

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

An Image Restoration Tool to Map Neuronal and Contractile Activity in the Intestinal Wall

Bárbara Nogueira da Silva



FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Mestrado em Bioengenharia

Supervisor: Pieter Vanden Berghe

Co-Supervisor: Paulo Aguiar

July 28, 2024

An Image Restoration Tool to Map Neuronal and Contractile Activity in the Intestinal Wall

Bárbara Nogueira da Silva

Mestrado em Bioengenharia

July 28, 2024

Abstract

The gastrointestinal tract is one of the most complex and dynamic systems in the human body, with its own independent nervous system, known as the Enteric Nervous System (ENS). It is responsible for digestion, nutrient uptake and interaction with the immune system. Through imaging techniques, we can obtain structural and temporal information on the ENS to better understand its function. However, current methods still have limitations when dealing with gut tissue due to its contraction that creates several recording artefacts. Current experimental approaches try to chemically or mechanically limit the movement, with potential loss of valuable information.

In this thesis, we aimed to develop a tool that aids in the analysis of recordings from live moving gut to extract as much information as possible. Such a tool could help research about the physiological and behavioral characteristics of this intricate system. To this goal, we present an analysis pipeline that incorporates a feature-based stitching algorithm to create a map of the moving tissue. These maps not only reflect structural information but also allow for the stabilization of the recording to extract temporal information such as contractile or neuronal activity.

According to the challenges found in recordings, adaptations to achieve a more general pipeline with parameters that can be tuned for specific data. Additional modules were developed to use before the stitching to improve its accuracy and after the stitching for extra results, to expand the amount of information extracted. The main pre-stitching modules include treating noisy low contrast images, warping to correct the distortions and clustering to upgrade the stitching process into an iterative one, making it more efficient.

Different datasets were collected experimentally with an increasing degree of complexity to test the adaptability of the proposed pipeline. Firstly, by automating image acquisition through scripts written for controlling a fluorescence widefield microscope, recordings with pre-defined movements on a fixed sample were created to have a real ground truth to compare with the feature-based result. Secondly, live moving tissue was recorded to test how warping can improve misalignment in stitching caused by distortion. Thirdly, recordings with activity were artificially distorted and then corrected through the pipeline to ensure that the activity measurements were reliable after processing. Finally, a highly complex example of intravital recordings collected through an abdominal window is used to demonstrate how the tool can help analyze these intricate datasets.

As results we present both a perceptual and mathematical quality assessment of the stitched maps. For simpler cases such as the fixed samples, the scores against ground truth data showed high success (99% pixel match, 79% feature match and 0.91 structural similarity). When working with live moving tissue, the parameters were lower (82% pixel match, 52% feature match and 0.68 similarity) while also showing success in correcting for distortions. With the increasing complexity of the recordings, more challenges were faced and the final maps were not as accurate. Nevertheless, we demonstrate a wide range of potential applications and possible improvements.

Finally, to promote all the functionalities of the developed pipeline, a guided interface was developed as a user friendly implementation to showcase the adaptability of these methods. This interface allows for the exploration of adjustable steps and parameters in the pipeline making it easier to adapt for widely different datasets, while also giving a better visualization of the specific modules' capabilities for a complete user experience.

Acknowledgements

The work developed in this thesis is a culmination of a year spent at KU Leuven in Belgium, an experience I will not forget. First I would like to thank Pieter Vanden Berghe, who welcomed me in his group as a master student and throughout this year taught me so much that goes way beyond science. It was a pleasure to follow your philosophical rants and, while trying not to get lost, learning so much from the wonderful way you think.

I would also like to thank everyone in LENS for letting me be a part of the group, for all the insights that have certainly made this work better and for hearing me complain about Belgian weather at least a hundred times. To Natalia, Milvia, Yi-Ning and Silvia I would like to leave a special thank you for all the warmth and light you brought to those colder days.

Thank you to my local supervisor back in Porto, Paulo Aguiar, for agreeing to accompany me on this journey and for all the support in getting this thesis finished.

To my friends, I have a lot of gratitude. To Sol, thank you for being an essential part of this last year and for all the times you had to listen to me go on and on about my work. By now, I am sure you know this thesis almost as well as I do and all your support made me strive to always do the best I could.

To Diana, Margarida and Hugo, a simple thank you is not enough for having been by my side these last five years. Certainly a lot could be said but from all the ups and downs that this experience has brought me, I know that the highs were only as memorable because they were spent with you and that the lows were overcome because we were together in them.

Lastly, I would like to thank my parents for their unconditional support. Everything that I have achieved, not only this past year but throughout my life, has been a result of the way you have always cared for me and encouraged me to keep trying even when I failed. Thank you for everything I have learned and still learn from you day by day.

Contents

1	Introduction	1
1.1	Context and Motivation	1
1.2	Main Goals	2
1.3	Document Structure	3
2	State of the Art	4
2.1	The Enteric Nervous System	4
2.2	Imaging Techniques	5
2.3	Extracting Information from Calcium Imaging	7
2.4	Stitching Algorithms	9
2.4.1	Registration Methods	10
2.4.2	Blending Methods	16
2.4.3	Relevance of Stitching for the Study of the ENS	16
3	Methodology	18
3.1	Data Acquisition and Experiment Planning	19
3.2	Algorithm Description	21
3.3	Quality Assessment	25
4	Results	26
4.1	Evaluation of Feature Detectors and Matchers	26
4.2	Added Modules	27
4.2.1	Pre-Processing	27
4.2.2	Warping	29
4.2.3	Clustering	30
4.3	Final Pipeline	31
4.4	Pipeline Validation: Artificial Recordings	32
4.5	Pipeline Validation: Controlled Movement on Fixed Tissue	34
4.5.1	Reoccurring Frames/Repetitive Movements	34
4.5.2	Out of Focus Frames	36
4.5.3	Matching Immunohistochemical Stainings	38
4.6	Pipeline Validation: Moving Tissue on Ring	39
4.7	Pipeline Validation: Activity Measurements	41
4.8	Pipeline Validation: Intravital Recordings	43
4.9	User Interface	45
5	Discussion and Conclusions	49
5.1	Challenges and Future Work	50

List of Figures

2.1	General structural organization of the ENS [4].	5
2.2	The three main stitching steps, adapted from [49].	10
2.3	Types of homography transformations.	11
2.4	Classification of registration methods used in stitching algorithms, adapted from [44, 50].	12
2.5	First and second stage of the SIFT algorithm [58] (a) for each octave of scale space the image is convolved with Gaussians and adjacent Gaussians are subtracted to produce the difference of Gaussians (b) extreme values are detected in the difference of Gaussian images.	14
2.6	Third and fourth stage of the SIFT algorithm [58] (a) each keypoint is described by gradient magnitude and orientation in a 16×16 region (b) a descriptor is created from the histograms of 4×4 the sub-regions.	15
2.7	Classification of the blending methods used in stitching algorithms, adapted from [44, 50]	16
3.1	Main goals and implementations steps developed in the thesis.	18
3.2	Experimental outline for collecting recordings of the ENS.	19
3.3	Experiment plan for controlled movement on fixed tissue trials.	19
3.4	Experiment plan for moving tissue in ring trials.	20
3.5	Experiment plan for activity measurements in corrected recordings.	20
3.6	Illustration of misalignment calculation from [82].	25
4.1	Example of artificial changes created for the evaluation of feature detectors and matchers: original frame vs. noisy (Gaussian and Salt&Pepper noise) and distorted versions.	26
4.2	Features matched between the original image and altered versions (in percentage).	27
4.3	Average distance of matched features between the original image and altered versions (in pixels).	27
4.4	Changes in a frame after different pre-processing steps (a) different values for brightness representing value added to pixels (b) different values for contrast representing scaling of pixels (c) filter strength for OpenCV non-local means denoising algorithm.	28
4.5	Difference in matching features (a) Raw data (b) Post pre-processing including denoising (filter=10), increasing brightness ($\alpha=8$) and increasing contrast ($\beta=2$).	28
4.6	Feature matching of two frames with one presenting deformation: initial result through Brute Force matching compared with result after excluding outliers through the RANSAC algorithm. The inliers keypoints allow for a homography transform estimation drawn on the left.	29

4.7	Demonstration of the warping correction process. First, a homography matrix is estimated from two frames considering one as the reference. Afterwards, the same matrix is used to warp the non-reference frame to the same plane as the reference one.	29
4.8	Comparison of the result of aligning two frames before and after warping correction.	29
4.9	Comparison of the result of aligning frames all at once or by an iterative process facilitated by similarity clustering.	30
4.10	Clustering results: number of distinctive maps created in each clustering method.	30
4.11	Clustering results: silhouette score for each clustering method.	31
4.12	Clustering results: time taken per iteration in each clustering method.	31
4.13	Final pipeline flowchart.	32
4.14	Artificial recordings created by cropping a consistent frame on a pre-defined ellipse path. Top row: representation of frames selected (in red) to create the recording. Bottom row: final map resulting from the stitching algorithm.	33
4.15	Match scores for artificial recordings.	33
4.16	Full field of view of fixed myenteric plexus on a fluorescence widefield microscope.	34
4.17	Full field of view of fixed myenteric plexus (grayscale) with specific regions that were imaged for the trials presented below. Square region represents the size of a single frame (with 10x magnification).	34
4.18	Illustrative example of one of the recordings results. (a) Map built from the individual frames. (b) Trajectory/movement programmed for the microscope stage. (c) Region in which the recording was taken. Supplementary material: original recording and trajectory followed on fixed	35
4.19	Two types of ground truth data that can be used to compare with the algorithm result. On top is the ground truth stitching obtained from the microscope coordinates. On the bottom is the full region through which the frames pass by.	35
4.20	Match scores in the different regions of the recordings. Both the matches with the ground truth from the microscope coordinates and the matches with the ground truth from the full region are presented.	35
4.21	Match scores summary for the reoccurring frames trial considering as ground truth the stitching from the microscope coordinates. Multiple iterations represents a recording in which the trajectory was run through multiple times in succession.	36
4.22	Example of maps obtained from recordings on the same region following the same trajectory but varying the focus to reproduce out of focus events. Supplementary material: original recording in focus , out of focus +50μm and out of focus +100μm	36
4.23	Difference in matching features with different levels of focus. Top row: keypoints and matches of a frame with itself, all features are matched. Middle row: keypoints and matches of a frame with the same region in +50 μ m focus. Bottom row: keypoints and matches of a frame with the same region in +100 μ m focus.	37
4.24	Matches scores for the out of focus trial considering as ground truth the in focus stitching from the microscope coordinates. Results are presented for the three regions. The out of focus values represent an average score of the three out of focus trials made in each region.	37
4.25	Illustration of the fixed sample with matching stainings. GFP labels enteric neurons and enteric glial cells, DAPI labels cell nuclei, HuC/D labels neurons and Tuj1 labels nerve fibers.	38
4.26	Maps created from recordings in different stainings: (a) DAPI (b) HuC/D (c) Tuj1.	38
4.27	Match scores for staining trials.	38

4.28	Examples of recordings from moving tissue on a ring. (a) First frame of recording (A: ileum ring, B-E: duodenum ring). (b) Map created from the alignment of original frames and trajectory/movement of the sequential frames (Note that midpanel C includes movement from the tissue and from the stage itself). Supplementary material: original recordings A, B, C, D and E	39
4.29	Match scores for moving tissue on ring, divided into recordings with small x&y shifts and large x&y shifts. Since the movement is not controlled and coordinates are not known, the scores are calculated for each individual frame and averaged for all the frames.	40
4.30	Illustrative example of the result of warping correction. Top row: original frames 12, 20 and 37 (out of 45). Bottom row: corresponding warped frames. The warping presented is sequential, which means every frame is warped against the one before it. Supplementary material: side by side of both recordings.	40
4.31	Comparison of maps created from original frames and from warped frames. The highlighted regions show the visual difference in the misalignment of the non-warped and warped version.	41
4.32	Scores for comparison of non-warped and warped maps averaged over all frames.	41
4.33	Illustrative example of the effect of activity in the feature matching process of the same region. (a) Represents some differences in activity with common sections. (b) Represents a complete difference in activity.	42
4.34	Example of a recording with activity present that was distorted artificially. Top row: frames 50 and 100 (out of 200) in the distorted recording. Bottom row: frames 50 and 100 (out of 200) in the warped corrected version. The warping correction was achieved using frame 0 as a reference for all subsequent frames. Supplementary material: original recording, distorted version and warped corrected version	42
4.35	Activity registered from the recordings including the original not corrected, original corrected and distorted corrected. (a) ROI1 (b) ROI2	43
4.36	Some examples of the distortion that occurs from the contraction of the gut in intravital recordings. The pair of frames are sequential, the red square represents the homography matrix and the points represent the keypoints identified and matched in both frames as well as their trajectory.	43
4.37	Illustration of the impact of using different modules of the pipeline. From left to right: inputting the 1000 frames with no pre-processing into the alignment algorithm; using pre-processing to denoise and improve brightness and contrast and therefore improve feature detection and matching; downsampling and clustering for a clearer composition.	44
4.38	Examples (A and B) of maps obtained from intravital recordings. (C) shows an example of the projection of frames in the final map. Supplementary material: original recordings A, B and C	44
4.39	Main design and organization of the User Interface.	45
4.40	Main stitching algorithm. (a) Top level window to select options for stitching including the detector, the thresholds and the order of drawing. If print parameters is selected, the final coordinates of the map will appear in the console. (b) View of the interface after stitching, original frames on the left and final map on the right.	46
4.41	Top level windows for pre-stitching modules and their respective options. (a) Pre-processing. (b) Warping Correction. (c) Clustering.	46

- 4.42 Extra results after stitching. (a) Top level window to select which results to present, if projection of frames is chosen, the type of projection can be selected. (b) View of the interface with the extra results which appear on extra tabs on the right. A graph of the x and y shifts is currently in view. 47
- 4.43 Other extra results that can be seen in their respective tabs. (a) Trajectory of the frames drawn in the map. (b) Projection of frames onto the final map, the result is a sequence of images with the same length as the original recording. Current view is the option 'Map background + Frame'. Supplementary material: projection result following other choices 1, 2, 3 or 4. 47
- 4.44 Module to match the final map with another image of a wider region. (a) Top level window with options for matching, a reference must be uploaded and the scale difference selected (1 if they are in the same magnification). (b) Result of matching represented with a rectangle in the region where the map was found. . . 48

List of Tables

4.1 Details on the artificial recordings including the size of ellipse path and singular frame. 33

Abbreviations

BRIEF	Binary Robust Independent Elementary Features
CNS	Central Nervous System
ENS	Enteric Nervous System
FAST	Features from Accelerated Segment Test
FT	Fourier Transform
GECI	Genetically Encoded Ca^{2+} Indicators
GEVI	Genetically Encoded Voltage Indicators
GFP	Green Fluorescence Protein
IHC	Immunohistochemistry
LENS	Laboratory for Enteric Neuroscience
MAP2	Microtubule-associated Protein 2
NCC	Normalized Cross Correlation
NMF	Non-negative Matrix Factorization
OAD	Open Application Development
OpenCV	Open Computer Vision Library
ORB	Oriented FAST and Rotated Brief
RANSAC	Random Sample Consensus
RMSE	Root Mean Square Error
ROI	Region of Interest
SIFT	Scale Invariant Feature Transform
SIQA	Stitching Image Quality Assessment
SSIM	Structural Similar Index Measure
SURF	Speeded Up Robust Features
TUJ1	Class III beta-tubulin

Chapter 1

Introduction

1.1 Context and Motivation

The Enteric Nervous System (ENS) is a complex and diverse system containing neurons and glia embedded in a ganglionated network present in the gastrointestinal wall. It is responsible for the regulation of gut function and does so fairly independently from the Central Nervous System (CNS). The neurons and glia interact with other cell types including epithelium, the interstitial cells of Cajal, muscle cells and the immune system to coordinate gut function, which includes mixing and moving intestinal contents, extracting essential nutrients and energy and protecting against caustic agents and pathogens. The interest in the ENS has been growing for the past decades and methods of studying it are constantly being developed and improved.

Different techniques exist to extract information from biological tissues, based on chemical and biochemical assays to determine the concentrations of ions and bioactive compounds, molecular biology to probe the expression of proteins and sequencing approaches to yield a complete overview of the cell and tissue state at any given time of isolation. However, when it comes to spatial relations in a tissue, microscopy is the technique of choice, as images can be generated with high resolution revealing the ultrastructure (electron microscopy), expression levels (immunohistochemistry) and even the activity (live imaging) of cells in their native and sometimes intact environment.

Especially with the advent of genetically encoded Ca^{2+} indicators, calcium imaging is one of the most common approaches to study neurons and, as such, it is also used to better understand the structure and function of the ENS. Continuous efforts to create better genetically encoded calcium indicators have allowed better visualization of the millions of neurons and glia in the ENS. Therefore, image analysis pipelines are essential to get further information from these images and typically include algorithms for motion correction, source extraction and spike inference.

Although live imaging is an extremely informative approach to revealing actual information about behaving cells, it comes with a number of complications which are inherent to the use of light and the collection of images. The wavelength of light determines the resolution and how deep one can penetrate, and the power of light has to be carefully managed not to kill the tissue,

yet get a sufficient signal in return. Moreover, even when the optical aspects are taken care of, live imaging can be seriously complicated by tissue movement, which makes the interpretation of the images, especially in low signal-to-noise conditions, not straightforward for the currently available analysis tools. Current approaches try to chemically or mechanically limit the movement, with potential loss or even alteration of results, which is not ideal and requires attention to assure that physiologically correct information is extracted. Without any physical or chemical restraints, the elaborate gut movement creates several recording artefacts, which make image analysis tasks highly demanding, or impossible.

Advances have been made to try to correct the movement artefacts in a series of microscopic images, but to the trained eye, it remains clear that still a lot of information from these recordings is lost. Moreover, when using high magnification lenses to collect sufficient light and detail from individual neurons, images are limited to only a small field of view and therefore do not reflect structural information of a larger region. A widely used approach to create these larger fields of view of static images is the use of stitching algorithms, already implemented in several fields ranging from surveillance to medical imaging. Stitching takes individual smaller images and creates a seamless composition that reflects a wide region, commonly called "panomara". With this method, it would be possible to create comprehensive maps from ENS recordings.

As expected, gut movement is undoubtedly the main obstacle to the good performance of stitching algorithms. In its most simple form, a series of spatially sequential images with static scenery and substantial overlap are aligned and arranged in a seamless composition. However, recordings obtained from the moving gut, such as those from intravital experiments, will introduce a plethora of challenges such as noise content, elements suffering deformation, regions that reoccur in front of the camera at different times, cells that become active and inactive during the recording, among many others. These changes, although easily perceived and understood from a researcher's perspective, pose an extreme burden on classically available image analysis tools. The development of an image analysis pipeline able to stitch together frames from different types of live moving gut recordings would majorly facilitate research on the intricate structural and functional characteristics of the ENS.

1.2 Main Goals

In this Master in Science, we address the problem of moving contractile scenery in microscopy recordings. We use fluorescent images recorded from the ENS embedded in a highly contractile tissue, the mouse intestinal wall. From these images we want to extract cellular information even though the cells change intensity and are displaced relative to the image frame as they move along with the muscle contractions that they coordinate.

According to the challenges previously discussed, the main aim is to develop a tool that could aid in the analysis of recordings from the live moving tissue. By improving image analysis techniques to cater to these specific challenges, there is less need to limit the movement of the region we are trying to analyze, which might be very pertinent in answering certain research questions.

Therefore, with a deeper look into an image analysis pipeline that facilitates the extraction of information from these recordings, the secondary goal is to help in reaching relevant scientific conclusions in the study of this intricate system.

To achieve the main goal of developing this tool, smaller analytical goals are laid out. A number of experiments are then defined with the intention of demonstrating the different capabilities of the developed algorithms for different specific cases.

Firstly, we aim to automate the acquisition of recordings from a fixed sample in order to create recordings following a specific trajectory. With this control we can obtain a real quality ground truth to evaluate the feature-based pipeline's accuracy on a stationary case. Secondly, by evolving to non-stationary samples, the goal is to record gut tissue actually contracting and evaluate the tool's ability of dealing with this relevant challenge. Thirdly, since activity quantification is so important in these type of recordings, we seek to assess if the results from our tool allow for neuronal activity measurements in a reliable way. Finally, with more complex datasets, such as those coming from intravital recordings, our final goal is to demonstrate the full capability of this tool and how it can be tuned for each widely different individual case, showing its adaptability.

To wrap up all its functionality, we also aim to create a guided user interface that allows the use of the wide range of tools and adjustable parameters present in the pipeline itself. With this interface, we can present the final algorithm as a ready-to-use application that aids in the analysis of this complex and interesting system.

1.3 Document Structure

This document is structured as follows: Chapter 2 presents a literature review on the ENS, the imaging techniques used to study it and a deeper analysis of stitching algorithms. Chapter 3 showcases the experimental designs used to obtain the data tested in this thesis, a description of the algorithms used in the pipeline and a review of quality assessment methods to evaluate our results. Chapter 4 showcases the most relevant results from this study including the evaluation of the different algorithms in the pipeline as well as a validation of the methods through real data. Finally, Chapter 5 presents a discussion of the results and main conclusions as well as possible implementations.

Chapter 2

State of the Art

2.1 The Enteric Nervous System

The gastrointestinal system is one of the most complex and dynamic and, while it is mainly responsible for digestion and nutrient uptake, it also has an important defensive function as it is continually exposed to stress factors from ingested content and from the diverse microbiota that inhabits the gut lumen. The gut has its own independent nervous system, known as the Enteric Nervous System (ENS), consisting of a meshed network of small groups of glia and neurons with different functions, including sensory, inter and motor neurons [1–3]. Studies on the gut-brain axis, referring to the bidirectional communication between the enteric and the central nervous system, have promising insights into associations between functional gastrointestinal and central nervous system disorders [4, 5].

The interest in gut function and its control by the ENS has tremendously increased over the past decades and extensive research is now ongoing on the ENS' structural and functional organization and its complex cellular and molecular characterization. Schneider et al. [6] cover the better documented ENS functions, such as control of bowel motility, epithelial secretion, and blood flow, as well as its interaction with the immune system and its potential relation to disease. A review published by Fung and Vanden Berghe [1] focuses on the functional circuitry of the ENS and how these circuits detect and process external stimuli. Most recently, Sharkey and Mawe [7] highlighted the latest developments in ENS research that contribute to further understanding of the physiology and pathophysiology of the system.

The ENS consists of a heterogeneous mixture of neurons and glia organized in two ganglionated and interconnected plexus layers, as shown in Figure 2.1. The myenteric plexus is situated between the circular and longitudinal muscle layers and is involved in coordinating gut movement. The submucosal plexus is located beneath the mucosal epithelium, with a role in controlling secretion and blood flow [1, 3]. Extensive connections among ganglia create a largely populated system that presents a complex cellular and molecular architecture. Functional, biochemical, histochemical, and immunohistochemical studies help characterize the different cell subclasses in the ENS

but there are still a lot of unknowns in the way they function and interact with each other, but also with other cells residing in the gut wall, like for instance cells from the immune system.

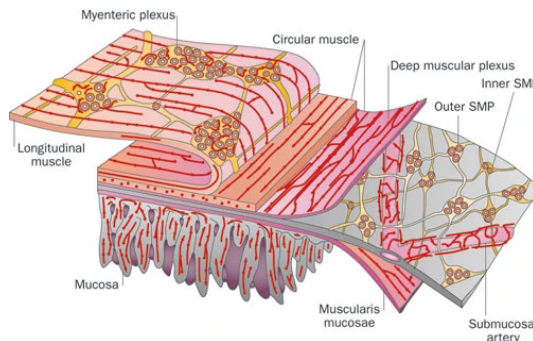


Figure 2.1: General structural organization of the ENS [4].

Traditional methods of investigation are often not adaptable enough when dealing with such a distinct, elaborate and complex system. In many analytical approaches, the spatial information is lost as extracts are made from tissue for further analysis. Using microscopy in fixed or living samples, this information can be retained to investigate the cellular organization and intricate molecular pathways. However, the fact that the contractile gut is constantly moving introduces an extra level of complexity and makes standard functional imaging very challenging. These problems can be mitigated with advances in microscopic imaging, genetic technologies [8] and image analysis.

