile for each layer. Each tile in a layer is averaged and all tile intensities divided by the average illumination. Stitching is achieved using a MATLAB script to call FIJI to perform linear blending of the processed images tiles. Section images are downscaled 50% from lossless TIFF files to lossy JPEG files to reduce computation time due to the non-linear time complexity of the quantification algorithm used downstream.

2.4 Image Registration

The acquired images require registration to a brain atlas in order to overlay anatomical regions. Registration is performed on the autofluorescence images due to improved visualisation of anatomical features. The morphological transform is then applied to the template to enable registration of the GFP fluorescence image. Both the Allen reference atlas and Paxinos and Franklin's stereotaxic mouse brain co-ordinates (Franklin and Paxinos 2008), supplemented with a custom colour ontology scheme, were used for automated and manual registration respectively.

2.4.1 Automated Registration

Automated registration to the Allen Reference Atlas was carried out using the aMAP (automated mouse atlas propagation) software based on the NiftyReg MRI Segmentation toolkit (Niedworok et al. 2016). The software has previously been validated against human raters for registration of adult mouse brain STP image data.

2.4.2 Manual Registration

Images were manually registered to Paxinos and Franklin's stereotaxic mouse brain coordinates using Adobe Illustrator CC 2017. Cortical regions were removed and subcortical regions aligned between section image and template. The template is then warped to more closely fit the forebrain. This atlas was chosen over the Allen reference atlas for manual registration due to the provision of anatomical segmentation template images for both hemispheres.

2.5 Quantification

2.5.1 Watershed Algorithm

A watershed algorithm implemented in MATLAB 2015b was used to detect cells in each section (Figure 2.2). The output of this algorithm is the total number of cells detected, the number of cells detected in each anatomical region, the number of cells detected in each anatomical region for each hemisphere, and the positions of these cells denoted by a co-ordinate representation of their centre of mass. Three images are used to count each slice: the Ch1 denoised stitched autofluorescence signal image, the Ch2 denoised stitched GFP signal image, and the registered Ch1 section with an overlaid colour ontology scheme.

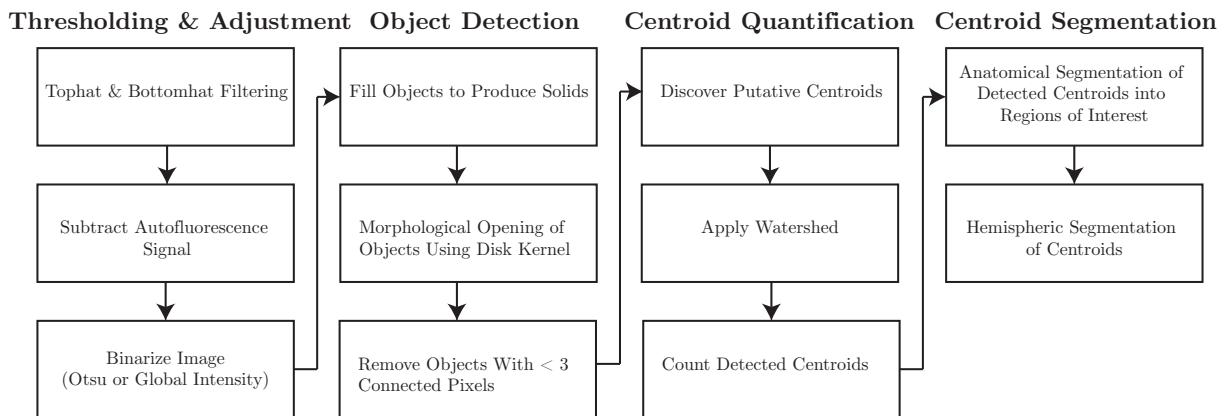


Figure 2.2: Overview of algorithm implementation employed to count cells. **Thresholding and Adjustment:** Contrast is enhanced using tophat and bottomhat filtering. Autofluorescence signal is subtracted from GFP fluorescence signal to remove residual background fluorescence. Image is converted from 8-bit grayscale to a binary 1-bit black and white image to enable object detection. **Object Detection:** Objects are filled to produce solid, continuous objects suitable for watershedding. Morphological opening performed to remove isthmi and protrusions from objects. Objects with fewer than 3 connected pixels removed to further remove noise and fluorescence not corresponding to cell somas. **Centroid Quantification** Putative centroids recognised using an extended-maxima transform to recognise regional maxima in the image. Watershed applied to define and enumerate detected centroids. **Centroid Segmentation** Centroids segmented into anatomical regions of interest using manually registered index images using the colour ontology scheme, and segmented into hemispheres depending on position relative to the image midline.

Thresholding and Adjustment

Top-hat filtering is performed on the GFP fluorescence signal image using a 10 pixel disk kernel - approximately the pixel radius of a cell - for contrast enhancement of bright objects on a dark background and to ensure even illumination across the section. Bottom-hat filtering is performed on another instance of the image to enhance black regions in the image. This is then subtracted from the top-hat filtered image to produce a contrast enhanced image. The autofluorescence signal is subtracted from this enhanced image to reduce residual background fluorescence.

An instance of the enhanced image is binarized to produce a black and white image using Otsu's method, a clustering-based image thresholding method (Otsu 1979), or a calculated global intensity threshold. Otsu's method assumes the image can be represented as a bimodal intensity distribution histogram, with each modal peak defining a class corresponding to foreground or background pixels, and seeks a threshold that maximises inter-class variance.

Object Detection

Objects are enhanced through image filling of holes - groups of pixels that cannot be reached by filling the background from the edge of the image. Morphological opening is performed using a 1 px disk kernel causing erosion of objects smaller than 1 px, thereby removing small protrusions and isthmi from the objects, and dilation of the remaining set. Objects containing fewer than 3 connected pixels are then removed as a final filtering step. The contrast enhanced image is masked using this binary image, preserving only pixels that correspond to non-zero pixels in the binary image.

Centroid Quantification

Putative centroids are discovered using an extended-maxima transform to detect regional maxima in the contrast enhanced image - regions of connected pixels with constant inten-

sity bounded by pixels of lower intensity values. Maxima below a threshold pixel intensity value of 20 are suppressed outputting a binary image. Morphological closing is performed with a 1 px disk kernel to fill in maxima objects.

To produce the image for watershedding, the contrast enhanced image is inverted and subjected to morphological reconstruction such that regional minima are only present wherever either of the two binary image masks, arising from object detection or putative centroid detection, are non-zero.

The watershed transform is applied to the resultant topographical relief. This locates catchment basins in the images through filling basins from the bottom with ‘water’. Different ‘waters’ (arising from disparate basins) are prevented from mixing and their interface locations used to define watershed lines corresponding to cell boundaries enabling quantification of individual and modestly overlapping cells.

Centroid Segmentation

To segment the detected centroids into their corresponding anatomical region, a colour ontology scheme was developed. This scheme was overlaid on the manually registered image for each section to create an index image. All of the detected centroids for a section are plotted on this image and assigned to an anatomical region of interest, dependent on the colour of the corresponding pixel in the index image.

Hemispheric segmentation is implemented by setting a threshold x -value that defines the section midline. Centroids with an x -position below the threshold are assigned to the left hemisphere, and centroids greater than the threshold are assigned to the right hemisphere.

2.5.2 Establishing a Ground-truth Dataset

Registered images were counted in FIJI and cell positions recorded manually as co-ordinate points. Sox14^+ interneurons in the dLGN were counted for each hemisphere.

Chapter 3

Results

3.1 Primary GFP Fluorescence Signal Sufficient For Cell Identifcation

$\text{Sox14}^{\text{gfp}/+}$ mice showed a strong primary fluorescence signal sufficient for identification of individual somas and axon fibres in the 50% downscaled JPEG file used for analysis (Figure 3.1).

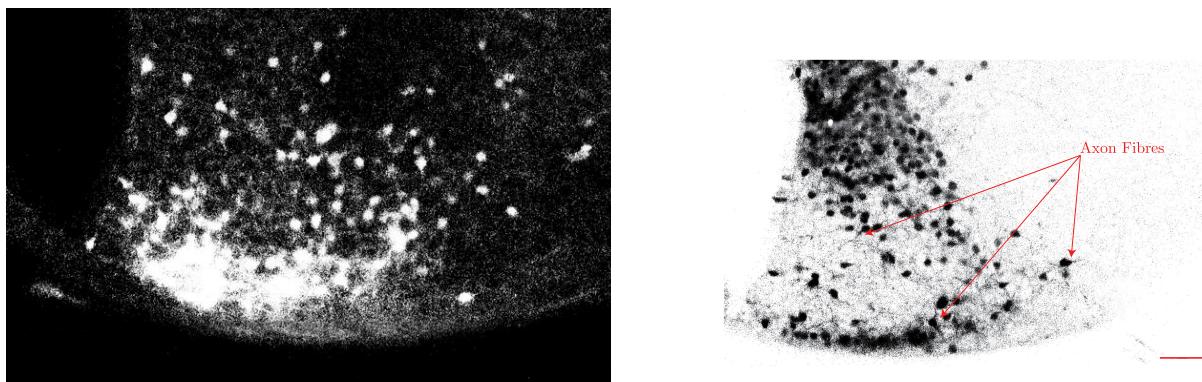


Figure 3.1: *Third ventricle cell distribution.* **Left panel:** Individual cell somas can be discerned in downscaled JPEG file used for analysis outside regions of saturating fluorescence. **Right panel:** Colour scheme inverted, enabling visualisation of axon fibres. **Scale bar:** 100 μm .

