Behavioral Analysis of Test Subjects Using Pose Detection

by

Ethan P. Pellegrini

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
(Artificial Intelligence)
in the University of Michigan-Dearborn
2023

Master's Thesis Committee:
Assistant Professor Jin Lu, Chair
Assistant Professor Niccolo Meneghetti
Assistant Professor Zhi Zhang

Acknowledgements

I would like to take this opportunity to express my gratitude to my advisor, Dr. Jin Lu, for spending so much time with me over the past year and guiding me through the process of writing this thesis. I’d also like to thank Dr. Zhi Zhang for providing us with this novel dataset and being so helpful and receptive to our questions. I would also like to extend my gratitude to my thesis committee for taking the time to evaluate this work. Finally, I’d like to thank my family and fiancée for being so loving and supportive while I’ve been attending school again.

Table of Contents

Acknowledgements ... ii

List of Figures ... iv

List of Tables ... vi

Abstract ... vii

Chapter 1: Introduction and Related Work ... 1

1.1 Introduction ... 1

1.2 Related Work ... 2

Chapter 2: Joint Tracking and Pose Analysis for Behavioral Classification ... 3

2.1 SLEAP Data Labeling ... 3

2.2 Embedding and Clustering ... 9

2.3 Fingerprinting ... 12

Chapter 3: Dataset and Results ... 16

3.1 Dataset ... 16

3.2 Results of Pipeline ... 17

3.2.1 Annotated Test Subject Video ... 17

3.2.2 Mean Wavelet Power During Each Behavior ... 21

Chapter 4: Conclusion and Future Work ... 25

4.1 Conclusion ... 25

4.2 Future Work ... 26

Appendix: Table of Expanded Behaviors for Future Work ... 27

References ... 28

List of Figures

Figure 1: The SLEAP Tool 4

Figure 2: Labeling a Frame of a Mouse Video with a Skeleton 6

Figure 3: Selecting Model Parameters for Model Training or Inference 7

Figure 4: Example of Labeled Frame with Predicted Skeleton Label 8

Figure 5: Embedding and Clustering Pipeline 9

Figure 6: PCA Estimation of Wavelet Decompositions 10

Figure 7: Emergent Clusters over Time 11

Figure 8: Fingerprinting 12

Figure 9: Joint Velocities over Time 13

Figure 10: Node Trajectories Transformed to the Frequency Domain 14

Figure 11:SHAM Male Performing Locomotion Behavior 18

Figure 12: SHAM Male Performing Investigating Behavior 18

Figure 13: SHAM Male Performing Pivoting Behavior 19

Figure 14: SHAM Male Performing Idling Behavior 19

Figure 15: SHAM Male Performing Climbing Behavior 20

Figure 16: SHAM Male Performing Dashing Behavior 20

Figure 17: SHAM Male Performing Locomotion Behavior 22

Figure 18: SHAM Male Performing Investigating Behavior 22

Figure 19: SHAM Male Performing Pivoting Behavior 23

Figure 20: SHAM Male Performing Idling Behavior 23

Figure 21: SHAM Male Performing Climbing Behavior 24

Figure 22: SHAM Male Performing Dashing Behavior 24

List of Tables

Table 1: Skeleton Node Names and Definitions 5

Table 2: Behaviors and their Descriptions 15

Abstract

This project seeks to design a machine learning behavioral analysis framework to discover the underlying mechanisms of neurodevelopmental disorders and the effectiveness of the treatment. Unique action patterns are hallmarks of neurodevelopmental disorders which could be caused by Traumatic Brain Injury (TBI), and vary considerably from individual to individual. In animal models, conventional behavioral phenotyping captures limited fine-scale variations. This project aims to design a deep learning model to characterize mouse movement on multiple timescales using the open-field test. The framework takes virtual markers from pose estimation to find behavior clusters and generate signatures of behavior classes. This tool measures spatial and temporal habituation to a new environment across minutes and days, different types of self-grooming, locomotion and gait. It also tests under task or social conditions which would reveal more information about behavioral dynamics and variability.

Chapter 1: Introduction and Related Work

1.1 Introduction

The field of machine learning has made significant strides in the behavioral sciences, finding its place as a tool for researchers to analyze a variety of test subjects. Prominently it is being used to quantify real-world behaviors in test subjects by attribution of a number of factors, i.e. joints, brain chemistry, or other stimuli. The goal of this project is to provide an approach to behavioral analysis for use with lab mouse test subjects in a novel dataset.

The dataset is very specifically of note here, as there have been a number of approaches that have targeted the lab mouse demographic before, but have either targeted mice with specific conditions or have been collected with an extensive array of sensors. The dataset in use for this project is general in nature but was collected only using a camera that included no built-in tracking or processing.

It should also be noted that this approach does not make use of any specific supervised learning methods to classify behaviors within these test subjects. Other approaches have utilized these methods because they allow a quantitatively accurate way to predict what the subject is doing given input features. With all of this in mind, the approach with this dataset has been to provide those who need to perform these analyses with the means to do so in a generalized manner and generate metrics that a researcher without a technical background could use.

1.2 Related Work

Pose tracking can be essential to behavioral analysis, and there have been numerous advances in annotation and video processing to aid in this. These tools that aid in the labeling and processing of videos, oftentimes require some degree of auxiliary sensors to be usable in a pipeline. As stated in Section 1.1 these sensors can provide a myriad of information, ranging from infrared sensors to sensors capturing brain activity.

Klibaite et al. utilized time series analyses and clustering techniques similar to the one found in this project, but also relied upon the mutation data to inform their analysis of the endophenotypes that appear in subjects exhibiting symptoms of autism. Additionally, the primary approach to categorize these behaviors was a deep learning approach. Hong et al. performs an analysis of similar phenotypes, but three-dimensional ocular sensing and depth tracking methods with centroid analysis. It also takes into account the genetic influences on social behavior. Lorbach et al. performs an analysis using deep learning methods to track and measure interactions between their test subjects based on sex and a number of social behaviors. Wiltschko et al. uses similar neural network methods and three-dimensional to track and predict the social behaviors associated with genetic disorders. In short, a general analysis of behaviors that utilize an unsupervised learning method to analyze behaviors and a unimodal sensor to collect data is an approach that has not been attempted yet.

Chapter 2: Joint Tracking and Pose Analysis for Behavioral Classification

This chapter outlines the process by which the behavioral analysis was constructed. The annotation tool, SLEAP, is introduced and the process for annotation is shown. The behavioral analysis pipeline developed for this project is then detailed in the following two sections, including projected outputs of each portion of the pipeline.

2.1 SLEAP Data Labeling

The SLEAP (Social LEAP Estimates Animal Pose) tool (Pereira et al., 2022) is a tool developed for use as a tool for researchers exploring behaviors in animals. The entire dataset for this project was labeled or inferred on using this tool and its functions