2.2 Imaging Techniques

Studying neuronal activity is essential to gain insights into how complex nervous systems function, how they develop and adapt over time, how they react to different stimuli and how this relates to behavior and disease. There are numerous ways to study neuronal activity, generally focused on either electrophysiological recordings or imaging, each with advantages and disadvantages. Electrophysiological based methods measure the electrical activity directly from individual or groups of cells providing high temporal resolution, but are limited in spatial information. Optical imaging poses an approach to obtain accurate spatial information but reports spiking activity in an indirect way, which might result in some discrepancies [9–11].

Fluorescence imaging is a widely used tool to monitor chemical changes by a process where a molecule absorbs light and emits it some nanoseconds later at a longer wavelength. In fluorescence microscopy, the emitted light is then detected and used to create digital images, allowing the study of cells and molecules in a more precise and targeted way. In neuroscience, this technique is often used to study the activity of neurons by using fluorescent dyes that bind to or compartmentalize in specific cell components or be influenced by specific cellular events, such as calcium concentration changes.

Calcium signaling plays an important and versatile role as a second messenger system in many different cell types. By tracking the changes in intracellular calcium, cellular activity can be inferred. In neurons, intracellular calcium rises as a consequence of action potential firing, and thus neuronal activity dynamics can be monitored using calcium sensitive fluorescent probes as a proxy. This technique, generally termed calcium imaging, has become increasingly important in research, with thousands of articles published on the topic. Originally, calcium imaging relied on small synthetic molecules that were introduced into the cell via injection or via ester loading. However, one key advancement was made when Green Fluorescent Protein (GFP) and its variants were introduced to genetically target specific cells in living systems [12, 13]. A first generation of these molecules was simply used to label cells (eg. eGFP, YFP), but soon after functional indicators were developed that were able to report on intracellular processes. These include Genetically Encoded Ca^{2+} Indicators (GECIs), which respond to changes in intracellular calcium concentration, and Genetically Encoded Voltage Indicators (GEVIs), which provide information on membrane potential changes [9].

Even though GEVI's yield a direct measurement of neuronal activity, the indirect Ca based methods have a competitive advantage as they rely on the high contrast of intracellular versus extracellular Ca concentrations. Advances are constantly made for engineered GECIs with higher sensitivity. GFP-based GCaMP sensors are one of the most common GECIs in neuroscience, consisting of a genetic fusion of eGFP, calmodulin (CaM) and a peptide sequence M13. First presented in 2001 by Nakai et al. [14] as an improved Ca^{2+} probe, several variants of GCaMP have been developed over the years [13, 15–17] to be either brighter and more sensitive or faster, tracking the changes in Ca^{2+} in a more precise way. These indicators have been essential in a wide range of research, allowing for both *in vivo* and *in vitro* measurements, as well as cell type-specific expression. Recently, ENS specific research showed that GCaMP sensors could provide relevant information on neuronal and glial behavior [18–20]. Compared to other techniques, these calcium indicators provide advantages in selectivity, signal-to-noise improvement and non-invasiveness.

Besides calcium imaging, immunohistochemistry (IHC) techniques have been used for more than three decades in studying the ENS and are still used recurrently [21–23]. IHC offers direct visualization of tissue antigens using labelled antibodies and is known for its versatility, simplicity and affordability making it appealing to several research fields. The downside of IHC is that it can only be applied to chemically fixed, and thus dead, tissues. Regarding the ENS, immunostaining allows the visualization of the complex distribution, morphology, and connectivity of neurons and other cellular components. To identify neurons, the primary targets are neuronal proteins such as class III beta-tubulin (TUJ1), microtubule-associated protein 2 (MAP2), and neurofilament protein. These markers help identify neuronal cell bodies, axons, and dendrites, enabling mapping of the spatial organization of neurons within the enteric ganglia. Additionally, neurotransmitter specific antibodies allow for the identification and characterization of specific neuronal subpopulations. These techniques can be used following a calcium imaging experiment to help identify the identity of the neurons that displayed activity.

In animal models of gastrointestinal diseases, immunostaining can also help report alterations in the ENS associated with certain conditions. By comparing the expression patterns of neuronal and glial markers between healthy and unhealthy tissues, researchers can pinpoint changes in neuronal morphology, neurotransmitter expression, and glial activation that may contribute to gastrointestinal symptoms and pathology.

2.3 Extracting Information from Calcium Imaging

Analysis of calcium imaging sequences involves extracting fluorescence intensity over time and interpreting these signals in the context of cellular activity. The size of calcium imaging datasets is constantly growing, requiring automated, scalable analysis pipelines to extract spatial and temporal information from the field of view. A typical calcium analysis pipeline involves three main steps: motion correction to remove artefacts, identification of the neurons' location and extraction of fluorescent traces and last spike inference to document the neuronal activity [24, 25].

1. **Motion Correction:** Even in a fairly static tissue like the brain, natural movement caused by heartbeat pulsation and blood flow creates artefacts in typical calcium imaging datasets. Depending on the extent of the movement, different methods have been developed to correct for it, trying to reduce the discrepancies between frames and facilitate source extraction. TurboReg [26] is a standard early template-based method, which first demonstrated accuracy in positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) images. It was later implemented as a plugin for ImageJ, a freeware and popular microscopy image analysis platform, for the alignment of an image to a certain target, and thus correcting lateral motion artefacts. Dubbs et al. [27] presented a faster implementation of TurboReg, using a Fourier transform approach.

Two types of motion can be identified: rigid and non-rigid. The motion is characterized as rigid when the whole area is uniformly transformed (such as single translation and/or rotation), or non-rigid when different parts of a single image present different movement profiles. Rigid motion correction usually takes advantage of the cross correlation between two frames to estimate what movement occurred between them [28]. As an alternative to correlation, tracking methods look for the specific position of different structures such as individual cells to assess the movement. However, these methods are often not enough to compensate for more complex movements.

In recent years, research efforts have been put into developing non-rigid motion correction algorithms. Pnevmatikakis and Giovannucci [29] developed NoRMCorre, an online algorithm for piece-wise rigid motion correction that works by splitting the image into overlapping patches to estimate the transformations pixel-wise, followed by upsampling to create a smooth motion field. Adaptations of this method were later proposed, including one for analysis of images recorded from the moving gut Barth et al. [30]. An alternative to these

methods is feature-based motion correction, by finding the most representative structures to estimate motion [31].

The current motion correction methods have the disadvantage of being computationally expensive and therefore not fast enough which can be an obstacle to implementing real-time analysis. Methods with computational improvement, in terms of speed, memory, and accuracy, have been published for single photon [32] and two photon imaging [33].

2. **Source Extraction:** Identifying singular cells is usually achieved by segmentation algorithms to isolate sources for further detailed analysis of the fluorescence changes in time. Segmentation is an extensive topic in itself and presents countless different computational approaches. In calcium imaging pipelines matrix factorization is commonly used as it presents a link between the spatial footprint of each source and the fluorescence activity. Approaches with matrix factorization include Independent Component Analysis [34] and Non-negative Matrix Factorization (NMF) [35–38]. Other approaches include clustering, dictionary learning and active contour models [24].

Instead of including source extraction in an analysis pipeline, there is also the option to manually draw regions of interest (ROIs) on image sets to extract the fluorescence traces, which is still widely used in many different studies such as in Desmet et al. [39].

3. **Spike Inference:** As fluorescence is not a direct measure of the spiking activity, the traces must be deconvolved to obtain approximate information on the spiking times. Berens et al. [40] characterized different spike inference methods, both supervised and unsupervised, showing that complex methods proved to have few advantages over simpler ones. Supervised methods are dependent on labelled data to train neuronal networks and can be subjected to poor training. Unsupervised methods can either be deterministic, such as sparse non-negative deconvolution, or probabilistic [38].

The development of analytical methods in calcium imaging goes hand in hand with the evolution of programming and computational capabilities. As such, image analysis packages have been developed to accommodate users' needs and aid in the advance of scientific research. A table of imaging analysis software for biology and neuroscience with a focus on calcium imaging is available in [41], with an emphasis on algorithms for cell extraction and reconstruction of activity traces, also including general calcium imaging packages. As an example, the CaImAn toolbox (large-scale Calcium Imaging Analysis) [38] offers algorithms for the three mentioned steps with batch processing for large datasets. It includes motion correction based on NoRM-Corre, source extraction using the Constrained Non-negative Matrix Factorization (CNMF) [36], spike deconvolution implemented with the Online Active Set method to Infer Spikes (OASIS) [42] and registration across multiple sessions.

Specifically for the analysis of cells within the ENS, the Gut Analysis Toolbox [43] was developed as a combination of different algorithms, proposed for routine image analysis tasks in ENS research. It takes 2D images of gastrointestinal wholemount preparations as input and provides

semi-automated neuronal and glial distribution analysis, alignment of images and extraction of traces based on classic time series extraction techniques.

2.4 **Stitching Algorithms**

Common calcium image analysis provides extensive and unique insights while still being restricted by the imaging setup. In microscopy, the acquisition of high resolution images is limited to a very small field of view but there is often a need to image over larger areas of tissue for clinical and research purposes. Therefore, to get a better and more reliable representation of a wider region, multiple images can be used together to create a seamless composition. Image stitching (also referred to as image mosaicing) is the technique used to align linked images to generate a single larger one. These stitched images, sometimes called panorama, have a wide range of applications such as in satellite imaging, virtual reality, visual effects, surveillance and medical imaging [44].

Stitching algorithms play an important role in creating these larger fields of view with structural information that otherwise would be difficult to comprehend from single images. There are multiple examples of the use of stitching algorithms in microscopy for different applications. Kose et al. [45] presents an approach for videos obtained with a handheld microscope moved manually over tissues, which creates a sequence of images of overlapping areas of tissue with cellular resolution. Examples of melanoma and non-melanoma skin cancers are used to demonstrate the potential of the algorithms in increasing the diagnostic capability of skin lesions by producing a large field of view, even when the sites have irregular topology. Thus, the use of real time video stitching can be extremely useful in guiding users in clinical procedures such as noninvasive screening and intraoperative consultation [46].

For typical microscopy setups, it is also very common to take advantage of stitching algorithms to produce a complete and comprehensive image of large biological samples in high resolution, by moving the microscope stage through the desired area to create a tiled scan. Since the physical coordinates from the stage are not reliable enough for direct reconstruction of the larger image at the submicrometer scale, tools have been developed to stitch a whole field of view from tiles, both in 2D and 3D, such as the one presented in Preibisch et al. [47], currently available as an ImageJ plugin. A similar approach was taken by Zhao et al. [48] to create an automatic image mosaic method for sequential scanning electron microscopy (SEM) images by computing translations between serial tiles with unknown overlap. These applications are typically referring to stationary samples (e.g. fixed samples that are imaged in separate tiles and then stitched together). In non-stationary cases, in which the sample is moving in front of the camera, the stitching process might generate temporal incoherence where different regions in the stitched image are not reflecting the instance in which they occurred. Depending on the application, both spatial and temporal dynamics can be extremely relevant and therefore the methods have to be careful of not producing discrepancies. Overall, stitching algorithms are recurrent and needed in microscopy applications, leading to a diverse range of research being published based on developments arising from specific biological questions.

A stitching algorithm is composed of three main steps: registration, re-projection and blending, as shown in Figure 2.2. The first step aims to estimate a correspondence between two images and align them to a common global coordinate. Secondly, the images are re-projected onto a larger surface creating the compositing frame. Finally, the warped images are blended to remove any visible artefacts by smoothing the seams, making the final image more visually appealing.

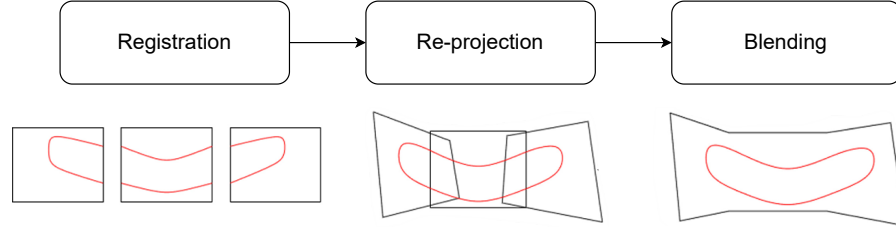


Figure 2.2: The three main stitching steps, adapted from [49].

Several review papers on the current stitching techniques take an in-depth look at the characteristics, mathematical background, main advantages and disadvantages of each one, as well as comparing their performances [44, 50, 51]. Different classifications have been proposed for these algorithms but the most common ones are according to the registration and/or blending methods.

2.4.1 Registration Methods

Registration is the first and possibly most important step of the stitching pipeline and it consists of aligning a set of images into one single coordinate system. It has extensive use in medical and research applications and choosing which approach to use is highly dependent on the type of image, the overlapping extent, the deformations and misalignment between images.

Some authors make the distinction of registration techniques in two groups, geometric and photometric [44, 49, 52]. In geometric registration, a transformation is estimated between two images to align them in the best way possible, presented as a homography matrix, H .

Homography methods are used for applications such as video stabilization for their strengths in estimating motion between two frames while being less affected by local distortions [53].

A 3×3 matrix allows the transformation of a point x from one image into a point x' in another image following the expression in Equation 2.1. Depending on the matrix's properties and degrees of freedom (DoF) - i.e. the number of unknowns that need to be solved for - it can represent several planar transformations such as translation, rotation, scaling, affine and projection. Figure 2.3 and Equations 2.2 - 2.5 showcase the different types of homography matrices and their respective bidimensional transformation.

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (2.1)$$

$$\textbf{Euclidean (3 DoF): rotation + translation} \rightarrow H = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & t_x \\ \sin(\alpha) & \cos(\alpha) & t_y \\ 0 & 0 & 1 \end{pmatrix} \quad (2.2)$$

$$\textbf{Similarity (4 DoF): euclidean + scale} \rightarrow H = \begin{pmatrix} s \times \cos(\alpha) & s \times -\sin(\alpha) & t_x \\ s \times \sin(\alpha) & s \times \cos(\alpha) & t_y \\ 0 & 0 & 1 \end{pmatrix} \quad (2.3)$$

$$\textbf{Affine (6 DoF): similarity + ratio + shear} \rightarrow H = \begin{pmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix} \quad (2.4)$$

$$\textbf{Projective (8 DoF): all transformations} \rightarrow H = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \quad (2.5)$$

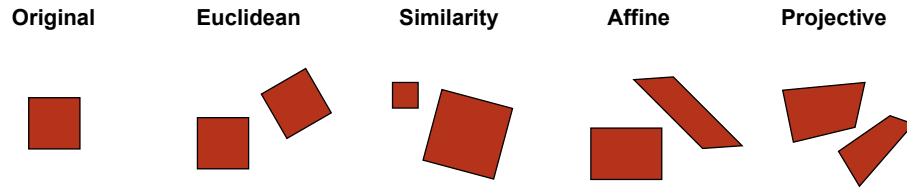


Figure 2.3: Types of homography transformations.

On the other hand, photometric registration takes into account differences introduced by camera and illumination variability by modelling a linear transformation for each channel separately. While geometric registration is essential in accounting for complex transformations and deformations in the images, photometric registration will deal with individual pixel values, making it suitable to correct color and light changes that occur locally. A possible approach would be to jointly estimate group wise geometric and photometric transformations, as shown in Bartoli [52].

More commonly, the registration techniques are classified as either direct methods, which use pixel information, or feature-based methods, which detect specific features in the images and match them for further alignment. Ghosh and Kaabouch [50] also considered the difference between spatial and frequency domain approaches, although there is a logical overlap for these two classifications as shown in Figure 2.4.

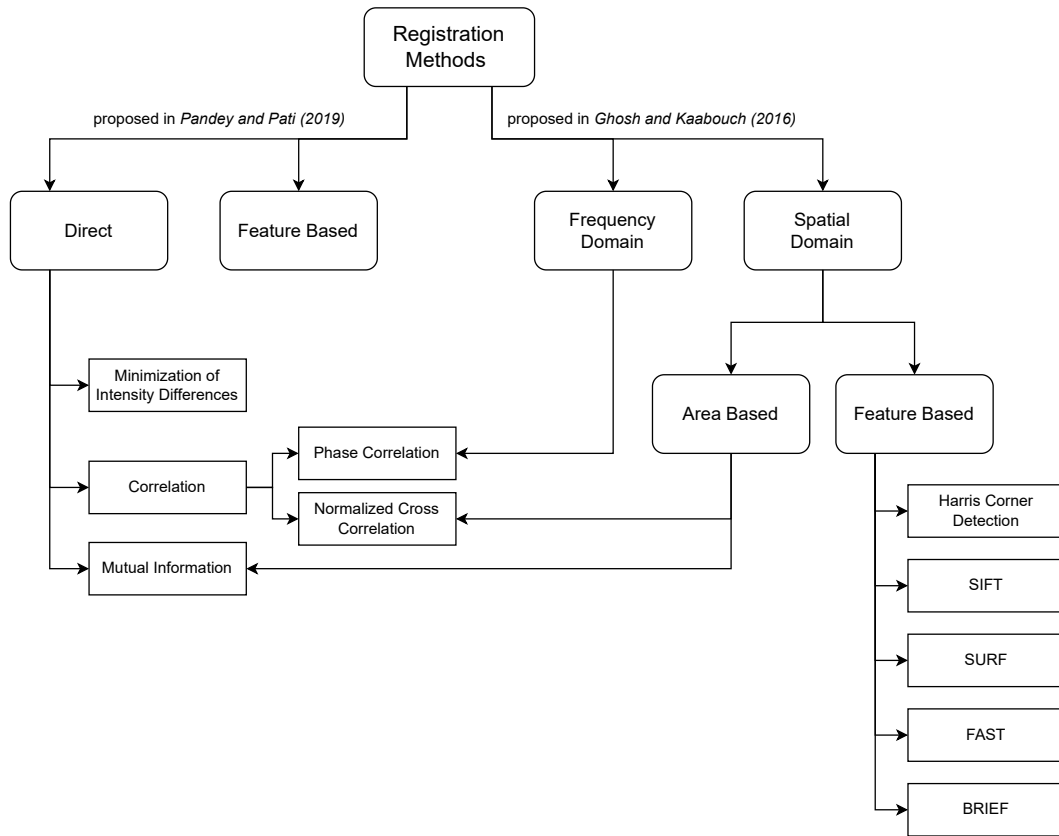


Figure 2.4: Classification of registration methods used in stitching algorithms, adapted from [44, 50].

Direct methods of registration aim to minimize discrepancies between matching pixels to optimize the final global image. One of the possible approaches is the minimization of intensity difference between two images where one image is moved over the other until the difference in the overlapping areas reaches a minimum value. Mutual information methods have a similar philosophy in that they search for the alignment that provides the highest value of shared information by the two images in terms of joint entropy, as proposed by Singh et al. [54] in a stitching application for face recognition.

The most common direct methods are based on correlation techniques, such as Normalized Cross Correlation (NCC), which has broad applications in image analysis and can be used to compute the similarity between windows in two images. A transformation can be determined by identifying which shift between the images results in a peak NCC value. Although these methods are simple, they tend to become slow when handling large images and require a significant overlap.

A different correlation approach, based on phase correlation, presents a very promising alternative to spatial based correlation methods and has proven better suited for registration tasks. In this case, the optimal transformation is computed with information from the frequency domain. This is possible by first calculating a Fourier Transform (FT) of the images and then estimating the normalized cross-power spectrum. The phase of this spectrum will be equivalent to the phase difference between the images [50]. Therefore, the registration can be done through the inverse

FT of the phase difference between two images since it produces a distinct peak at the overlap areas [44], which does not always happen with NCC. By focusing on the spectrum of the images, low frequencies do not contribute disproportionately to estimating the transformation.

Direct methods can be very reliable since they take into account all the information of the image. However, this also makes them less robust to point variations that can occur from light changes, zoom, distortions, movement of the elements to be registered, etc. To overcome these problems, feature-based approaches were introduced where specific or prominent elements are taken from every image, allowing a more precise comparison between them by focusing on the relevant information only. A detailed description of different features applied in stitching algorithms was presented by Ma et al. [55].

A feature is any type of relevant information about an image that can be used to solve a computational task. These can vary from corners and points to lines and gradients, chosen according to the specific images and goal of the detection. Feature detection is typically one of the first steps in computer vision tasks as it allows a better description of the image and its elements. The feature-based registration pipeline consists of three main stages:

1. **Detection:** identifying the specific features or keypoints that are relevant in describing the image. Typically a keypoint has a relevant position in the image and is stable under local and global changes.
2. **Description:** for each keypoint detected, a descriptor vector is created that reflects the local appearance around it in an ideally invariant way. A feature descriptor algorithm acts as a numerical fingerprint to differentiate two features.
3. **Matching:** the descriptors from two images are compared to identify similar features. The accuracy of matching algorithms depends largely on the type of feature detectors and descriptors used. Alternatively, instead of feature matching feature tracking can be used, where the search is confined to a small neighbourhood around the detected features.

An extensive review of the current findings in feature detection and matching for computer vision can be found in Szeliski [56]. The most prominent and commonly used feature detection and description strategies in stitching are as follows:

- **Harris Corner Detection:** a corner detection operator often used to extract corners and edges first introduced in 1988 by Harris et al. [57]. Corners represent points where different directed edges meet and they can provide relevant information about the structures present in an image as they are generally invariant to translation, rotation and light changes. Focusing on the edges and corners of different elements also reduces the amount of data to be processed in image analysis tasks.

- **Scale Invariant Feature Transform (SIFT)**: proposed in 2004 by Lowe [58], the SIFT algorithm was a major breakthrough in the field of feature detection and description, with several variants and improved versions published afterwards [59]. As described by Lowe, the algorithm consists of four major steps:

1. *Scale-space extrema detection*: a difference-of-Gaussian (DoG) is implemented over different scales to identify points that are invariant to both scale and rotation (Fig.2.5a).
2. *Keypoint localization*: features are selected based on their stability, removing low contrast error-prone ones (Fig.2.5b).
3. *Orientation assignment*: each keypoint gets a dominant orientation based on local gradients by picking the direction of the highest peak of a gradient histogram. From this step, the data has been transformed relative to the orientation, scale and location of each keypoint, making it invariant (Fig.2.6a).
4. *Keypoint descriptor*: local image gradients are measured at the selected scale. Each keypoint is described by gradient magnitude and orientation in a 16×16 region, divided into sub-blocks of 4×4 size. For each sub-block, 8 bin orientation histograms are created, resulting in a 4×4 descriptor with 8 orientations presented as a final vector of 128 elements (Fig.2.6b).

The final representation should be invariant to significant changes in local shape distortion and change in illumination, making SIFT features very attractive for a variety of image analysis tasks [60–62].

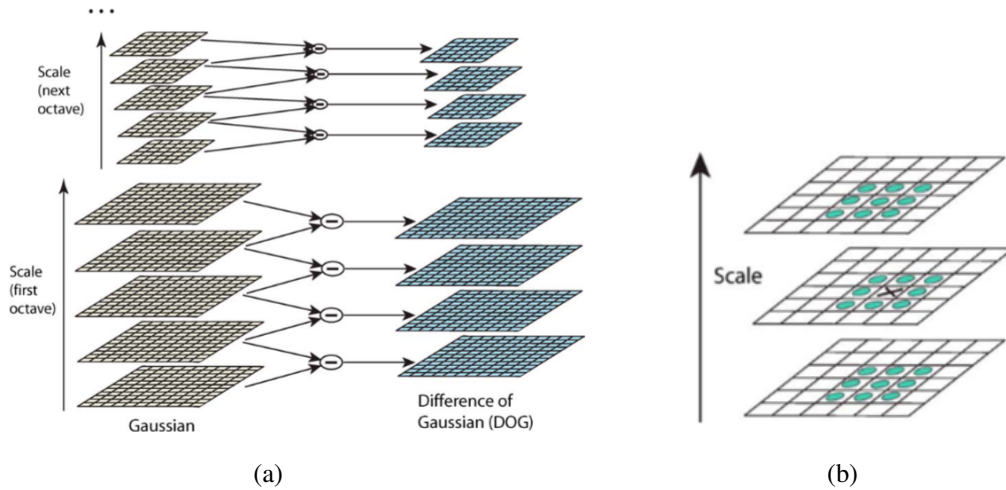


Figure 2.5: First and second stage of the SIFT algorithm [58] (a) for each octave of scale space the image is convolved with Gaussians and adjacent Gaussians are subtracted to produce the difference of Gaussians (b) extreme values are detected in the difference of Gaussian images.