3.2 Distribution of Sox14^+ Interneurons

The TissueCyte platform enables complete brain-wide 3D reconstruction of Sox14^+ expression (Figure 3.2). Sox14^+ interneurons are not restricted to the thalamus; expression can be observed in hypothalamic, hippocampal and cerebellar regions. A symmetrical distribution of Sox14^+ interneurons across both hemispheres is observed.

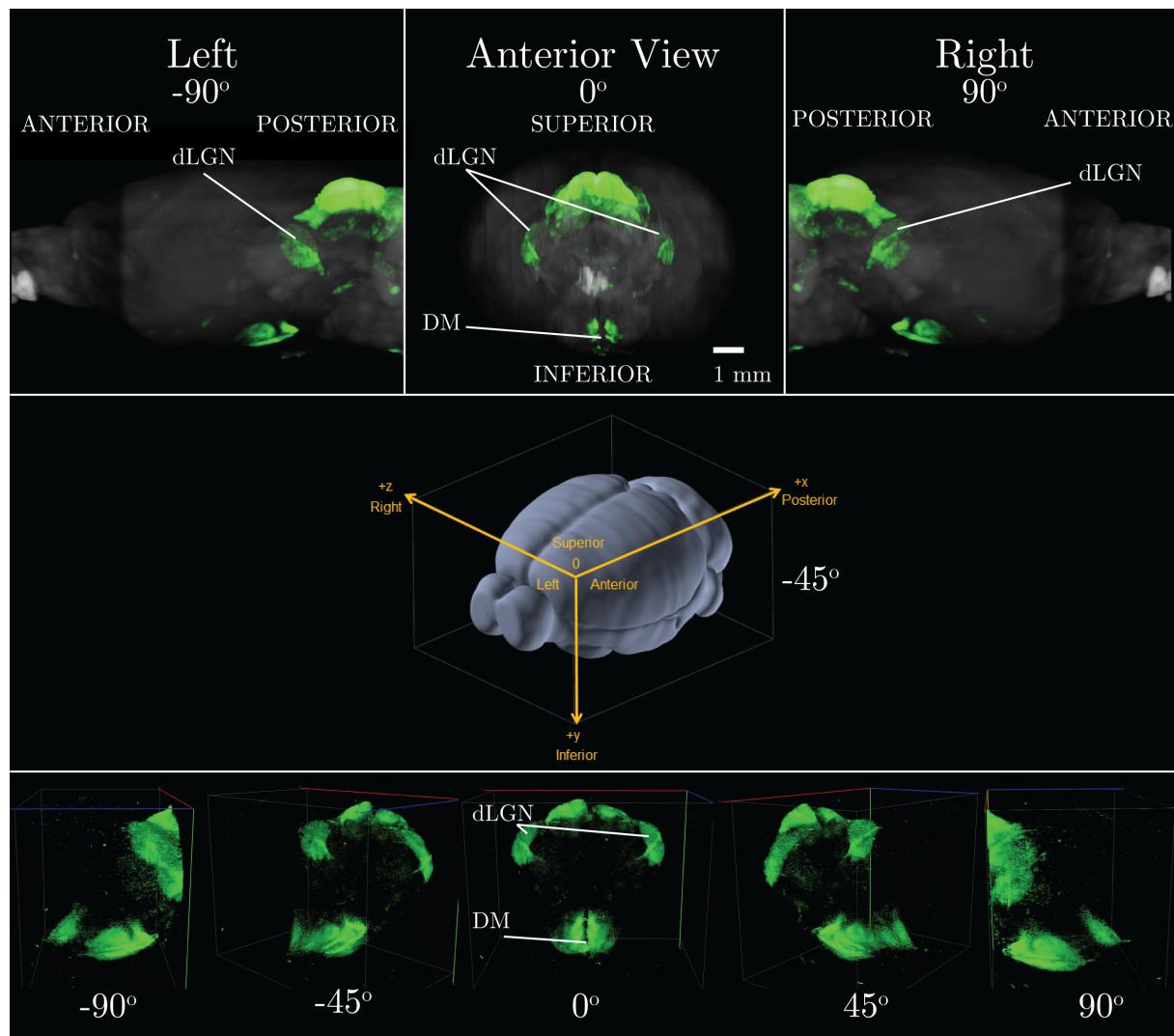


Figure 3.2: 3D Reconstruction of whole $\text{Sox14}^{gfp/+}$ mouse brain using serial two photon tomography. **dLGN** shows symmetrical distribution of $\text{Sox14}^{gfp/+}$ interneuron population across both hemispheres. **DM** and surrounding tissues display strong, saturating fluorescence signal. **Top row:** full brain reconstruction using autofluorescence signal. GFP signal overlaid in green. **Middle panel:** common reference space used to define orientation. Taken from Allen Brain Atlas. **Bottom row:** GFP fluorescence visualisation only. **dLGN:** dorsal lateral geniculate nucleus; **DM** Dorsomedial nucleus of the hypothalamus.

3.3 Image Registration

A $200\ \mu\text{m}$ region of the thalamus was selected for registration and quantification. Due to the optical sectioning performed in $5\ \mu\text{m}$ increments, this corresponded to a region of 40 images. Of these, 10 images with a $20\ \mu\text{m}$ spacing - approximately the diameter of a cell soma - were selected to overcome oversampling of cells. The region analysed is shown in Figure 3.3.

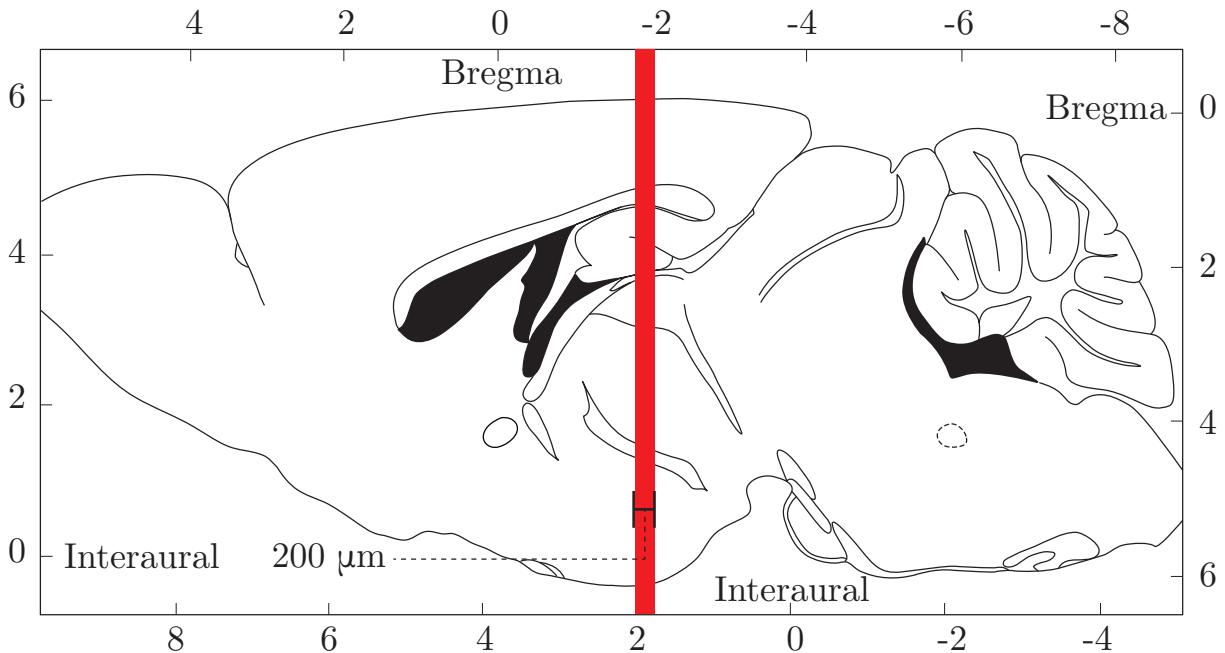


Figure 3.3: *Stereotaxic coordinates (mm) for adult mouse brain. Red line: width indicates $200\ \mu\text{m}$ region where the 10 thalamic sections analysed are derived from. Images analysed are spaced $20\ \mu\text{m}$ apart, the approximate width of a cell soma, to overcome oversampling of cells due to images being taken in $5\ \mu\text{m}$ optical sections. All coordinate measurements relative to bregma. Figure adapted from Franklin and Paxinos 2008.*

3.3.1 Automated Registration

Automated registration of the autofluorescence signal to Allen Reference Atlas produces accurate cortical alignment. However, the Sox14^+ interneuron populations of interest are located in the forebrain. The automated registration procedure produces inaccurate registration of these regions and is hence insufficient to enable accurate anatomical segmentation of detected cells (Figure 3.4).