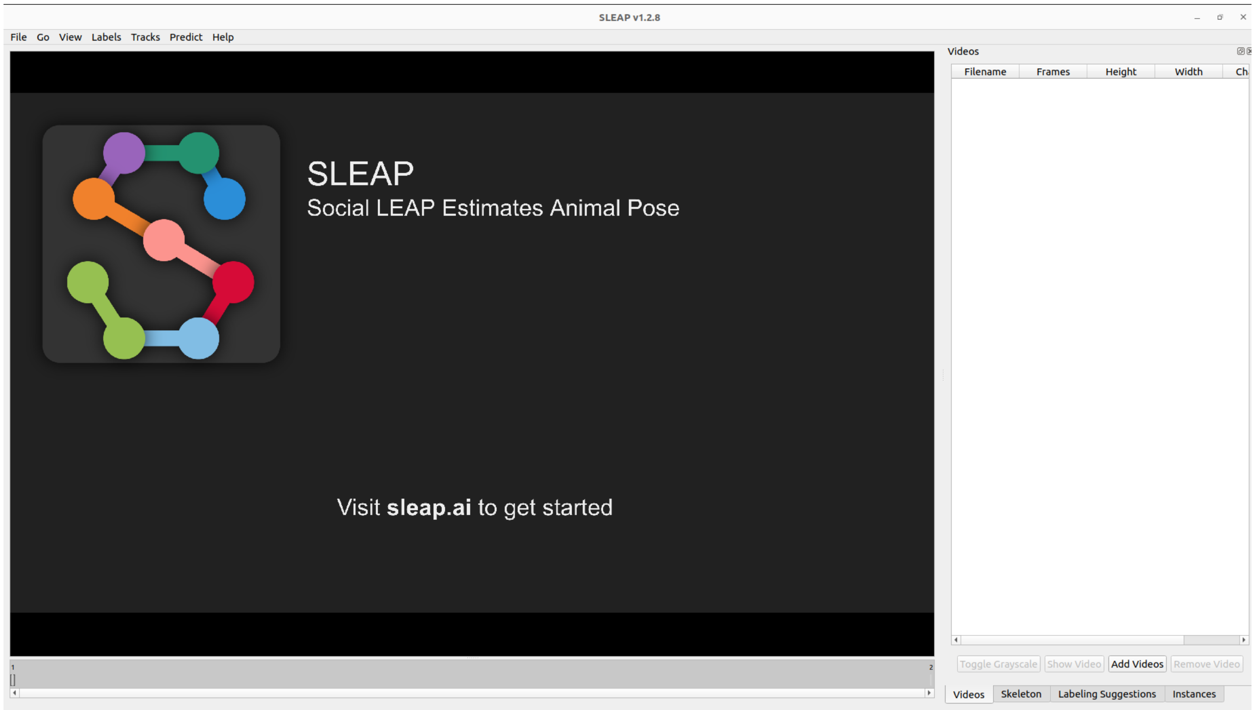

Figure 1: The SLEAP Tool

The first step in the labeling process is to create a skeleton that would represent the joints on the mouse's body. The skeleton is an abstract data structure that holds data about itself and its connections to the other joints. For example, a mouse's head would connect to its body, and its body to its tail. This would be described in this structure. Manual labeling was done by physically placing nodes on the appropriate body parts for each mouse. Each node that was observed and its respective definition can be found in Table 1. These nodes are the ones that were used in this project, however, the skeletons that contain them can be fully customized to suit a project's needs.

Appendage	Description
Nose	The tip of the mouse’s head
Head	The center of mass of the head of the mouse
Front Left Foot	The front left foot of the mouse from the top-down perspective
Front Right Foot	The front right foot of the mouse from the top-down perspective
Body	The center of mass for the mouse as a whole
Back Left Foot	The back left foot of the mouse from the top-down perspective
Back Right Foot	The back right foot of the mouse from the top-down perspective
Tail Base	Where the tail connects to the body of the mouse
Tail Tip	The end of the tail

Table 1: Skeleton Node Names and Definitions

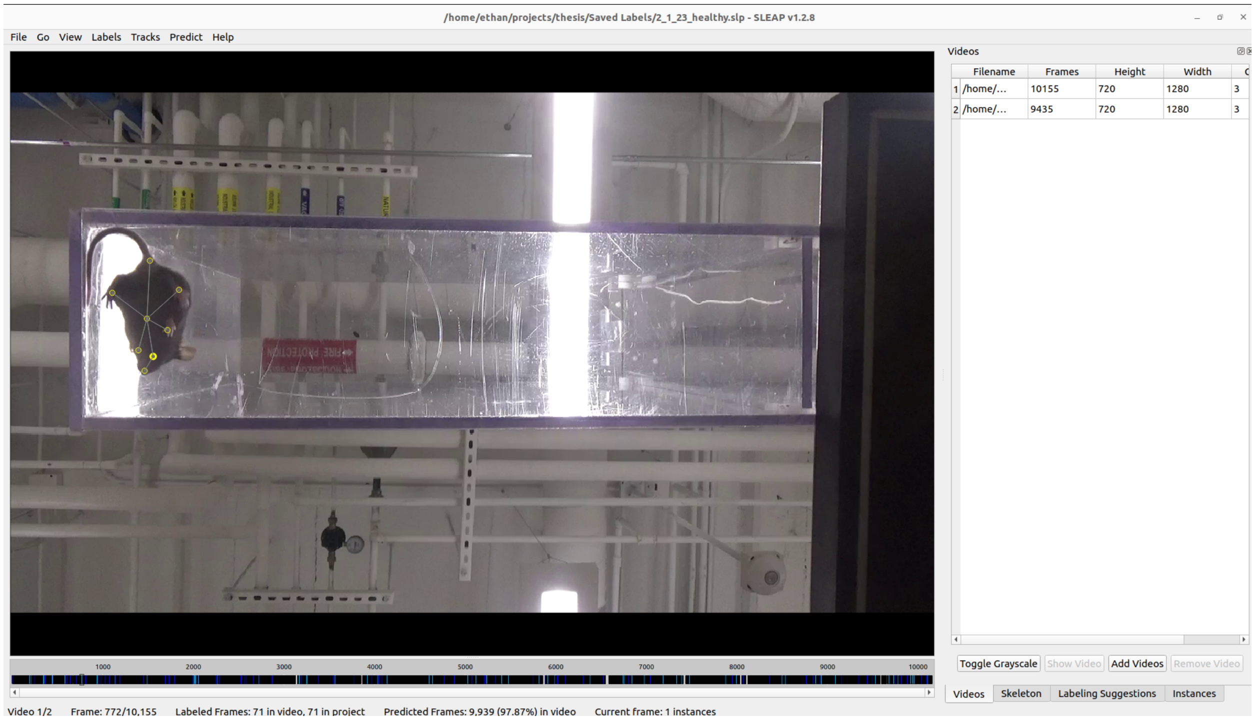

Figure 2: Labeling a Frame of a Mouse Video with a Skeleton

After the data has been labeled, the second step of this process can begin. This process involves choosing a model and model parameters to begin the learning process. These parameters include batch size, epochs, learning rate, and any augmentation parameters.

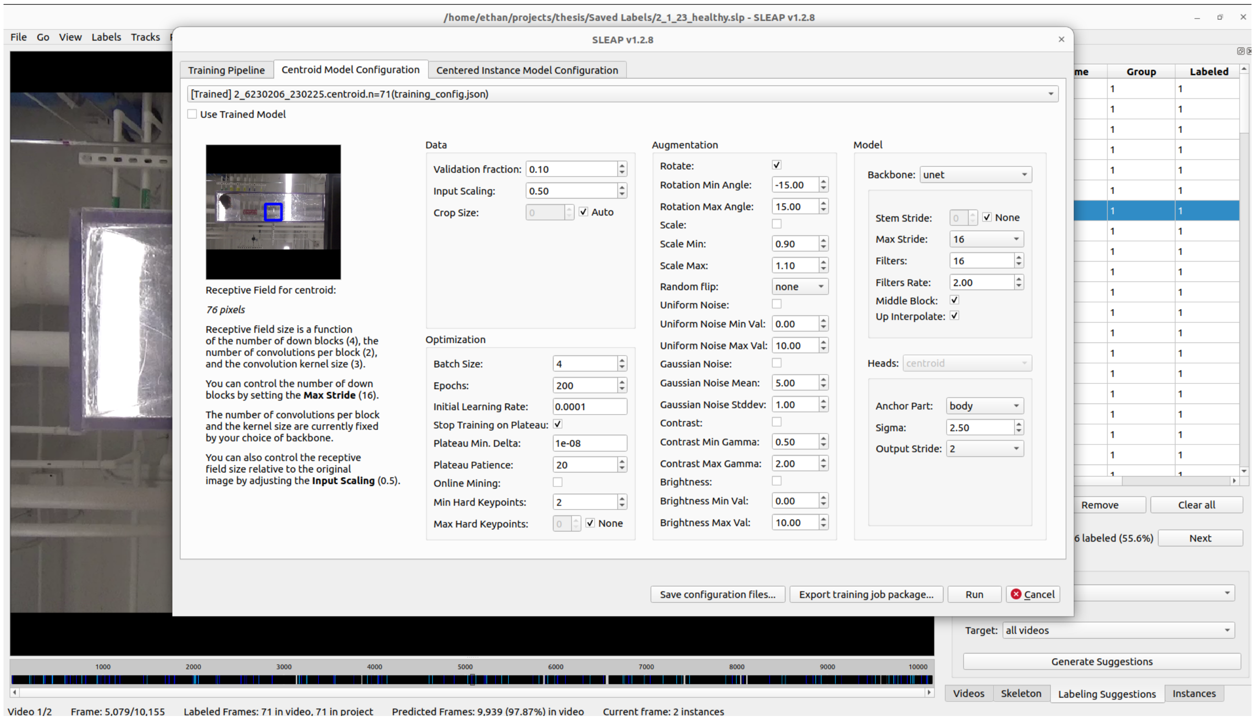

Figure 3: Selecting Model Parameters for Model Training or Inference

Training of the U-Net neural network is now ready to begin. Prior to this, the training configuration can be saved for future uses. Once training is complete, the dataset can now be predicted on and using the model that was now trained on. Predictions can be done on the rest of the dataset that was not labeled or any number of videos that are loaded into the SLEAP tool, the caveat being that it will only be accurate on videos of the same animal type.