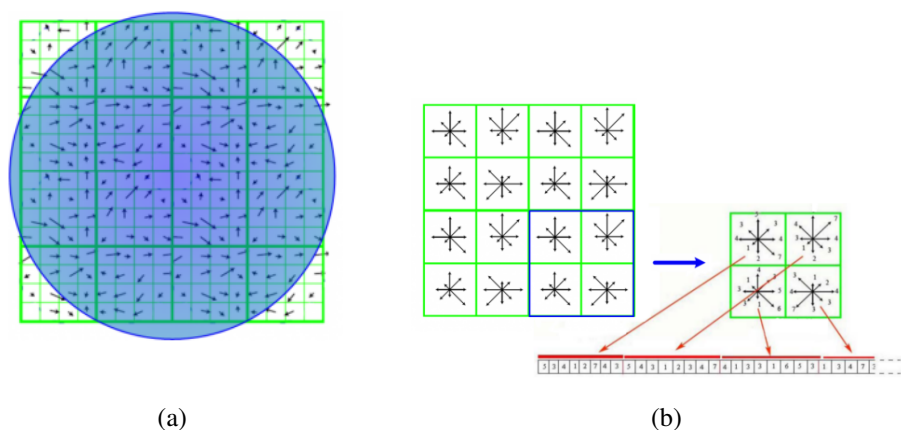


Figure 2.6: Third and fourth stage of the SIFT algorithm [58] (a) each keypoint is described by gradient magnitude and orientation in a 16×16 region (b) a descriptor is created from the histograms of 4×4 the sub-regions.

- **Speeded Up Robust Features (SURF):** SURF is a scale-space detector similar to SIFT with significant improvements in computation speed first described in 2006 [63, 64]. This is achieved by using a Hessian matrix approximation for the detection of interest keypoints and a distribution-based descriptor. By simplifying the methods, not only is the computation of the keypoints faster but it also makes the comparison between them easier, improving the performance of matching algorithms.
- **Features from Accelerated Segment Test (FAST):** a corner detection method developed for real-time video applications that can not be based on computationally expensive tasks [65, 66]. A machine learning method was developed to improve the performance of the algorithm by first detecting corners on a set of training images and creating a decision tree to be used for fast detection in other images.
- **Binary Robust Independent Elementary Features (BRIEF):** this algorithm converts the neighborhood around keypoints into a binary feature vector [67]. To reduce computational time, it is easier to have a dimensionality reduction of the descriptors which usually involves computing the whole descriptor before processing it. BRIEF proposes direct computing of binary strings from image patches, creating a more efficient descriptor.
- **Oriented FAST and Rotated Brief (ORB):** the algorithm was presented as an alternative to SIFT and SURF. It is built on FAST keypoint detectors and BRIEF descriptors, taking advantage of their fast and accurate performance and low computational cost [68].

2.4.2 Blending Methods

The final step of the stitching algorithm, known as blending, is essential to obtain a successful final mosaic. An appropriate blending algorithm must be chosen when piecing together the previously aligned images to achieve uniformity in color, illumination and other aspects of the composite image. Inconsistencies in the final image can result from a variety of factors, including variations in camera exposure, changes in scene illumination, moving objects between frames, etc. Therefore, selecting the right blending algorithm helps reduce the appearance of these discrepancies.

There are two main approaches to blending techniques, transition smoothing and optimal seam finding (Fig. 2.7). The former refers to methods that create a transition between several images by modifying the overlapping region, while the latter tries to detect the optimal boundary that minimizes the chance of a visible seam. Hybrid blending combines multiple methods, exploiting their advantages to reduce different artefacts, with the disadvantage of being more computationally expensive [44].

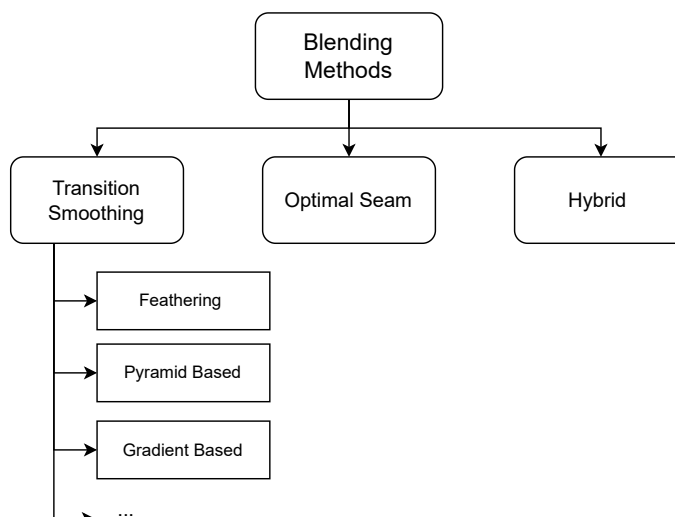


Figure 2.7: Classification of the blending methods used in stitching algorithms, adapted from [44, 50]

2.4.3 Relevance of Stitching for the Study of the ENS

To further understand the structural and functional characteristics of the ENS, imaging approaches are essential. However, as described in Sections 2.1 and 2.2, the constant movement of the gut makes it difficult to obtain clear and precise *in vivo* images. Many methods resort to limiting the movement by applying pharmacological agents to paralyze the tissue or manually stretching to prevent the largest contractions, with the danger of impacting signalling and mechanical properties. Recently, Barth et al. [30] proposed a method to do calcium imaging in the moving gut based on the registration of images recorded from moving tissue, without paralytics but still pinned into

a dish. The methods use a variation of the NoRMCorre algorithm for movement correction associated with muscle contractions and identification of out-of-focus events to improve the registration process.

Being able to correct the complex movement artefacts from recordings of the living gut takes us one step forward in expanding and improving image analysis methods that are already applied to other types of images, with less complex problems. One extremely relevant approach in studying large systems is stitching, which allows the visualization of cellular details on a large scale providing additional structural information. Graham et al. [69] presented a 3D visualization of human colon ENS without sectioning the tissue by stitching stacks of images to cover large regions. However, as the preparation of the tissue consisted of chemically fixating it, contraction artefacts were not considered. Although very relevant for reconstructing large 3D structures, the algorithms do not require dealing with low signal-to-noise images of a moving scenery as is often the case with live imaging.

A method for stitching images obtained from live moving gut has not yet been presented. However, such a method would be useful to create a larger field of view of interest regions and extract structural information. Additionally, the stitched image could also provide a means of stabilizing the recording to extract temporal information such as contractile or neuronal activity. Taking advantage of the current knowledge of movement correction algorithms, calcium imaging analysis and stitching algorithms, it is extremely relevant and timely to develop such an approach.

Chapter 3

Methodology

The following chapter details the experimental and computational methods used throughout the development of this thesis. The work and data here presented were developed at the Laboratory for Enteric Neuroscience (LENS) at KU Leuven.

As a main goal in this thesis, a pipeline was created to be used as a tool that aids in the quantitative analysis of recordings from the moving gut. The pipeline presented was developed considering different methods typically used in stitching algorithms, movement correction and live image analysis. According to the challenges found in recordings of the moving gut, adaptations were created to achieve a more general algorithm with several parameters that can be tuned for specific data. The whole pipeline was implemented in Python, taking advantage of widely used libraries such as the Open Computer Vision Library (OpenCV).

The main part of the pipeline is a stitching algorithm to align the frames and draw a final map that includes all spatial information. Additional modules were created to use before the stitching to improve its accuracy, and after the stitching for extra results to expand the amount of information, including temporal, we can extract from the recordings, as presented in Figure 3.1. The pre-stitching modules are optional, which means they can be more useful for certain types of recordings, but not necessary for others. With this in mind, a user interface was built to accommodate different trials for different challenges, further explored in Chapter 4. Therefore, the stitching algorithm can work on its own for simple recordings while the most complex cases might benefit from the added modules.

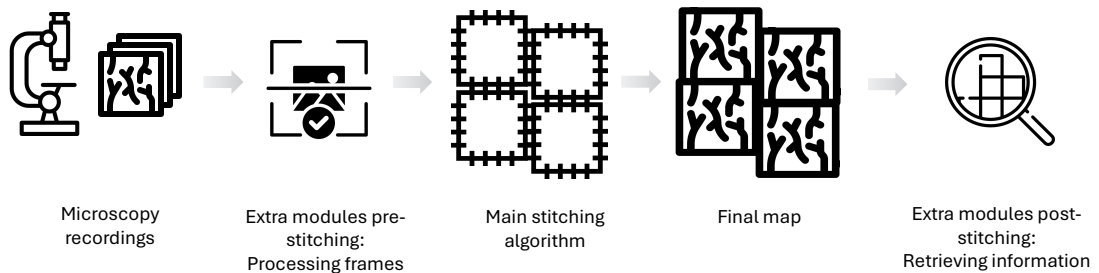


Figure 3.1: Main goals and implementations steps developed in the thesis.

3.1 Data Acquisition and Experiment Planning

The idea behind generating a pipeline with a main algorithm and extra modules was to create an adaptable way of dealing with different recordings with a wide range of complexity and problems. This way, the pipeline is not restrictive to a very specific kind of data. The experimental outline for collecting the recordings is presented in Figure 3.2.

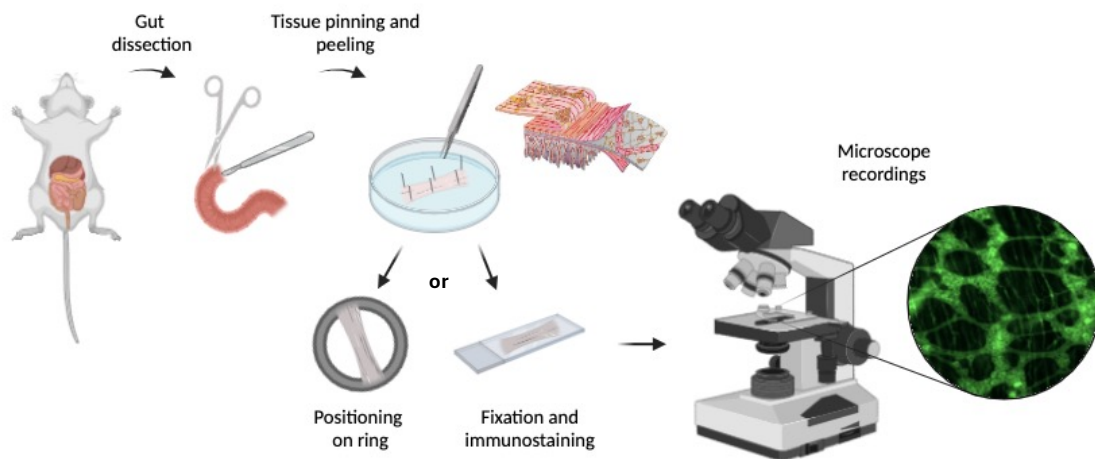


Figure 3.2: Experimental outline for collecting recordings of the ENS.

In order to assess the quality of the pipeline, several experiments were designed to showcase the different abilities of the algorithms first separately and later combined. Besides that, these tests were made in recordings with increasing complexity which escalates the pipeline requirements to process it. We aim to first prove that the most basic capabilities of the algorithm are achieved and then show how it can be tuned to work for more complicated data.

The first recordings evaluated were obtained from tissue that was chemically fixated with paraformaldehyde to avoid any spontaneous movement. In these tissues the fluorescent signals are still present but the cells are no longer alive. From these tissues a consecutive image series was created that mimics natural movement by moving the microscope stage in a pre-determined trajectory. By controlling the movement precisely, the alignment of the frames is known and therefore can be used as a ground truth for assessment of the maps created from the feature based algorithm (Figure 3.3). The automation was possible through Python scripts we developed in the Open Application Development (OAD) programming environment in ZEN for automatic image acquisition, which allowed full control of experimental setups such as the stage x and y position, focus and channels.

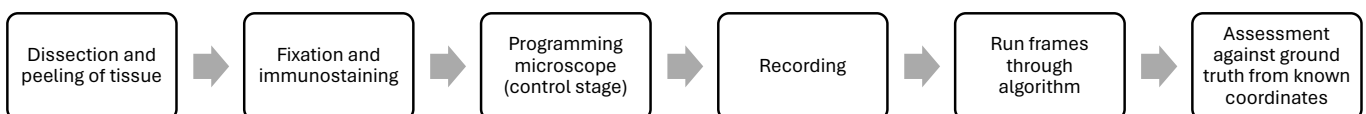


Figure 3.3: Experiment plan for controlled movement on fixed tissue trials.

The second stage of testing aimed to demonstrate how the pipeline can deal with uncontrolled physiological movement inherent to the tissues. This type of movement involves complex distortions that the additional modules try to correct to improve the main stitching algorithm. To obtain recordings from freely moving tissue, after dissection the tissue was positioned on small inox with a rubber o-ring keeping it in place. In this case no chemical fixative (which would kill the tissue) or pharmacological agents (to relax the muscle) were added (Figure 3.4).

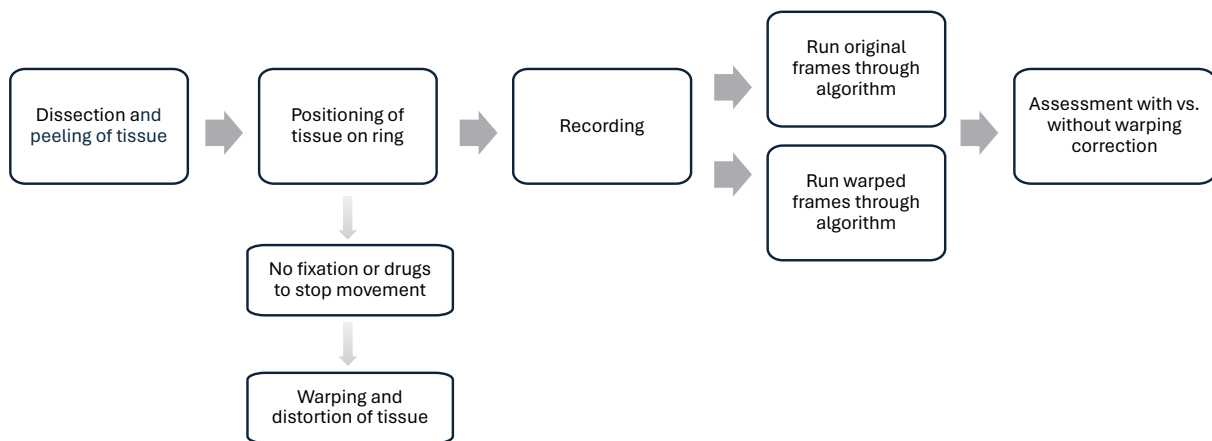


Figure 3.4: Experiment plan for moving tissue in ring trials.

One of the most important and common measurements made in these kind of systems is neuronal activity, reflected in intensity changes of the Ca indicator GCaMP. Therefore, it is of extreme importance that whatever kind of processing made in the recordings does not significantly affect these values. To test if this was the case, recordings with activity present were changed artificially to mimic distortions that could arise from movement. Afterwards, the distorted recordings were corrected and aligned through the pipeline and activity was measured in the same ROIs in the original and corrected recordings (Figure 3.5).

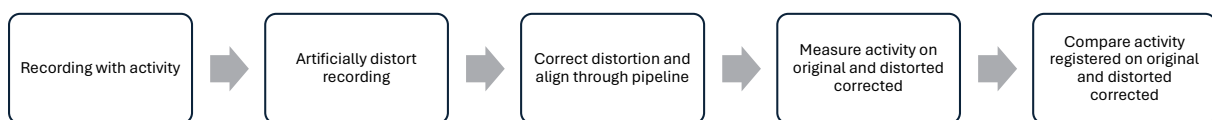


Figure 3.5: Experiment plan for activity measurements in corrected recordings.

The final data presented in this thesis comes from intravital recordings made through an abdominal window, which are part of the innovative and remarkable work performed by Dr. Candice Fung in the LENS group. Due to the elaborate nature of these experiments, the recordings obtained are incredibly intricate as a consequence of the gut movement in real time. These recordings were made available to test the tools developed in this master thesis, but also to demonstrate how relevant these analysis can be in assisting other lines of research.

3.2 Algorithm Description

The algorithms behind the different steps in the pipeline are as follows:

- **Pre-Processing:** The first step implemented includes basic tasks of image pre-processing such as denoising, brightness and contrast changes. Pre-processing is a critical step in image analysis, playing a foundational role in enhancing the quality and interpretability of raw image data. By performing these operations, the overall performance of different algorithms can be improved [70].

In general, noise introduced during the acquisition process can affect feature extraction methods by decreasing the ability to extract important features such as edges, corners, and shapes. Features are typically highly distinctive and invariant, but noisy images might lessen the amount of correct matches between different frames [71]. Besides noise, studies have shown that other image characteristics such as brightness also have an impact on tasks like object detection [72].

Other options in the pre-processing stage include vignette correction to correct edge artefacts and a down-sampling option that can, in cases with a large number of frames, make the algorithm faster and more efficient.

- **Warping:** The second step has the main goal of correcting discrepancies in the final stitched image caused by distortion and movement. In general, warping is fundamental in image stitching, since it addresses challenges related to geometric alignment and perspective correction. By transforming and aligning the images accurately, warping ensures that the final stitched image is seamless, visually coherent, and free of noticeable distortions.

The warping operation takes two frames, a reference and a frame to be warped, and calculates a homography matrix through the matched features between them. To compute the homography, the algorithm uses the Random Sample Consensus (RANSAC) method, an iterative way of estimating the parameters of a mathematical model while excluding outliers from the data. The matrix is then used to transform the frame into a new perspective.

Two different approaches are available: the reference can either be a fixed frame or it can change in each iteration. The choice of which to use is highly dependent on the amount of horizontal and vertical shift observed in the group of frames to be stitched. If most of the frames can be correctly warped to a singular reference, such as a frame unaffected by contraction, the first approach is the most appropriate. On the other hand, if the recording shows an incremental warping of the frames combined with major shifts in the area presented in front of the camera, the second approach yields better results.

- **Clustering:** The third step consists of clustering the images based on their similarity. The images are clustered into groups and each of them goes through the stitching process. This results in smaller maps which then are stitched together themselves to compose the final larger structural map.

The purpose of applying clustering before the stitching algorithm is twofold. In the case of long recordings with repetitive back and forward movement, certain regions will appear in front of the camera at multiple distinctive times. By clustering these frames based on their content instead of analysing them in temporal order, the stitching process becomes more efficient by first representing smaller regions and then finally combining them into a larger map. Additionally, if there is a fast transition or some artefact that makes it impossible to find overlap through all small regions, the information will still be kept in the form of multiple maps instead of one final map missing specific regions.

Clustering images based on similarity involves grouping images so that those within the same cluster are more similar to each other than to those in other clusters. To determine the best approach to include in the main pipeline, two different methods were implemented and tested: similarity matrix-based clustering and a combination of pre-trained models for feature extraction with clustering algorithms.

- **Similarity matrix:** The similarity matrix, also described as a distance matrix, is a two-dimensional square symmetrical matrix containing the similarities between each pair of images. The clustering process aims to minimize the distance between pairs within a cluster, while any pair from different clusters has a maximum distance [73].

The distance measure reflects the similarity between two images and several different metrics serve this purpose. Indexes like Root Mean Square Error (RMSE) and Peak Signal-to-Noise Ratio (PSNR) give a direct measure of how different two images are, while more complex metrics like Structural Similar Index Measure (SSIM) measure the perceptual difference between images [74], as presented in Equations 3.1 and 3.2.

$$RMSE = \sqrt{\frac{1}{M * N} \sum_{i=0, j=0}^{M-1, N-1} [Image_1(i, j) - Image_2(i, j)]^2} \quad (3.1)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3.2)$$

where μ_x and μ_y are the averages of x and y respectively; σ_x and σ_y are the variances of x and y ; σ_{xy} is the covariance of x and y , $c_1 = (k_1L)^2$ and $c_2 = (k_2L)^2$ are two variables that stabilize the division, L is the dynamic range of the pixel values (typically given by $2^{\text{bitsperpixel}} - 1$), and $k_1 = 0.01$ and $k_2 = 0.03$ by default.

After the similarity matrix is built, it can be used as an input for a clustering algorithm. Both hierarchical and spectral methods were tested. The goal of hierarchical clustering techniques is to create a sequence of nested divisions from a given dataset. There are two types of partitions: divisive (top-down) and agglomerative (bottom-up).

Agglomerative clustering starts with each data point establishing its unique cluster. Once the data points are clustered together in a single cluster, the algorithm then iteratively merges the closest pairings of clusters. Divisive hierarchical clustering, on the

other hand, begins with every data point being a part of a single cluster. Once each data point is given to a distinct cluster, the algorithm divides bigger clusters into smaller ones [73].

Spectral clustering uses the characteristics of eigenvalues and eigenvectors of matrices generated from data and can capture complex structures, making it helpful for grouping non-linearly separable data. However, this method is highly sensitive to the choice of similarity measure and needs very suitable similarity matrices to produce reliable results [75].

- **Pre-trained models for feature extraction:** Some approaches to image clustering are based on feature selection and subsequent intricate representations of them, such as bags of features (from SIFT, SURF, etc.), to prepare them for clustering algorithms [76, 77]. Recently, using features taken from pre-trained models as a substitute for raw picture data has significantly improved clustering findings [78]. These models are typically trained on large datasets and can be used as feature extractors by removing the final classification layer and using the output of one of the earlier layers (often the last fully connected or pooling layer).

This approach was tested using the VGG16 and ResNet50 convolutional neural networks, considered state of the art for image recognition tasks. The features extracted were then clustered through the KMeans algorithm.

- **Stitching:** The main stitching algorithm implemented is feature-based since the extracted features are more likely to remain invariant to changes such as intensity or deformation. Both SIFT and ORB detectors, along with the Brute Force matcher, are implemented through their OpenCV functions. Typically, an alignment algorithm analyzes the frames sequentially and adds them to the panorama expanding it step by step. However, in this case, global stitching methods are more appropriate, in which the frames are registered not only with their time-adjacent frames but with spatial-adjacent frames acquired anytime during the recording. This allows for continuous correction of accumulated spatial errors and deformations [46].

To achieve this global alignment method, the algorithm was adapted from *stitch2D*, a Python script that stitches a two-dimensional grid of tiles into a mosaic, originally developed for stitching images collected in the Department of Mineral Sciences at the Smithsonian National Museum of Natural History, available in [79].

The original script offers the option of using a structured mosaic, where it is possible to specify how the tiles should be arranged. Since this is not the case in the recordings explored in this thesis, the adaption of the algorithm was directly from the unstructured mosaic class. In this approach, the algorithm takes the individual frames without any presupposed information on their spatial arrangement and aligns them based on iterative feature matching. This is a much more complex task compared to the structured stitching approach, which is why the steps before this one are so valuable in improving its efficacy.

The output of this stitching step in the pipeline is the final map obtained from the frames, as well as metadata containing the coordinates of each frame that was placed.

- **Extra results after final map:** The final map provides valuable structural information. Nevertheless, it lacks information about the physiological processes that occur during the recording, such as movement and activity. Therefore, in the last step of the pipeline, there is the opportunity to combine multiple details to draw relevant conclusions.

Firstly, by taking the coordinates of the frames, the trajectory of the movement is drawn into the final map itself. It is also possible to obtain x and y shift representations along time, that help to quantify and possibly reveal certain patterns in the movement.

Secondly, the original frames are projected onto the final map, creating a temporal representation of the recording. This projection can be analyzed temporally as a recording itself where the stabilization in space allows for a variety of measurements, such as activity, which could not be made in the original recording. Additionally, the projection of frames can be turned into a tridimensional representation of the movement, where the frames are layered in their specific position along the z axis (standing for time).

If all the frames are used the alignment then all the frames will have specific coordinates attributed and the extra results explained above are built from them. However this may not always be the case if there are certain frames which can not be aligned through their features. In this case, the final map is made up of only a percentage of frames and the coordinates do not reflect all the information needed for extra results.

To deal with this issue, a template matching algorithm is used, which looks for smaller parts of an image which match a template image. By finding the place on the map that corresponds to the individual frame, specific coordinates can be attributed to the frames that did not have them from the alignment process. The implementation was made through OpenCV's template matching function.

Lastly, the opportunities for using the template matching method were expanded by looking into the option to identify where a single frame or even the entire final map is located in a larger region. For example, if a recording is made in a tissue and this tissue is later fixed, the method can be of use to find back the regions observed in the recordings. If the images are taken at different magnifications, the algorithm takes a scale factor and compensates for these differences.

