CLASSIFICATION OF SPINAL CORD INJURY IN RATS BY MACHINE LEARNING OF GAIT PARAMETERS

by

Perver Atilla İNCE

B.S., in Biomedical Engineering, Faculty of Technology, Pamukkale University, 2021

Submitted to the Institute of Biomedical Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Biomedical Engineering

Boğaziçi University
2025

CLASSIFICATION OF SPINAL CORD INJURY IN RATS BY MACHINE LEARNING OF GAIT PARAMETERS

APPROVED BY:

Prof. Dr. Burak Güçlü .....
(Thesis Advisor)

Prof. Dr. Esin Öztürk Işık .....

Prof. Dr. Gülgün Şengül .....

DATE OF APPROVAL: 15 January 2025

ACKNOWLEDGMENTS

Reaching the end of this thesis is, quite frankly, one of the greatest thanks I can give myself. So, dear readers and committee members, I ask for your kind indulgence—I might get a bit carried away here. But hey, if I can complete a thesis, surely you can spare a few more pages, right?

I would like to thank myself for reminding me of what I have accomplished when I doubt myself, for discovering friendships I could never have imagined in moments of solitude, and for always finding solutions in the face of challenges, even when it felt like none existed. I thank myself for carrying on when it would have been easier to stop, for believing in the possibility of something greater, and for proving that resilience is not just enduring the storm, but learning to dance in the rain. I thank to my family. I may not have learned everything from them, but they taught me how to learn everything. They may not fully understand this thesis as I do, but I thank them for understanding how to give meaning to a life that made this achievement possible. I'm not sure if there are enough words to thank them for the effort it takes to create and sustain a family, but for making me who I am, whatever light you see in me, its source was always them. To my aunt, whose quiet inspiration and example as a dedicated professor continue to motivate me, I am grateful. Her wisdom and a single book she once shared during my undergraduate years have left a lasting mark on my journey.

When I first met her, I was just a young man, uncertain about his goals and far from imagining himself as an engineer. Together, we not only graduated as biomedical engineers but also embarked on a journey of love and companionship. What began as being a classmate turned into a personal odyssey, as the person who started this path as my girlfriend is now my fiancée. While I stand at the threshold of becoming a master in biomedical engineering, she has been, and always will be, the master of my heart. For three years, she stood by me through every sleepless night, every moment of doubt, and every mountain of stress. She was my anchor when the waves of this

demanding journey threatened to pull me under and my light when the path ahead seemed too dark to see. My dearest love, thank you—for inspiring me to strive for a better version of myself and for giving meaning to everything I've learned. You are not just my partner; you are my greatest friend and the one who makes every step of this life a little brighter, a little better with your boundless joy and radiant smile.

I would also like to thank Berfuğ Yaren Karaharman, my laboratory companion, whose dedication and effort in collecting the data for this study laid the groundwork for my research. Beyond the data, your unwavering support and camaraderie made this journey not just manageable, but truly rewarding. I would also like to thank Mara and Denisa, our laboratory interns from Romania, whose meticulous efforts were instrumental to this study. Working alongside you both was not only productive but also enriching. Lastly, I extend my heartfelt gratitude to Prof. Dr. Burak Güçlü, my advisor, whose high expectations and relentless pursuit of excellence pushed me beyond my limits. Even though the journey often required revisiting and refining my work, it taught me the value of resilience and the art of continuous improvement. Your guidance, while challenging, has shaped not just this thesis, but also my approach to learning and growth. Sharing the tangible embodiment of your artistry and opening the doors to your personal library were gestures that resonated deeply. Lastly, your unmatched expertise in MATLAB, cultivated through years of dedication, was a constant source of inspiration during our collaboration. Witnessing your mastery was a privilege and a benchmark for what it means to excel both technically and creatively.

This work was supported by TÜBİTAK grant no: 221N399 as part of the European Union's FLAG-ERA JTC 2021 project RESCUEGRAPH and H2020 Graphene Flagship Core 2. I also extend my gratitude to Pulseven Teknoloji Yatırım Dış Ticaret A.Ş. for their financial support and to Verimli Danışmanlık ve Ticaret Ltd. for providing the equipment necessary for this study. Their contributions were invaluable, and I am truly appreciative of their support.

ACADEMIC ETHICS AND INTEGRITY STATEMENT

I, Perver Atilla İNCE, hereby certify that I am aware of the Academic Ethics and Integrity Policy issued by the Council of Higher Education (YÖK) and I fully acknowledge all the consequences due to its violation by plagiarism or any other way.

Name :

Signature:

Date:

ABSTRACT

CLASSIFICATION OF SPINAL CORD INJURY IN RATS BY MACHINE LEARNING OF GAIT PARAMETERS

Spinal cord injury (SCI) represents a critical neurological condition with high morbidity, significantly impacting sensory motor functions. This thesis introduces a novel approach to classify the time after SCI and rats which recieved neuromodulation therapy by using machine learning (ML) based analysis of gait parameters and locomotor scores. Utilizing data derived from video imaging of rat locomotion, this study evaluates the accuracy and feasibility of the approach compared to previous literature. Key parameters analyzed include stance duration, swing duration, stride distance, limb duty factor, and paw area, alongside traditional metrics like Basso, Beattie and Bresnahan (BBB) locomotor rating scale and Von Frey (vF) withdrawal thresholds.

The integration of markerless pose estimation tool, DeepLabCut (DLC), allowed for detailed extraction of gait parameters, overcoming challenges associated with lateral plane-focused methodologies. By using images from lateral and bottom views, footsteps were associated with animals' body movement and the gait pattern was extracted by deep-learning methods. Gait parameters, BBB scores and vF thresholds were input as features to Ensemble Learning including various ML methods to predict time after injury and neuromodulation treatment. In animals with SCI, gait parameters and BBB scores from both fore-hindlimbs, allowed prediction of post injury class (6 end-points) with %43 accuracy. The same feature set predicted treatment with %71 accuracy. Results showed that the methodology can effectively differentiate between treatment groups and time after SCI. However, additional work is needed to improve accuracy. Parameter extraction by deep-learning also provides an accessible and cost-effective solution for SCI research.

Keywords: Spinal Cord Injury, Gait Analysis, Machine Learning, Classification

ÖZET

SIÇANLARDA YÜRÜME PARAMETRELERİ VE MAKİNE ÖĞRENMESİ İLE OMURİLİK YARALANMASININ SINIFLANDIRILMASI

Omurilik yaralanması (OY), duyu ve motor fonksiyonlarını önemli ölçüde etkileyen, yüksek morbiditeye sahip kritik bir nörolojik durumdur. Bu tez, yürüme parametrelerinin ve lokomotor skorlamalarının makine öğrenimi tabanlı analizini kullanarak sıçanlarda OY sonrası geçen süreyi ve nörostimülasyon tedavi alımını sınıflandırmak için yeni bir yaklaşım sunmaktadır. Sıçan hareket görüntülerinden edinilen verilerle sınıflandırma doğruluğu ve yöntemin uygulanabilirliği açısından literatürdeki mevcut çalışmalarla karşılaştırmaktadır. Analiz edilen temel parametreler arasında Basso, Beattie and Bresnahan (BBB) lokomotor derecelendirme skalası ve Von Frey (vF) geriçekme eşiği gibi ölçümlerin yanı sıra basma fazı, salınım fazı, adım uzunluğu, basma fazı oranı ve pati alanı yer almaktadır.