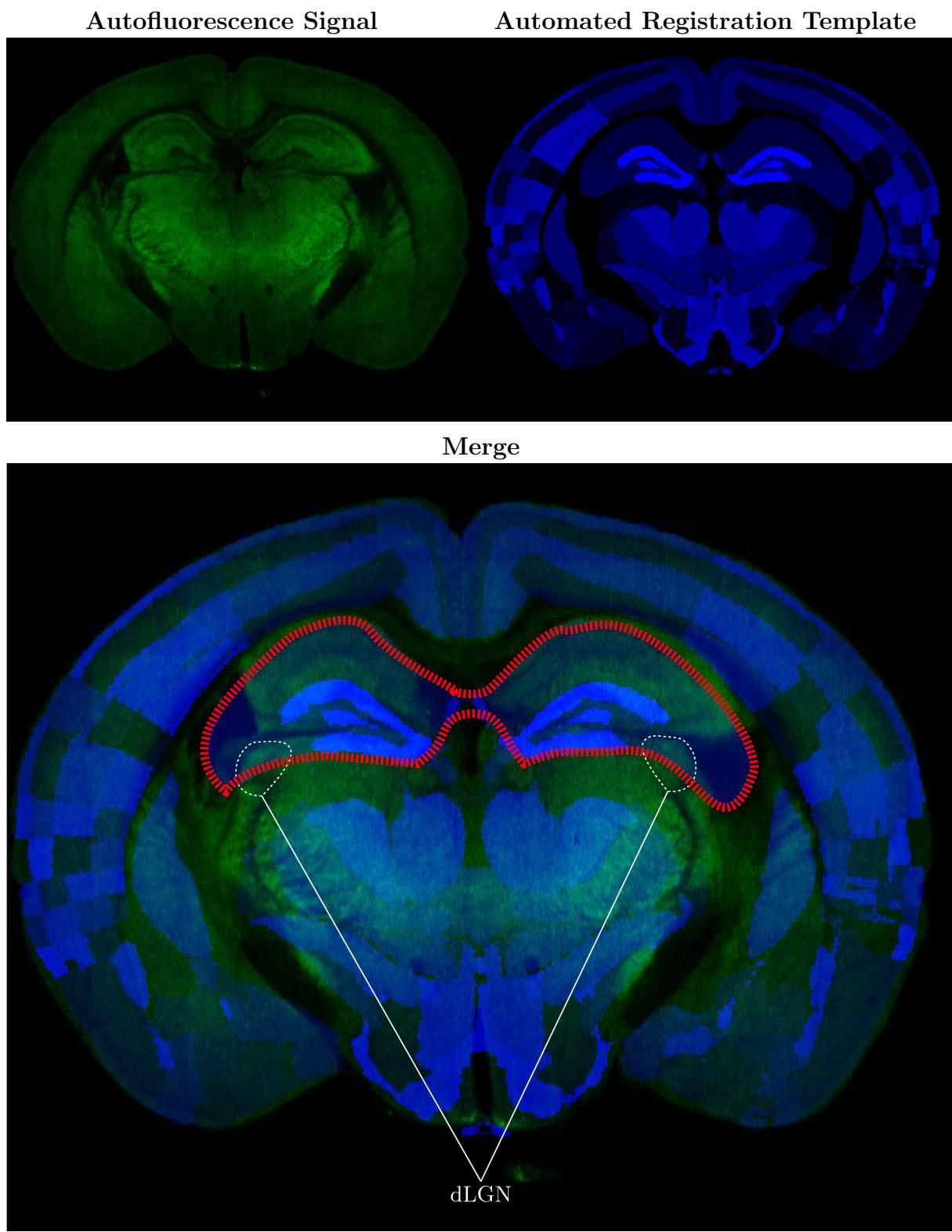


Figure 3.4: *Visualisation of automated registration of coronal sections. Automated registration produces reasonable cortical alignment. Subcortical alignment accuracy is insufficient to allow accurate anatomical segmentation of cellular distribution.* **Green:** autofluorescence intensity signal. **Blue:** registered anatomical atlas; anatomical regions defined using intensity ontology scheme from Allen reference atlas. **Red:** Outline of template hippocampus.

3.3.2 Manual Registration

Manual registration of images yields an improvement in forebrain alignment accuracy. Due to structural disparity between the existing atlas and the P20 mouse brain, the scope of the analysis must be restricted. Alignment of subcortical regions is focussed on the dLGN to yield a set of images with sufficient uniformity in registration to allow downstream analysis. The dLGN was selected due to its sparse distribution of the Sox14gfp/+ interneuron population and its proximal location on the perimeter of the thalamus providing reduced susceptibility to age-related rearrangement of structural boundaries.

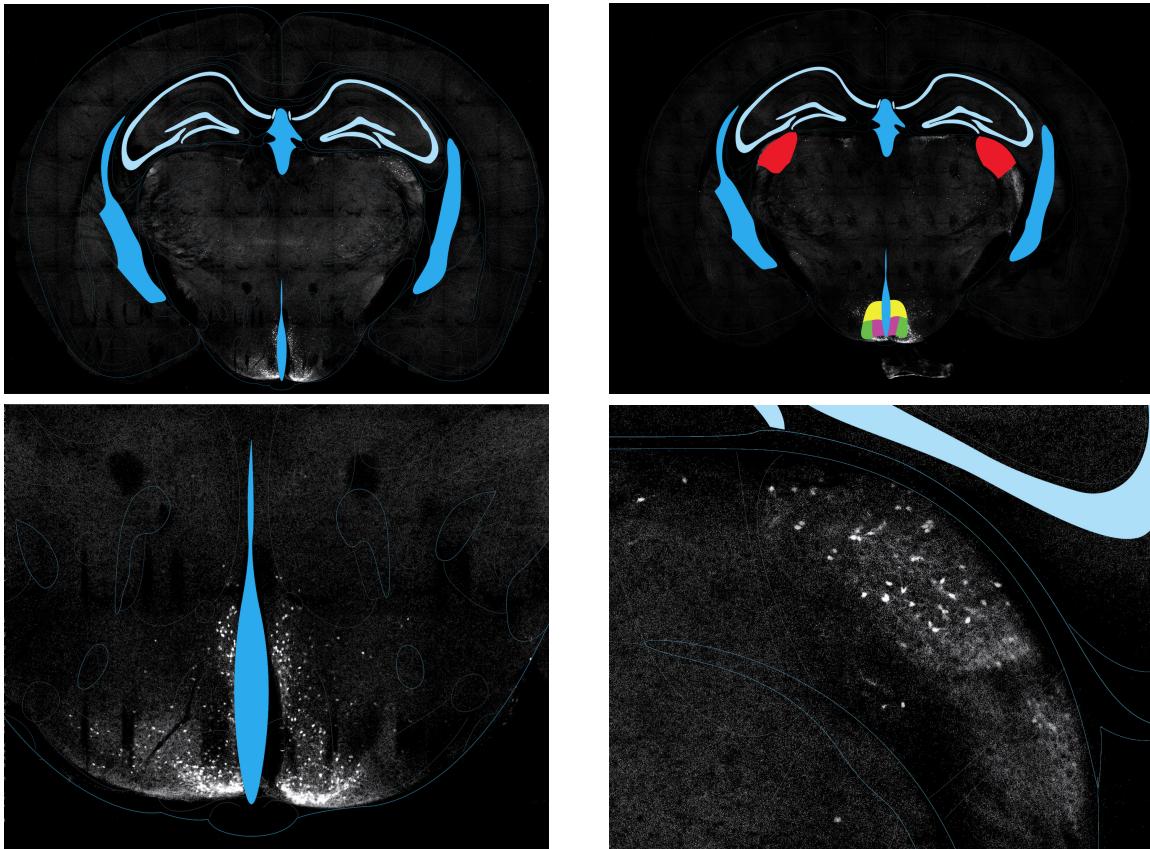


Figure 3.5: *Example of a section image manually registered to Paxinos and Franklin's stereotaxic mouse brain co-ordinates. **Top left:** manually registered image with outlines used for manual counting. Alignment of dLGN prioritised to provide uniform dataset for counting. **Top right:** manually registered image with colour ontology scheme overlaid used as an index for anatomical segmentation of cells. **Bottom left:** close up of ventricular misalignment in the hypothalamus. **Bottom right:** close-up of dLGN (right hemisphere) showing regional segmentation. Blue: ventricles; light blue: granular cell layer of the hippocampus; red: dLGN; yellow: dorsomedial nucleus of hypothalamus; green & purple: arcuate nucleus of hypothalamus (ArcMP & ArcLP, respectively).*

3.4 Evaluation of Automated Cell Counts

3.4.1 Thresholding & Cell Segmentation

Image Intensity Distribution

The intensity distribution of the TissueCyte image dataset produces a unimodal histogram, typical of fluorescence microscopy images selectively illuminating small objects in a large imaging area (Figure 3.6). It was postulated that manually defining a global intensity threshold would produce more accurate counts than conventional Otsu thresholding, a thresholding technique suited for bimodally distributed images.