To consolidate this pipeline into a usable tool, the final step was to develop a guided user interface that allowed the complete automatization and visualization of the steps mentioned beforehand. With this interface, it is possible to achieve the full adaptability potential of the pipeline while also making it easy to use. The user interface was created in Python using the *tkinter* package, built with custom commands to create and manipulate interface widgets.

3.3 Quality Assessment

To validate the use of the pipeline, different datasets were used to create maps that need to be further evaluated. Ideally if we have access to an image that perfectly represents what the final stitched map should be, that could be used as a ground truth for typical data comparison. Having these two images, one could use typical methods such as a pixel by pixel comparison, feature matching or similarity index like SSIM. However, reference stitched images are often unavailable so other methods must be developed [80].

Stitching image quality assessment (SIQA) is a widely studied topic with many different methods proposed [81–83]. These can be divided into two groups: full reference SIQA and no reference SIQA. In full reference SIQA, the individual images needed to create the panorama are used to assess the final stitched image. In no reference SIQA, specific characteristics such as structural consistency and perceptual quality of the stitched image are used to quantify its quality without considering any prior information.

In full reference SIQA, the features from individual frames and from the final map can be used to estimate quality parameters relevant to this work such as misalignment and structural distortion.

For misalignment, we take the matched features between two adjacent input frames and search for their corresponding feature match in the stitched image. Then, we calculate the spatial difference between these corresponding points. Ideally, with no misalignment, the distance should be close to zero since matched feature points in two frames should represent the same region in the stitched image, as seen in Figure 3.6.

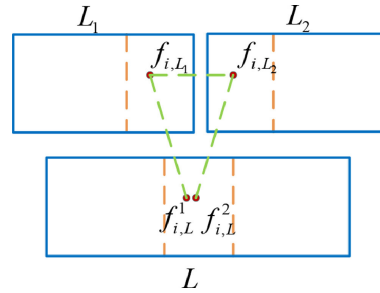


Figure 3.6: Illustration of misalignment calculation from [82].

$$MIS = \frac{1}{N_i} \sum_{i=1}^{N_i} \|f_{iL}^1 - f_{iL}^2\|^2 \quad (3.3)$$

For structural distortion, we take the homography matrix between the input frames and the stitched image to establish a transformation that can reflect the amount of deformation in the final map. We followed the same approach as in [83] which calculates the structural distortion from the values on the matrix as:

$$ST = \underbrace{(h_{11} - 1)^2 + (h_{22} - 1)^2 + h_{12}^2 + h_{21}^2}_{\text{Scale and Rotation}} + \underbrace{h_{31}^2 + h_{32}^2}_{\text{Deformation}} + \underbrace{(h_{11} - h_{22})^2 + (|h_{12}| + |h_{21}|)^2}_{\text{Uniformity}} \quad (3.4)$$

Chapter 4

Results

4.1 Evaluation of Feature Detectors and Matchers

The following chapter presents a summary of the major results found during the development of this work, both from artificially generated and actual recorded data. Firstly, taking into account that the feature-based stitching algorithm is the main part of the pipeline, different feature detectors and matchers were evaluated. With this goal, several frames were artificially altered as shown in Figure 4.1. The original image was then matched with the altered one using either the SIFT or ORB detector and the Brute Force or FLANN based match. Figure 4.2 represents the percentage of features from the original image matched with the altered image and Figure 4.3 represents the average distance of these matched features. When matching an image to itself, there will be 100% matches with 0 average distance. Therefore, the combination with a higher percentage and lower distance, in this case SIFT+BF, proves to be the most reliable.

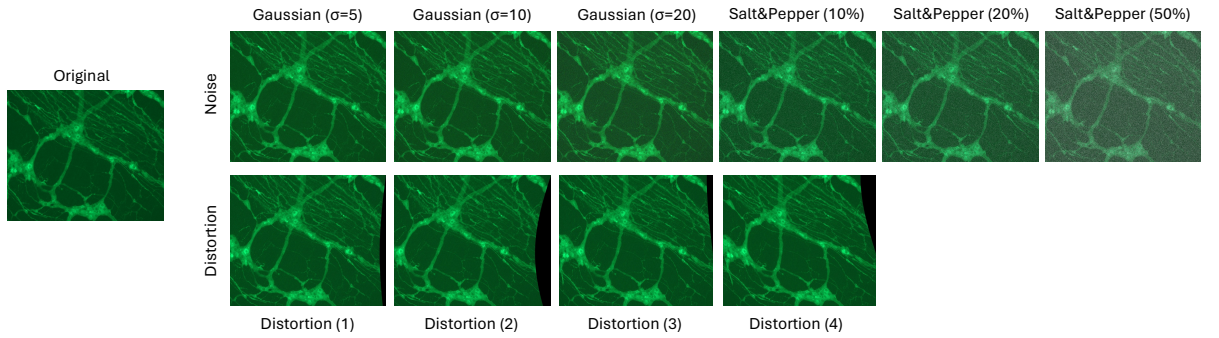


Figure 4.1: Example of artificial changes created for the evaluation of feature detectors and matchers: original frame vs. noisy (Gaussian and Salt&Pepper noise) and distorted versions.

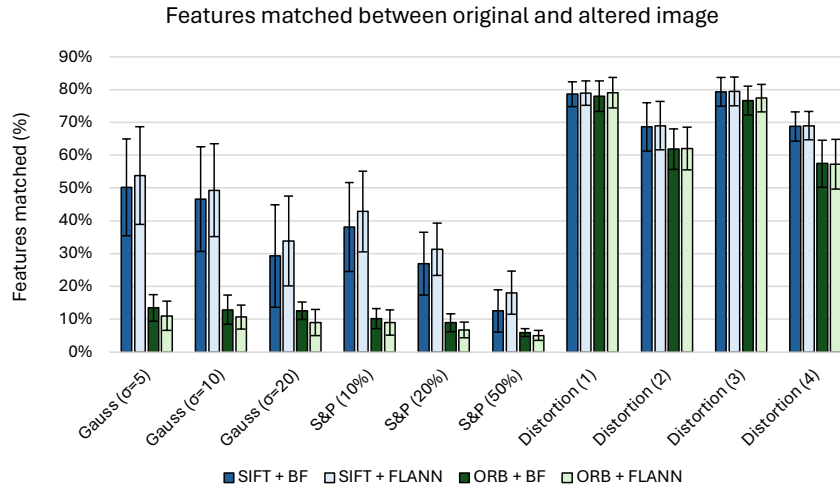


Figure 4.2: Features matched between the original image and altered versions (in percentage).

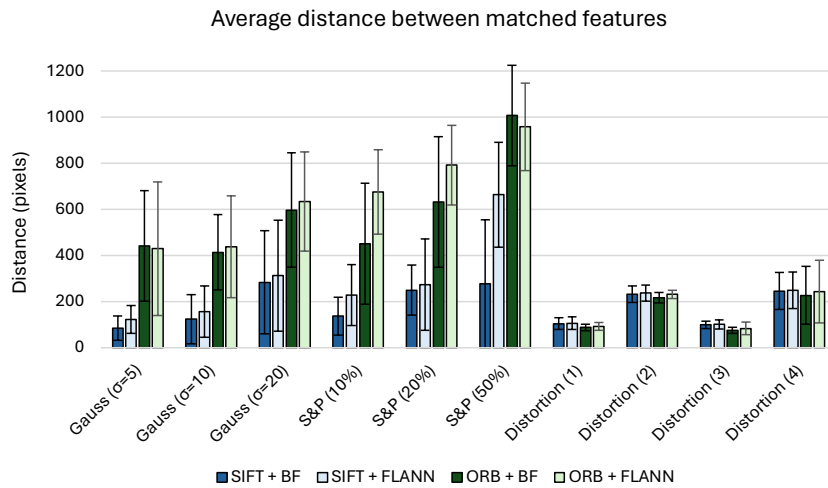


Figure 4.3: Average distance of matched features between the original image and altered versions (in pixels).

4.2 Added Modules

This section aims to illustrate what the function and importance is of the pre-stitching modules.

4.2.1 Pre-Processing

As mentioned in Chapter 3, some pre-processing can greatly affect the amount and reliability of features detected in the frames. As an example, Figure 4.4 shows what visual changes occur in the images when denoising, increasing brightness and contrast.

More importantly, Figure 4.5 shows how these changes can impact the feature matching process. In Figure 4.5a, a raw data version, only two matches were obtained and they were spatially incorrect. On the other hand, in Figure 4.5b, enough matches are identified to allow for a correct alignment of the frames.

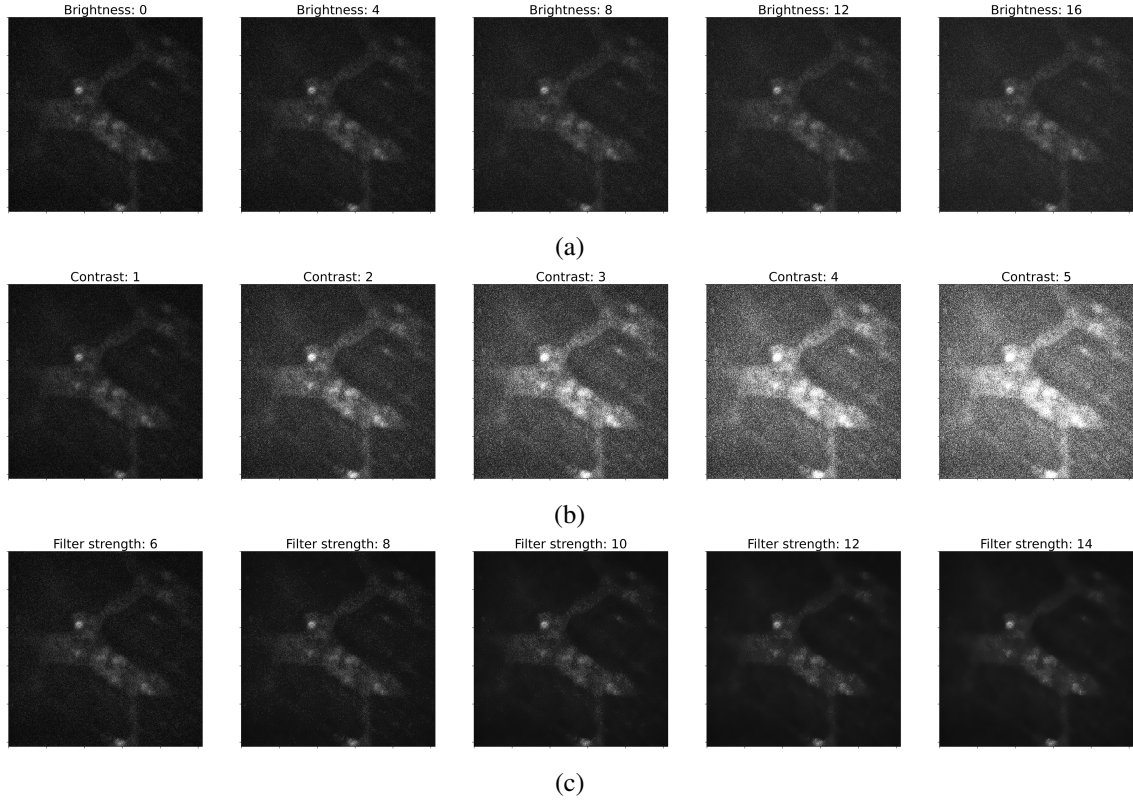


Figure 4.4: Changes in a frame after different pre-processing steps (a) different values for brightness representing value added to pixels (b) different values for contrast representing scaling of pixels (c) filter strength for OpenCV non-local means denoising algorithm.

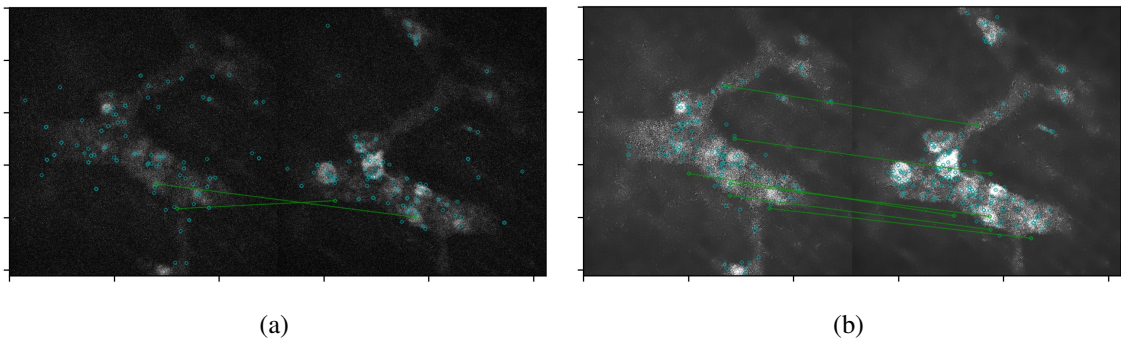


Figure 4.5: Difference in matching features (a) Raw data (b) Post pre-processing including denoising (filter=10), increasing brightness ($\alpha=8$) and increasing contrast ($\beta=2$).

4.2.2 Warping

The warping module aims to correct for distortions that are a consequence of movement. To better demonstrate this process, Figures 4.6 to 4.8 showcase the different steps. Firstly, matches are found between two frames and outliers are excluded according to the RANSAC algorithm. With the inliers, a homography matrix is calculated and then applied on the frame to be transformed. By doing this transformation, the alignment of the distorted frames is now improved.

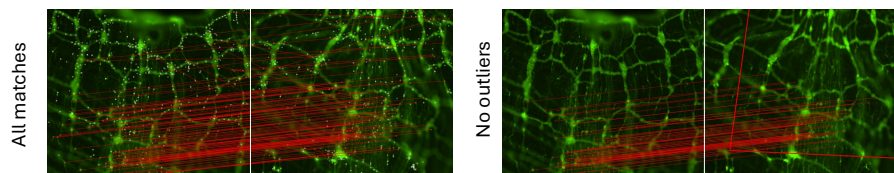


Figure 4.6: Feature matching of two frames with one presenting deformation: initial result through Brute Force matching compared with result after excluding outliers through the RANSAC algorithm. The inliers keypoints allow for a homography transform estimation drawn on the left.

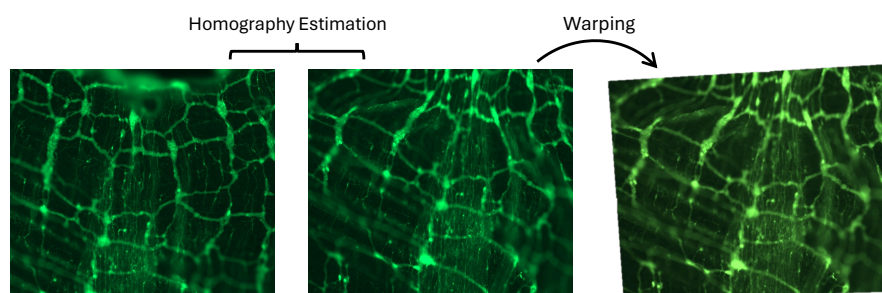


Figure 4.7: Demonstration of the warping correction process. First, a homography matrix is estimated from two frames considering one as the reference. Afterwards, the same matrix is used to warp the non-reference frame to the same plane as the reference one.

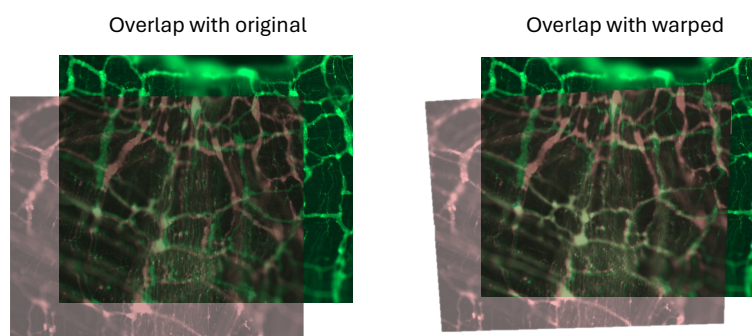


Figure 4.8: Comparison of the result of aligning two frames before and after warping correction.

4.2.3 Clustering

As explained in Chapter 3, the clustering module aims to facilitate an iterative stitching process to ensure that the most information possible is present in the final map. A demonstration of this process and its advantage is shown in Figure 4.9. Recordings with fast transitions were chosen to further analyze the benefits of clustering, as well as different clustering methods. The results for this analysis are presented below in Figures 4.10-4.12. When searching for the best method, it was important to choose one that gave high distinctive information for the final map, high silhouette score and low processing time. Therefore, the VGG16 feature extraction with Kmeans clustering was included in the pipeline.

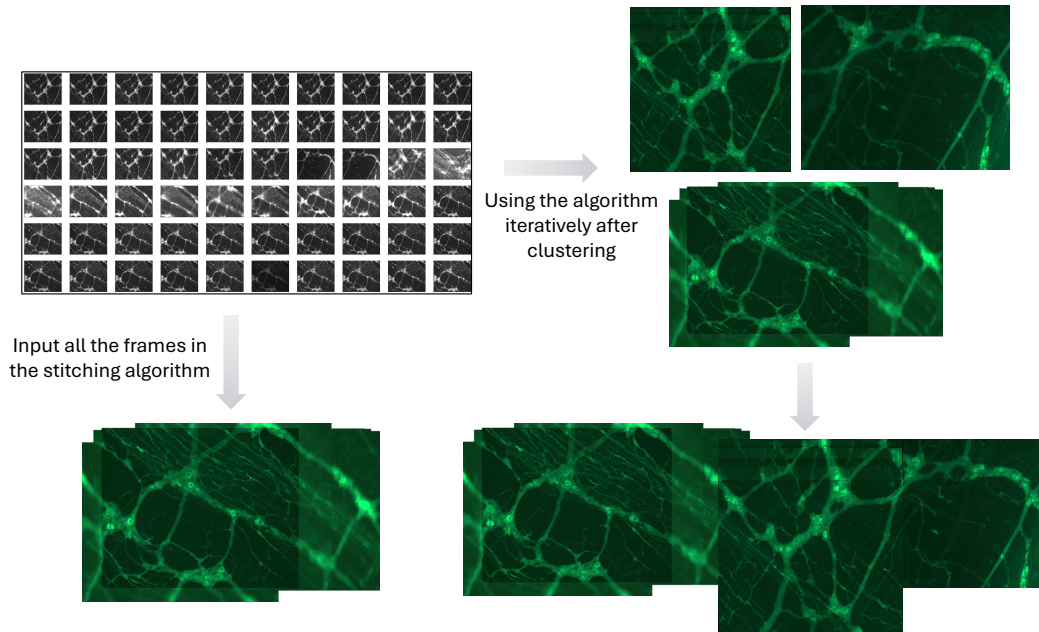


Figure 4.9: Comparison of the result of aligning frames all at once or by an iterative process facilitated by similarity clustering.

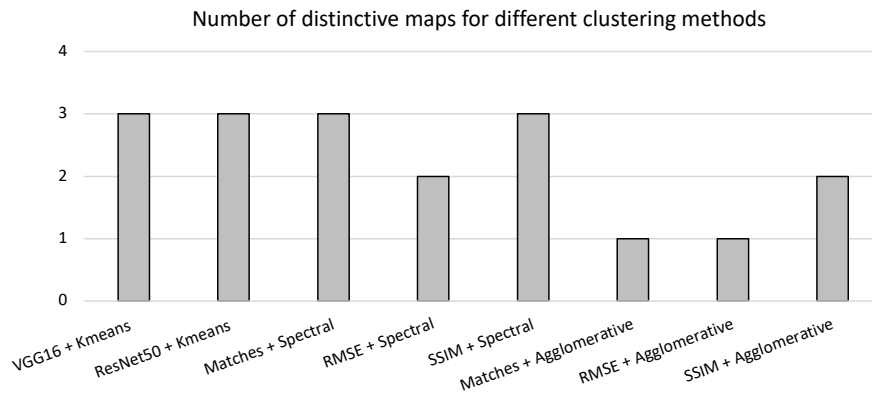


Figure 4.10: Clustering results: number of distinctive maps created in each clustering method.

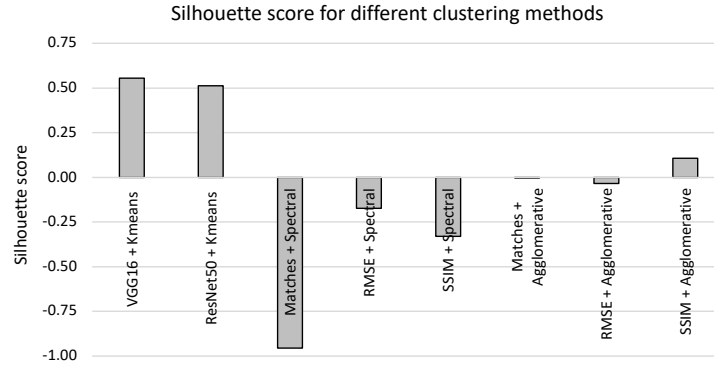


Figure 4.11: Clustering results: silhouette score for each clustering method.

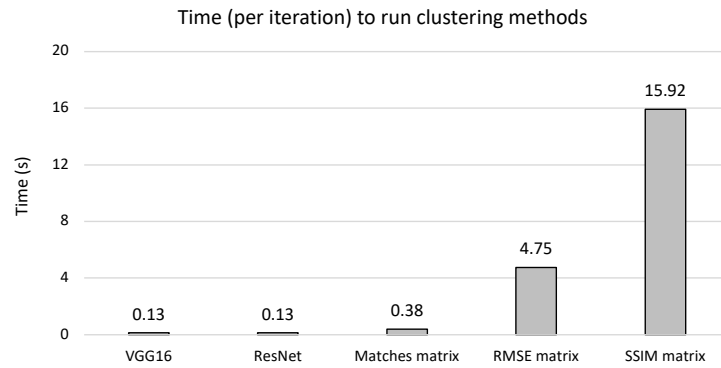


Figure 4.12: Clustering results: time taken per iteration in each clustering method.

4.3 Final Pipeline

According to the modules explored in the above sections, the final pipeline was built up to the result presented by the flowchart in Figure 4.13. It is important to notice a few details that can be modified by user input, the more general of them being if or not the pre-stitching modules are used. Inside each module there are also certain inputs:

- **Pre-processing:** the values attributed to the denoising filter strength, brightness and contrast can be modified.
- **Warping:** there is the option to perform a sequential warping or against a singular frame, in which case it is possible to choose which frame.
- **Clustering:** the number of clusters can be pre-defined, if not it will be chosen by the algorithm according to the best silhouette score.
- **Stitching:** two thresholds are set, one to filter out the best matches and one to choose how many matches two images have to have to be aligned together.

The filtering of matches follows Lowe [58] ratio test that checks if the distances between the two best matches from a singular keypoint are sufficiently different, i.e. if the best match found is distinctive enough to be considered a match.

After alignment, when drawing the final map, the order of setting the tiles can be defined as sequential in time which means the first frames are placed first (or the inverse) or the order can be at feature level by placing the feature rich frames on top (or the inverse).