İşaretleyicisiz poz tahmin aracı olan DeepLabCut (DLC) ile sadece yanal görüntü odaklı metotlarla ilişkili zorlukların üstesinden gelinerek yürüme parametreleri çıkarılmıştır. Daha sonra yürüme parametreleri, BBB skorları ve vF eşik değerleri farklı algoritmalar kullanan Kollektif Öğrenme yöntemine öznitelik olarak verilmiş ve yaralanma sonrası süre ve deneklerin nöromodülasyon tedavisi alıp almadıkları öngörülmüştür. OY olan hayvanlarda ön ve arka ayaklardan elde edilen yürüme parametreleri ve BBB skorlarıyla yaralanma sonrası 6 farklı hedef sınıfı %43 doğrulukla bulunmuştur. Aynı öznitelik kümesiyle tedavi ise %71 doğrulukla öngörülmüştür. Sonuçlar, yöntemin tedavi gruplarını ve OY süresini sınıflandırmada etkili olduğunu göstermiştir. Ancak doğruluğu arttırmak için ek çalışmalar gerekmektedir. Derin öğrenme ile parametre çıkarımı OY araştırmaları için erişilebilir, uygun maliyetli çözüm sağlamaktadır.

Anahtar Sözcükler: Omurilik Yaralanması, Yürüme Analizi, Makine Öğrenmesi, Sınıflandırma

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii
ACADEMIC ETHICS AND INTEGRITY STATEMENT v
ABSTRACT vi
ÖZET vii
LIST OF FIGURES x
LIST OF TABLES xviii
LIST OF ABBREVIATIONS xix
1. INTRODUCTION 1
1.1 Motivation and Aims 3
1.2 Outline 4
1.3 Novelty 5
2. BACKGROUND 6
2.1 Spinal Cord Injury 6
2.2 Gait Analysis in Humans 9
2.3 Gait Analysis in Rats 11
2.4 Machine Learning in Gait Analysis 13
3. MATERIALS AND METHODS 18
3.1 Experiment and Data Collection 18
3.2 Deep Learning for Video Analysis 20
3.2.1 Point-Oriented 21
3.2.2 Area-Oriented 24
3.3 Preprocessing for Classification 25
3.4 RatSuite Code Group 30
3.4.1 Data Wizard 30
3.4.2 Range Finder 30
3.4.3 Area Master 31
3.4.4 Gait Guru 32
3.4.5 Gait Mapper 32
3.4.6 Step Cycler 35

3.5 Machine Learning for Classification 36
3.5.1 Features and Response 38
3.5.2 Data Splitting 40
3.5.3 Validation Scheme 40
3.5.4 Feature Ranking 41
3.5.5 Cost Matrix 41
3.5.6 Machine Learning Model 43
3.6 Performance Metrics 43
4. RESULTS 45
4.1 Optimization of Data Collection 45
4.1.1 LtoR, RtoL Movement 45
4.1.2 Increasing Video Quality 45
4.1.3 Bottom View Perspective 45
4.2 Optimization of Preprocessing 48
4.3 Classification Data 48
4.4 Classification Accuracies 66
4.4.1 Classification in Normal Format 66
4.4.2 Classification in Extended Format 70
5. DISCUSSION 73
5.1 Main Conclusions 73
5.2 Previous Studies 74
5.3 Limitations 75
5.4 Future Work 75
5.5 Conclusion 76
REFERENCES 77

LIST OF FIGURES

Figure 2.1	A figure from the study of G. Garcia-Alias et al. (2010) [1].	7
Figure 2.2	A figure from the study of Timotius et al. (2021) [2]. Prog.: progesterone, Veh.: vehicle, dpi: days post injury.	15
Figure 2.3	A figure from the study of Aljovic et al. (2022) [3].	16
Figure 3.1	Figure showing the walking platform of the parkour. This frame was taken from a video of a rat.	20
Figure 3.2	Paws labels as shown in the figure. Label name, which is denoted by F in the image, varies according to the paw.	21
Figure 3.3	At the 100,000 iteration limit, the loss value up to the maximum iteration.	22
Figure 3.4	Point-Oriented labeling.	22
Figure 3.5	Figure showing flow of X and Y coordinates as the color-coded paws moving along the track. Since it moves on the horizontal axis, the Y coordinates remain constant while the X coordinates change with each stride. Stance, stride duration and stride distance can be observed. We used discrete color code for each paw but the DLC, in continous manner shows the color scale. As can be seen on the right legend, it is indicates which paw corresponds to which color code. In this study, Turkish equivalents were used for the labels.	23
Figure 3.6	Trajectory plot showing the placement of time-independent paw labels in the coordinate system along the gait. In an ideal Trajectory Graph we would not expect to see following lines. This shows that the model follows the paw even in the absence of contact. Legend on the right is same as Figure 3.5.	24
Figure 3.7	A frame showing Laplacian of Gaussian (LoG) method being used.	24
Figure 3.8	A frame showing Prewitt Operator (Prewitt) method being used.	24

Figure 3.9 A frame showing Sobel Operator (Sobel) method being used. While we can detect edges along the rats gait, it does not provide any useful information about contact. Uneven ground and non-uniform lightning factors make it difficult. 25

Figure 3.10 The scatter in the first frames shows the model's difficulty in following the paw. In an ideal figure we would expect to see no jumping lines. Mind the legend on the right, is changed. Instead of Turkish equivalents of paw labels, to ease the process of operators, Romanian equivalents, from laboratory interns, are used 2 for each paw; one for digits and one fore heel. SS: stanga spate (left hindlimb), SF: stanga fata (left forelimb), DS: dreapta spate (right hindlimb), DF: dreapta fata (right forelimb). 25

Figure 3.11 The jumps in the Figure 3.10 are clearly visible on the Y axis in here. Legend on the right is same as Figure 3.10. 26

Figure 3.12 "+" signs are human labels. "X" and "O" signs are model labels. "O" for high likelihood, "X" for low likelihood markings. A frame from the 7-label model on a video. It can be seen that the right rear paw (at the top) was not in contact and the model did not recognize it. Both high likelihood and low likelihood markings can be seen on the human labels on the right front paw (top left). On the left front paw (bottom), it can be seen that the paw was recognized close to contact, but not yet with high enough likelihood to be marked. 26

Figure 3.13 A frame from the 9-label model on a video. Almost perfect labeling with high likelihood values at the moment all paws are in contact. A label on the left front paw (far left) has a low likelihood for some reason. There may have been insufficient, incorrect labeling, or simply the model was not trained well enough. 27

Figure 3.14 A frame from the 9-label model on a video. Almost perfect labeling with high likelihood values at the moment all paws are in contact. A label from the right front paw (top right) jumps from the left front paw (bottom) and an another label receives a low likelihood value. Figure 3.13 reasons apply here too. 27

Figure 3.15 Figure showing a successful model. X and Y coordinates did not jump. The start and end of the Stride along the Gait are successfully recognized. Romanian equivalents are still in use as can be seen on legend at right. However as label size increased, for label count to each paw is also increased, legend grows. As digits and corners of paw started to label, order is given as Figure 3.2. 28

Figure 3.16 Paws and labels on paws are successfully separated from each other. Legend on the right is same as Figure 3.15. 28

Figure 3.17 Flow chart of Data Wizard. 31

Figure 3.18 Flow chart of Range Finder. 31

Figure 3.19 The user is instructed by the terminal to draw a line between the 2 markers. After the operation, the coversion ratio is saved in the workspace as a variable as it appears on the terminal. 32

Figure 3.20 Flow chart of Area Master. 33

Figure 3.21 Flow chart of Stride Distance. At this stage Gait Guru code group started to work. 33

Figure 3.22 Flow chart of calculation the stance, swing and stride durations. 34

Figure 3.23 Flow chart of calculating the maximum area of each paw while rats gait. 35