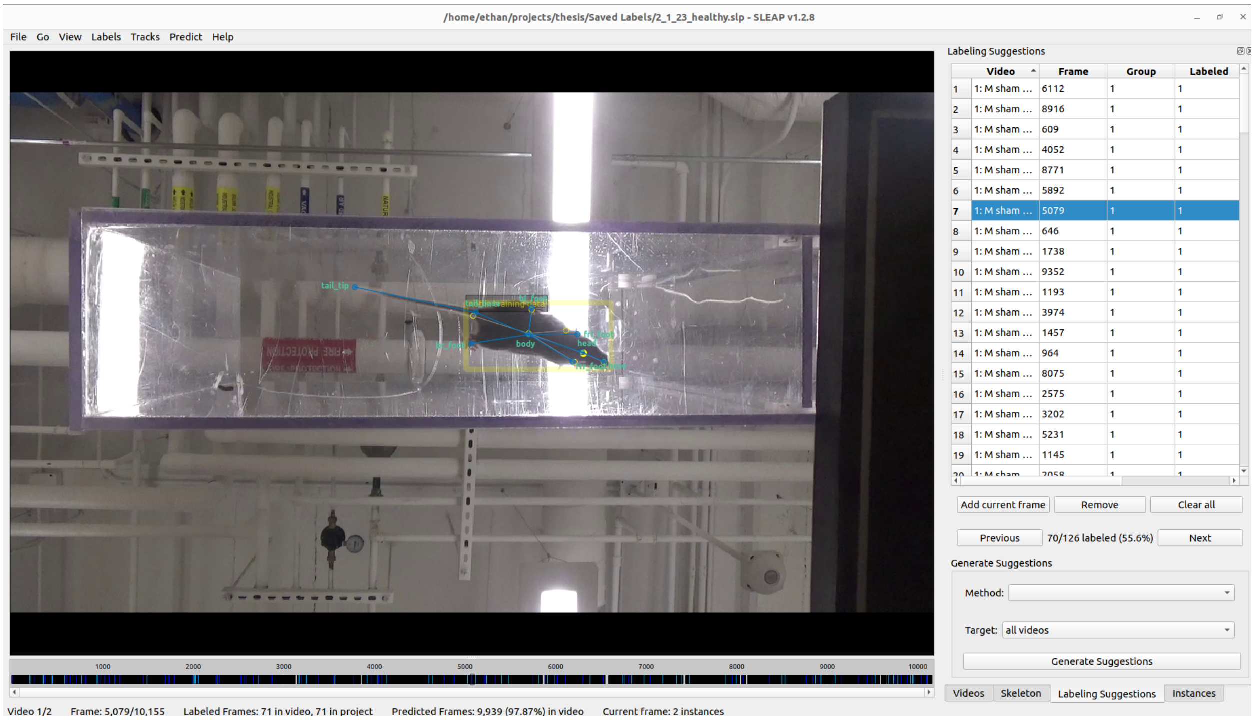

Figure 4: Example of Labeled Frame with Predicted Skeleton Label

After prediction has finished the predicted joint trajectories can be exported into an H5 file, which contain the joint names, joint locations, and occupancy matrices for each video.

2.2 Embedding and Clustering

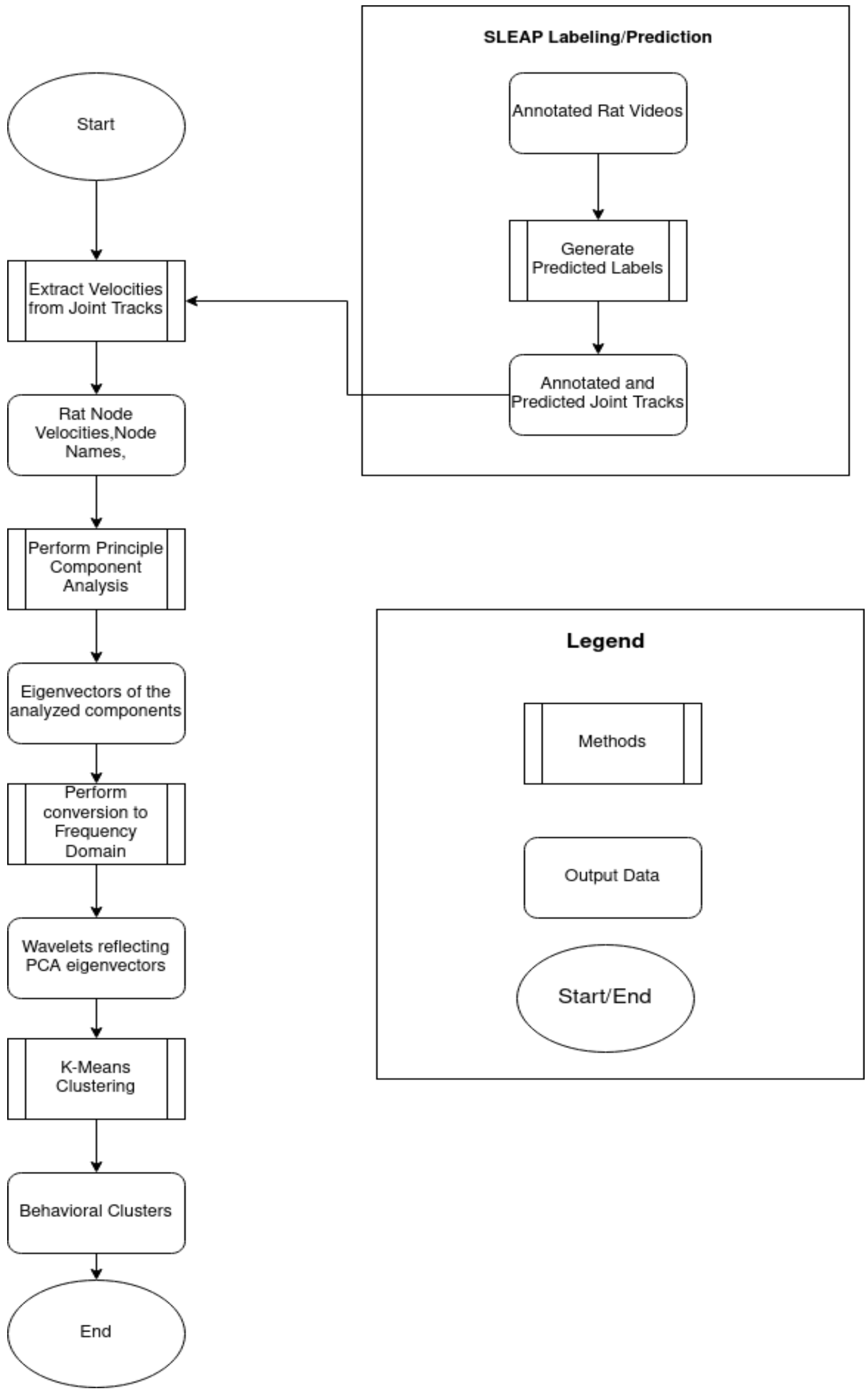

Figure 5: Embedding and Clustering Pipeline

With the skeleton information now available to the beginning of the pipeline, the analysis of this data can begin.

This pipeline begins with an analysis of the occupancy of the nodes over time, after which a Principle Component Analysis (PCA) is employed to gain an understanding of the variance of the dataset. Normally, this would provide a greater understanding of the data, but these variances are imperative to the next step in the pipeline.