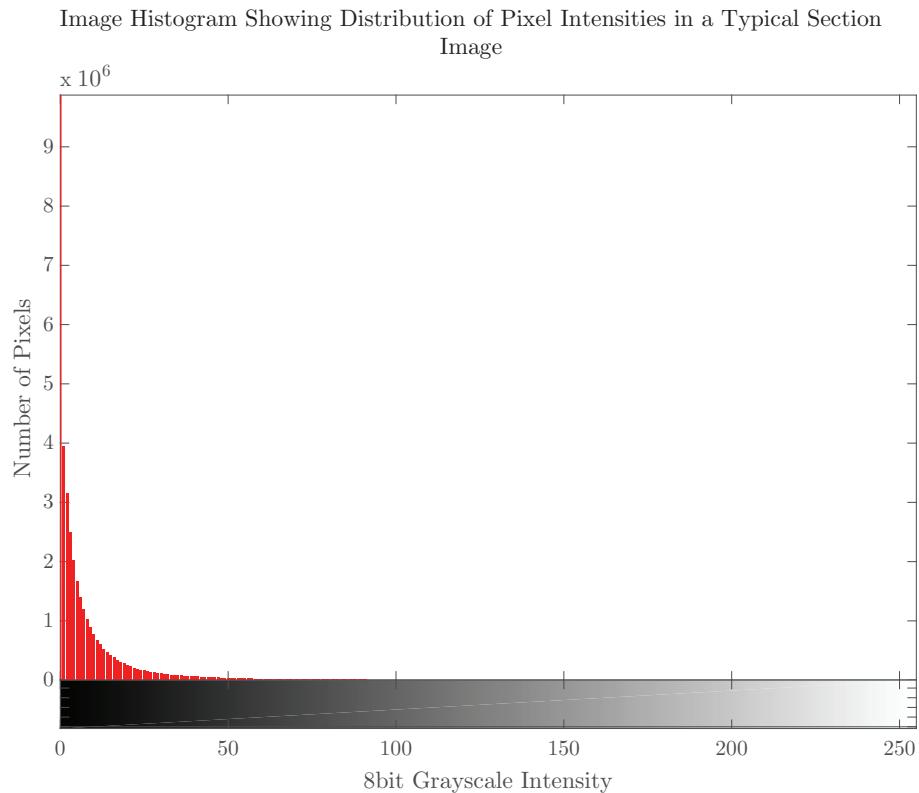


Figure 3.6: *Image histogram showing distribution of pixel intensities in a typical section image. X axis shows pixel intensity on an 8-bit integer scale (0-255) from black to white. Y axis shows the number of pixels $\times 10^6$ in each intensity bin. Section images show a unimodal intensity distribution - typical of fluorescence micrographs due to selective illumination of small objects in a large imaging area. Unimodal distribution prevents use of conventional image segmentation methods that seek to find the optimum threshold between two classes in a bimodal distribution. Instead, a global intensity threshold must be used to segment images.*

Defining a Global Intensity Threshold

In order to confirm whether global intensity thresholding would provide better counting performance, counts were performed on all registered images over a range of threshold values. The total combined counts for the dLGN, as well as the hemispheric counts for the dLGN in each hemisphere across all 10 images are given in Figure 3.7.

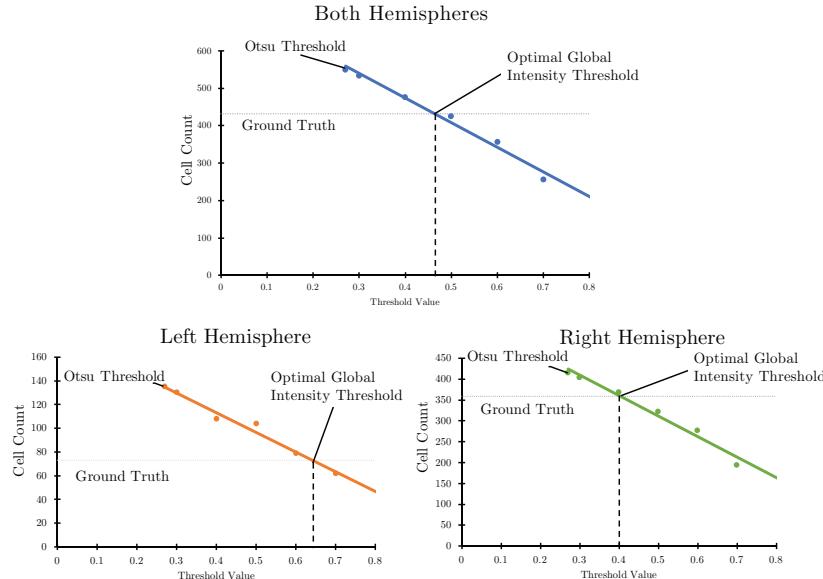


Figure 3.7: Total cell counts for varying binarization thresholds across 10 thalamic sections for dLGN across both hemispheres (**top panel**), left hemisphere (**bottom left panel**), right hemisphere (**bottom right panel**). **Horizontal line** shows total ground truth counts for each condition. **Dashed line** shows optimal threshold derived from intersection of automated and ground truth counts. **Otsu threshold** yielded a binarization threshold of 0.27. Thresholds selected: 0.65 (left hemisphere), and 0.4 (right hemisphere).

Global Intensity Thresholding Improves Algorithm Performance

Following determination of suitable intensity thresholds for each hemisphere, an increase in counting accuracy had to be observed for individual images. As thresholds were calculated using total counts across all 10 images, counts for each image were collected using Otsu thresholding and the calculated global intensity threshold to ensure a consistent improvement in accuracy for each image. A generalised improvement in accuracy is seen for both hemispheres across the images using the calculated thresholds (Figure 3.8).

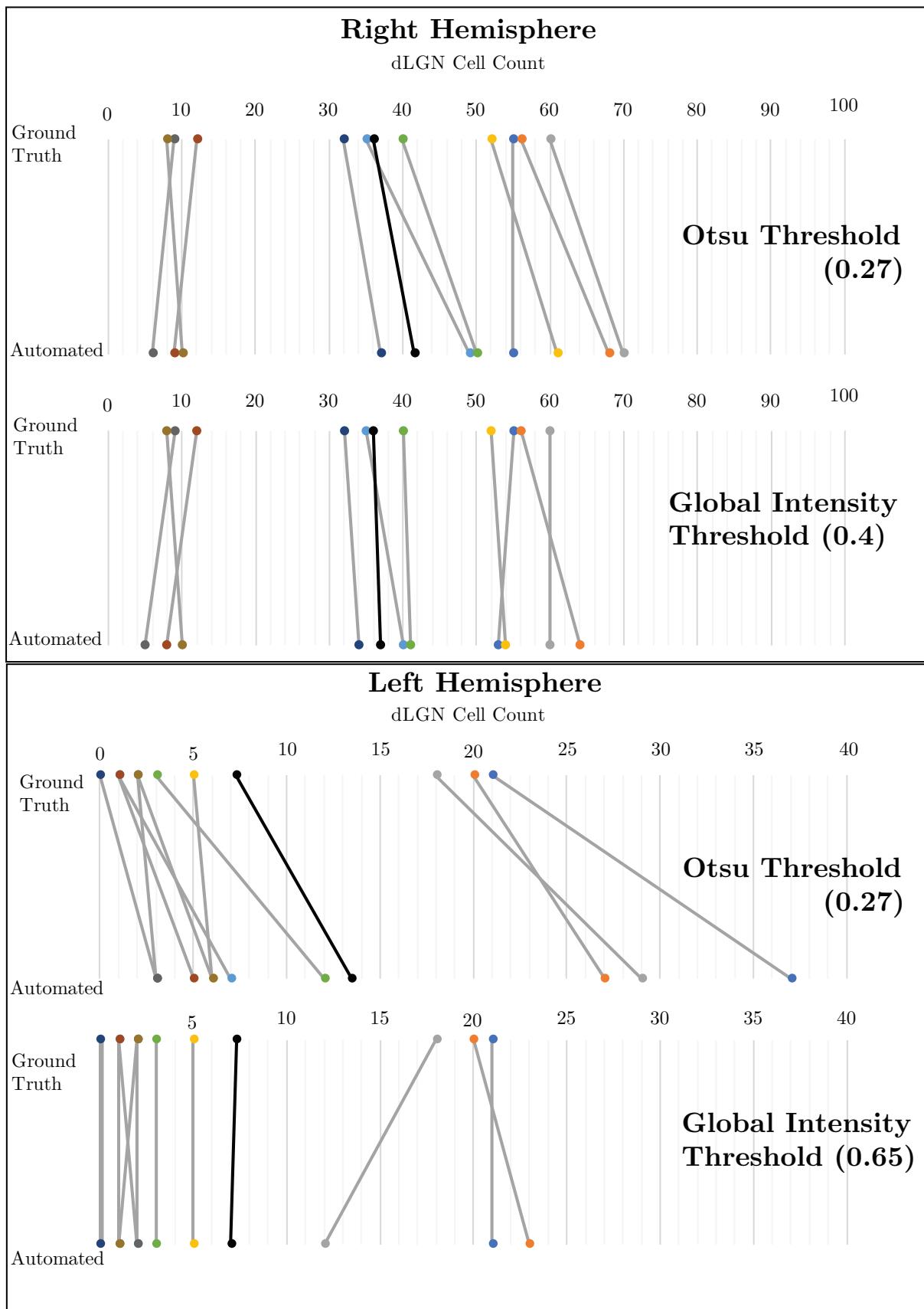


Figure 3.8: *Manually defined global intensity thresholding improves counting algorithm accuracy than conventional bimodal thresholding techniques for images with unimodally distributed pixel intensities. Slopechart plots of manual and automated cell counts for dLGN using Otsu and optimal Global intensity threshold. Each pair of connected points corresponds to the counts obtained from one thalamic section image. Top panel: right hemisphere; lower panel: left hemisphere. Black line shows mean for each plot.*

3.4.2 Centroid Detection

Following evaluation of the optimal binarization threshold, it must be demonstrated that the counts produced actually correspond to cells in the images. Detected centroid positions were plotted on the original image. A comparison of centroid identification in regions of dense and sparse Sox14^+ interneurons populations is given in Figure ??.