Figure 3.24 Gait map, showing in which frames the rat was in stance phase. 35

Figure 3.25 Gait cycle, showing the order in which the rat takes steps. 36

Figure 3.26 Flow chart of Gait Mapper and Gait Cycler. 37

Figure 3.27 Figure showing gait parameters, BBB, and vF scores extracted from the videos of rats 79 and 78 at the designated DPO dates in Normal Format. In the normal format, each column represents a feature, while DPOs and their corresponding data are listed as rows. This presentation highlights the straightforward alignment of data features across multiple time points. 39

Figure 3.28 Figure showing gait parameters, BBB, and vF scores extracted from the videos of rats 79, 78, and 77 at the designated DPO dates in Extended Format. In the extended format, features at each DPO are represented as columns, which allows for proportional scaling of the dataset. Despite including more rats, the dataset is visually expanded. Injury and treatment classifications are presented in their respective columns, and the figure is displayed in three consecutive parts for clarity. 39

Figure 3.29 The validation accuracy of the model was evaluated under varying proportions of training data during the model-building phase. As the proportion of training data decreased, validation accuracy tended to decline. This decline can be attributed to the reduced volume of data available for the model to learn patterns effectively, as a larger percentage of the dataset is allocated to testing. These findings highlight the critical role of sufficient training data in optimizing model performance and ensuring robust generalization. 40

Figure 3.30 The relationship between model accuracy and the number of folds in k-means cross-validation was analyzed during the model-building phase. The general trend indicates that accuracy improves as the number of folds in k-means cross-validation increases. This is because increasing the number of folds reduces bias by allowing the model to train on a greater variety of data subsets. However, this also enhances generalization by providing a more robust evaluation of the model's performance across diverse data splits. These observations underscore the trade-off between model bias and generalization as the k-means cross-validation configuration is adjusted. 41

Figure 3.31 A screenshot of the window used for Feature Ranking. The results obtained with the ANOVA method can be seen. 42

Figure 4.1 3 axis, showing Loss-Iteration-Training Duration correlation. 46

Figure 4.2 Correlation heatmap of parameters in training the paw tracking model. 47

Figure 4.3 Distributions of all variables in the dataset containing Sham and SCI rats, colored according to End-point Classes, in the study. Both of the classes contains both neuromodulated and non-neuromodulated rats. Styled according to Injury Classes. Circles (•): sham, squares (■): SCI models. "dpo": Day post Operation, "area": Area, "stance_dur": Stance Duration, "stride_dis": Stride Distance, "stride_dur": Stride Duration, "swing_dur": Swing Duration, "limb_duty": Limb Duty Factor, "speed": Speed, "bbb": Basso, Beattie and Bresnahan scoring, "vonFrey": Von Frey, "neuromod": Neuromodulation, "injury": Injury. 49

Figure 4.4 Continuation of the Figure 4.3. 51

Figure 4.5 Continuation of the Figure 4.4. 52

Figure 4.6 Continuation of the Figure 4.5. 53

Figure 4.7 Distributions of all variables in the dataset containing Sham and SCI rats, colored according to Treatment Classes, in the study. Sized according to End-point Classes. Orange colored (•): neuromodulated, Blue colored (•): non-neuromodulated. Parameters are the same as Figure 4.3. 54
Figure 4.8 Continuation of the Figure 4.7. 55
Figure 4.9 Continuation of the Figure 4.8. 56
Figure 4.10 Continuation of the Figure 4.9. 57
Figure 4.11 Distributions of all variables in the dataset containing only SCI model rats, colored according to Treatment Classes, in the study. Orange colored (•): neuromodulated, Blue colored (•): non-neuromodulated. Parameters are the same as Figure 4.3. 58
Figure 4.12 Continuation of the Figure 4.11. 59
Figure 4.13 Continuation of the Figure 4.12. 60
Figure 4.14 Continuation of the Figure 4.13. 61
Figure 4.15 Distributions of all variables in the dataset containing only Sham rats, colored according to End-point Classes, in the study. Orange colored (•): neuromodulated, Blue colored (•): non-neuromodulated. Parameters are the same as Figure 4.3. 62
Figure 4.16 Continuation of the Figure 4.15. 63
Figure 4.17 Continuation of the Figure 4.16. 64
Figure 4.18 Continuation of the Figure 4.17. 65

Figure 4.19 Prediction of DPO, using dataset with containing fore and hindlimb records. Dataset formatted as Normal. All groups include groups receiving and not receiving neuromodulation. "Gait+BBB+vF": Full feature set comprising gait parameters used in this study (area, stance duration, stride distance, stride duration, swing duration, limb duty factor, speed), BBB scoring, and Von Frey (vF) scoring, "Gait+vF": Gait parameters and Von Frey scoring, excluding BBB scoring, "Gait+BBB": Gait parameters and BBB scoring, excluding Von Frey scoring. "Gait": Gait parameters only, "BBB+vF": Combination of BBB and Von Frey scoring, excluding gait parameters, "vF": Von Frey scoring only, "BBB": BBB scoring only. For each feature set, results are categorized into three subject groups: “Sham+SCI”: combined injured and non-injured subjects, “SCI”: injured subjects, “Sham”: noninjured subjects. A chance level line is set at %16, representing the minimum accuracy required for random prediction across six endpoint classes (Pre-Op, D1, D7, D14, M1, M1.5, M2). This level serves as a baseline for evaluating model performance. 66

Figure 4.20 Prediction of DPO, using dataset with containing only hindlimb records. Dataset formatted as Normal. All groups include groups receiving and not receiving neuromodulation. Axis are same as Figure 4.19. 67

Figure 4.21 Prediction of treatment, using dataset with containing fore and hindlimb records. Dataset formatted as Normal. All groups include groups receiving and not receiving neuromodulation. A chance level line is set at %50, representing the minimum accuracy required for random prediction across two treatment classes (neuromodulated, non-neuromodulated). This level serves as a baseline for evaluating model performance. Axis are same as Figure 4.19. 68

Figure 4.22 Prediction of treatment, using dataset with containing only hindlimb records. Dataset formatted as Normal. All groups include groups receiving and not receiving neuromodulation. Axis are same as Figure 4.19. 69

Figure 4.23 Prediction of DPO of models supported by the toolbox, using dataset with containing fore and hindlimb records. Dataset formatted as Normal and contains only SCI model rats. All groups include groups receiving and not receiving neuromodulation. Bars are colored according to feature sets as stated in Figure 4.19. 70

Figure 4.24 Prediction of treatment of models supported by the toolbox, using dataset with containing fore and hindlimb records. Dataset formatted as Normal and contains only SCI model rats. Bars are colored according to feature sets as stated in Figure 4.19. 70

Figure 4.25 Prediction of treatment, using dataset with containing fore and hindlimb records. Dataset formatted as Extended. Axis are same as Figure 4.19. 71

Figure 4.26 Prediction of treatment of models supported by the toolbox, using dataset with containing fore and hindlimb records. Dataset formatted as Extended and contains only SCI model rats. Bars are colored according to feature sets as stated in Figure 4.19. 72

LIST OF TABLES

Table 3.1	Summary of animals in each group. SCI: spinal cord injury, Sham: not injured, NS+: neuromodulated, NS-: non-neuromodulated.	19
Table 3.2	Number of rows of the data sets used for model training. Each row corresponds to one record.	38
Table 3.3	Default cost matrix. Tested to reinforce True Positive and False Negative predictions.	42
Table 3.4	Customized cost matrix. The confusion matrix output of the model was analyzed and weights were distributed in effort to correct False Positive predictions.	42