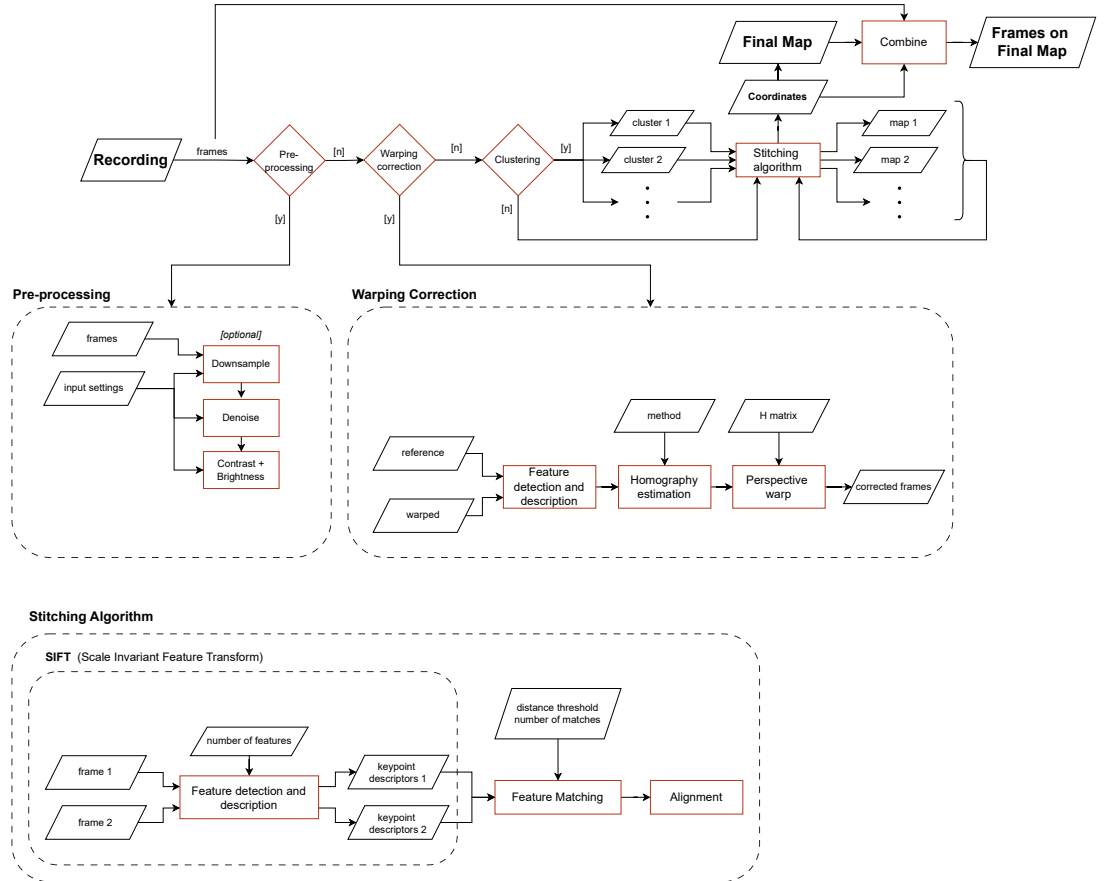


Figure 4.13: Final pipeline flowchart.

4.4 Pipeline Validation: Artificial Recordings

As a trial before applying the pipeline to real recordings, artificial versions were created by taking large image of a sample and running a region selection, along a specific path, through that image to obtain a selection of frames. The ellipse paths used for this experiment are detailed in Table 4.1 and shown in Figure 4.14. The assessment was then made against the original region and the scores were registered including a Pixel Match, SIFT Match and SSIM score in Figure 4.15.

Table 4.1: Details on the artificial recordings including the size of ellipse path and singular frame.

Recording	1	2	3	4	5	6
Ellipse size (relative to width and height of original)	$w/2 \times h/2$	$w/2 \times h/2$	$w/1.5 \times h/1.5$	$w \times h$	$w/2 \times h/2$	$w/1.5 \times h/1.5$
Frame size (pixel)	1920×1600	1200×1000	1200×1000	1200×1000	600×500	600×500

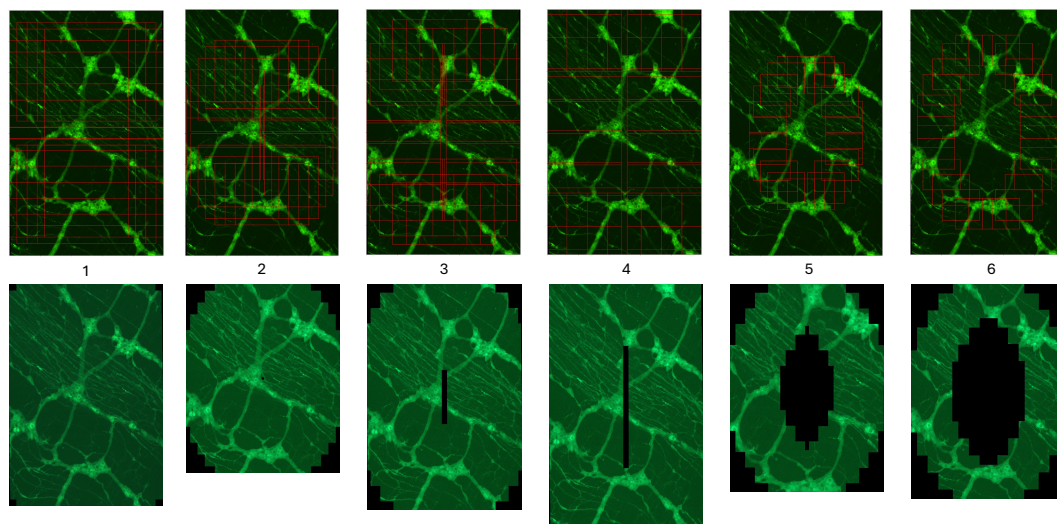


Figure 4.14: Artificial recordings created by cropping a consistent frame on a pre-defined ellipse path. Top row: representation of frames selected (in red) to create the recording. Bottom row: final map resulting from the stitching algorithm.

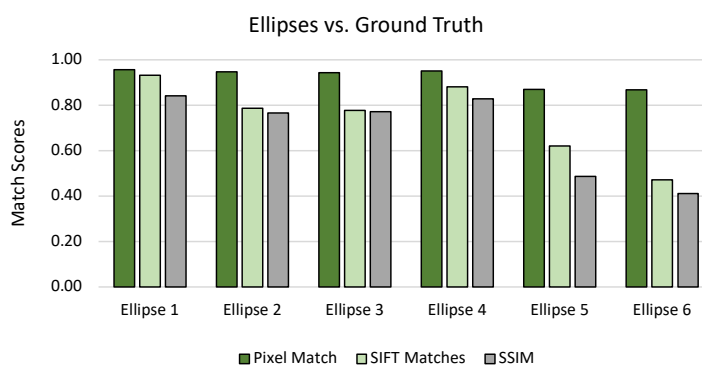


Figure 4.15: Match scores for artificial recordings.

4.5 Pipeline Validation: Controlled Movement on Fixed Tissue

The first data acquired, as explained in Figure 3.3, came from a piece of fixed tissue. Figure 4.16 shows a whole field of view of the tissue. For testing purposes to ensure the robustness of the methods three subregions of the tissue were defined. In the following section they are referred to as regions a, b and c as presented in 4.17. With the control of the microscope stage, two different trials were made on this sample.

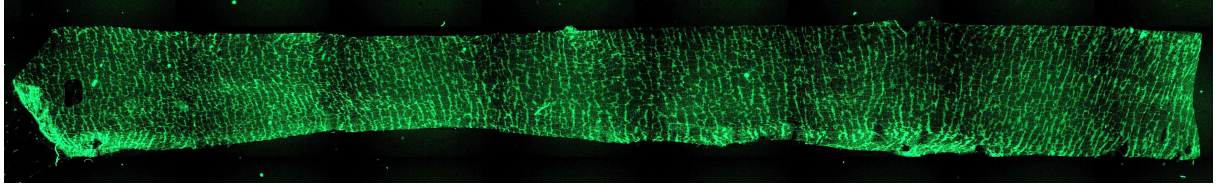


Figure 4.16: Full field of view of fixed myenteric plexus on a fluorescence widefield microscope.

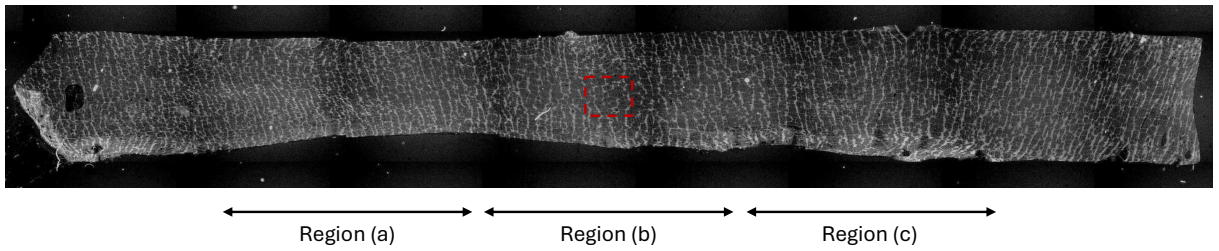


Figure 4.17: Full field of view of fixed myenteric plexus (grayscale) with specific regions that were imaged for the trials presented below. Square region represents the size of a single frame (with 10x magnification).

4.5.1 Reoccurring Frames/Repetitive Movements

Firstly, to ensure the algorithm was accurate for recordings with repetitive back-and-forth movements leading to reoccurring frames, an "8" trajectory was defined (Figure 4.18). The recordings were made with only one iteration through the trajectory or multiple iterations, in three different regions of the tissue (as shown in Figure 4.17).

For the matching scores, two approaches can be made, either considering as ground truth stitching obtained from the microscope coordinates which we know to be the true coordinates, or considering as ground truth the region in which it was recorded, as explained in Figure 4.19. Both scores are shown in Figure 4.20 where we can see that in general by comparing to the region itself, the match scores tend to be lower.

Lastly, Figure 4.21 shows a summary of this trial considering as ground truth the stitching from the microscope coordinates and adding a new variable to the group. Trajectory match is calculated based on the relative error between the trajectory obtained from the coordinates given by our pipeline and the trajectory from the ground truth coordinates. It shows that throughout the three regions and with multiple iterations, the match scores stay consistent.

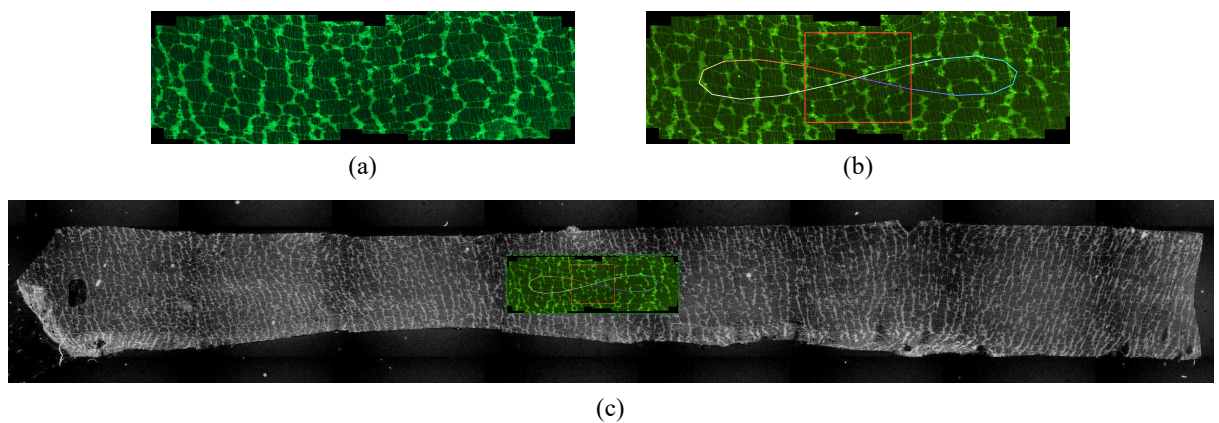


Figure 4.18: Illustrative example of one of the recordings results. (a) Map built from the individual frames. (b) Trajectory/movement programmed for the microscope stage. (c) Region in which the recording was taken. Supplementary material: [original recording](#) and [trajectory followed on fixed](#).

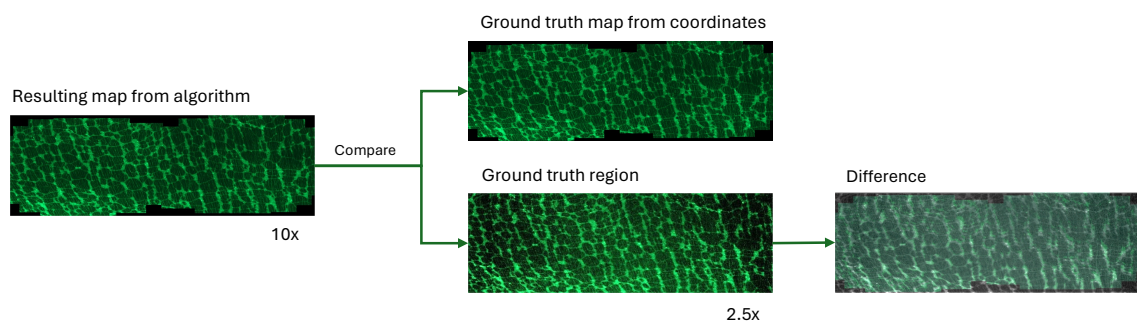


Figure 4.19: Two types of ground truth data that can be used to compare with the algorithm result. On top is the ground truth stitching obtained from the microscope coordinates. On the bottom is the full region through which the frames pass by.

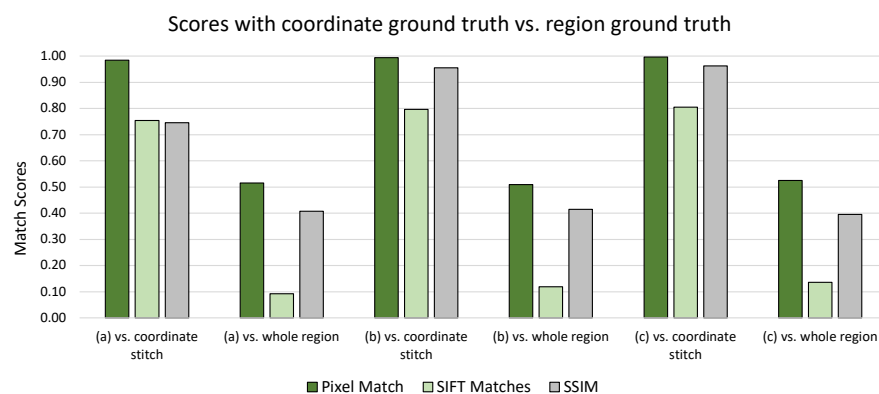


Figure 4.20: Match scores in the different regions of the recordings. Both the matches with the ground truth from the microscope coordinates and the matches with the ground truth from the full region are presented.

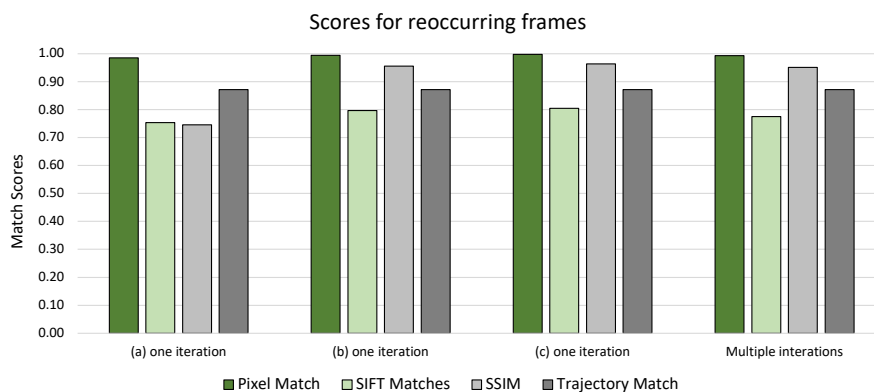


Figure 4.21: Match scores summary for the reoccurring frames trial considering as ground truth the stitching from the microscope coordinates. Multiple iterations represents a recording in which the trajectory was run through multiple times in succession.

4.5.2 Out of Focus Frames

The second trial aimed to evaluate the accuracy of the pipeline with out of focus events, which occur typically when something is moving. To achieve this, a single line was programmed into the microscope stage and specific changes in focus were input for some frames. Figure 4.22 shows an example of the type of maps obtained. Once again, the trial was repeated in the same three different regions (as shown in Figure 4.17).

When a frame is out of focus, its features change with a tendency to decrease in number since the structures are not as well defined. To illustrate this phenomenon, Figure 4.23 shows the difference in match number with an in focus frame and with an out of focus frame.

Finally, the match scores are presented in Figure 4.24. The most impact is seen in SIFT matches when comparing in focus with out of focus, which is consistent with what Figure 4.23 above demonstrates.

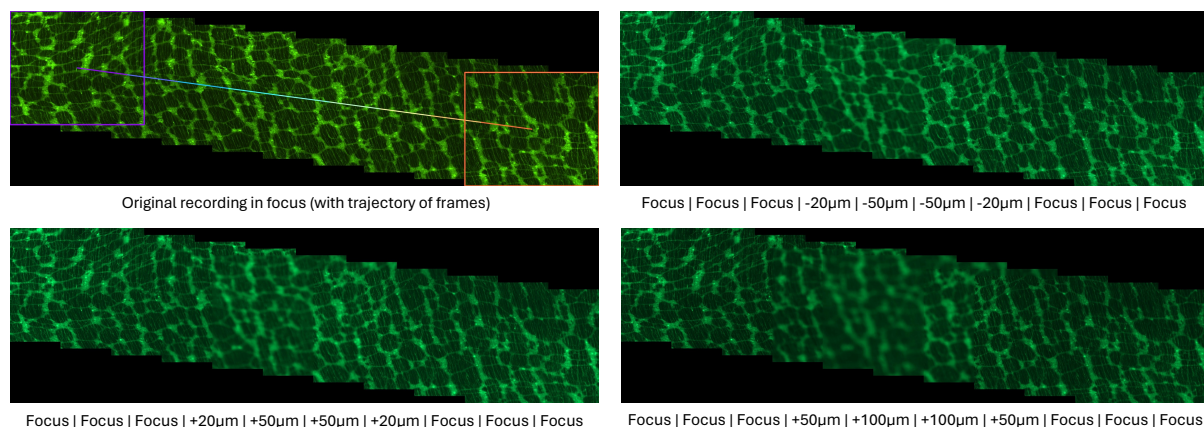


Figure 4.22: Example of maps obtained from recordings on the same region following the same trajectory but varying the focus to reproduce out of focus events. Supplementary material: [original recording in focus](#), [out of focus +50μm](#) and [out of focus +100μm](#).

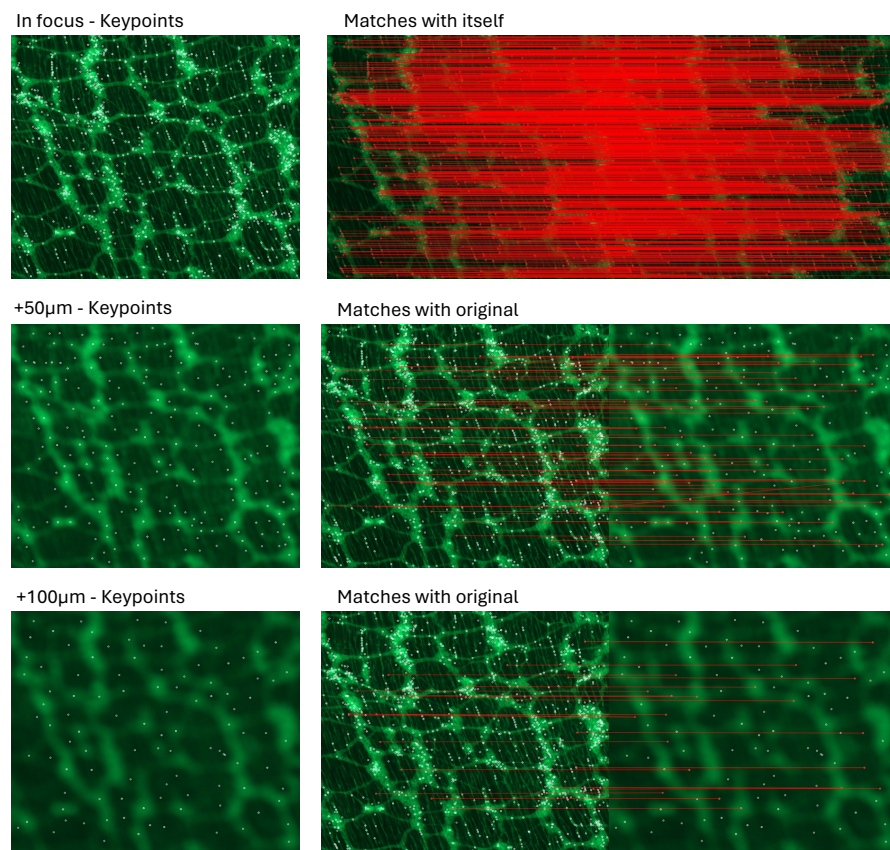


Figure 4.23: Difference in matching features with different levels of focus. Top row: keypoints and matches of a frame with itself, all features are matched. Middle row: keypoints and matches of a frame with the same region in $+50\mu\text{m}$ focus. Bottom row: keypoints and matches of a frame with the same region in $+100\mu\text{m}$ focus.

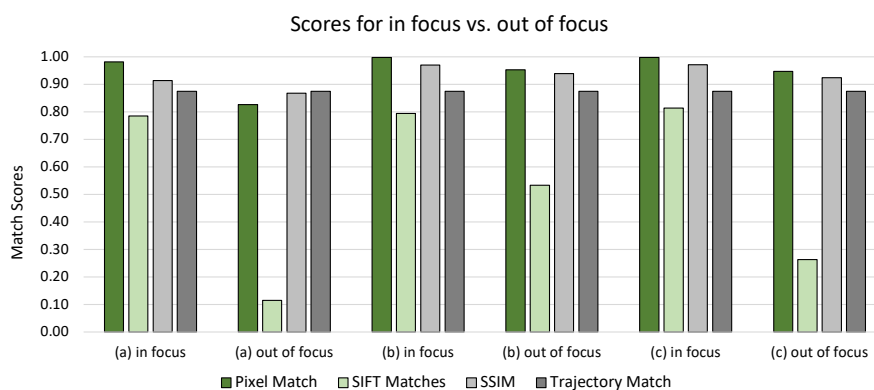


Figure 4.24: Matches scores for the out of focus trial considering as ground truth the in focus stitching from the microscope coordinates. Results are presented for the three regions. The out of focus values represent an average score of the three out of focus trials made in each region.

4.5.3 Matching Immunohistochemical Stainings

Finally, a number of histological and immunohistochemical stainings were made to identify nuclei, neurons and neuronal fibers in the tissue (Figure 4.25). To this end DAPI, and antibodies against HuC/D and TUJ1 were used. These are stainings commonly used to study the ENS. Even though the stainings label markedly different cellular structures, the overall features (e.g. grouping of cells and organisation) is similar. Therefore, it was interesting to test whether the algorithm would also be able to stay consistent for these frames. As long as the features throughout the frames are consistent, a map can be created (some examples in Figure 4.26).

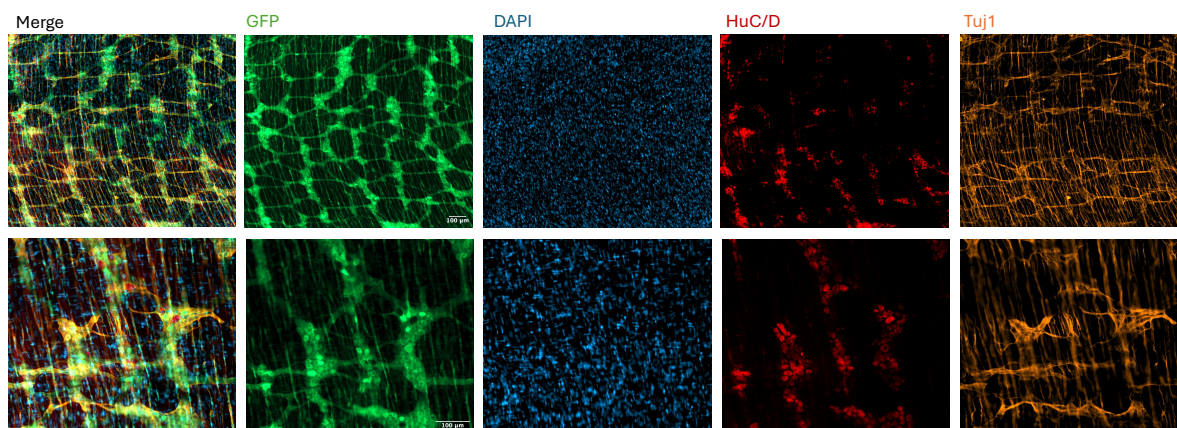


Figure 4.25: Illustration of the fixed sample with matching stainings. GFP labels enteric neurons and enteric glial cells, DAPI labels cell nuclei, HuC/D labels neurons and Tuj1 labels nerve fibers.