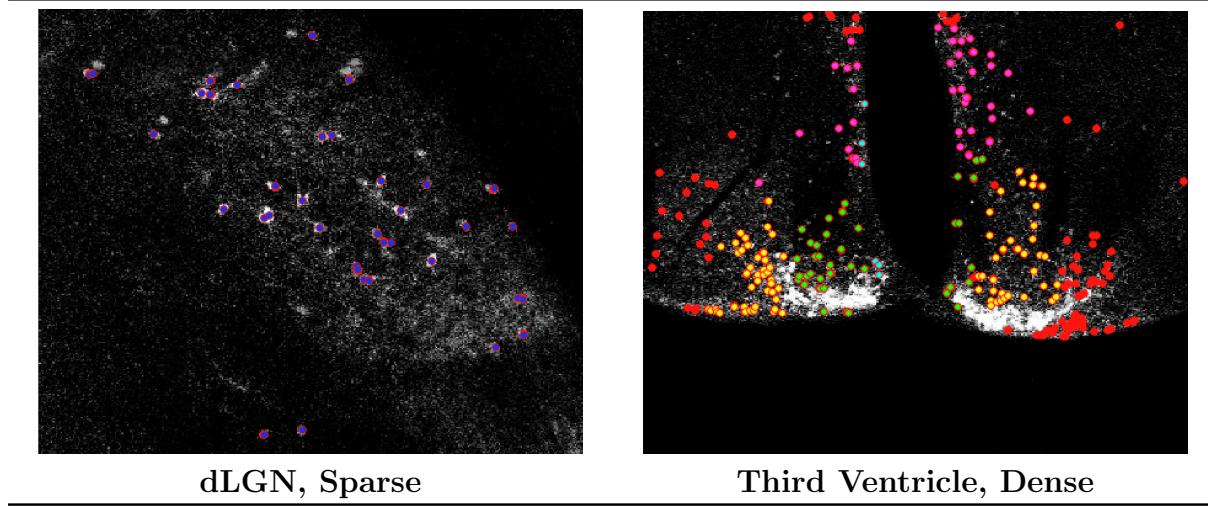


Figure 3.9: *Comparison of centroid detection in dense and sparse regions.* **Sparse regions:** cells are identified by counting algorithm with good accuracy. Small clusters of overcounting can be observed and may be reduced. **Dense regions:** cells cannot be counted accurately due to saturating fluorescence. Individual cell somas cannot be discerned manually or boundaries defined computationally using standard computer vision toolkits. Development of a fluorescence area-based approximation method or a full 3D rendering of each cell volume from data with greater z-resolution may enable estimation of cell counts for these regions. Colour of point corresponds to anatomical region: **Red:** detected centroid; **blue:** dLGN, **yellow:** ArcMP, **green:** ArcLP; **magenta:** DM; **cyan:** ventricule

Sparse Regions

Regions with a sparse population of Sox14^+ interneurons lend themselves to accurate cell soma detection. Small clusters of cells can result in modest overcounting, however improved thresholding greatly reduces this.

Dense Regions

Regions with a dense population of Sox14^+ can not be counted accurately using this method. Where the fluorescence signal is saturating, individual cell somas cannot be accurately discerned for manual or automated counting. As a result, the watershedding method detects these regions of seemingly continuous somas as a single very large cell.

3.4.3 Anatomical Segmentation of Centroids

Anatomical segmentation of centroids detected by watershedding is limited by accuracy of manual registration to anatomical template.

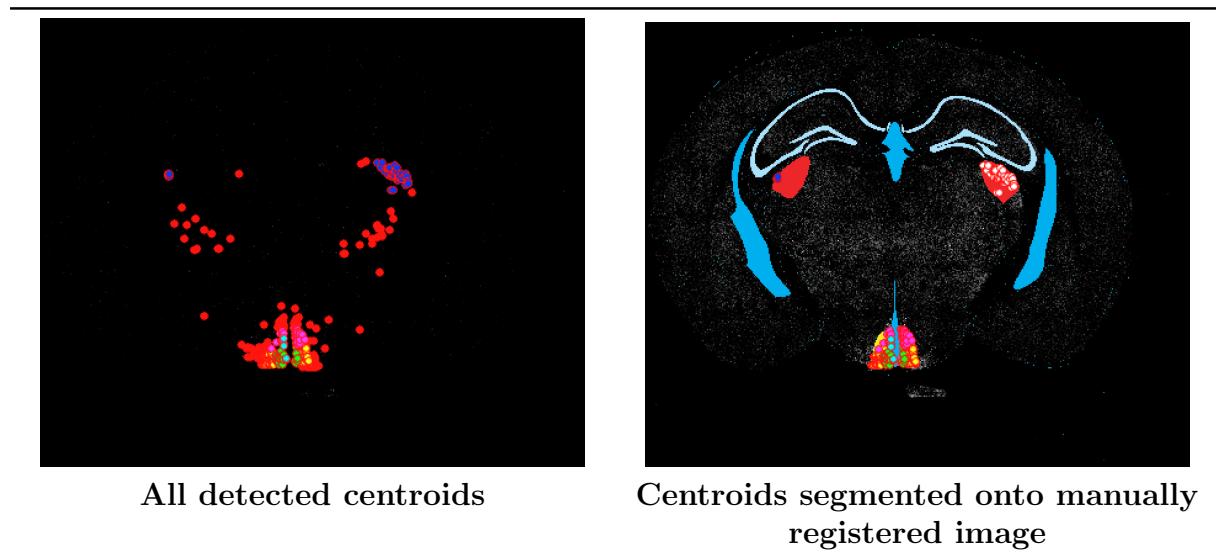


Figure 3.10: *Visualisation of Centroid Segmentation.* **Left panel:** All detected centroids in a section. **Red points:** detected centroid, **blue points:** centroid corresponding to dLGN **Right panel:** All detected centroids corresponding to an anatomical region defined in the colour ontology scheme. **Blue:** ventricles, **red structure:** dLGN, **magenta:** DM, **yellow:** ArcMP. **green:** ArcLP.

Chapter 4

Discussion

4.1 Summary of Thesis Achievements

1. Manually registered forebrain of a P20 mouse to an adult anatomical atlas.
2. Demonstrated quantification of fluorescently-labelled target cells within separate anatomical regions from serial two-photon tomographs of whole mouse brains.
3. Algorithm requires iterative thresholding refinement based on background fluorescence to produce accurate counts.

4.2 Manual Registration

Accurately aligning the stereotaxic co-ordinates to both the third ventricular region and dLGN proved unsuccessful. This difficulty arises from aligning an adult mouse template to images taken from a 20 day old mouse. The emergence of new structures and rearrangement of internal structural boundaries will reduce accuracy of anatomical segmentation. However, we focussed on accurately segmenting and counting the DLGn as this allowed the greatest overall subcortical alignment and the fluorescence signal is saturating surrounding the third ventricle, preventing reliable counting. Furthermore, nuclei deeper

in the thalamus were less likely to be defined accurately due to structural boundary rearrangements. This has no bearing on the counting in terms of cellular identification, which has been shown to be good for sparse regions, but limits the accuracy of assigning detected centroids to an anatomical region.

Verification of the manual segmentation by an external assessor, such as a histology expert is lacking. If this cannot be pursued, an alternative is to apply the counting algorithm to an accurately segmented mouse brain closer in age to the atlas. However, for the study of Sox14⁺ interneurons, this proves difficult as expression of Sox14 is suppressed postnatally. In the case that the algorithm still performs well, the problem of manual registration can be relegated to being specific to this work due to the age of the animal imaged.

4.2.1 Hemispheric Asymmetry of Cell Counting

The symmetrical distribution of Sox14⁺ interneuron population across hemispheres (Figure 3.2) is not reflected in the quantification of both hemispheres; the right hemisphere shows a significantly larger interneuron population across the images analysed. Whilst modest asymmetry in cell numbers may be expected, the explanation for this discrepancy is most likely due to the sample preparation procedure. Embedding the sample requires manual removal of a portion of the cerebellum to provide a flat surface for the sample to stand on. If this cut is not made accurately, or an air bubble is trapped beneath the sample, the sample will be positioned within the agarose block off-center. As a result, the imaging plane will be at an angle that is no longer perpendicular to the central axis (Figure 4.1). This suggests an explanation for the different hemispheric threshold requirements obtained (Figure 3.7); the right hemisphere will have greater background fluorescence due to the underlying Sox14⁺ interneurons and therefore require more stringent thresholding to accurately segment individual cells against the background.