LIST OF ABBREVIATIONS

SCI	Spinal Cord Injury
DLC	DeepLabCut
DPO	Day Post Operation
DPI	Day Post Injury
BBB	Basso, Beattie, and Bresnahan (Locomotor Rating Scale)
vF	von Frey (Withdrawal Threshold Rating Scale)
LoG	Laplacian of Gaussian
ALS	Amyotrophic Lateral Sclerosis
MS	Multiple Sclerosis
FPS	Frames Per Second
3D	Three-Dimensional
GAITOR	The Gait Analysis Instrumentation and Technology Optimized for Rodents
Agatha	Automated Gait Analysis Through Hues and Areas
ALMA	Automated Limb Motion Analysis
PrAnCER	Paw-Print Analysis of Contrast-Enhanced Recordings
BoS	Base of Support
RPP	Relative Paw Position
ML	Machine Learning
RF	Random Forest
AI	Artificial Intelligence

1. INTRODUCTION

With a high rate of morbidity, SCI is a serious and fatal condition of the neurological system [4]. The impact of SCI on individuals' lives underscores the urgency of research in this field. Studies conducted in animal models, particularly rats, have played a pivotal role in advancing our understanding of SCI. Rats are chosen for these investigations for several reasons, such as their cost-effectiveness and the availability of comparative data in the existing literature [5]. Furthermore, rats exhibit physiological responses to SCI that are closely similar to human reactions; in rats and humans alike, the development of cysts cranial and caudal to the site of injury is common, making them valuable substitutes for experimental purposes [6]. Gait Analysis is a crucial tool in examining the immobilizing effect of SCI on the individual and the effectiveness of treatment. One of the primary objectives of SCI research in rats is to assess the effects of these injuries on gait parameters of gait analysis such as contact area, stance duration, swing duration, stride duration, stride distance, limb duty factor, base of support (BOS), and relative paw position (RPP). Animals with SCI may exhibit a variety of motor impairments, including shorter stride distance, decreased swing speed, a wider BOS, and reduced regularity of gait. These impairments can make it difficult for animals to walk and can lead to secondary complications such as muscle atrophy and joint pain. Research in gait analysis of SCI in rats began with the simple method of painting the rats' paws and observing their movements on paper. However, this method had limitations, including the challenge of ensuring consistent paint marks and the changing characteristics of the paint over time, preventing it from advancing significantly. The Basso, Beattie, and Bresnahan (BBB) scores are widely used in assessing the effectiveness of surgical operations and therapeutic interventions in SCI rats. However, the challenge lies in the need for more sensitive tools to evaluate gait recovery accurately, as current analyses often concentrate on the BBB score. The BBB score primarily focuses on hindlimb function [7], potentially limiting comprehensive gait analysis. Additionally, the site of injury within the spinal cord can impact the BBB score, with dorsal region injuries showing fewer score variations [2]. Furthermore, the subjectivity

involved in BBB scoring, despite the requirement for trained evaluators, introduces an element of uncertainty to the assessments [8],[2],[9]. For these reasons, alternatives that are cost-effective and have low human error and resource requirements are being sought. As computer technologies have advanced, scientists have begun to utilize these advances in the study of gait analysis in SCI. With the start of digital data creation and transmission and the development of computer vision, the accuracy of research has increased. Automated gait analysis systems, such as the CatWalk, have revolutionized the field by providing quantitative and reproducible metrics of rodent gait [10],[11]. Using high-speed cameras and a glass walkway, records paw placements, calculating parameters such as stride distance, swing time, BOS, and the regularity index. These systems minimize observer bias and enhance data reliability. However, certain limitations, such as the system's dependency on the rat's coat texture and the high cost of equipment, pose challenges for widespread adoption. VICON, one of the effective uses of computer vision, can directly transfer physiological parameters from the physical domain to the digital domain for interpretation, which, despite their accuracy, requires extensive space and significant financial investment. Recent trends in this field have introduced markerless pose estimation tools, such as DeepLabCut (DLC), which utilize deep learning techniques to track not only rats but also various moving objects [12]. DLC offers the advantage of non-intrusive and automated motion tracking, making it a valuable tool in gait analysis, and it has the potential to significantly reduce the financial and resource burdens associated with traditional gait analysis methods. DLC not only enables precise and automated tracking of movement but also provides the opportunity for multi-object tracking and the analysis of complex behavioral data, making it a versatile and cost-effective alternative for researchers in the field. The machines are not only used to detect the paws of walking rats and extract gait parameters from the detected paws. The researchers also used the machines to classify rats into specific classes. For this, the machines learned [3],[13]. The high expenses, resource demands, and expertise necessary for operating sophisticated gait analysis tools such as CatWalk and MOCAP systems such as VICON pose significant barriers. Certainly, the challenges surrounding gait analysis extend beyond the cost constraints of advanced tools. Certain existing classification methods might lack the requisite precision due to measurement uncertainties, variability in underlying parameters, or methodological

limitations yet prove significant value in research. These restrictions may limit their use in particular situations considering the subject's physiological attributes—such as height, gender, and weight—vary and there is genetic variation, emphasizing the fact that the problem goes beyond simple financial considerations. To handle the complexity of gait analysis in SCI research, it is necessary to take into account the general accessibility, accuracy, and application of current methods in addition to their financial expenses. Here we are proposing a cheaper, easy to implement and interpret and has broader application frame for gait analysis that can be classified according to certain classes.

1.1 Motivation and Aims

The impact of SCI on individuals' lives highlights the urgency of conducting research in this area. In this context, we aim to utilize gait analysis as a vital tool to assess the effects of SCI on gait parameters in rats. Traditional gait analysis tools, such as CatWalk or VICON, present significant challenges for many laboratories due to their high cost, resource demands, and required expertise. Additionally, current SCI locomotor rating scale (e.g., BBB scoring) and the limited application of certain classification techniques further emphasize the need for more accurate, affordable, and accessible alternatives. Our primary objective is to introduce an innovative approach to gait analysis and classification in SCI rodent models. This involves the use of markerless pose estimation tools, specifically DLC, which are widely utilized in rodent imaging but are often limited to lateral plane analysis. To overcome this limitation, we propose acquiring bottom view images and extracting unique data from these perspectives, which are typically overlooked in traditional methods. Through this approach, we aim to develop a classification model that effectively captures and differentiates time after injury and neuromodulation treatment across the gait cycle. By doing so, we intend to improve the accuracy of assessing gait function in SCI rat models. Additionally, we aim to offer a low-cost, customizable, and open-source program, making it easier and more affordable for researchers to study SCI. Our goal is to create an optimized classification model based on bottom view imaging gait data, demonstrating that it is

possible to achieve results comparable to professional tools without requiring complex and expensive equipment. We expect that the outcomes of this new approach will align closely with those of traditional methodologies. At the same time, we recognize potential limitations in analyzing and classifying rat gait. To address these challenges, we aim to offer a simpler yet effective alternative, tailored for research facilities with limited resources or capacity.

1.2 Outline

The purpose of the thesis is to classify time after SCI and neuromodulation treatment in rats by ML of gait parameters and BBB scores and vF thresholds. The thesis begins with an introduction that provides an overview of SCI and its impact on gait, emphasizing the significance of rat models for studying SCI. The motivation, aims, and novel aspects of the proposed approach, such as using markerless pose estimation tools for gait analysis, are also outlined in this section. Following the introduction, the background chapter explores fundamental concepts, including SCI, traditional gait analysis methods in humans and rats, and the integration of ML into gait analysis. The materials and methods chapter details the experimental setup, including the acquisition of bottom view imaging data, preprocessing workflows, and the implementation of ML algorithms for classification. This chapter also covers the methodological innovations in extracting gait parameters and optimizing the data pipeline. The results section presents the findings of the study, including the optimization of data collection, preprocessing methods, and the accuracy of classification models in differentiating gait parameters. It highlights the advantages of the bottom view imaging approach and its impact on improving classification accuracy. The discussion chapter evaluates the results in light of previous studies, addresses the limitations of the current approach, and proposes future directions for enhancing gait analysis in SCI research. Finally, the conclusion summarizes the key findings, discusses their implications for the field, and outlines the broader contributions of the thesis to advancing accessible and accurate SCI research methodologies.