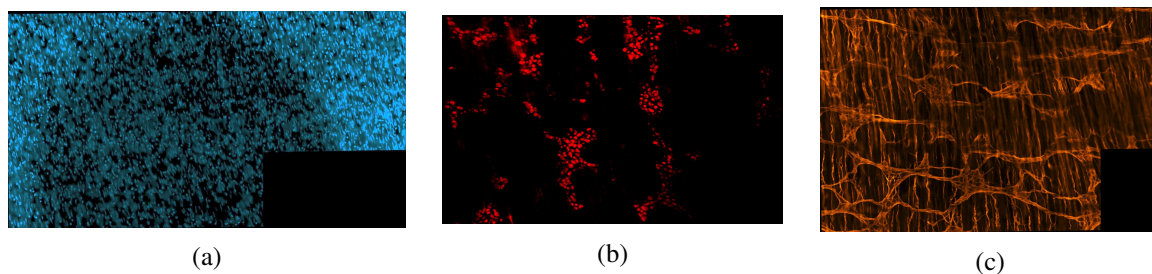


Figure 4.26: Maps created from recordings in different stainings: (a) DAPI (b) HuC/D (c) Tuj1.

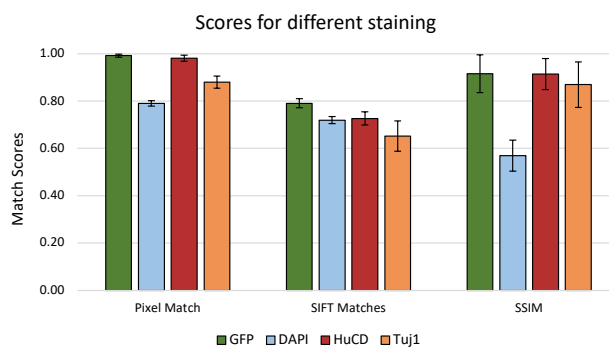
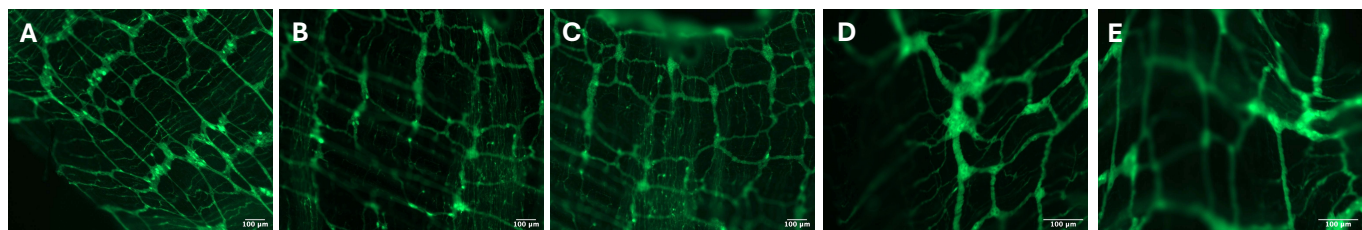


Figure 4.27: Match scores for staining trials.

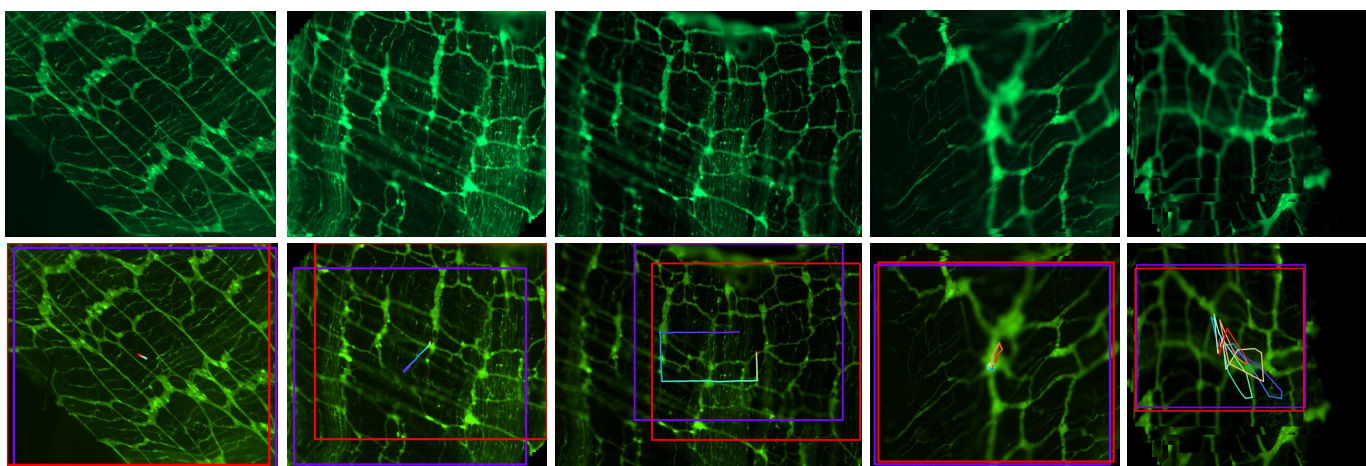
4.6 Pipeline Validation: Moving Tissue on Ring

The second dataset acquired, as explained in Figure 3.4, came from a piece of tissue positioned on a ring in order to capture its movement. Several recordings were made, with some examples presented in Figure 4.28 along with their maps and trajectories as well as the matching scores in Figure 4.29. Since there are no ground truth coordinates in this case, the match scores were obtained against each frame and averaged. Depending on the magnification used, the recording showed a larger or smaller x&y shift whenever it moved. With larger shifts and more complicated contractions, the maps are harder to align, which creates discontinuities in the edges of the frames. To correct this displacement, warping can help to project the frames onto a common plane, making it easier to align (Figure 4.30).

Maps were created from the original and warped versions and compared, as shown in Figure 4.31, and summarized in match scores in Figure 4.32a. The match scores do not give a clear distinction between the two, therefore other parameters such as the misalignment and distortion were also explored (following equations 3.3 and 3.4). As can be seen from Figures 4.32b and 4.32c these parameters were clearly lower for the warped versions.



(a)



(b)

Figure 4.28: Examples of recordings from moving tissue on a ring. (a) First frame of recording (A: ileum ring, B-E: duodenum ring). (b) Map created from the alignment of original frames and trajectory/movement of the sequential frames (Note that midpanel C includes movement from the tissue and from the stage itself). Supplementary material: original recordings A, B, C, D and E.

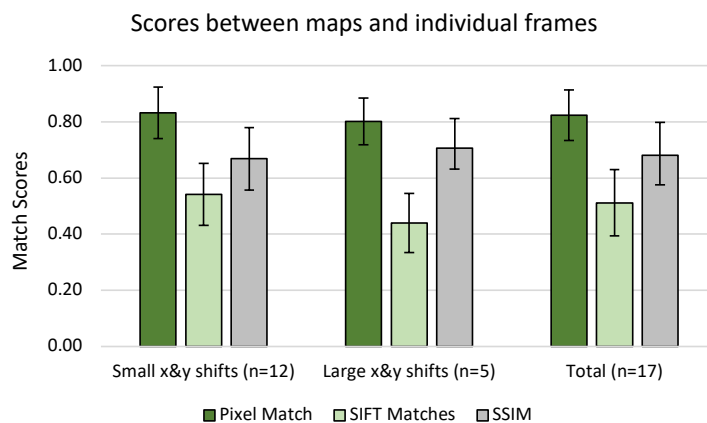


Figure 4.29: Match scores for moving tissue on ring, divided into recordings with small x&y shifts and large x&y shifts. Since the movement is not controlled and coordinates are not known, the scores are calculated for each individual frame and averaged for all the frames.

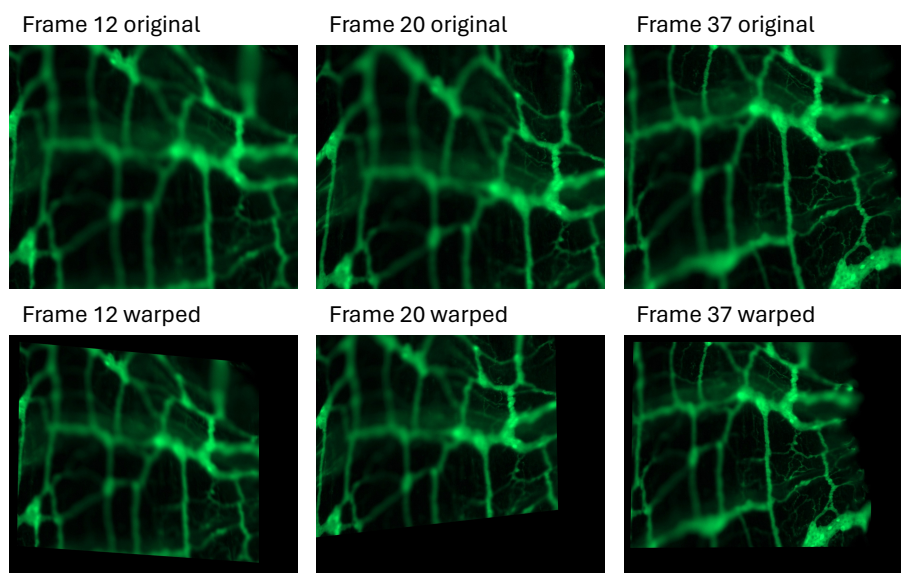


Figure 4.30: Illustrative example of the result of warping correction. Top row: original frames 12, 20 and 37 (out of 45). Bottom row: corresponding warped frames. The warping presented is sequential, which means every frame is warped against the one before it. Supplementary material: [side by side](#) of both recordings.

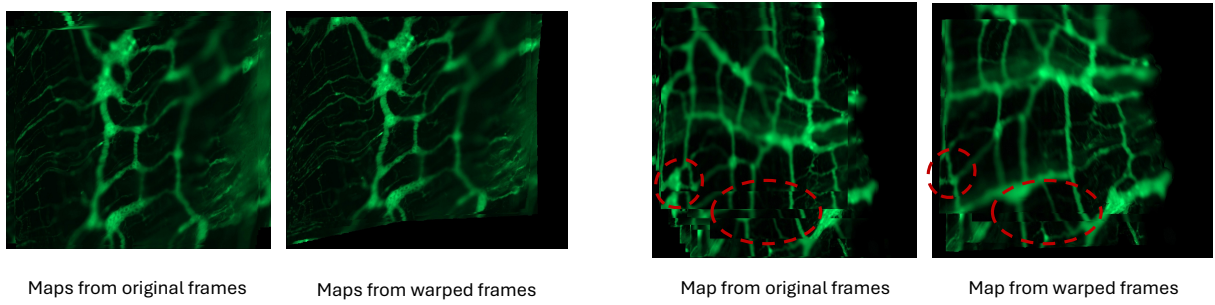
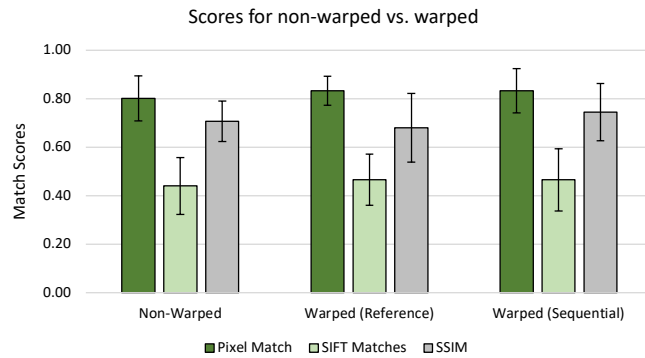
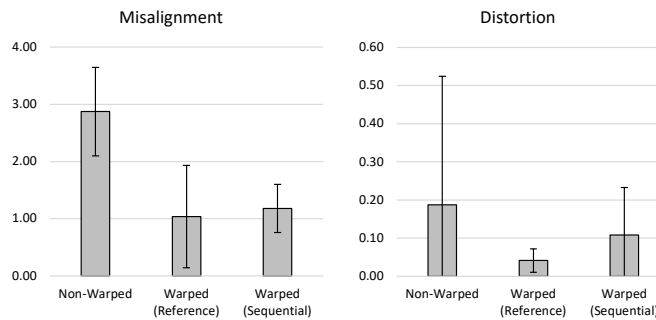


Figure 4.31: Comparison of maps created from original frames and from warped frames. The highlighted regions show the visual difference in the misalignment of the non-warped and warped version.



(a)



(b)

(c)

Figure 4.32: Scores for comparison of non-warped and warped maps averaged over all frames.

4.7 Pipeline Validation: Activity Measurements

In a recording with repetitive movement, we can have reoccurring regions that sometimes have cells active and other times not. It is important that the algorithm still identifies it as the same region, which usually happens if other distinctive structures are present on the surroundings, as shown in Figure 4.33.

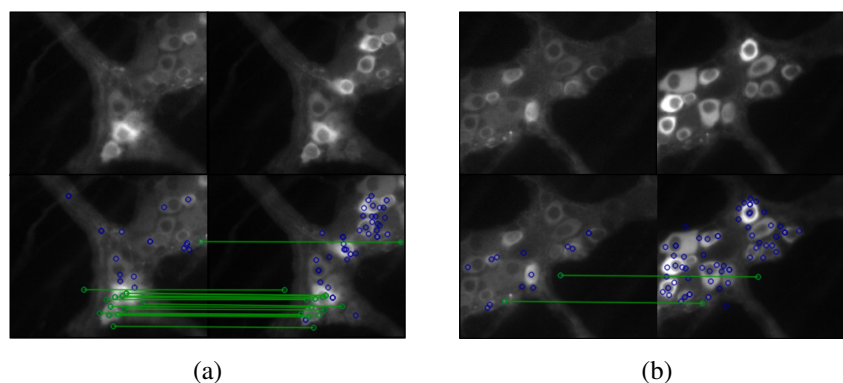


Figure 4.33: Illustrative example of the effect of activity in the feature matching process of the same region. (a) Represents some differences in activity with common sections. (b) Represents a complete difference in activity.

To study activity measurements in maps resulting from the pipeline, we followed the experiment as described in Figure 3.5. The results from the artificial distortion and the correction that was applied afterwards are presented in Figure 4.34 and the activity measurements for two ROIs are shown in Figures 4.35a-4.35b. The grey line for the original recording demonstrates the need for registration since the slight up and down movement on the original recording causes an erroneous quantification. When comparing the original corrected to the distorted correction, there is a clear similarity, which is ideal since we do not want the pipeline to affect possible measurements done in its results.

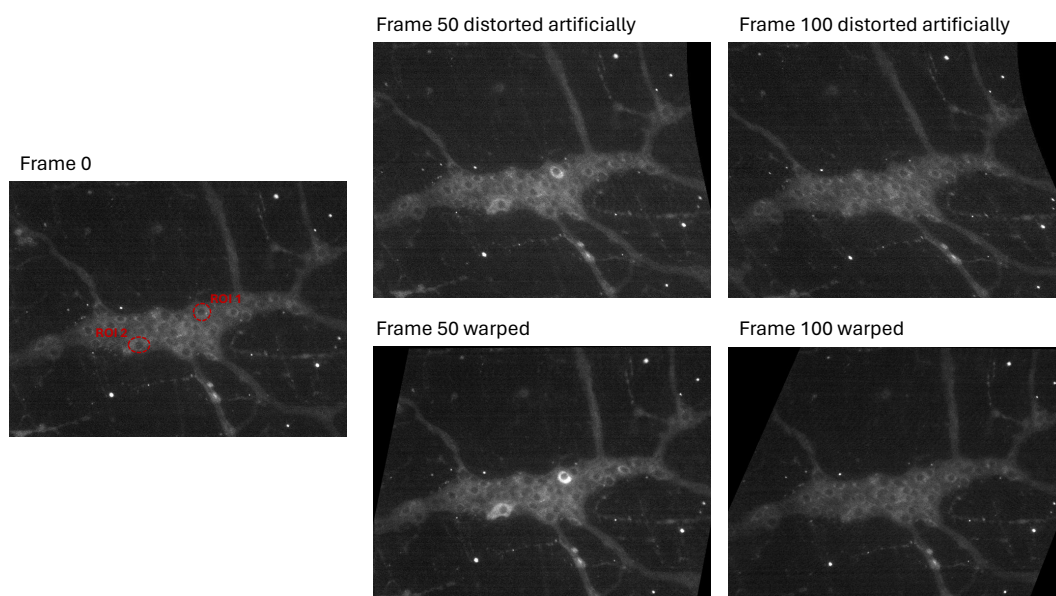


Figure 4.34: Example of a recording with activity present that was distorted artificially. Top row: frames 50 and 100 (out of 200) in the distorted recording. Bottom row: frames 50 and 100 (out of 200) in the warped corrected version. The warping correction was achieved using frame 0 as a reference for all subsequent frames. Supplementary material: [original recording](#), [distorted version](#) and [warped corrected version](#).

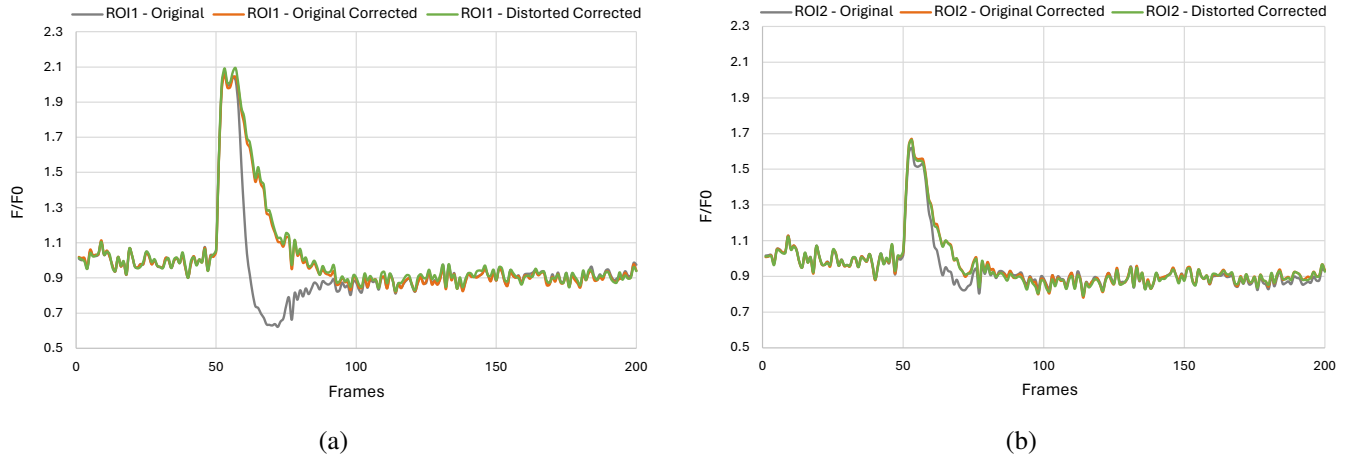


Figure 4.35: Activity registered from the recordings including the original not corrected, original corrected and distorted corrected. (a) ROI1 (b) ROI2

4.8 Pipeline Validation: Intravital Recordings

Finally, to showcase a very complex use of the pipeline, intravital recordings of the intestine imaged through an abdominal window were used. These recordings bring to light any type of issue previously discussed from noise to low contrast, out of focus events and repetitive movement, distortion from the contracting gut (shown in Figure 4.36) and cell activity.

The careful use of specific modules and tuning of parameters is essential for these recordings. In Figure 4.37 we demonstrate how using certain modules can drastically improve the result of the final map. Pre-processing in this case is essential to guarantee feature detection and matching, while warping and clustering help to ensure the alignment is done correctly and as seamlessly as possible. Some examples of maps from these recordings can be explored in Figure 4.38.

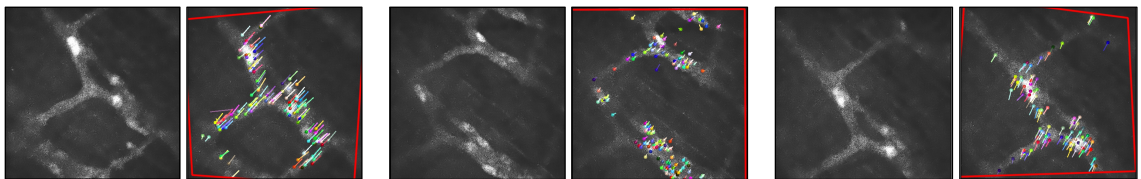


Figure 4.36: Some examples of the distortion that occurs from the contraction of the gut in intravital recordings. The pair of frames are sequential, the red square represents the homography matrix and the points represent the keypoints identified and matched in both frames as well as their trajectory.

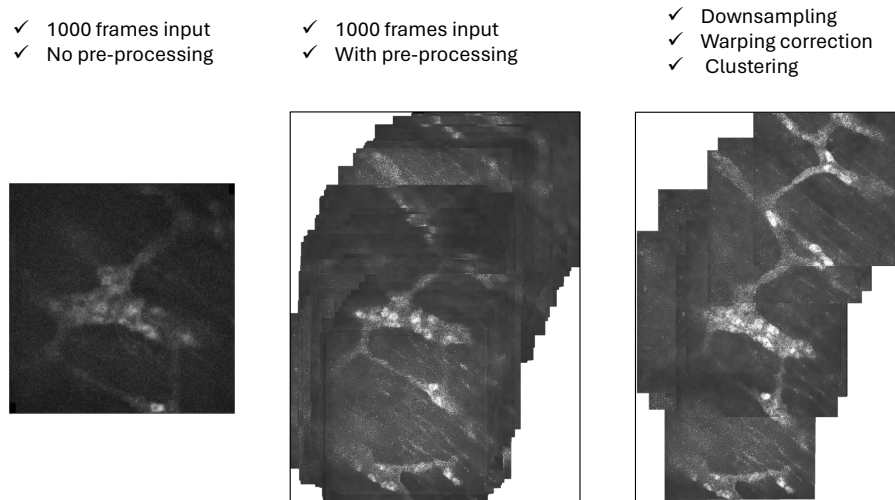


Figure 4.37: Illustration of the impact of using different modules of the pipeline. From left to right: inputting the 1000 frames with no pre-processing into the alignment algorithm; using pre-processing to denoise and improve brightness and contrast and therefore improve feature detection and matching; downsampling and clustering for a clearer composition.

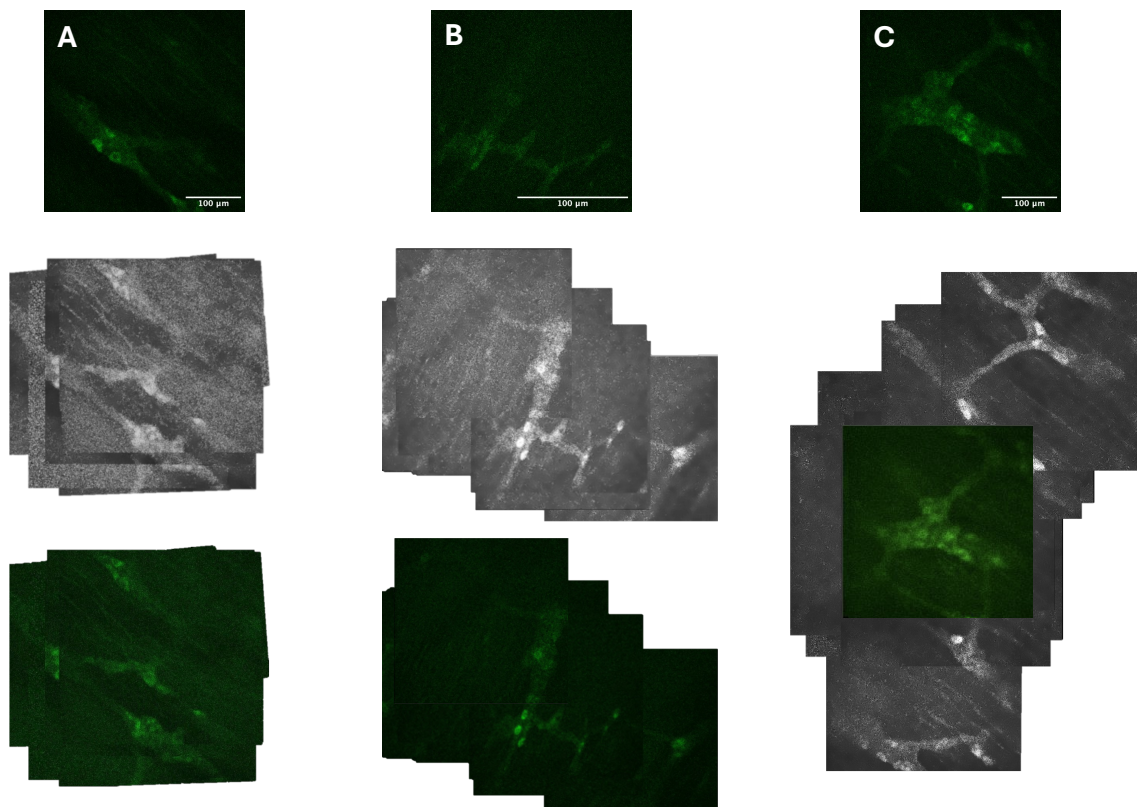


Figure 4.38: Examples (A and B) of maps obtained from intravital recordings. (C) shows an example of the projection of frames in the final map. Supplementary material: original recordings [A](#), [B](#) and [C](#).