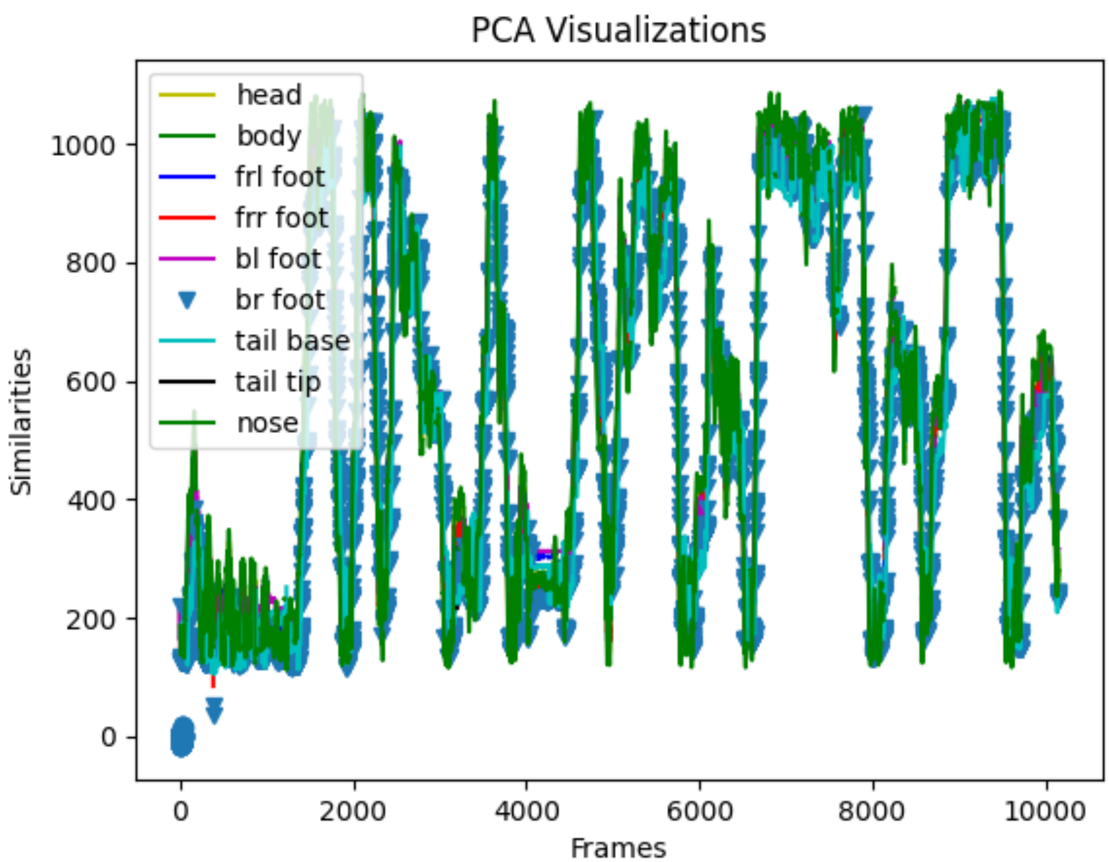

Figure 6: PCA Estimation of Wavelet Decompositions

Following the PCA, we convert the variance to the frequency domain in the form of a discrete wavelet transform to capture the changes in posture over time. These changes over time are then clustered using K-Means Clustering to assign these changes in posture to behavioral clusters. These behavioral clusters outline the behaviors that the test subject exhibits over time.

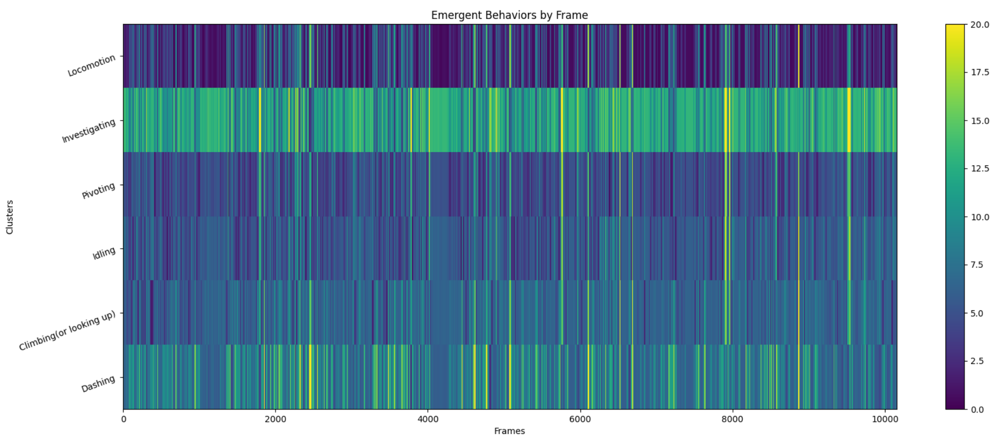

Figure 7: Emergent Clusters over Time

2.3 Fingerprinting

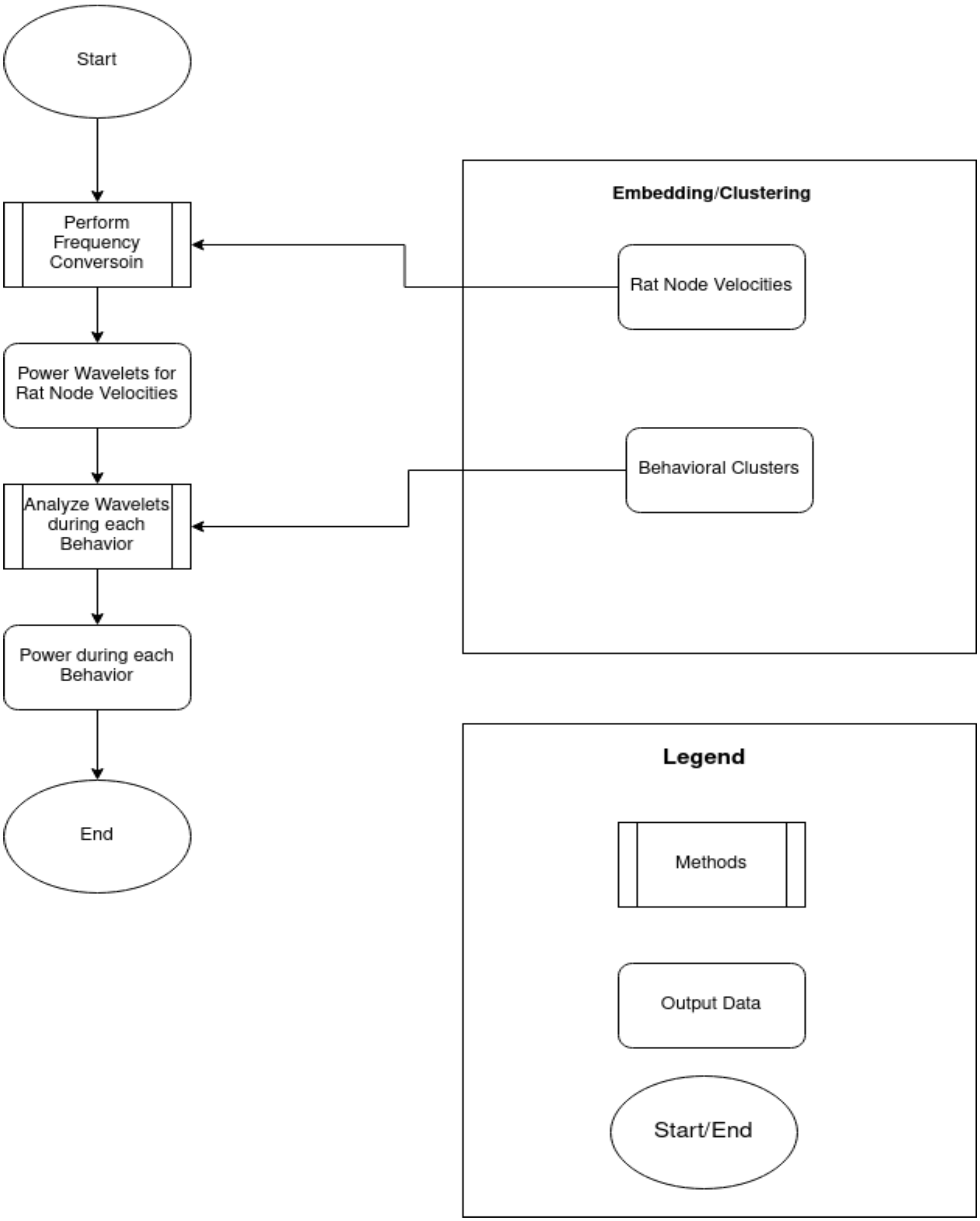

Figure 8: Fingerprinting

The second half of this pipeline is used to compare the joints captured in the video against the behavioral clusters generated in the previous half of the pipeline. This process is begun by performing an analysis of the predicted joints and capturing their trajectories over time.