However, this disparity cannot be directly attributed to errors in sample preparation from the 200 μm section of images alone. The observed asymmetry may be an accurate

reflection of the distribution of Sox14⁺ interneurons in this region. As a result, I analysed the whole image stack and determined that there is a consistent lag of ~150 μm in the appearance of structures between hemispheres. Examples of this are given in the appendix.

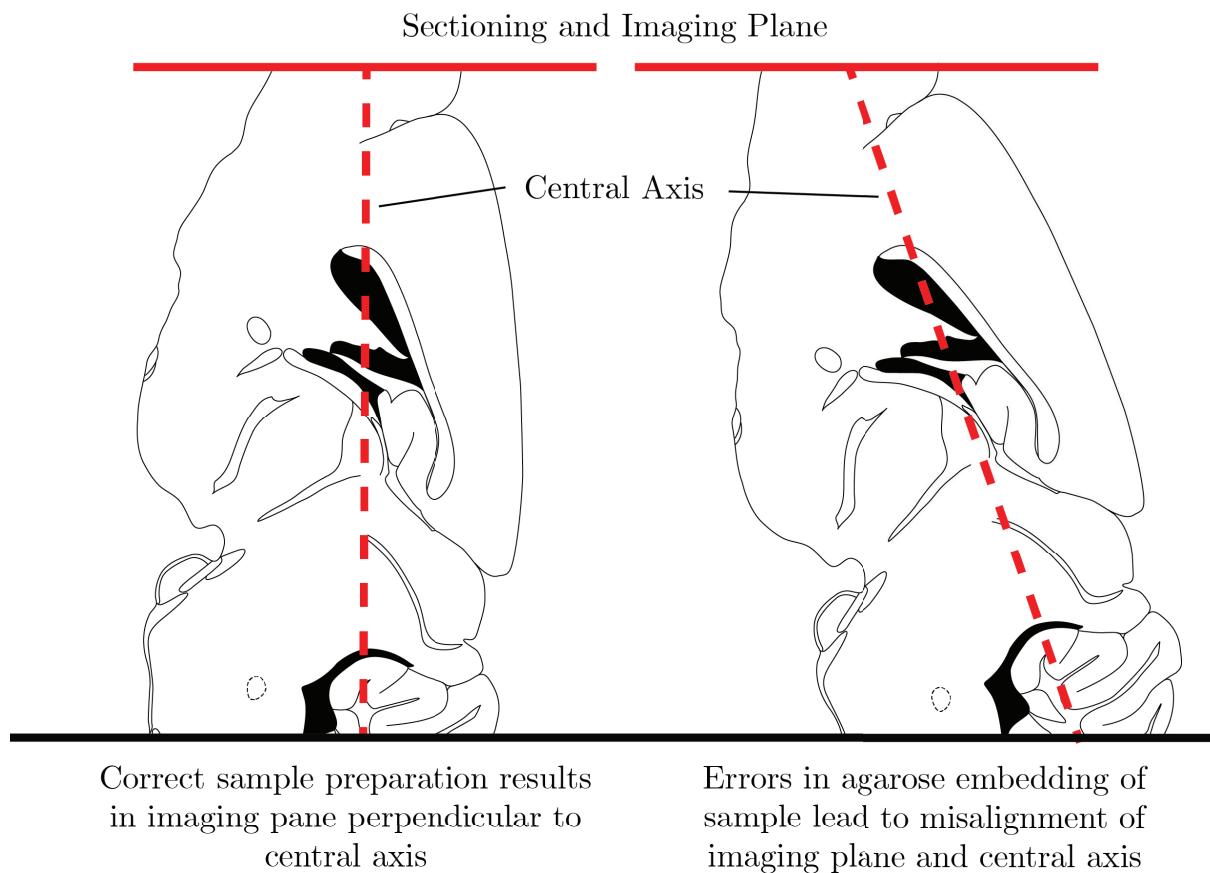


Figure 4.1: *Incorrect sample preparation results in imaging plane no longer being perpendicular to central axis. Misalignment of imaging and central axes likely result in the need for disparate hemispheric binarization thresholds due to the underlying cellular populations providing unequal fluorescent backgrounds. Registration of young mouse brain may be possible with true coronal sections aligned with the central axis. Alignment problem can be overcome by reslicing of data on a tilt, however this is limited by z-resolution to enable accurate interpolation of images between layers.*

It should be noted that imaging off-axis may also provide an explanation for the difficulty in automated registration of the young mouse brain to the adult atlas. This may in fact be possible but the off-axis imaging does not produce the true coronal sections required for accurate registration. Despite this, previous work in the group has shown difficulty in automated registration of images of young mouse brains without such a strong mis-

alignment (data not shown). This suggests age-related structural disparity may still be a limitation of registration of young mouse brains to present anatomical atlases. In any case, manual removal of the cerebellum for embedding provides a potential source of error in the imaging procedure and is a limitation on the technique considering the precision required.

A possible solution to this alignment problem is to reslice the data on a tilt so as to interpolate slices aligned with the central axis of brain rather than analysing slices aligned with the imaging axis. However, the validity of this is entirely dependent on the z -resolution of the dataset to enable interpolation of gaps between image layers.

4.3 Algorithm Improvement

Despite showing good cell identification in sparse regions, the counting algorithm suffers from overcounting of cells in some images. This overcounting has two linked principal sources: the thresholding and binary segmentation of the images, and localised overcounting of cell clusters. These clusters are often found on fluorescence artefacts in the image that do not conform to typical cellular morphology in regions with significant background fluorescence, preventing accurate cellular segmentation.

In order to refine the parameters to enable accurate counting, the ground truth counts have to be collected manually. Cells in sparse regions can be counted with sufficient ease without the use of this algorithm, and cells in dense regions cannot be accurately counted manually or using this algorithm. In its current iteration, the algorithm requires significant refinement to increase global accuracy and quantify dense cellular populations.

4.3.1 Thresholding

The largest room for improvement in counting accuracy lies in the thresholding of the images. This will allow for increased identification of currently unidentified cells and prevent clusters of overcounting through more accurate segmentation. Furthermore, the optimal thresholds were defined iteratively for each hemisphere thereby limiting throughput.

Due to the unimodal intensity distribution of fluorescence microscopy images, a number of traditional segmentation algorithms fail. This unimodality is often inherent in fluorescence micrographs; as images are of selectively illuminated objects, which typically do not comprise the majority of the image, against a large background.

Three possibilities exist to circumvent this problem: develop a custom thresholding algorithm, possibly at the cost of universality to other imaging platforms and increased time complexity; take advantage of the *TissueCyte* system multispectral imaging capabilities by labelling cells of interest with a combination of fluorophores enabling a consensus of cellular location to be established; use another imaging system, such as an improved fMOST platform, capable of avoiding background fluorescence.

Thresholding Algorithms for Unimodal Fluorescence Micrographs

Previous work on establishing a suitable binarization threshold for fluorescence micrographs of cells has yielded little improvement. In fact, a comparison of segmentation algorithms deployed on fluorescence micrographs of mouse fibroblast cells demonstrated that no segmentation algorithm could consistently out-perform manual segmentation (Dima et al. 2011). Thus, the required counting resolution of this tool must be considered. If manual adjustments to thresholding are required to produce accurate counts, perhaps this tool is more useful as an exploratory analytical tool in its current iteration. Furthermore, many of the thresholding algorithms tested greatly increase the time complexity of the counting algorithm.

4.3.2 Overcounting of Clusters

Deploying a rule within the algorithm describing a minimum spacing between centroids derived from the width of a typical cell soma may allow less stringent thresholding. Such an addition could enable the use of a single global intensity threshold in counting multiple anatomical regions.

4.3.3 Alternative Approaches to Counting Dense Regions

Area Approximation

Counts for dense regions could be approximated using a method that determines the area of fluorescent objects in each image. However, this approximation would again be very dependent on determining a suitable threshold as noise and background fluorescence will contribute significantly, but unevenly across each image. Furthermore, such a method assumes a degree of uniformity in cell size which may not hold. Such a method could be validated using the counting tool developed here to establish a linear relationship between fluorescence area and cell number for sparse regions, which could then be applied to previously uncountable dense regions.

3D Volume Reconstruction of Cells

3D reconstruction of each cell would likely yield the most accurate method of counting by defining a minimum and maximum volumes that an object can occupy in order to be considered a cell. Furthermore, morphological restrictions could be applied to further improve accuracy of detection. This method would, however, greatly increase imaging time as greater z -resolution is required to produce accurate reconstructions. Computational time would be greatly increased as well as such an algorithm would have a much greater time complexity than the algorithm presented here.

4.4 Applications and Further Work

The significance of this work to the fields of neurodegeneration and healthy aging lies in the ability to study cellular population dynamics. The benefit, however, is predominantly based on the quantitative output of the algorithm describing cell numbers. Considering that this method has the fundamental ability to identify sparse populations of fluorescently-labelled objects from an image, this work has both quantitative and qualitative applications highly relevant to the field of connectomics.