4.9 User Interface

The development of this pipeline was translated into a user friendly interface that allows for further exploration of the several modules and extra results available. As a tool, the pipeline presented in this thesis does not work the same for every type of recording. This way, every step is easily manipulated by the user to achieve the best results for their specific recording.

From the menu on the side, the user can upload the frames (takes any folder with formats .png, .jpeg, .tiff). Afterwards, they can either immediately align the frames using the stitching algorithm or apply some pre-stitching modules before. Whatever operation the user decides to perform the frames will be continuously updated to their most recent state until they are finally aligned.

For the pre-stitching modules, extra windows will appear with the corresponding options and the results will appear in an extra tab in the "Frames" side. Similarly, for the extra results post-stitching, extra windows will appear for selection of which result the user wants to obtain and these will appear in an extra tab in the "Results" side. All these results can also be saved as images if necessary.

The output console on the lower part serves as a communication window for the steps being taken, any error message if something is not working and for printing additional resulting parameters if the user wishes to access them. Every optional input has a default value that will run as such if not changed. Therefore, the tool can be used as simply or as complexly as desired, within the possibility of the algorithm. Figure 4.39 shows the organization and functions of the main window and Figures 4.40 to 4.44 illustrate the other functions available in the interface.

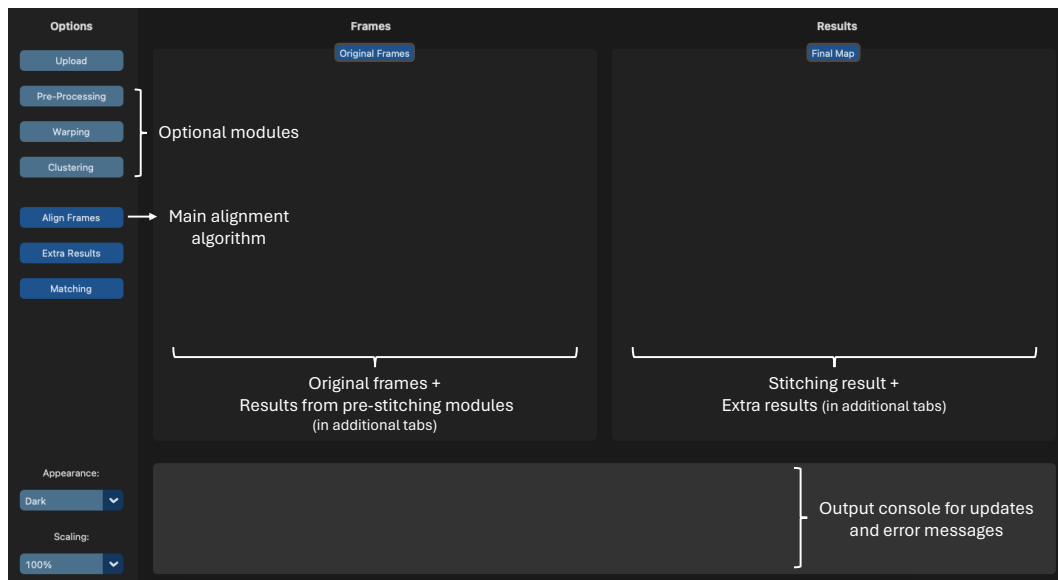


Figure 4.39: Main design and organization of the User Interface.

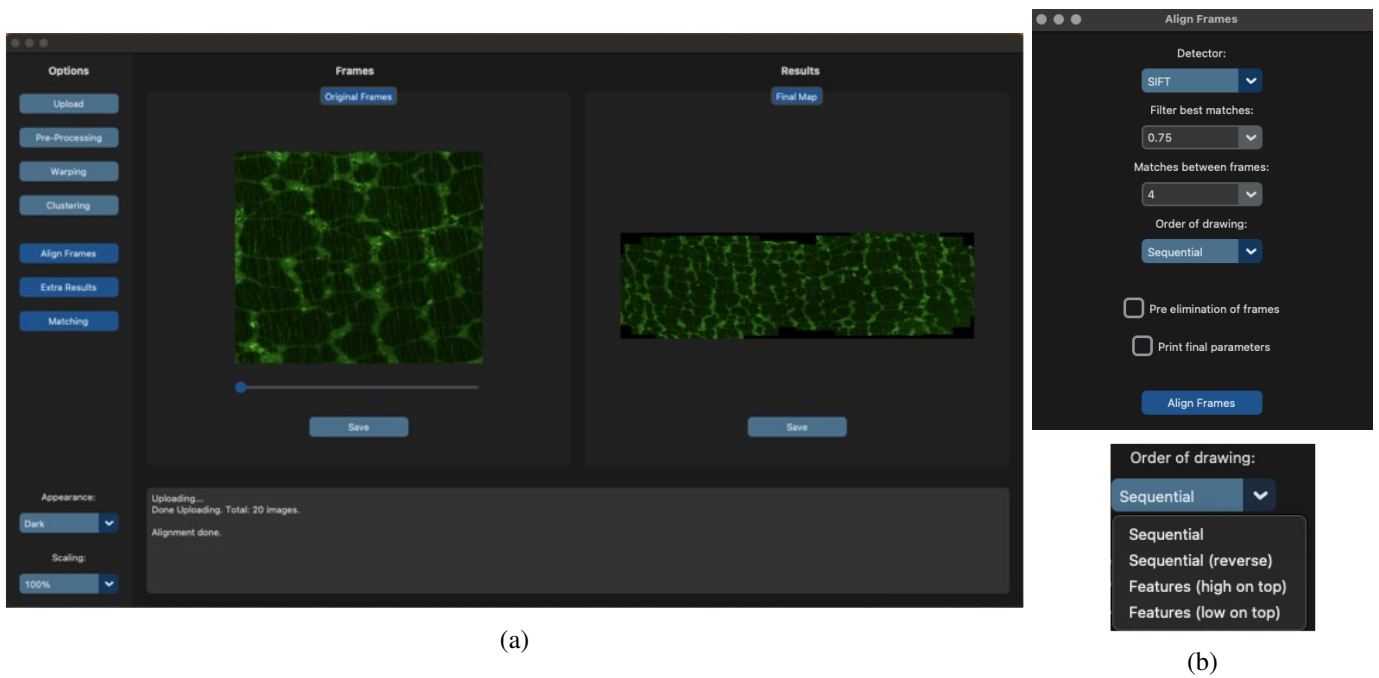


Figure 4.40: Main stitching algorithm. (a) Top level window to select options for stitching including the detector, the thresholds and the order of drawing. If print parameters is selected, the final coordinates of the map will appear in the console. (b) View of the interface after stitching, original frames on the left and final map on the right.

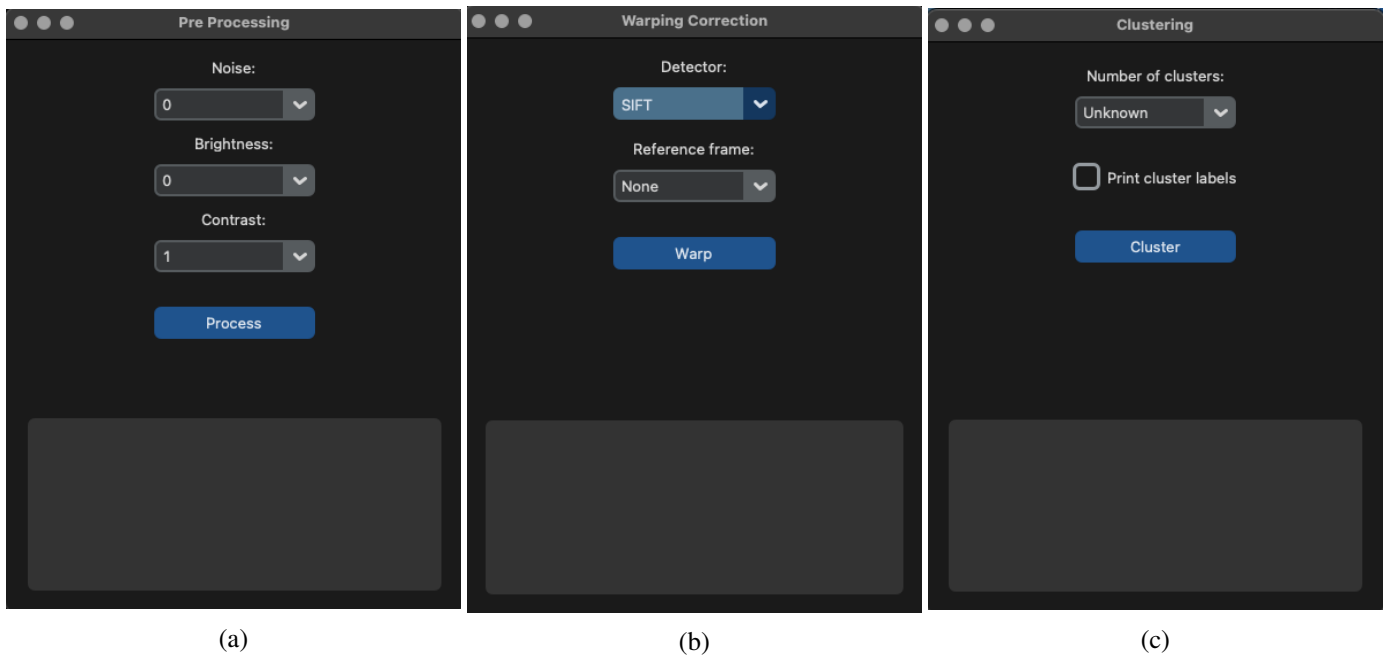


Figure 4.41: Top level windows for pre-stitching modules and their respective options. (a) Pre-processing. (b) Warping Correction. (c) Clustering.

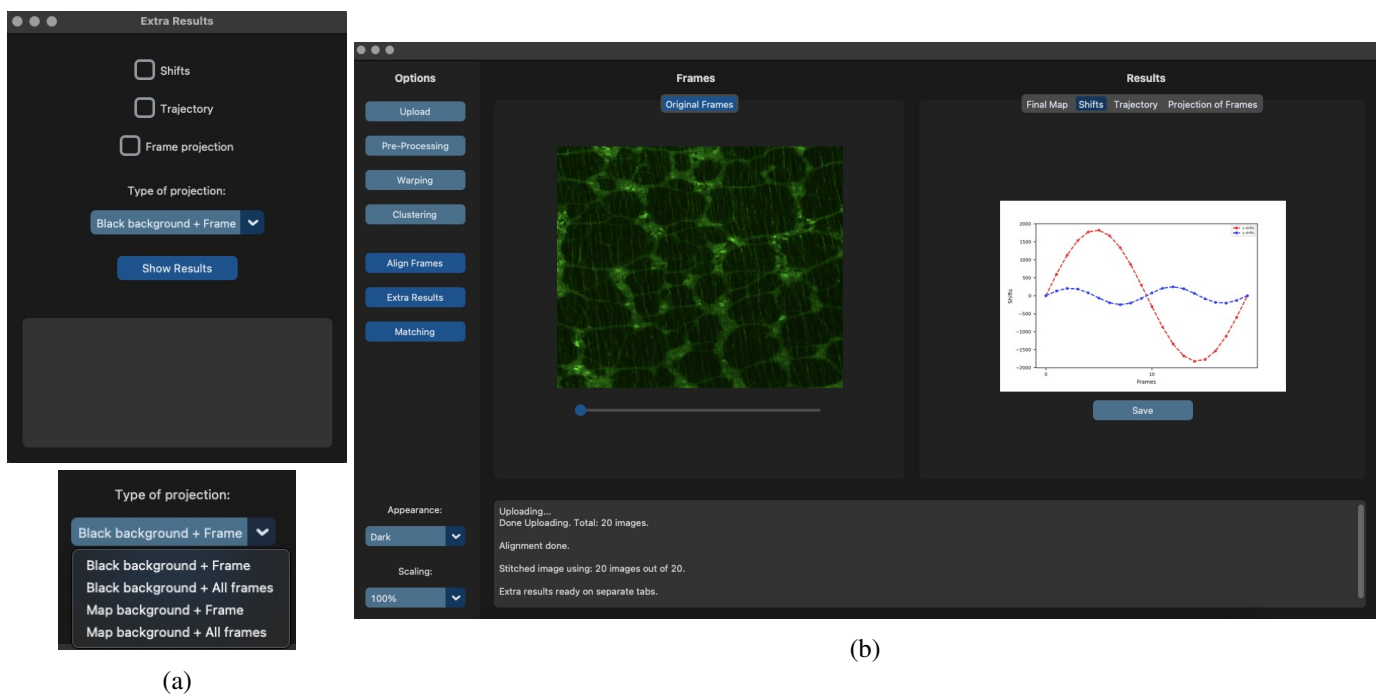


Figure 4.42: Extra results after stitching. (a) Top level window to select which results to present, if projection of frames is chosen, the type of projection can be selected. (b) View of the interface with the extra results which appear on extra tabs on the right. A graph of the x and y shifts is currently in view.

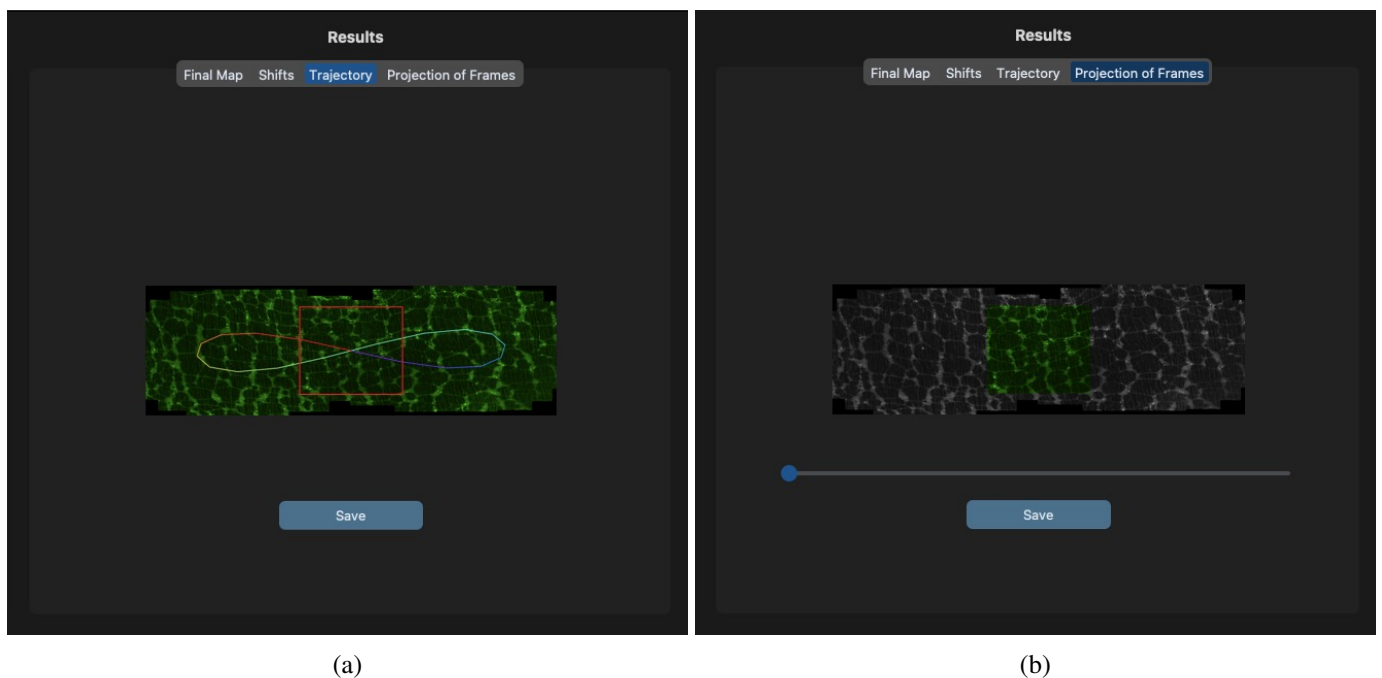


Figure 4.43: Other extra results that can be seen in their respective tabs. (a) Trajectory of the frames drawn in the map. (b) Projection of frames onto the final map, the result is a sequence of images with the same length as the original recording. Current view is the option 'Map background + Frame'. Supplementary material: projection result following other choices 1, 2, 3 or 4.

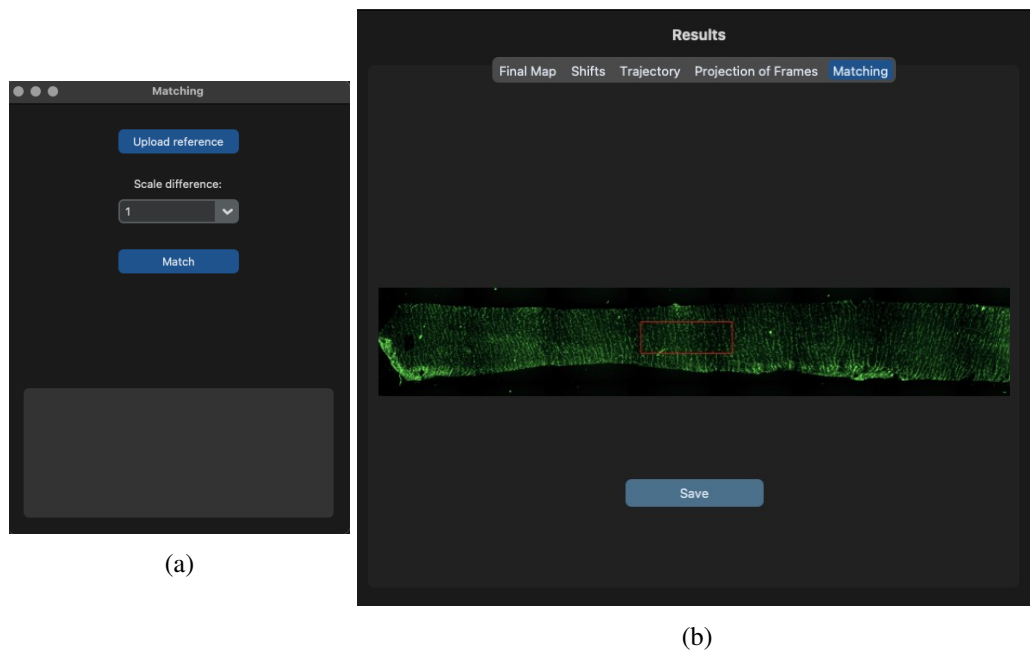


Figure 4.44: Module to match the final map with another image of a wider region. (a) Top level window with options for matching, a reference must be uploaded and the scale difference selected (1 if they are in the same magnification). (b) Result of matching represented with a rectangle in the region where the map was found.

Chapter 5

Discussion and Conclusions

In this thesis, we aimed to develop a tool for quantitative imaging of the contracting gut and explore its use in different types of challenges as well as validate it through various recordings. For validation of our pipeline, different experiments were outlined to obtain recordings with increasing complexity in order to tackle several challenges.

In the first dataset, in which fixed samples were imaged, the goal was to enable microscope automation in order to create recordings precisely controlled so that the coordinates of each frame in the map were known as a ground truth. The scores obtained on average for the total set of recordings were 99% pixel match, 49% SIFT matches, 88% trajectory match and 0.91 SSIM (the value of SSIM varies from -1 to 1 where 1 indicates perfect similarity). The lower value on SIFT matches can be easily explained by the inclusion of out of focus recordings, to which SIFT is very sensitive. As shown in the results, out of focus frames will have fewer interest keypoints and therefore fewer matches. Without the out of focus images, the SIFT score was about 79%. Additionally, these tests showed that the feature based stitching algorithm utilized all the frames available in the recordings, including out of focus and reoccurring ones. As a first step, this allowed us to verify the validity of the stitching algorithm on feature rich frames with no gut movement.

Afterwards, a different experimental preparation of the gut tissue allowed us to prepare recordings where the intrinsic movement of the muscle layer was present. For this dataset, the goal was to explore how the pre-stitching module like warping could aid in the better alignment of distorted frames. Since there was no absolute ground truth, the scores were obtained using the individual frames as a reference and averaging them over the total number of frames. The match scores for all the recordings, including the ones with large and small shifts were 82% pixel match, 52% SIFT matches and 0.68 SSIM. Although the scores might reflect a somewhat good alignment, the visual analysis of the maps created showed that recordings with large shifts and more distortions gave rise to misaligned maps with a lot of discontinuities in their edges. Hence the first assumption that warping would be needed proved to be true.

After warping the recordings, new maps were created and compared with the original. In the match scores there was not a significant difference, so other parameters were explored according to the literature including misalignment and structural distortion. For misalignment, the non-warped

maps had an average of 2.87 while the warped maps had an average of 1.10. For structural distortion the same pattern was found with the non-warped maps presenting an average value of 0.19 and for the warped maps this average was 0.07. These parameters corroborated the visual perception that the warping could improve not only the cleanliness of the final map but also maintain the correct structural information of the region.

The third test was created to guarantee that measurements such as activity can still be extracted from in-time ordered maps, which eventually result from the pipeline. By distorting recordings artificially and then correcting them, we simulated a challenge that can occur in real moving recordings. In the evaluation of the activity traces for different ROIs, either extracted from the original or from the corrected distorted images, we showed that they were nearly identical and that if the distortion is well corrected, measurements such as this are still possible and reliable.

Lastly, with the intravital data we wanted to show a potential extremely challenging application of this pipeline when working with recordings that are noisy, highly distorted and moving uncontrollably. This data was crucial in showcasing the need for a more curated pipeline with extra steps, including the need for pre-processing to improve feature detection and the need for clustering and downsampling in recordings with a large number of frames that show repetitive movements. Only by using a combination of all these steps can we get a cleaner, more readable and, above all, more correct final map.

The incorporation of different data types in this work allowed us to showcase different abilities. However, for the ones such as intravital recordings, the results were not as pristine as others which was expected due to their complex nature.

The work developed throughout this thesis was a combination of different image analysis techniques all gathered together with one goal: how can we facilitate the analysis of recordings with such levels of complexity and how can we extract the most information out of them. As with every developed tool, each step of the pipeline has its importance but also has their disadvantages and flaws that could be upgraded with time and with improved algorithms. However, we consider that presenting a line of thought that integrates all these different aspects into one working tool brings us one step forward in the future analysis of recordings like the ones shown here or even in other similar circumstances.

5.1 Challenges and Future Work

As with any work, there is always a margin for improvement and in this particular case, we can divide the pipeline into its modules and improve each one in order to make the whole tool more efficient. Looking at every step of this pipeline as a necessary one for its success in dealing with complex datasets, future work could take many routes in improving small steps like feature detection or higher levels such as the preparation of the data before stitching and the stitching algorithm itself.

In the warping trials, it was noticed that if we have very few matches between frames this might result in an incorrect homography estimation which will consequently provide a wrong

transformation to the frame and damage its alignment in the final map or make it impossible to align at all. More trials should be made to better understand the different methods of estimating homography as well as finding a way to filter out what are considered incorrect matrices to avoid wrong transformations.

For the clustering, the approach with the VGG16 was chosen as the most adequate for the dataset tested but it might work differently in other types of data so other methods could be further implemented and tested to guarantee that the separation of frames according to their similarity helps the stitching process. Other types of partitioning of the frames could also be tested.

Regarding the stitching itself, numerous improvements in stitching algorithms have been published and are still being studied. Therefore, that main step in the pipeline could always be improved even to the point of incorporating tasks performed by the pre-stitching modules presented here.

At an even higher level, we could try changes in the pipeline itself like allowing for warping inside each cluster or for different types of pre-processing among the frames. To deal with very complex distortions we could also treat the frames as several different patches which could improve its alignment, but this will make the pipeline also more computationally expensive.

Finally, for the extra post-stitching modules, other results could be designed taking into account all the information we can gather from the algorithms presented.

Nevertheless, we consider the suggested solution to be an elegant and highly adaptable way of dealing with such different characteristics. By creating something that flexible one could lose the accuracy against very particular methods for one type of data while in return winning the advantage of the method not being overly specific.