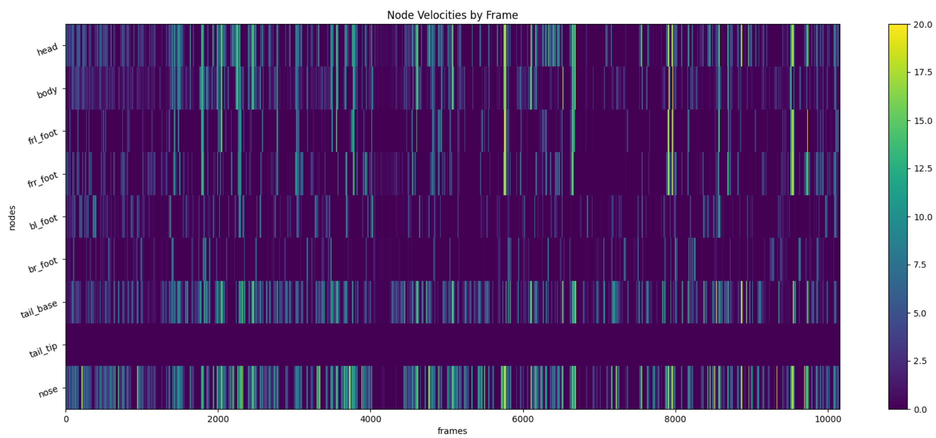

Figure 9: Joint Velocities over Time

These trajectories are taken and converted into the frequency via a discrete wavelet decomposition, similar to the way that the Principle Component Analyses were in the previous half of the pipeline. We want to capture the changes in each of these trajectories over time for a correlation with the behavioral clusters.

Figure 10: Node Trajectories Transformed to the Frequency Domain

We next want to perform an analysis of what each of the joints look like with respect to each behavioral cluster. For example, if the Back Right Foot is active when the mouse is exhibiting the Pivoting behavior but not the Investigating behavior, this section would outline that. Each behavior and their descriptions can be seen in Table 2.

Behavior	Description
Locomotion	Mouse is steadily moving across the enclosure
Investigating	Mouse is looking around its current space, not moving from its position
Climbing(or Looking Up)	Mouse is trying to climb the walls of the enclosure or looking up at the ceiling
Pivoting	Mouse is turning in-place
Idling	Mouse is sitting in one spot. This encompasses all sedentary behaviors: sitting, grooming, etc
Dashing	mouse moves fast across the enclosure

Table 2: Behaviors and their Descriptions

Generating these clusters is currently not an automated process. K-Means Clustering is only an unsupervised learning algorithm, meaning that the data, and therefore the generated clusters, have no labels attached to them. This means that the end user needs to label the clusters manually before producing any sort of output annotation or behavioral metrics surrounding each input. The user currently receives a video annotated with the cluster number, which means that Table 2's definitions must be specific and concise. This process is typically not very difficult and the behaviors make themselves apparent quite easily.

Chapter 3: Dataset and Results

This chapter contains the description of the dataset and the outputs from the pipeline that was described above. It also includes a short examination of the dataset used in this project against datasets that have been used in other projects, defining what makes this novel. The second section in this chapter presents the analysis that is generated from this novel dataset.

3.1 Dataset

The dataset used in this project was entirely novel, collected at the University of Michigan-Dearborn by Dr. Zhi "Elena" Zhang's lab. The unique feature of this dataset that isn't typically present in other datasets or approaches is that it was taken using exclusively an optical sensor. Other approaches use infrared sensors that are used to capture the movement in the joints and track using those sensors, allowing the pipelines to get an preprocessed input instead of raw data. There have also been approaches that have collected data about brain activity and therefore passing more descriptive data into their pipelines.

While these auxiliary sensors provide some layers of robustness and information, they limit the accessibility of producing quality data. To account for this robustness, the data being used must be annotated by the user. The classifier model can make accurate predictions with one hundred fifty labeled samples. With the methods that have been described there may be some reduction in precision surrounding the generated behaviors, but the accessibility of this unimodal dataset makes up for it.

3.2 Results of Pipeline

The purpose of this pipeline is to provide the end user a quantitative, illustrated understanding of the subjects that they are trying to test. For this to happen, the pipeline outputs two primary artifacts: an annotated video and a mean power spectrum by behavior. An example of what these would look like follows below.

3.2.1 Annotated Test Subject Video

The first output of this pipeline is a video of the test subject annotated with the behavior that it's performing at any given frame. The current framework only supports the six behaviors listed in Table 2 for this, but this functionality will be expanded upon in the future (See Section 4.2 for further details). Figures 11-16 outline the annotated videos that are generated for the user and what each behavior would look like within the video.