Connectomics seeks to construct a representation of all synaptic connections in a neural system. Such a representation can either be a 3D reconstruction of the tissue volume or a graph, where nodes represent neurons and edges synapses. Owing to the complexity and density of connections within neural tissue, only the 302 neuron and 7000 synapse strong connectome of *C. elegans* has been completed (White et al. 1986). Expanding beyond the worm is a significant challenge due to issues of scale and non-stereotyped neuronal connections.

4.4.1 Distributed Validation

Following validation of the counting algorithm for Sox14⁺ interneurons in the thalamus, the next steps are to validate the robustness of the algorithm for cell types with disparate morphologies. Furthermore, this validation does not have to be limited to murine or even brain tissues as the TissueCyte system is capable of imaging intact rodent organs or sections of human tissue, provided there exists an anatomical atlas for registration.

4.4.2 mGRASP Synaptic Counting

Mammalian GFP Reconstitution Across Synaptic Partners (mGRASP) involves expression of split GFP carrier fragments at synaptic terminals, enabling reconstitution of the

GFP signal at synapses (Kim et al. 2012). Theoretically, the counting method developed here can be used in conjunction with mGRASP to count sparsely labelled synaptic connections. This would certainly require parameter optimization to produce reliable identification and subsequent quantification of synapses as the fluorescence signal will be weaker and fluorescent objects smaller. As the TissueCyte is capable of multispectral image acquisition, mGRASP can be used in conjunction with additional genetic tools to express a second fluorophore, for instance mCherry, in target neurons. Thus, a target cell-type can be imaged with the the positions of its synaptic connections apparent.

4.4.3 Rabies Virus Tracing

Glycoprotein (G)-deleted rabies virus (RVdG) is a viral tracer capable of retrograde trans-synaptic transmission. The use of such a tracer in conjunction with the TissueCyte system enables neuronal circuit reconstruction. Furthermore, the question of quantifying the number of synaptic partners, a number that varies hugely depending on the cell-type in question, can be addressed. Elucidation of circuit connectivity rules are vital to fully understanding network function.

Returning to Cajal’s cerebellar Purkinje cell (Figure 1.1), the connectivity rules describing these cells are currently unknown. Purkinje cells are the sole output of motor co-ordination in the cerebellum, receiving inputs from $\sim 200,000$ granule cell via parallel fibres. Granule cells receive their inputs from mossy fibres, however, the rules describing this connectivity are unknown. The question remains whether each Purkinje cell receives input via parallel fibres from 200,000 mossy fibres (dense coding) or a million mossy fibres (sparse coding). The brainbow mouse and geometrical analysis of these cells have given clues that this coding may be a mixture of dense and sparse. Rabies virus tracing of mossy fibres, quantified with an algorithm such as this, coupled with functional recordings, could help to elucidate such cerebellar connectivity rules in an empirical manner.

Appendix A

Appendix

A.1 Coronal Sections Acquired are Off-Axis



Figure A.1: *Visualisation of delayed structure emergence in the olfactory bulb (anterior region of the brain). 150 μm in the z-axis between first and last image*

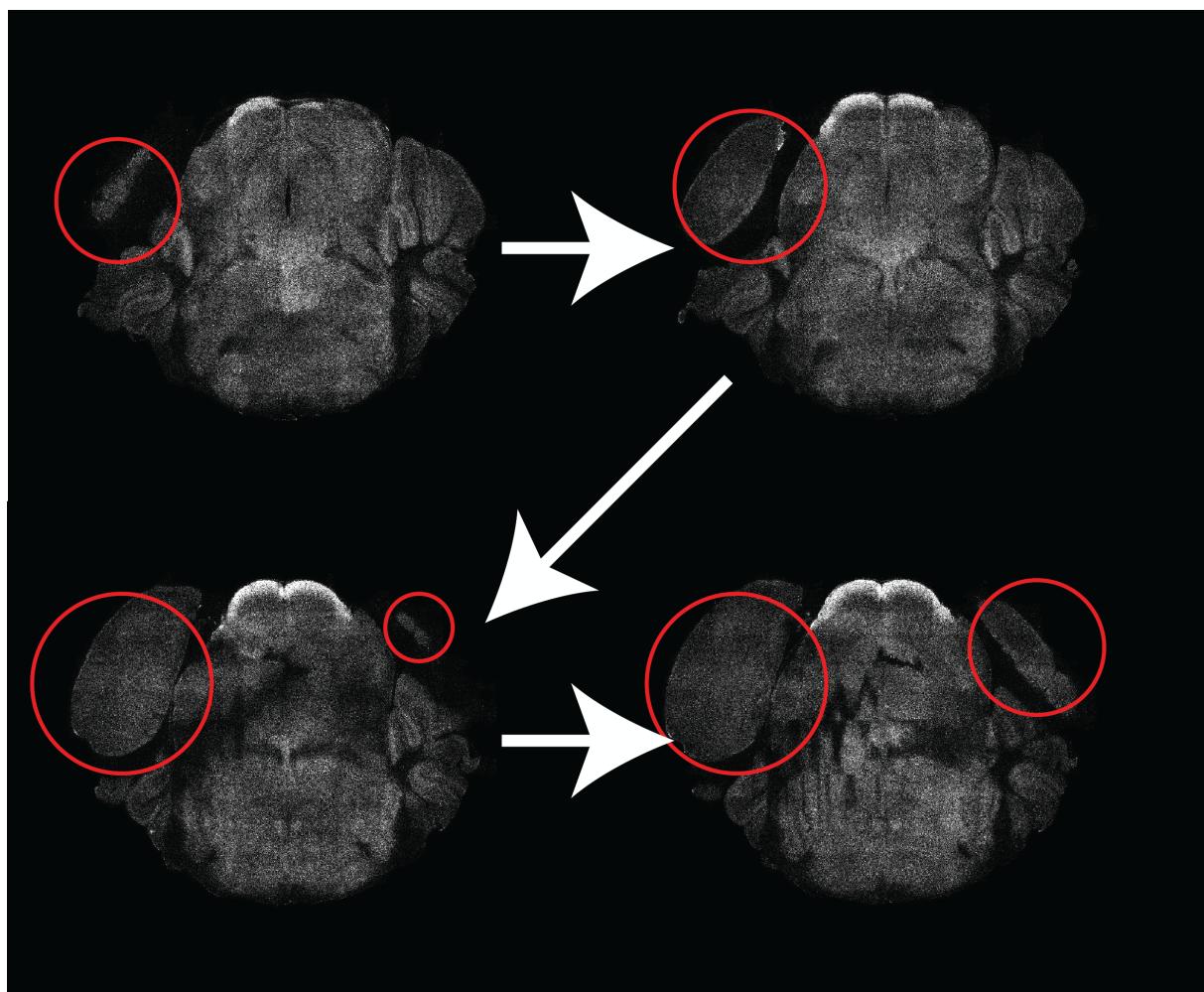


Figure A.2: *Visualisation of delayed cortical emergence in the hindbrain (posterior region of the brain). 150 μm in the z-axis between emergence of structure in each hemisphere*

Appendix B

List of Abbreviations

aMAP - Automated Mouse Atlas Propagation

ArcMP/ArcLP - Arcuate Nucleus of the Hypothalamus

dLGN - Dorsal Lateral Geniculate Nucleus

DM - Dorsomedial Nucleus of the Hypothalamus

EM - Electron Microscopy

FIJI - FIJI Is Just ImageJ

fMOST - Fluorescence Micro-Optical Sectioning Tomography

GFP - Green Fluorescent Protein

JPEG - Joint Photographic Experts Group

LSFM - Light Sheet Fluorescence Microscopy

mGRASP - Mammalian GFP Reconstitution Across Synaptic Partners

NAS - Network Attached Storage

P20 - Postnatal Day 20

PBS - Phosphate Buffered Saline

PFA - Paraformaldehyde

PMT - Photomultiplier Tube

RVdG - Glycoprotein (G)-deleted Rabies Virus

SSA - Serial Section Analysis

STP - Serial Two-Photon Tomography

TIFF - Tagged Image File Format

Appendix C

Glossary

ArcMP/ArcLP - Arcuate nucleus of the hypothalamus, clusters of neurons located adjacent to the third ventricle and median eminence.

Allen Brain Atlas - Project to map genomic information to neuroanatomy data to produce a brain-wide map of gene expression funded by the Allen Brain Institute.

Anatomical Segmentation - Process of dividing a tissue into disparate anatomical regions

Atlas - Collection of histological maps used to define anatomical regions.

Axial Resolution - Specifies minimum distance between two objects to be discerned as separate in the axis of the beam.

Binarization - Conversion of a colour or multi-bit grayscale image to a binary black or white image using a threshold value.

Bottom-hat Filtering - Transform to detect dark objects of a given size (determined by a kernel) on a bright background.

Bregma - Location on the skull where the coronal and saggital sutures intersect.

Camera Lucida - Instrument uses a prism to reflect rays of light such that an image can be traced on paper.