1.3 Novelty

We are proposing a innovative contributions, centering around the application of advanced deep learning and machine learning techniques in the analysis and classification of gait parameters in SCI rat models. By employing DLC, we were able to convert raw coordinates into meaningful gait parameters such as stride distance and paw area. This approach simplifies and automates gait analysis, offering a robust alternative to traditional methods. With these extracted parameters, we utilized machine learning to predict key outcomes in SCI models. Specifically, the study classifies the time after injury and the application of neuromodulation therapy with machine learning. By integrating this dual-layered deep learning and machine learning we are presenting a novel framework for analyzing and classifying SCI recovery in rats. This novel methodology not only advances the utility of deep learning in biomedical research but also bridges the gap between computational analysis and clinical applicability, offering a cost-effective and scalable alternative to existing tools. By extracting hidden parameters and using predictive modeling, we provide a basis for improving the accuracy and efficiency of SCI research.

2. BACKGROUND

2.1 Spinal Cord Injury

SCI is a complex and debilitating condition that disrupts the communication pathways between the brain and the body. These injuries can lead to partial or complete loss of motor, sensory, and autonomic functions below the injury site, deeply affecting an individual's quality of life. SCI is classified into complete injuries, which result in the total loss of function, and incomplete injuries, where some degree of functionality remains. The extent of damage and recovery potential depends on the injury's location and severity [4].

SCI occurs in two distinct phases. The primary phase is the result of the initial mechanical trauma, causing direct cellular damage, disruption of neural pathways, and local vascular injuries. This is followed by the secondary phase, where processes such as inflammation, oxidative stress, neural excitotoxicity, and scar tissue formation worsen the damage. These secondary mechanisms create significant barriers to recovery by amplifying the initial damage and limiting the spinal cord's regenerative capacity [14],[8].

Animal models, particularly rats, are widely used to study SCI due to their structural and functional similarities to the human spinal cord. Controlled injury methods, such as the weight-drop and impactor techniques, allow researchers to induce injuries of varying severities, mimicking the clinical presentation of SCI in humans. These models have been essential in understanding the progression of SCI and testing potential of therapeutic strategies [4],[15].

Figure 2.1 shows the different outcomes of a rat with SCI at different designated day post operation (DPO) dates. Figures A and B are a contusion SCI and Sham; immunostained transverse image of the T8 spinal segment. The same injury model

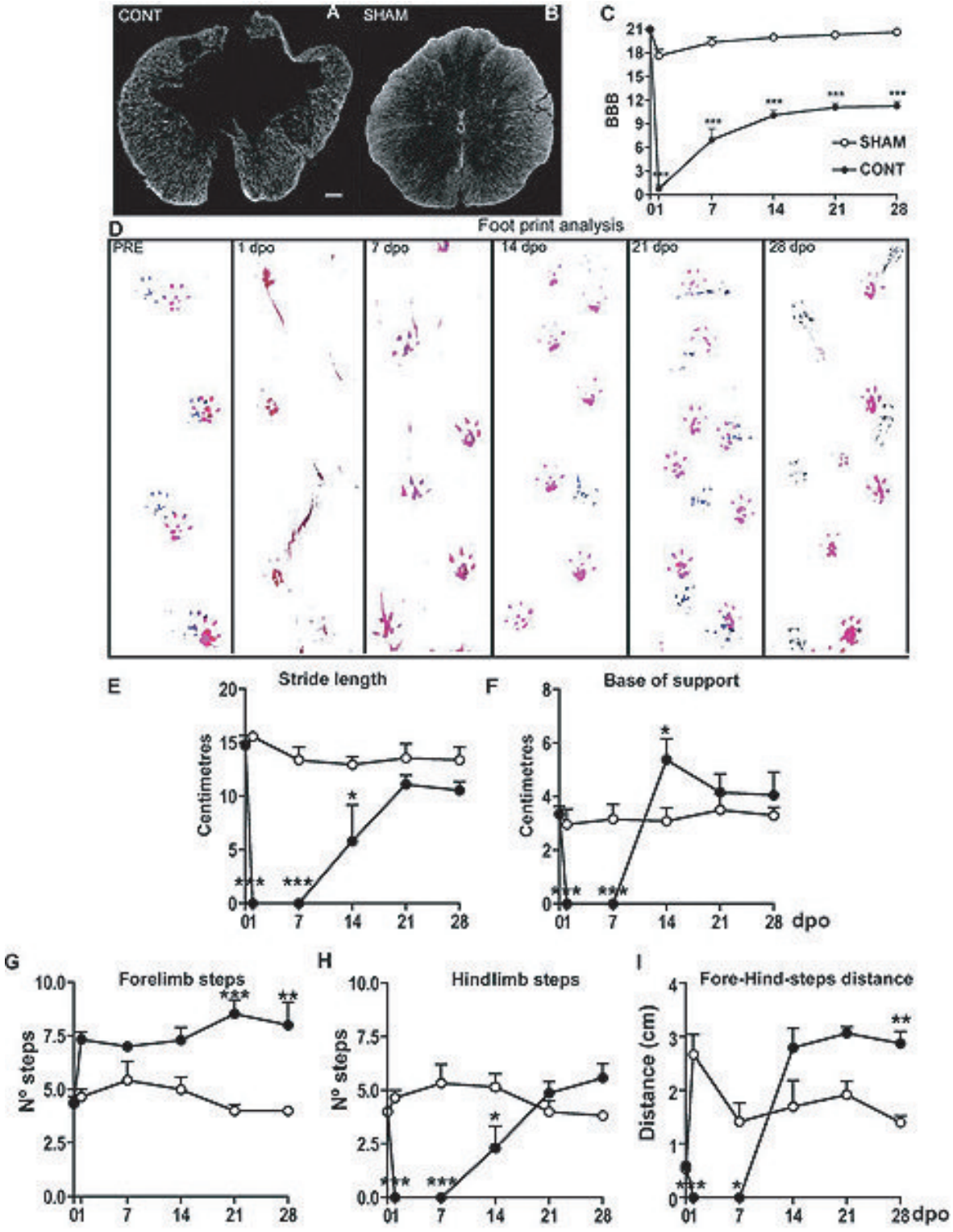


Figure 2.1 A figure from the study of G. Garcia-Alias et al. (2010) [1].

was used for the rats in this study. Figure C shows the changes of the BBB score at the reported DPO dates. The recovery curve of the SCI model rat follows a much lower score, while the sham recovered very quickly from the mild damage caused by the operation. Figure D shows the front paw print analysis at the same DPO dates. It can be seen that the rat is driving its paw especially at DPO 1 (2nd column). Figure E shows the measurements of stride distance and Figure F shows the BOS parameters at the DPO dates between the sham and SCI models. The stride distance of the sham rats has very small fluctuations, while the SCI model rats did not take a step until DPO 14. The researchers were only able to collect BOS data at DPO 14.

The assessment of functional recovery after SCI often focuses on locomotion, as it is a key indicator of neural and motor function. Traditional tools like BBB scale provide subjective evaluations of movement patterns, such as hindlimb coordination and weight-bearing. More advanced systems, such as CatWalk, offer quantitative analysis of gait parameters, including stride distance, paw placement, and interlimb coordination. These tools are essential for detecting subtle changes in recovery and evaluating the efficacy of experimental treatments [2],[8].