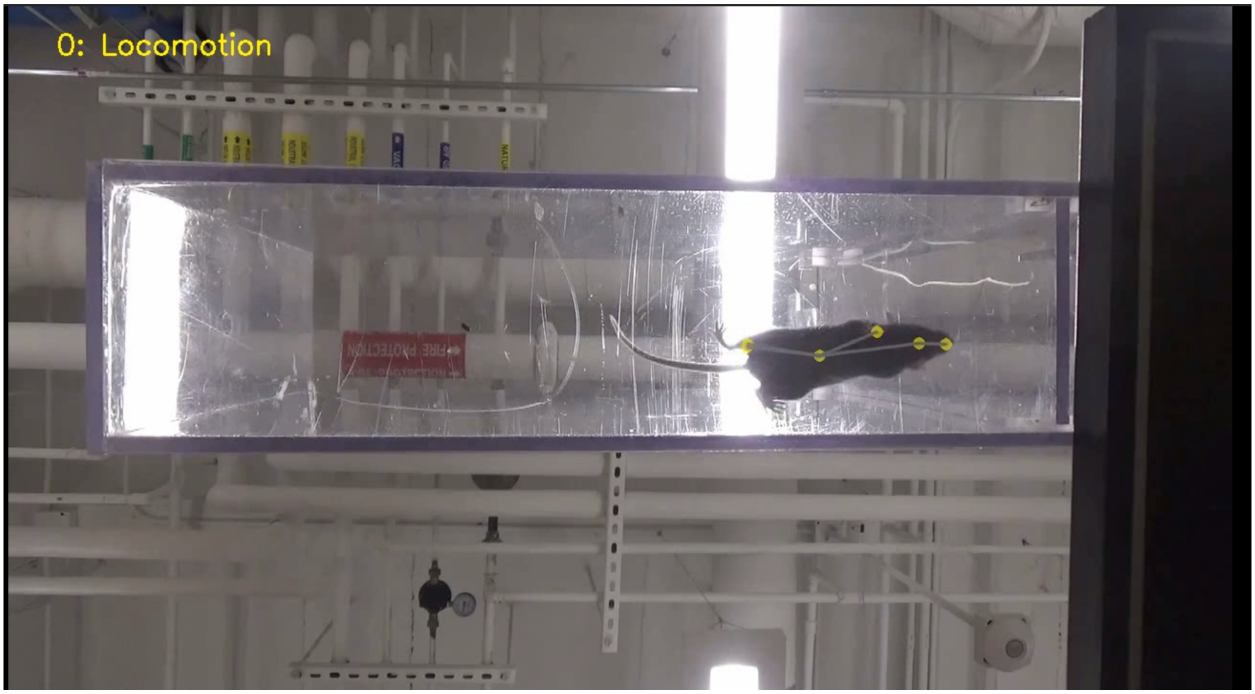

Figure 11:SHAM Male Performing Locomotion Behavior

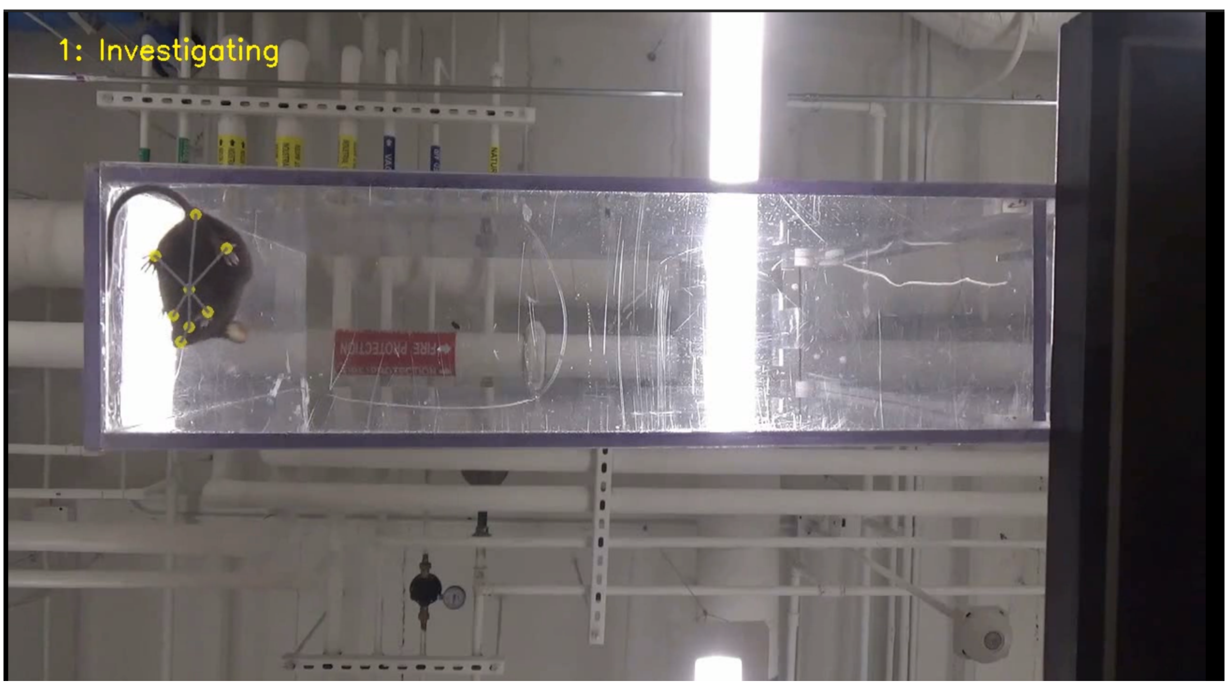

Figure 12: SHAM Male Performing Investigating Behavior

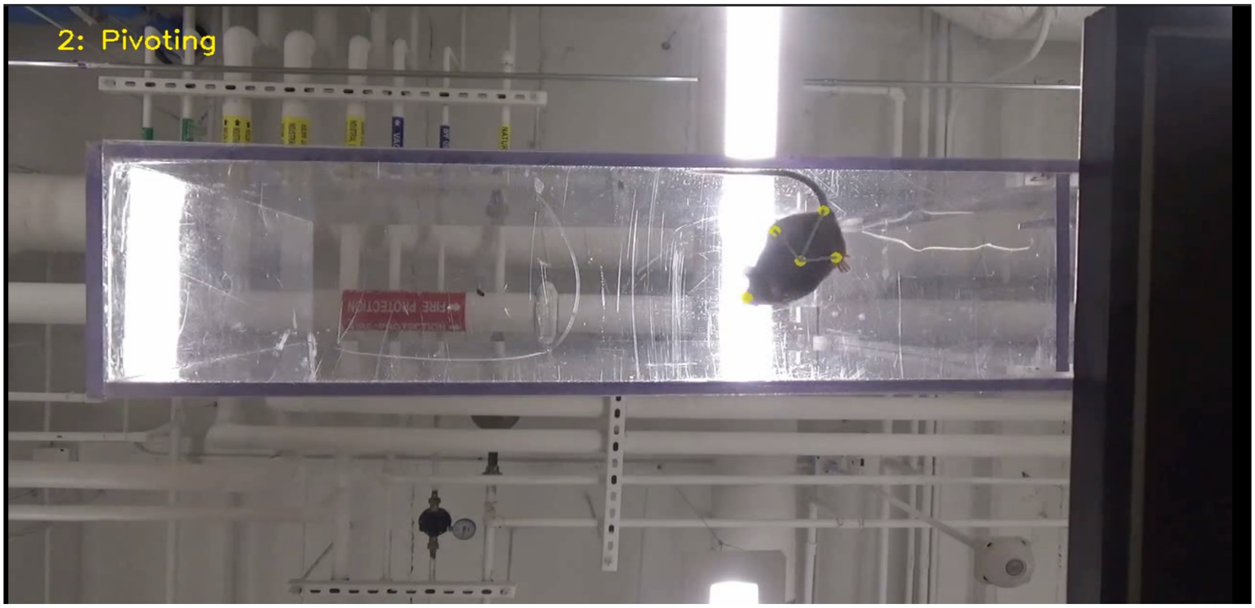

Figure 13: SHAM Male Performing Pivoting Behavior

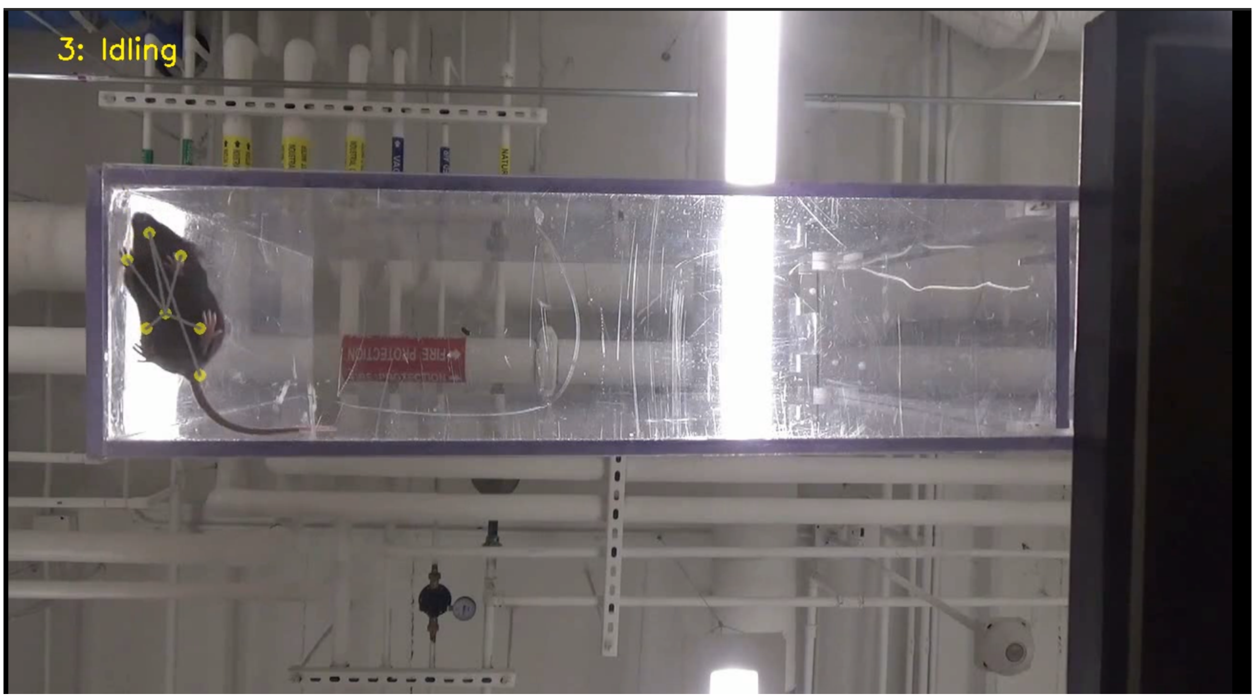

Figure 14: SHAM Male Performing Idling Behavior

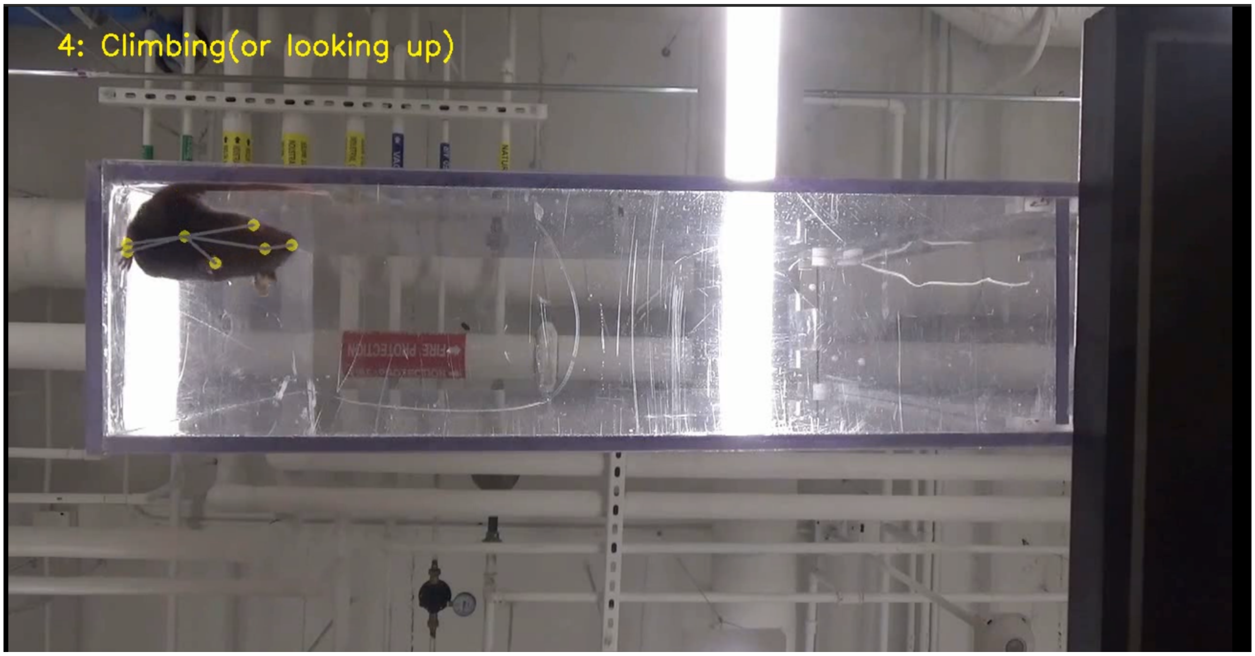

Figure 15: SHAM Male Performing Climbing Behavior

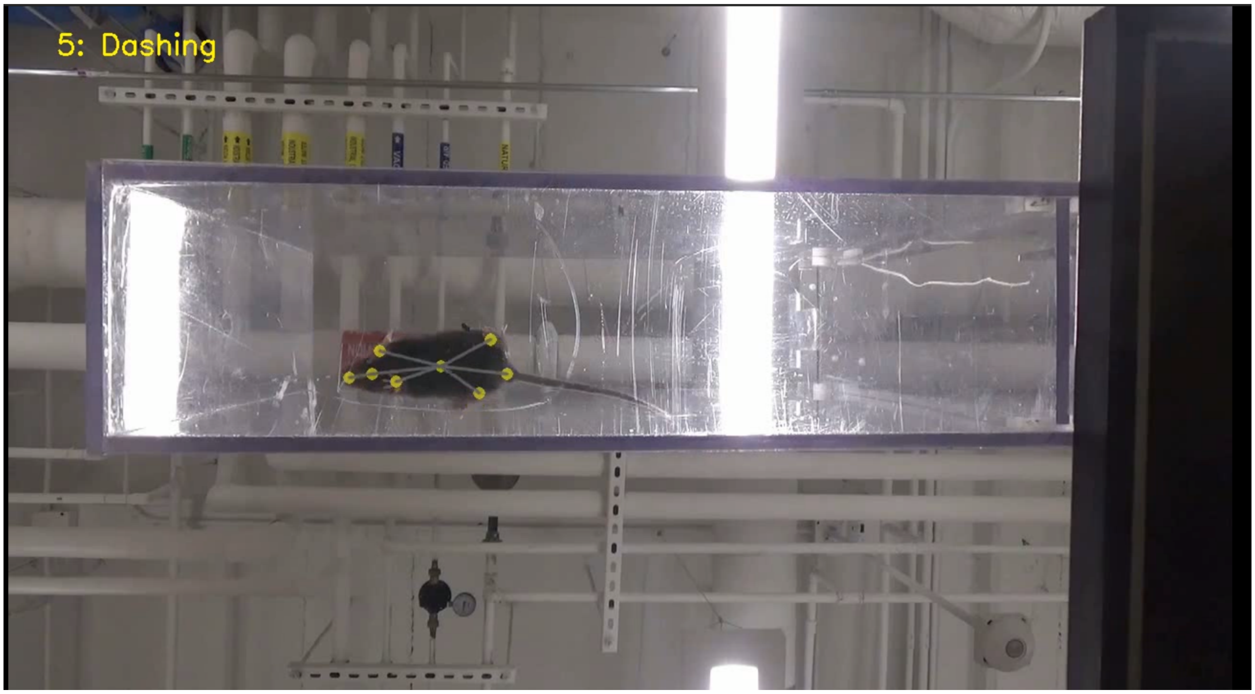

Figure 16: SHAM Male Performing Dashing Behavior

3.2.2 Mean Wavelet Power During Each Behavior

Outside of an annotated video representation of what the test subject is doing at all times, the end user should be provided with a more granular analysis of the behaviors that their subjects are exhibiting. An examination of those behaviors can be seen in Figures 17-22. We can see from examining the heatmaps for the Locomotion, Dashing, and Climbing behaviors that the SHAM Male seems to have consistent movement patterns across all frames where that is present. However, upon examination of the Investigating, Pivoting, and Idling behaviors, we see that joint velocities are less constant and tend towards more sporadic movements. This is consistent with the descriptions of these behaviors- the behaviors that are primarily associated with movement have consistent movement, while the ones that involve the subject thinking or examining its environment have spontaneous movements.