Centroid - Putative cell object detected by counting algorithm

Catchment Basin - Local minima produced from a watershed transform. Correspond to a region of high intensity in the original image used to define cellular positions.

Clearing - Electrophoretic removal of lipids from tissue, thereby reducing optical scattering and increasing permeability to antibodies

Connectome - Comprehensive map of connectivity in the brain. Connectomes can be classified as macro, meso and micro in scale.

Disk Kernel - Small matrix representation of a circle with a given radius used to probe digital images.

Diamond Knife - Sharp diamond-edged knife used where a trenchant and long-lasting blade is required.

dLGN - Thalamic relay centre in the visual pathway. Receives major input from the retina via the optic nerve and projects to the occipital lobe.

Fixation - Preservation process designed to prepare tissue for imaging. Prevents decay of sample and provides a degree of structural support through hardening.

FlyEM - Project seeking to produce a complete connectome for *Drosophila melanogaster*. Recently acquired whole EM volume of an adult female fly at synaptic resolution over the course of three years. EM volume dataset will require large concerted effort in order to yield a completed connectome.

Forebrain - Anterior region of the brain containing the thalamus and hypothalamus.

Galvanometer Mirror - Mirror attached to galvanometers used to scan the incident beam across the imaging sample.

Global Intensity Threshold - See: Thresholding

Granule Cell - Receive mossy fibre inputs and parallel fibres synapse onto Purkinje cells in the cerebellar microcircuit

Golgi Stain - Staining method employing potassium dichromate and silver nitrate. Results in selective labelling of neurons in their entirety, enabling visualisation of structural organisation and fine morphological features such as dendrites.

Image Segmentation - Partitioning foreground pixels (objects of interest) from background pixels.

Index Image - Section image registered to an anatomical atlas with an overlaid colour ontology scheme. Provides an index from which detected cells can be assigned to an anatomical region.

Interaural - Plane defining midline between the ears.

Interneuron - Neurons exclusive to the central nervous system that transmit impulses between other neurons. Both local (short axons) and relay (long axon projections) exist.

Light Sheet - Thin beam of light with a large lateral span used for illumination of whole samples at once.

mGRASP - Genetic labelling technique for synaptic identification. Involves expression of split GFP carriers in synaptic partners, resulting in GFP signal reconstitution at synapses.

Microtome - Instrument enabling thin sections of tissue to be cut, typically as part of sample preparation for imaging.

Morphological Closing - Dilation followed by erosion of an object (using the same kernel), resulting in enlarged boundaries of foreground objects.

Morphological Opening - Erosion followed by dilation of an object (using the same kernel), resulting in reduced boundaries of foreground objects.

Mossy Fibre - Primary sensory input into the cerebellum.

Mouse Brain Architecture Project - Project seeking to understand the mouse connectome at mesoscale resolution.

Nucleus (Neuroanatomy) - Cluster of neurons. Neurons within a nucleus typically have similar connectivity and functionality.

Optical Sectioning - Process by which microscopes can be produces images of focal planes within the tissue sample imaged.

Optical Window - Region of electromagnetic spectrum in which light is not strongly absorbed by an object. This window is near infra-red for biological tissues.

Parallel Fibre - Long granule cell axon in the cerebellum synapsing onto Purkinje cells.

PMT - Photomultiplier tube are vacuum tubes used to collect low-level light signals through successive amplification.

Photobleaching - Photochemical alteration of a fluorophore such that it is no longer able to fluoresce.

Photodamage - Photochemical damage to a sample .

Primary Fluorescence - Fluorescence signal directly arising from expression of a fluorophore without the use of fluorescently-tagged antibodies.

Purkinje Cell - Sole output of motor-coordination from the cerebellum

Ramon y Cajal - Spanish neuroanatomist renowned for applying the double impregnation adaption of the Golgi stain to visualise neuronal morphologies and his scientific artistry.

Registration - Process of transforming multiple datasets into a unified coordinate system.

Sox14 - Transcription factor and marker of interneurons in the thalamus.

Section - Thin sheet of tissue or its corresponding image.

Spatial Resolution - Dependent on the number of pixels used to construct the image.

Stereotaxic Co-ordinates - Co-ordinate space used to define anatomical regions for surgical procedures.

Stitching - Process of assembling image tiles to form a mosaic representation of the entire imaging area.

Thalamus - Centrally located brain structure that relays sensory information (e.g. from the sensory and motor cortices) to the cerebral cortex.

Thresholding - An arbitrary value used to define pixel intensities that correspond to black or white during binarization. Threshold can be expressed as a fraction of intensity space (0-1) or absolute to the intensity/colour scheme used (0-255).

Time Complexity - Time taken for an algorithm to compute the output as a function of the input size.

Top-hat Filtering - Transform to detect bright objects of a given size (determined by a kernel) on a dark background.

Voxel - Each of an array of elements of volume that constitute a notional three-dimensional space in computational image representation.

Watershed - Ridge separating two catchment basins arising from a Watershed Transform.

Watershed Transform - Algorithm that treats an image like a topographic map to define local minima (catchment basins) and watershed ridge lines.

Z-registration - Process of aligning serial images in the Z-dimension to produce a stack of images that form a digital representation of the object imaged.

Bibliography

- [1] Bohland, Jason W. et al. “A Proposal for a Coordinated Effort for the Determination of Brainwide Neuroanatomical Connectivity in Model Organisms at a Mesoscopic Scale”. In: *PLOS Computational Biology* 5.3 (Mar. 2009), pp. 1–9. DOI: [10.1371/journal.pcbi.1000334](https://doi.org/10.1371/journal.pcbi.1000334).
- [2] Dima, Alden A. et al. “Comparison of segmentation algorithms for fluorescence microscopy images of cells”. In: *Cytometry Part A* 79A.7 (2011), pp. 545–559. ISSN: 1552-4930. DOI: [10.1002/cyto.a.21079](https://doi.org/10.1002/cyto.a.21079).
- [3] Duan, Huiling et al. “Age-related Dendritic and Spine Changes in Corticocortically Projecting Neurons in Macaque Monkeys”. In: *Cerebral Cortex* 13.9 (2003), p. 950. DOI: [10.1093/cercor/13.9.950](https://doi.org/10.1093/cercor/13.9.950).
- [4] Franklin, K.B.J. and Paxinos, G. *The Mouse Brain in Stereotaxic Coordinates*. Academic Press, 2008. ISBN: 9780123694607.
- [5] Gage, Gregory J., Kipke, Daryl R., and Shain, William. “Whole Animal Perfusion Fixation for Rodents”. In: 65 (2012), e3564. ISSN: 1940-087X. DOI: [10.3791/3564](https://doi.org/10.3791/3564). URL: <https://www.jove.com/video/3564>.
- [6] Gómez-Isla, Teresa et al. “Profound Loss of Layer II Entorhinal Cortex Neurons Occurs in Very Mild Alzheimer’s Disease”. In: *Journal of Neuroscience* 16.14 (1996), pp. 4491–4500. ISSN: 0270-6474.
- [7] Kim, Jinhyun et al. “mGRASP enables mapping mammalian synaptic connectivity with light microscopy”. In: *Nature Methods* 9.1 (Jan. 2012). Article, pp. 96–102. ISSN: 1548-7091. DOI: [10.1038/nmeth.1784](https://doi.org/10.1038/nmeth.1784).

- [8] Niedworok, Christian J. et al. “aMAP is a validated pipeline for registration and segmentation of high-resolution mouse brain data”. In: *Nature Communications* 7 (July 2016). Article, DOI: 10.1038/ncomms11879.
- [9] Osten, Pavel and Margrie, Troy W. “Mapping brain circuitry with a light microscope”. In: *Nature Methods* 10.6 (June 2013). Review, pp. 515–523. ISSN: 1548-7091. DOI: 10.1038/nmeth.2477.
- [10] Otsu, Noboyuki. “A Threshold Selection Method from Gray-Level Histograms”. In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (Jan. 1979), pp. 62–66. ISSN: 0018-9472. DOI: 10.1109/TSMC.1979.4310076.
- [11] Ragan, Timothy et al. “Serial two-photon tomography for automated ex vivo mouse brain imaging”. In: *Nature Methods* 9.3 (Mar. 2012), pp. 255–258. ISSN: 1548-7091. DOI: 10.1038/nmeth.1854.
- [12] Sallee, Christopher J. and Russell, David F. “Embedding of Neural Tissue in Agarose or Glyoxyl Agarose for Vibratome Sectioning”. In: *Biotechnic & Histochemistry* 68.6 (1993), pp. 360–368. DOI: 10.3109/10520299309105642.
- [13] Sunkin, Susan M. et al. “Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system”. In: *Nucleic Acids Research* 41.Database issue (Jan. 2013), pp. D996–D1008. ISSN: 0305-1048. DOI: 10.1093/nar/gks1042.
- [14] White, J. G. et al. “The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*”. In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 314.1165 (1986), pp. 1–340. ISSN: 0080-4622. DOI: 10.1098/rstb.1986.0056.
- [15] Zheng, Zhihao et al. “A Complete Electron Microscopy Volume Of The Brain Of Adult *Drosophila melanogaster*”. In: *bioRxiv* (2017). DOI: 10.1101/140905.