Advances in technology have also introduced innovative methods for analyzing movement. For example, markerless kinematic tracking systems and ML algorithms are increasingly used to capture detailed gait parameters with minimal disruption to the natural behavior of animals. These methods reduce errors associated with traditional marker-based systems and provide more comprehensive data for evaluating recovery [12].

Researching SCI in animal models helps us better understand the mechanisms behind the injury and helps guide human rehabilitation approaches. The parallels between rat and human response to SCI make it possible to translate findings from preclinical studies into practical applications. By combining insights from traditional evaluation techniques with modern tools like ML, researchers are paving the way for more accurate and efficient assessments of gait and recovery. To fully understand these connections, it is essential to explore how gait analysis is performed in humans before

comparing it to approaches used in animals. This progression highlights the evolving nature of gait analysis and its potential to revolutionize SCI research and rehabilitation.

2.2 Gait Analysis in Humans

The study of human gait has evolved significantly, moving from simple observations to advanced, technology-driven methods. This progress combines biomechanics, engineering, and clinical sciences to deepen understanding of human locomotion. Historically, Aristotle's work in 350 BCE linked joint movements to muscles, while Giovanni Borelli's Renaissance studies integrated anatomy and physics. In the 19th century, Marey and Muybridge used photography to provide empirical insights into movement. By the 20th century, marker-based motion capture systems like those by Davis et al. (1991) utilized video processing and reflective markers, revolutionizing gait analysis through advancements in computing [16].

Modern gait analysis employs systems that capture, measure, and interpret biomechanical data. The gait cycle, starting with one heel strike and ending with the next, consists of the stance phase (foot on the ground) and swing phase (foot in the air), together offering detailed insights into movement patterns [17]. Key measurements include kinematics, such as joint angles, kinetics like forces, and spatiotemporal features including stride length, cadence, and walking speed [17],[16].

Technologies used today include wearable sensors, marker-based systems, markerless motion capture, and force platforms. Wearable devices, equipped with accelerometers and gyroscopes, provide compact, real-time data [18]. Marker-based systems deliver precise 3D analyses but require controlled settings [16]. Recent advancements in markerless systems, powered by deep learning tools like OpenPose, offer non-intrusive, cost-effective alternatives. Force platforms measure ground reaction forces, aiding in kinetic assessments [17].

Applications of gait analysis span medicine, rehabilitation, sports, and security.

It aids in diagnosing neurodegenerative diseases like Parkinson's, where subtle signs such as reduced stride length and freezing of gait provide early warnings [17]. In rehabilitation, personalized therapy plans address deficits in post-stroke or spinal injury recovery [19]. Sports science uses biomechanical evaluations to optimize performance and prevent injuries (Xu et al., 2024). Additionally, unique gait patterns are utilized in biometric security systems [18].

Emerging technologies, including artificial intelligence (AI) and ML, are transforming the field by automating feature extraction and enhancing diagnostic precision [20],[21]. Real-time data analysis from wearable devices enables remote monitoring, expanding accessibility [16]. Challenges remain, such as variability in gait due to age and health, but affordable, user-friendly systems are a growing priority. AI-driven diagnostics and predictive models offer potential for more accurate assessments and treatments.

Gait analysis is particularly transformative in neurodegenerative diseases like Parkinson's, Alzheimer's, and ALS. Subtle changes in gait can signal early stages of these conditions, guiding timely interventions. For example, dual-task walking combining cognitive and motor tasks—reveals the connection between cognitive decline and motor impairments in Alzheimer's. In ALS, gait tracking helps monitor disease progression and informs care decisions.

Future developments aim to refine diagnostic tools and expand accessibility through smartphone-based systems and wearable technologies. Gait features may also serve as digital biomarkers, offering non-invasive ways to track disease and treatment effects. By integrating advanced technologies and interdisciplinary research, gait analysis is poised to revolutionize clinical practices, enhancing patient care and deepening understanding of human movement.

2.3 Gait Analysis in Rats

Gait analysis in rats is a key tool in neuroscience research, especially for studying SCI and other neurological conditions. This method combines behavioral observations, kinematic measurements, and computational techniques to evaluate locomotor function. It helps researchers understand movement patterns, test therapies, and link functional recovery to molecular and histological changes. Developments in automated systems and computational tools have extended the field and allowed for better comprehension of locomotion. Rats are widely used in neuroscience due to their cost-effectiveness and physiological similarities to humans. These models are essential for SCI, testing regenerative therapies, and exploring neural repair mechanisms.

Gait analysis also plays a critical role in studying neurodegenerative diseases like ALS and Parkinson's disease [22],[3]. By combining kinematic and dynamic analyses, such as joint trajectories and swing velocities, researchers can track recovery patterns and correlate functional outcomes with molecular changes. The ability to detect subtle gait changes enables early diagnosis and timely interventions [14],[5]. From painting paws [23],[24], to training machines to recognize paws [25] Since the first locomotor performance assesment studies by W. Sperry in the 1940s [26],[27],[28], many different methods have been used for gait analysis.

The Basso, Beattie, and Bresnahan (BBB) scale, developed in 1995, remains a widely used method for assessing rat locomotion [7],[9]. It scores hindlimb movement, weight support, and coordination during free movement. Despite its simplicity, the BBB scale is subjective and prone to variability, highlighting the need for quantitative approaches. However, its standardization has ensured its continued use in preclinical research.

Rats were walked on paper for a while, with their paws treated with ink or other marking agents. Recording the steps, which form the backbone of the rats' gait parameters paved the way for new research. However, the mess and inaccuracy created by such fluid substances not only made the analysis difficult, but also rendered

the studies inadequate due to the volatility of some substances and their insufficient tracing potential [29].

Developing photographic technologies were also used to prevent this. Mice paws on photographic films were treated with photo devoloping inks [30]. In this way, the messiness of the ink and the inadequate traces were relatively avoided with a technological approach. However, the process is cumbersome, costly and can be disrupted by small errors, which has led to the search for solutions.

With the development of computers and computer technologies, researchers have started to take advantage of these developments in gait analysis. With the digitalization of data, they started to perform parametric inference calculations on computers with more accurate and digitally unbiased calculations [10],[11]. Again, paws painted with ink were carried out on paper, but this time the paper was transferred to the computer using scanners to improve the accuracy of the calculations. As with almost every new technology, it was not immediately embraced by the majority and had to wait for methods to evolve and become more feasible.

As developing photographic films were digitized and digital cameras became widespread, research also benefited from these developments. By collecting recordings, this time directly from digital cameras, researchers were able to store the rats' walks so that they could be analyzed again and again later on the computer [31]. The digitization of the data, as well as its digital transmission, sped up the process, but the calculations were still not automated, which took longer and could lead to biased results because they were open to human interpretation.

As technology continues to evolve, researchers looking to optimize and automate the process have been working to improve existing methods while other methods continue to get better. Systems and technologies such as CatWalk, Digigait, Treadscan, some of which we still hear about today, and Footprints, which we no longer hear about but which is a timeless breakthrough [32], have come to light [33]. Sophisticated and complex, these systems could provide a plug-and-play approach in some cases, while

in others they required additional arrangements. The reflection of some rats' fur and the lack of a uniform computational technique for each rat, limited systems usefulness. Although it provided high performance, its high cost and specialized use also limited its widespread use. Although the current literature is dominated by studies using these or similar systems and these have the most reliable outputs, almost every study points to the same drawbacks.