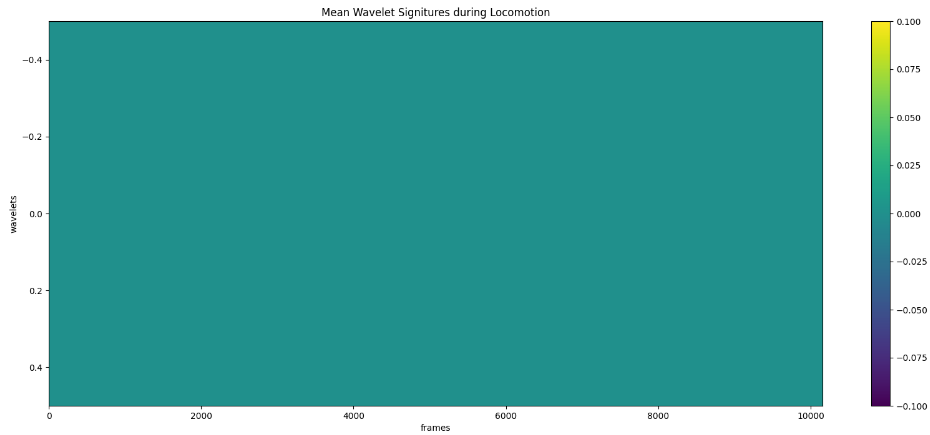

Figure 17: SHAM Male Performing Locomotion Behavior

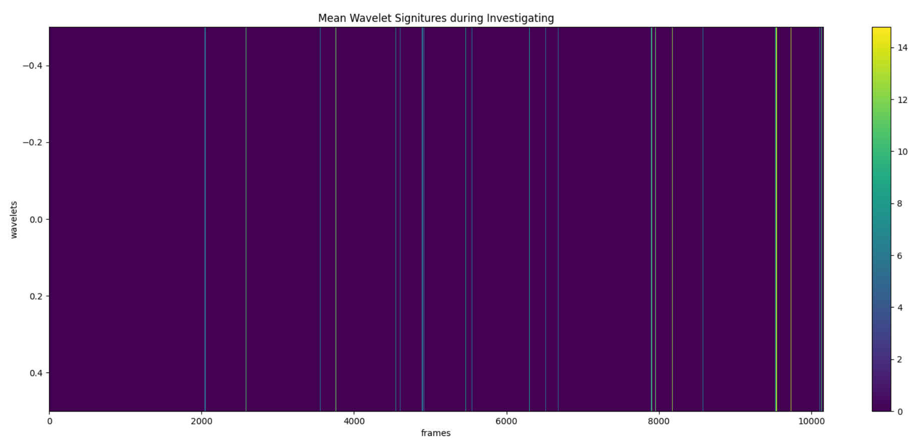

Figure 18: SHAM Male Performing Investigating Behavior

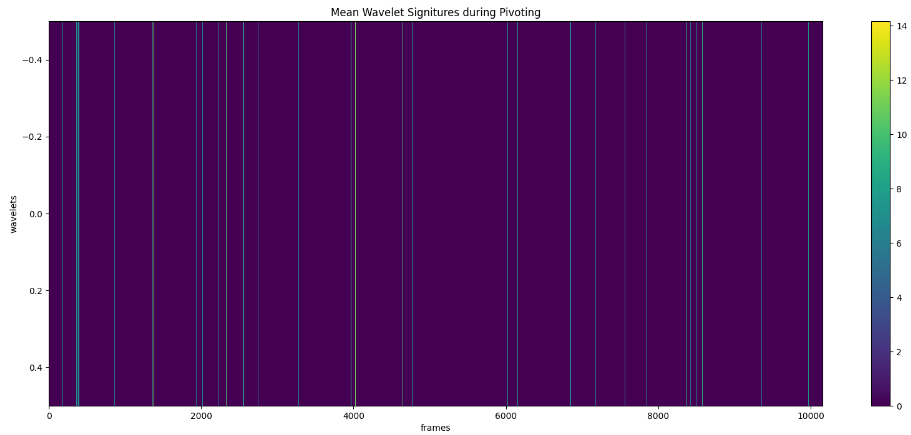

Figure 19: SHAM Male Performing Pivoting Behavior

Figure 20: SHAM Male Performing Idling Behavior

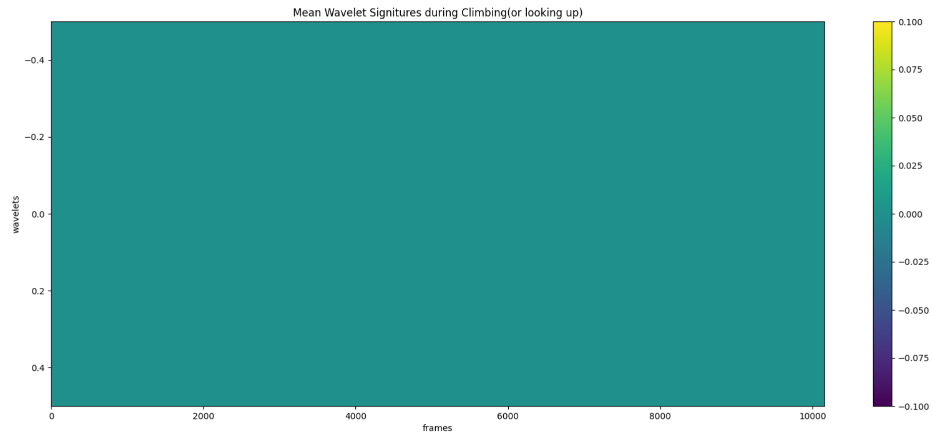

Figure 21: SHAM Male Performing Climbing Behavior

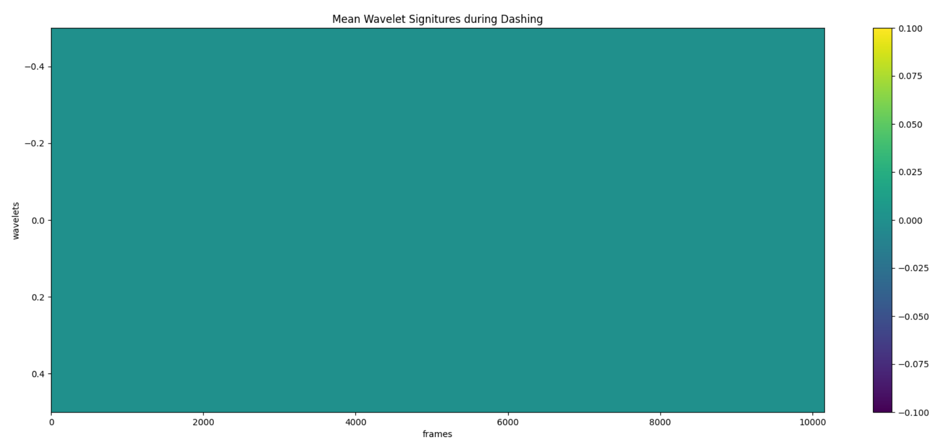

Figure 22: SHAM Male Performing Dashing Behavior

Chapter 4: Conclusion and Future Work

This section outlines the conclusions that can be drawn from the work that has been done in Chapters 2 and 3. It also describes the future work that could be done to further utilize the behavioral analysis tool to create a more robust, versatile tool.

4.1 Conclusion

In this thesis, a pipeline was developed to perform a behavioral analysis on a completely novel dataset of test subjects. This pipeline used a variety of machine learning techniques and data transformations to form a behavioral analysis given unimodal data, which is comparable to other multimodal approaches to these types of tasks. Even given the unimodal approach, comparably illustrative results to some of the other methods cited in this project can be noted. The image segmentation classifier model performed with 85% accuracy given only 150 samples to learn from. While the clustering method lacked the precision of some of the other contemporary methods, the accessibility of the dataset more than makes up for it. This approach offers a viable way that researchers with limited means to collect data can perform an accurate analysis of their test subjects and produce artifacts that provide a quantitative measure of why each behavior was categorized

4.2 Future Work

As stated in the previous sections, this pipeline currently contains two main parts- the first being Embedding and Clustering followed by the Fingerprinting portion. For this pipeline to work, it requires user input- which can lead to more inconsistent results in the behavioral analysis data. With this in mind, the first step of work to be done is to automate the labeling of behavioral clusters after the K-Means. There are a number of different approaches that could be used to do this, but the first and safest approach would be to start with a multiclass image classifier. Once the clusters are consistently being produced, the behavioral analysis that follows should be not only accurate, but explainable.

The second major push would be to integrate the SLEAP tool into the pipeline in an automated manner. The SLEAP tool currently remains agnostic of the pipeline, only providing the annotated input videos that are used to extract the node velocities. Ideally, this would not be the case. The user should be able to pass a video and have a model run inference on it to produce those node trajectories for the first half of the pipeline without any manual annotation.

Finally, this tool should be able to function on the command line exclusively. Most of the target users for this tool will likely be unfamiliar with Python and a number of the libraries used in this project, so removing them from that is very important. This push would be converting the pipeline to accept a number of parameters- dataset path, output annotation path, type of model for inference on dataset, and more. With all of these future changes, this project could be an extremely versatile tool at a researcher's disposal.

Appendix: Table of Expanded Behaviors for Future Work

Behavior	Present in Current System	Description
Locomotion	Yes	Mouse is steadily moving across the enclosure
Investigating	Yes	Mouse is looking around its current space, not moving from its position
Climbing	Yes	Mouse is trying to climb the walls of the enclosure or looking up at the ceiling
Pivoting	Yes	Mouse is turning in-place
Idling	Yes	Mouse is sitting in one spot. This encompasses sedentary behavior
Dashing	Yes	Mouse moves fast across the enclosure
Grooming	No	Mouse cleans itself in-place, different than idling because of the actions that the mouse does while sitting there
Looking Up	No	Mouse is investigating, looking upward. Could be at the researcher, the ceiling of the enclosure
Slow Explore	No	Investigating and exploring while walking towards a spot slowly
Fast Explore	No	Investigating and exploring while walking towards a spot quickly

References

Hong, W., Kennedy, A., Burgos-Artizzu, X. P., Zelikowsky, M., Navonne, S. G., Perona, P., & Anderson, D. J. (2015, September 22). Automated measurement of mouse social behaviors using depth sensing, video tracking, and Machine Learning. *Proceedings of the National Academy of Sciences of the United States of America*. Retrieved April 8, 2023, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4586844/

Klibaite, U., Kislin, M., Verpeut, J. L., Bergeler, S., Sun, X., Shaevitz, J. W., & Wang, S. S.-H. (2022, March 12). Deep phenotyping reveals movement phenotypes in mouse neurodevelopmental models - molecular autism. *BioMed Central*. Retrieved April 8, 2023, from https://molecularautism.biomedcentral.com/articles/10.1186/s13229-022-00492-8

Lorbach, M., Kyriakou, E., Poppe, R., van Dam, E., Noldus, L., Veltkamp, R. Learning to recognize rat social behavior: Novel dataset and cross-dataset application. *Journal of Neuroscience Methods, Volume 300*. 2018. Pages 166-172. ISSN 0165-0270. https://doi.org/10.1016/j.jneumeth.2017.05.006.

Pereira, T. D., Tabris, N., Matsliah, A., Turner, D. M., Li, J., Ravindranath, S., Papadoyannis, E. S., Normand, E., Deutsch, D. S., Wang, Z. Y., McKenzie-Smith, G. C., Mitelut, C. C., Castro, M. D., D'Uva, J., Kislin, M., Sanes, D. H., Kocher, S. D., Wang, S. S.-H., Falkner, A. L., … Murthy, M. (2022, April 4). Sleap: A deep learning system for multi-animal pose tracking. *Nature News*. Retrieved April 8, 2023, from https://www.nature.com/articles/s41592-022-01426-1

Weber, R. Z., Mulders, G., Kaiser, J., Tackenberg, C., & Rust, R. (2022, October 15). Deep learning-based behavioral profiling of rodent stroke recovery - BMC biology. *BioMed Central*. Retrieved April 8, 2023, from https://bmcbiol.biomedcentral.com/articles/10.1186/s12915-022-01434-9

Wiltschko, A. B., Johnson, M. J., Iurilli, G., Peterson, R. E., Katon, J. M., Pashkovski, S. L., Abraira, V. E., Adams, R. P., & Datta, S. R. (2015, December 16). Mapping sub-second structure in mouse behavior. *Neuron*. Retrieved April 8, 2023, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4708087/