References

- [1] Candice Fung and Pieter Vanden Berghe. Functional circuits and signal processing in the enteric nervous system. *Cellular and Molecular Life Sciences*, 77(22):4505–4522, November 2020. doi: 10.1007/s00018-020-03543-6.
- [2] Nick J. Spencer and Hongzhen Hu. Enteric nervous system: sensory transduction, neural circuits and gastrointestinal motility. *Nature Reviews Gastroenterology & Hepatology*, 17(6):338–351, June 2020. doi: 10.1038/s41575-020-0271-2.
- [3] John B. Furness. The enteric nervous system and neurogastroenterology. *Nature Reviews Gastroenterology & Hepatology*, 9(5):286–294, May 2012. ISSN 1759-5045, 1759-5053. doi: 10.1038/nrgastro.2012.32.
- [4] Meenakshi Rao and Michael D Gershon. The bowel and beyond: the enteric nervous system in neurological disorders. *Nature reviews Gastroenterology & hepatology*, 13(9):517–528, 2016.
- [5] Marilia Carabotti, Annunziata Scirocco, Maria Antonietta Maselli, and Carola Severi. The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Annals of Gastroenterology*, 28(2):203–209, 2015.
- [6] Sabine Schneider, Christina M. Wright, and Robert O. Heuckeroth. Unexpected roles for the second brain: Enteric nervous system as master regulator of bowel function. *Annual Review of Physiology*, 81(Volume 81, 2019):235–259, 2019. doi: 10.1146/annurev-physiol-021317-121515.
- [7] Keith A. Sharkey and Gary M. Mawe. The enteric nervous system. *Physiological Reviews*, 103(2):1487–1564, April 2023. doi: 10.1152/physrev.00018.2022.
- [8] Werend Boesmans, Marlene M. Hao, and Pieter Vanden Berghe. Optogenetic and chemogenetic techniques for neurogastroenterology. *Nature Reviews Gastroenterology & Hepatology*, 15(1):21–38, 2018. ISSN 1759-5045, 1759-5053. doi: 10.1038/nrgastro.2017.151.
- [9] Mallikarjunarao Ganesana, Scott T. Lee, Ying Wang, and B. Jill Venton. Analytical Techniques in Neuroscience: Recent Advances in Imaging, Separation, and Electrochemical Methods. *Analytical Chemistry*, 89(1):314–341, 2017. doi: 10.1021/acs.analchem.6b04278.

- [10] Richard H. Roth and Jun B. Ding. From Neurons to Cognition: Technologies for Precise Recording of Neural Activity Underlying Behavior. *BME Frontiers*, 2020:7190517, 2020. ISSN 2765-8031. doi: 10.34133/2020/7190517.
- [11] Ziqiang Wei, Bei-Jung Lin, Tsai-Wen Chen, Kayvon Daie, Karel Svoboda, and Shaul Druckmann. A comparison of neuronal population dynamics measured with calcium imaging and electrophysiology. *PLOS Computational Biology*, 16(9):e1008198, 2020. doi: 10.1371/journal.pcbi.1008198.
- [12] Christine Grienberger and Arthur Konnerth. Imaging Calcium in Neurons. *Neuron*, 73(5): 862–885, March 2012. doi: 10.1016/j.neuron.2012.02.011.
- [13] Tsai-Wen Chen, Trevor J. Wardill, Yi Sun, Stefan R. Pulver, Sabine L. Renninger, Amy Bao-han, Eric R. Schreiter, Rex A. Kerr, Michael B. Orger, Vivek Jayaraman, Loren L. Looger, Karel Svoboda, and Douglas S. Kim. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, 499(7458):295–300, July 2013. doi: 10.1038/nature12354.
- [14] Junichi Nakai, Masamichi Ohkura, and Keiji Imoto. A high signal-to-noise Ca²⁺ probe composed of a single green fluorescent protein. *Nature Biotechnology*, 19(2):137–141, 2001. doi: 10.1038/84397.
- [15] Lin Tian, S Andrew Hires, Tianyi Mao, Daniel Huber, M Eugenia Chiappe, Sreekanth H Chalasani, Leopoldo Petreanu, Jasper Akerboom, Sean A McKinney, Eric R Schreiter, Cornelia I Bargmann, Vivek Jayaraman, Karel Svoboda, and Loren L Looger. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods*, 6(12): 875–881, 2009. doi: 10.1038/nmeth.1398.
- [16] Hod Dana, Yi Sun, Boaz Mohar, Brad K. Hulse, Aaron M. Kerlin, Jeremy P. Hasseman, Getahun Tsegaye, Arthur Tsang, Allan Wong, Ronak Patel, John J. Macklin, Yang Chen, Arthur Konnerth, Vivek Jayaraman, Loren L. Looger, Eric R. Schreiter, Karel Svoboda, and Douglas S. Kim. High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nature Methods*, 16(7):649–657, 2019. ISSN 1548-7091, 1548-7105. doi: 10.1038/s41592-019-0435-6.
- [17] Yan Zhang, Márton Rózsa, Yajie Liang, Daniel Bushey, Ziqiang Wei, Jihong Zheng, Daniel Reep, Gerard Joey Broussard, Arthur Tsang, Getahun Tsegaye, Sujatha Narayan, Christopher J. Obara, Jing-Xuan Lim, Ronak Patel, Rongwei Zhang, Misha B. Ahrens, Glenn C. Turner, Samuel S.-H. Wang, Wyatt L. Korff, Eric R. Schreiter, Karel Svoboda, Jeremy P. Hasseman, Ilya Kolb, and Loren L. Looger. Fast and sensitive GCaMP calcium indicators for imaging neural populations. *Nature*, 615(7954):884–891, March 2023. doi: 10.1038/s41586-023-05828-9.
- [18] Werend Boesmans, Michiel A. Martens, Nathalie Weltens, Marlene M. Hao, Jan Tack, Carla Cirillo, and Pieter Vanden Berghe. Imaging neuron-glia interactions in the enteric nervous

- system. *Frontiers in Cellular Neuroscience*, 7, 2013. ISSN 1662-5102. doi: 10.3389/fncel.2013.00183.
- [19] Grant W. Hennig, Thomas W. Gould, Sang Don Koh, Robert D. Corrigan, Dante J. Heredia, Matthew C. Shonnard, and Terence K. Smith. Use of Genetically Encoded Calcium Indicators (GECIs) Combined with Advanced Motion Tracking Techniques to Examine the Behavior of Neurons and Glia in the Enteric Nervous System of the Intact Murine Colon. *Frontiers in Cellular Neuroscience*, 9, 2015. doi: 10.3389/fncel.2015.00436.
- [20] Zhiling Li, Marlene M Hao, Chris Van den Haute, Veerle Baekelandt, Werend Boesmans, and Pieter Vanden Berghe. Regional complexity in enteric neuron wiring reflects diversity of motility patterns in the mouse large intestine. *eLife*, 8:e42914, February 2019. doi: 10.7554/eLife.42914.
- [21] A.-S. Desmet, C. Cirillo, and P. Vanden Berghe. Distinct subcellular localization of the neuronal marker HuC/D reveals hypoxia-induced damage in enteric neurons. *Neurogastroenterology and Motility*, 26(8):1131–1143, August 2014. doi: 10.1111/nmo.12371.
- [22] Yanfen Jiang, Hui Dong, Lars Eckmann, Elaine M. Hanson, Katherine C. Ihn, and Ravinder K. Mittal. Visualizing the enteric nervous system using genetically engineered double reporter mice: Comparison with immunofluorescence. *PLoS ONE*, 12(2):e0171239, February 2017. doi: 10.1371/journal.pone.0171239.
- [23] Phillip A. Baker, Matthew D. Meyer, Ashley Tsang, and Rosa A. Uribe. Immunohistochemical and ultrastructural analysis of the maturing larval zebrafish enteric nervous system reveals the formation of a neuropil pattern. *Scientific Reports*, 9(1):6941, 2019. doi: 10.1038/s41598-019-43497-9.
- [24] Eftychios A Pnevmatikakis. Analysis pipelines for calcium imaging data. *Current Opinion in Neurobiology*, 55:15–21, April 2019. doi: 10.1016/j.conb.2018.11.004.
- [25] Carsen Stringer and Marius Pachitariu. Computational processing of neural recordings from calcium imaging data. *Current Opinion in Neurobiology*, 55:22–31, April 2019. doi: 10.1016/j.conb.2018.11.005.
- [26] P. Thevenaz, U.E. Ruttimann, and M. Unser. A pyramid approach to subpixel registration based on intensity. *IEEE Transactions on Image Processing*, 7(1):27–41, 1998. doi: 10.1109/83.650848.
- [27] Alexander Dubbs, James Guevara, and Rafael Yuste. moco: Fast Motion Correction for Calcium Imaging. *Frontiers in Neuroinformatics*, 10, 2016. doi: 10.3389/fninf.2016.00006.
- [28] David S. Greenberg and Jason N.D. Kerr. Automated correction of fast motion artifacts for two-photon imaging of awake animals. *Journal of Neuroscience Methods*, 176(1):1–15, 2009. doi: 10.1016/j.jneumeth.2008.08.020.

- [29] Eftychios A. Pnevmatikakis and Andrea Giovannucci. NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *Journal of Neuroscience Methods*, 291:83–94, November 2017. doi: 10.1016/j.jneumeth.2017.07.031.
- [30] Bradley B. Barth, Emily R. Redington, Nitisha Gautam, Nicole A. Pelot, and Warren M. Grill. Calcium image analysis in the moving gut. *Neurogastroenterology & Motility*, 35(12): e14678, August 2023. doi: 10.1111/nmo.14678.
- [31] Weiyi Liu, Junxia Pan, Yuanxu Xu, Meng Wang, Hongbo Jia, Kuan Zhang, Xiaowei Chen, Xingyi Li, and Xiang Liao. Fast and Accurate Motion Correction for Two-Photon Ca²⁺ Imaging in Behaving Mice. *Frontiers in Neuroinformatics*, 16:851188, April 2022. doi: 10.3389/fninf.2022.851188.
- [32] Mingkang Li, Changhao Liu, Xin Cui, Hayoung Jung, Heecheon You, Linqing Feng, and Shaomin Zhang. An Open-Source Real-Time Motion Correction Plug-In for Single-Photon Calcium Imaging of Head-Mounted Microscopy. *Frontiers in Neural Circuits*, 16:891825, June 2022. doi: 10.3389/fncir.2022.891825.
- [33] Victoria A. Griffiths, Antoine M. Valera, Joanna Yn. Lau, Hana Roš, Thomas J. Younts, Bóris Marin, Chiara Baragli, Diccon Coyle, Geoffrey J. Evans, George Konstantinou, Theo Koimtzis, K. M. Naga Srinivas Nadella, Sameer A. Punde, Paul A. Kirkby, Isaac H. Bianco, and R. Angus Silver. Real-time 3D movement correction for two-photon imaging in behaving animals. *Nature Methods*, 17(7):741–748, July 2020. doi: 10.1038/s41592-020-0851-7.
- [34] Eran A. Mukamel, Axel Nimmerjahn, and Mark J. Schnitzer. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron*, 63(6):747–760, 2009. ISSN 0896-6273. doi: 10.1016/j.neuron.2009.08.009.
- [35] Ryuichi Maruyama, Kazuma Maeda, Hajime Moroda, Ichiro Kato, Masashi Inoue, Hiroyoshi Miyakawa, and Toru Aonishi. Detecting cells using non-negative matrix factorization on calcium imaging data. *Neural Networks*, 55:11–19, 2014. ISSN 0893-6080. doi: 10.1016/j.neunet.2014.03.007.
- [36] Eftychios A. Pnevmatikakis, Daniel Soudry, Yuanjun Gao, Timothy A. Machado, Josh Merel, David Pfau, Thomas Reardon, Yu Mu, Clay Lacefield, Weijian Yang, Misha Ahrens, Randy Bruno, Thomas M. Jessell, Darcy S. Peterka, Rafael Yuste, and Liam Paninski. Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron*, 89(2):285–299, January 2016. doi: 10.1016/j.neuron.2015.11.037.
- [37] Marius Pachitariu, Carsen Stringer, Mario Dipoppa, Sylvia Schröder, L. Federico Rossi, Henry Dalglish, Matteo Carandini, and Kenneth D. Harris. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv*, 2017. doi: 10.1101/061507.
- [38] Andrea Giovannucci, Johannes Friedrich, Pat Gunn, Jérémie Kalfon, Brandon L. Brown, Sue Ann Koay, Jiannis Taxidis, Farzaneh Najafi, Jeffrey L. Gauthier, Pengcheng Zhou,

- Baljit S. Khakh, David W. Tank, Dmitri B. Chklovskii, and Eftychios A. Pnevmatikakis. CaImAn an open source tool for scalable calcium imaging data analysis. *eLife*, 8:e38173, January 2019. ISSN 2050-084X. doi: 10.7554/eLife.38173.
- [39] An-Sofie Desmet, Carla Cirillo, Jan Tack, Wim Vandenberghe, and Pieter Vanden Berghe. Live calcium and mitochondrial imaging in the enteric nervous system of parkinson patients and controls. *Elife*, 6:e26850, 2017.
- [40] Philipp Berens, Jeremy Freeman, Thomas Deneux, Nicolay Chenkov, Thomas McColgan, Artur Speiser, Jakob H. Macke, Srinivas C. Turaga, Patrick Mineault, Peter Rupprecht, Stephan Gerhard, Rainer W. Friedrich, Johannes Friedrich, Liam Paninski, Marius Pachitariu, Kenneth D. Harris, Ben Bolte, Timothy A. Machado, Dario Ringach, Jasmine Stone, Luke E. Rogerson, Nicolas J. Sofroniew, Jacob Reimer, Emmanouil Froudarakis, Thomas Euler, Mirosław Román Rosón, Lucas Theis, Andreas S. Tolias, and Matthias Bethge. Community-based benchmarking improves spike rate inference from two-photon calcium imaging data. *bioRxiv*, 2018. doi: 10.1101/177956.
- [41] Biafra Ahanonu. Table of imaging analysis methods and software, September 2023. URL [10.5281/zenodo.8349533](https://doi.org/10.5281/zenodo.8349533).
- [42] Johannes Friedrich, Pengcheng Zhou, and Liam Paninski. Fast online deconvolution of calcium imaging data. *PLOS Computational Biology*, 13(3):1–26, 03 2017. doi: 10.1371/journal.pcbi.1005423.
- [43] Luke Sorensen, Adam Humenick, Sabrina S.B. Poon, Myat Noe Han, Narges Sadat Mahdavian, Ryan Hamnett, Estibaliz Gómez de Mariscal, Peter H. Neckel, Ayame Saito, Keith Mutundwe, Christie Glennan, Robert Haase, Rachel M. McQuade, Jaime P.P. Foong, Simon J.H. Brookes, Julia A. Kaltschmidt, Arrate Muñoz-Barrutia, Sebastian K. King, Nicholas A. Veldhuis, Simona E. Carbone, Daniel P. Poole, and Pradeep Rajasekhar. Gut analysis toolbox: Automating quantitative analysis of enteric neurons. *bioRxiv*, 2024. doi: 10.1101/2024.01.17.576140.
- [44] Achala Pandey and Umesh C. Pati. Image mosaicing: A deeper insight. *Image and Vision Computing*, 89:236–257, September 2019. doi: 10.1016/j.imavis.2019.07.002.
- [45] Kivanc Kose, Mengran Gou, Oriol Yélamos, Miguel Cordova, Anthony M. Rossi, Kishwer S. Nehal, Eileen S. Flores, Octavia Camps, Jennifer G. Dy, Dana H. Brooks, and Milind Rajadhyaksha. Automated video-mosaicking approach for confocal microscopic imaging in vivo: an approach to address challenges in imaging living tissue and extend field of view. *Scientific Reports*, 7(1):10759, September 2017. doi: 10.1038/s41598-017-11072-9.
- [46] Chengbo Yin, Linpeng Wei, Kivanc Kose, Adam K. Glaser, Gary Peterson, Milind Rajadhyaksha, and Jonathan T. C. Liu. Real-time video mosaicking to guide handheld in vivo microscopy. *Journal of Biophotonics*, 13(6):e202000048, 2020. doi: 10.1002/jbio.202000048.

- [47] Stephan Preibisch, Stephan Saalfeld, and Pavel Tomancak. Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics*, 25(11):1463–1465, June 2009. doi: 10.1093/bioinformatics/btp184.
- [48] Binglu Zhao, Kaidi Zhang, Peng Liu, and Yuhang Chen. Large-scale time-lapse scanning electron microscopy image mosaic using a smooth stitching strategy. *Microscopy Research and Technique*, 86(8):929–942, 2023. doi: 10.1002/jemt.24334.
- [49] David Capel. *Image mosaicing*. Springer, 2004. doi: 10.1007/978-0-85729-384-8.
- [50] Debabrata Ghosh and Naima Kaabouch. A survey on image mosaicing techniques. *Journal of Visual Communication and Image Representation*, 34:1–11, January 2016. doi: 10.1016/j.jvcir.2015.10.014.
- [51] Dushyant Vaghela and Prof Naina. A review of image mosaicing techniques. *arXiv preprint arXiv:1405.2539*, 2014. doi: 10.48550/arXiv.1405.2539.
- [52] Adrien Bartoli. Groupwise geometric and photometric direct image registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(12):2098 – 2108, 2008. doi: 10.1109/TPAMI.2008.22.
- [53] Weiqing Yan, Yiqiu Sun, Wujie Zhou, Zhaowei Liu, and Runmin Cong. Deep Video Stabilization via Robust Homography Estimation. *IEEE Signal Processing Letters*, 30:1602–1606, 2023. doi: 10.1109/LSP.2023.3329436.
- [54] Richa Singh, Mayank Vatsa, Arun Ross, and Afzel Noore. A mosaicing scheme for pose-invariant face recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(5):1212–1225, 2007. doi: 10.1109/TSMCB.2007.903537.
- [55] Xiaomin Ma, Ding Liu, Jian Zhang, and Jing Xin. A fast affine-invariant features for image stitching under large viewpoint changes. *Neurocomputing*, 151:1430–1438, March 2015. doi: 10.1016/j.neucom.2014.10.045.
- [56] Richard Szeliski. Feature detection and matching. In *Computer Vision: Algorithms and Applications*, pages 333–399, Cham, 2022. Springer International Publishing. ISBN 978-3-030-34372-9. doi: 10.1007/978-3-030-34372-9_7.
- [57] Chris Harris, Mike Stephens, et al. A combined corner and edge detector. In *Alvey vision conference*, volume 15, pages 10–5244. Citeseer, 1988.
- [58] David G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, November 2004. doi: 10.1023/B:VISI.0000029664.99615.94.
- [59] Jian Wu, Zhiming Cui, Victor S. Sheng, Pengpeng Zhao, Dongliang Su, and Shengrong Gong. A Comparative Study of SIFT and its Variants. *Measurement Science Review*, 13(3): 122–131, 2013. doi: 10.2478/msr-2013-0021.

- [60] Jun Luo, Yong Ma, Erina Takikawa, Shihong Lao, Masato Kawade, and Bao-Liang Lu. Person-Specific SIFT Features for Face Recognition. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, volume 2, pages II–593–II–596, April 2007. doi: 10.1109/ICASSP.2007.366305.
- [61] Yan Liu, Jiawei Tian, Rongrong Hu, Bo Yang, Shan Liu, Lirong Yin, and Wenfeng Zheng. Improved Feature Point Pair Purification Algorithm Based on SIFT During Endoscope Image Stitching. *Frontiers in Neurorobotics*, 16:840594, 2022. doi: 10.3389/fnbot.2022.840594.
- [62] Surbhi Gupta, Kutub Thakur, and Munish Kumar. 2D-human face recognition using SIFT and SURF descriptors of face’s feature regions. *The Visual Computer*, 37(3):447–456, 2021. doi: 10.1007/s00371-020-01814-8.
- [63] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3):346–359, 2008. ISSN 1077-3142. doi: 10.1016/j.cviu.2007.09.014. Similarity Matching in Computer Vision and Multimedia.
- [64] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer Vision – ECCV 2006*, pages 404–417, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. doi: 10.1007/11744023_32.
- [65] Edward Rosten and Tom Drummond. Machine learning for high-speed corner detection. In *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9*, pages 430–443. Springer, 2006.
- [66] Deepak Geetha Viswanathan. Features from accelerated segment test (fast). In *Proceedings of the 10th workshop on image analysis for multimedia interactive services, London, UK*, pages 6–8, 2009.
- [67] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: Binary robust independent elementary features. In *Computer Vision – ECCV 2010*, pages 778–792, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. doi: 10.1007/978-3-642-15561-1_56.
- [68] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In *2011 International Conference on Computer Vision*, pages 2564–2571, 2011. doi: 10.1109/ICCV.2011.6126544.
- [69] Kahleb D. Graham, Silvia Huerta López, Rajarshi Sengupta, Archana Shenoy, Sabine Schneider, Christina M. Wright, Michael Feldman, Emma Furth, Federico Valdivieso, Amanda Lemke, Benjamin J. Wilkins, Ali Naji, Edward J. Doolin, Marthe J. Howard, and Robert O. Heuckeroth. Robust, 3-Dimensional Visualization of Human Colon Enteric Nervous System Without Tissue Sectioning. *Gastroenterology*, 158(8):2221–2235.e5, 2020. doi: 10.1053/j.gastro.2020.02.035.

- [70] Kiran Maharana, Surajit Mondal, and Bhushankumar Nemade. A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1):91–99, June 2022. doi: 10.1016/j.gltp.2022.04.020.
- [71] Sidheswar Routray, Arun Kumar Ray, and Chandrabhanu Mishra. Analysis of various image feature extraction methods against noisy image: SIFT, SURF and HOG. *2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, February 2017.
- [72] José A. Rodríguez-Rodríguez, Ezequiel López-Rubio, Juan A. Ángel Ruiz, and Miguel A. Molina-Cabello. The impact of noise and brightness on object detection methods. *Sensors*, 24(3):821, January 2024. doi: 10.3390/s24030821.
- [73] Jasmine Irani, Nitin Pise, and Madhura Phatak. Clustering techniques and the similarity measures used in clustering: A survey. *International journal of computer applications*, 134(7):9–14, 2016.
- [74] Umme Sara, Morium Akter, and Mohammad Shorif Uddin. Image quality assessment through fsim, ssim, mse and psnr—a comparative study. *Journal of Computer and Communications*, 07(03):8–18, 2019. doi: 10.4236/jcc.2019.73002.
- [75] Seyoung Park and Hongyu Zhao. Spectral clustering based on learning similarity matrix. *Bioinformatics*, 34(12):2069–2076, February 2018. doi: 10.1093/bioinformatics/bty050.
- [76] Takayuki Fukui and Toshikazu Wada. *Commonality Preserving Image-Set Clustering Based on Diverse Density*, page 258–269. Springer International Publishing, 2014. doi: 10.1007/978-3-319-14249-4_25.
- [77] Banu Wirawan Yohanes. Images similarity based on bags of sift descriptor and k-means clustering. *Techne: Jurnal Ilmiah Elektroteknika*, 18(02):137–146, October 2019. doi: 10.31358/techne.v18i02.217.
- [78] Joris Guérin, Stéphane Thiery, Eric Nyiri, Olivier Gibaru, and Byron Boots. Combining pretrained cnn feature extractors to enhance clustering of complex natural images. *Neuro-computing*, 423:551–571, January 2021. doi: 10.1016/j.neucom.2020.10.068.
- [79] Adam Mansur. stitch2d, 2022. URL <https://github.com/adamancer/stitch2d>.
- [80] Krzysztof Okarma and Mateusz Kopytek. Improved combined metric for automatic quality assessment of stitched images. *Applied Sciences*, 12(20):10284, 2022.
- [81] Hayat Ullah, Sitara Afzal, and Imran Ullah Khan. Perceptual quality assessment of panoramic stitched contents for immersive applications: a prospective survey. *Virtual Reality & Intelligent Hardware*, 4(3):223–246, 2022.

- [82] Weiqing Yan, Guanghui Yue, Yuming Fang, Hua Chen, Chang Tang, and Gangyi Jiang. Perceptual objective quality assessment of stereoscopic stitched images. *Signal Processing*, 172:107541, 2020.
- [83] Chongzhen Tian, Xiongli Chai, and Feng Shao. Stitched image quality assessment based on local measurement errors and global statistical properties. *Journal of Visual Communication and Image Representation*, 81:103324, 2021.