Researchers using their competencies to solve these problems started to develop their own methods with open source solution approaches with high accuracy, low cost and general applicability. In particular, software such as Agatha and Gaitor started to do what sophisticated systems do at a fraction of the price [34]. The fact that the code was open-source and customizable was a big plus and allowed for wide use. However, again, since this type of software is not commercially sold, it appeals to a more niche audience and some technological competence in installation and use is expected. In addition, the need for more automated and human bias-free solutions for frame-by-frame video analysis, which consumes the largest amount of time resources in gait analysis, continued [29],[12].

2.4 Machine Learning in Gait Analysis

Researchers have rolled up their sleeves to create a solution to this problem and combine ML with computer vision. Especially in terms of ease of use and performance, solutions such as DLC and VisualGaitLab (VGL), which integrate the deep learning systematics of ML into research [25],[35].

DLC alone does not provide sufficient parametric calculations for gait analysis, but VGL also uses DLC's infrastructure for model training [31]. Researchers who wanted and tried to address both perspectives independently of each other developed their own methods with software such as Automated Limb Motion Analysis (ALMA) and PrAnCER [29],[3]. Another point where software such as PrAnCER differed was the ability to classify both parameters, paws and the condition of the mice using this

quantitative information. PrAnCER was able to classify the paws by analyzing the video, but it did not provide clear information about the condition of the rat.

Timotius et al. (2021) focused on analyzing recovery in SCI rat models, especially for different types and severities of thoracic and cervical injuries. Instead of using machine learning for predictions, they applied Linear Discriminant Analysis (LDA) to create a scoring system called the SCI Gait Index. This index combined nine key gait parameters collected from the CatWalk system to better understand recovery. They found that this method was more sensitive than the traditional Basso, Beattie, Bresnahan (BBB) score in detecting recovery differences. While their approach highlighted the potential of combining gait parameters for analysis, they didn't report prediction scores like accuracy or precision, as their goal was to enhance recovery assessment, not make predictions. In Figure 2.2 precision can be seen as it compared with BBB score. Especially in the subfigure e, developed method shows where it shines, as it can discriminate the young from the old while BBB lacks. This figure combinations of gait parameters (pLDA) and BBB scores for vehicle and young groups in their Studies done in the paper.

As mentioned above Aljovic et al. (2022) ALMA Toolbox, uses Random Forest (RF) to classify recovery stages after SCI and predict disease onset in multiple sclerosis (MS) models. The toolbox is built on DeepLabCut, which allows markerless motion tracking. ALMA achieved impressive results, with %98 accuracy for identifying early recovery stages (baseline to 3 days post-injury) and %86 accuracy for predicting MS onset a day before symptoms appeared. The system highlighted key gait features like stride distance and joint angles, providing researchers with an easy-to-use tool for kinematic analysis that is affordable and highly sensitive. In the Figure 2.3 they are showing the model prediction accuracies (g-right) as well as significant features (g-left) and rats posture in determined DPI points (f). The high predictive capability of the study cannot be ignored, but I would like to point out that the predictions were made between groups of 2. In the subfigure g-right, confusion matrix tables of the random forest method is shown. In the subfigure f, as in the Figure 2.1 figure, the mouse is seen to be dragging in the early phases of the injury, but now at DPI 21 the recovery

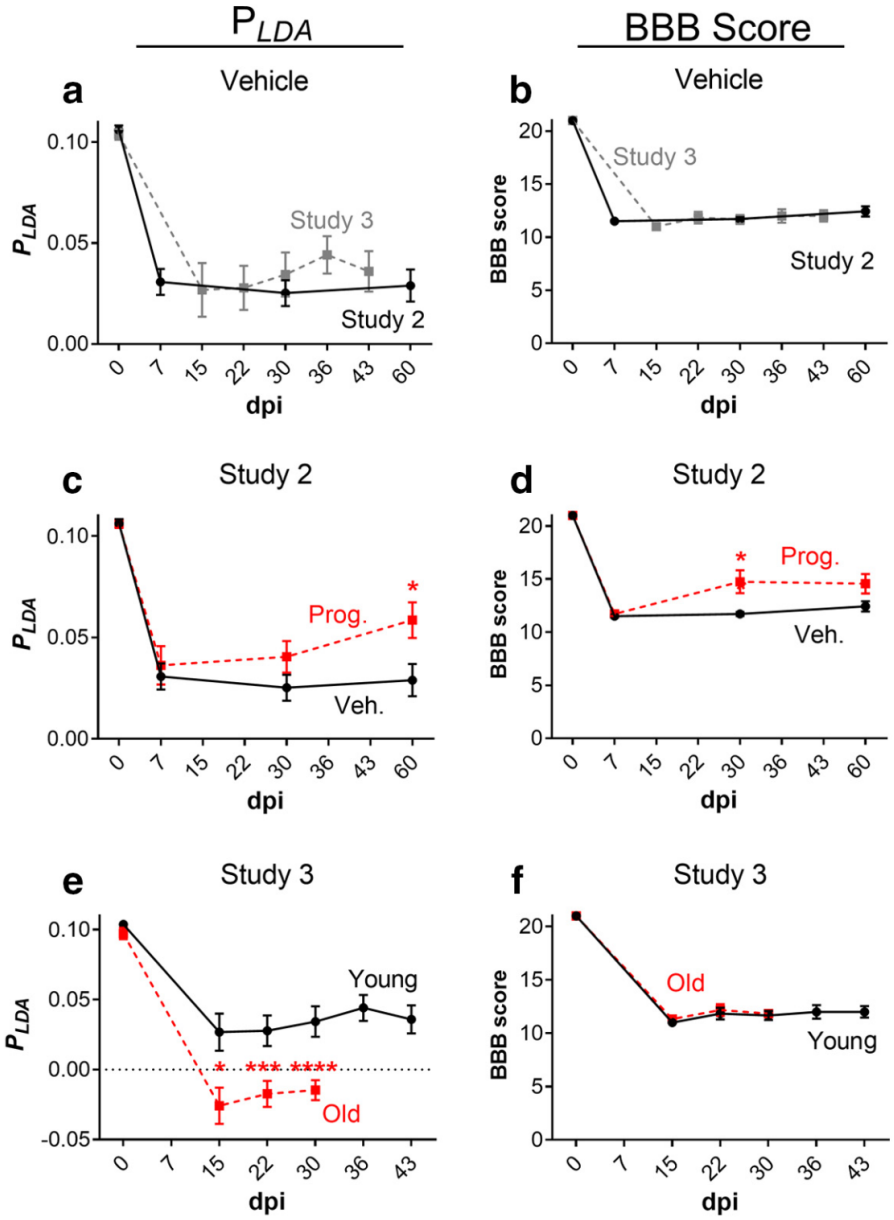


Figure 2.2 A figure from the study of Timotius et al. (2021) [2]. Prog.: progesterone, Veh.: vehicle, dpi: days post injury.

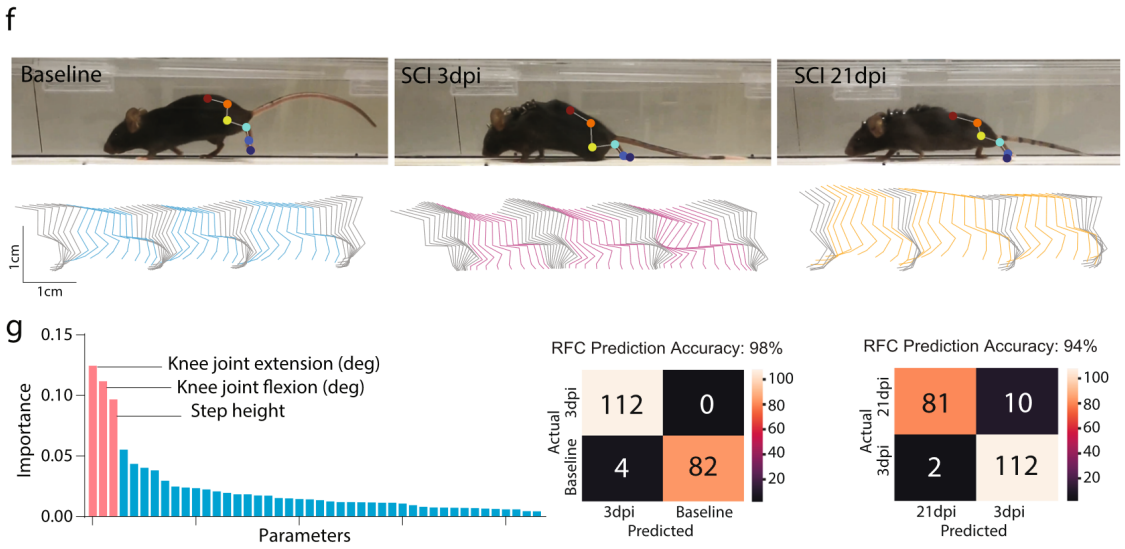


Figure 2.3 A figure from the study of Aljovic et al. (2022) [3].

is almost complete.

Huang et al. (2021), they studied gait in J20 transgenic mice, a model for Alzheimer's disease. Using RF, they classified the mice into their respective genotypes with about %93 accuracy. Their work identified age-specific gait features, such as trunk tilt and shoulder movement, which helped distinguish between the J20 and wild-type mice. The team also used regression models to predict locomotion metrics like running speed and dragged steps with high reliability. This study showed how kinematic gait analysis can detect motor deficits early, even before cognitive symptoms appear, highlighting its potential for early diagnosis in neurodegenerative diseases.

As a locomotor performance assessment tool, which is the origin of gait analysis, these tools offered incredible opportunities to investigate the conditions that affect locomotor performance [36],[37]. By examining the rat models in which gait analysis is performed with these classification methods, the models developed can predict the neurodegenerative and SCI groups to which the rat belongs [38]. In gait analysis, from the beginning to the end, from recording the mouse's gait, extracting the necessary gait parameters from the gait, processing these gait parameters and making predictions on the mouse using these outputs, ML is benefited at one point or another and the process

is accelerated. However, these case-specific approaches miss the broader application aspect over the head and limit its use. In some cases, mice are expected to have a specific phenotype; they must be within certain species, sex and weight restricted ranges; and data can be collected on a specific locomotion limit of the mouse, only some time after the operation or at a specific injury severity. Although the process has been shortened, its cumbursome has been reduced and its areas of use have been expanded since the early days, whether it is the track on which the mice will be walked, the camera on which the walking mice are recorded, the software that will analyze the rat's gait or classify the analysis outputs, at some point, the high cost and high technological expertise as well as the limited usage area continue to show itself. In addition to the need for research to move away from being sophisticated enough to become niche and to be applicable in a more general framework, the need for interpretability continues to manifest itself [2],[8].

3. MATERIALS AND METHODS

3.1 Experiment and Data Collection

The data that's been used for classification in this study were extracted from videos of rats, obtained within the framework of the TÜBİTAK [Grant 221N399] within European Union's FLAG-ERA JTC 2021 project GRAFIN Acronym RESCUEGRAPH and carried out under the supervision of Prof. Dr. Burak Güçlü. These videos were captured by Berfuğ Yaren Karaharman, a master's student at the Institute of Biomedical Engineering, with whom I collaborated as part of the mentioned project. The videos and output data from these videos, originally used within the project mentioned above, been served as a resource for this thesis. The study was approved by the Boğaziçi University Institutional Animal Experiments Local Ethics Committee (HADYEK). Study conducted and video recording are obtained at the Boğaziçi University Vivarium. The aforementioned study is related to trancutaneous electrical stimulation treatment method applied to SCI in rats. In this study, sham (non-injured) and SCI model groups were established. These groups were also divided into groups receiving neurostimulation and not receiving neurostimulation. This thesis uses all groups of rats. This thesis utilized 63 rats and data from the aforementioned study. Among them, 47 were injured (SCI), and 16 were sham-operated controls. Out of the 47 SCI model rats, 24 received neuromodulation therapy, while 23 did not. In the sham group, 8 rats received neuromodulation, and 8 did not. Most of the injured animals had moderate injuries, with some having mild injuries. This setup allowed us to analyze the effects of injury severity and neuromodulation therapy on gait parameters. The distribution and the injury, treatment classes of the animals can be seen in the Figure 3.1.

The recorded videos were obtained at 120 FPS with a width of 1920 pixels and a height of 1080 pixels. The videos are of varying length depending on the condition and performance of the rat. The videos contain 2 perspectives, lateral plane and bottom view, of the rat running on the transparent platform according to the plan of

Table 3.1
Summary of animals in each group. SCI: spinal cord injury, Sham: not injured, NS+: neuromodulated, NS-: non-neuromodulated.

SCI		Sham	
NS+	NS-	NS+	NS-
24	23	8	8

the experiment mentioned above. Lateral plane and bottom view perspectives were obtained with a mirror placed under the gait platform at a 45-degree angle. There are markers on the platform at one centimeter intervals. These markers are visible in the videos and are used as a reference for calibration. Each rat was walked on the track at designated day post operations (DPO) according to the experimental plan. These DPOs were before operation (Pre-Op), 1 day post operation (D1), 7 days post operation (D7), 14 days post operation (D14), 1 month post operation (M1), 1.5 months post operation (M1.5), 2 months post operation (M2). The technical structure dimensions of the track on which the rats are walked are as follows:

- Width (Parkour): 14 cm
- Length (Parkour): 150 cm
- Width (Mirror): 25.5 cm
- Height (Glass panel to top): 30 cm
- Height (Glass panel to bottom): 20 cm
- Distance (Camera to parkour): 87 cm
- Thickness (Frame): 3.5 cm
- Angle (Mirror placement): 45°

In the Table 3.1 the walking platform of the parkour can be seen. The markers mentioned in the Section 3.1 can be seen in the top and middle parts of the frame. The bottom view image is obtained with the mirror placed at 45°. Videos were recorded using the GoPro HERO4 Black v05.00 High Definition Action Camera. For the development of the ML model for paw tracking, we utilized a computer system with a CPU running at 2.8 GHz ES and Intel Core Skylake-X architecture, equipped with 2 NVIDIA GeForce RTX 3060 GPUs, 32 GBytes of DDR4 memory, and running on Windows 10 Pro x64, located in the Tactile Reasearch Laboratory of Boğaziçi University Biomedical Engineering Institute.



Figure 3.1 Figure showing the walking platform of the parkour. This frame was taken from a video of a rat.

3.2 Deep Learning for Video Analysis

The current version of DeepLabCut at the time of model creation, version 2.3.5, is available via the software's GitHub repository. A total of 1260 labelled frames from 35 videos acquired within the scope of the project were used. The image extraction process employed the k-means (image-based clustering) method. We opted for the kmeans clustering method because it down-samples the video and clusters frames using k-means, treating each frame as a feature vector. Frames from different clusters were selected to ensure feature diversity.

In each frame, a minimum of 1 and a maximum of 9 labels were used depending on the contact state of the paw. These labels were specific to the point on the rat's paw that they represented. 4 of these labels are on the 4 end edges to form a quadrilateral on the paw, and the remaining 5 labels represent the tip of each digit. Depending on the phase of the rat's paw in the gait cycle, such as heel-strike or toe-off, we placed the labels on the appropriate areas depending on the contact surface. 2 operators assisted in labeling the frames. Paws were labeled as shown in Figure :paw$_l abels$

We used the ResNet_50 neural network architecture and 100.000 iterations by keeping the default setting to train the model with the labelled images. We left the

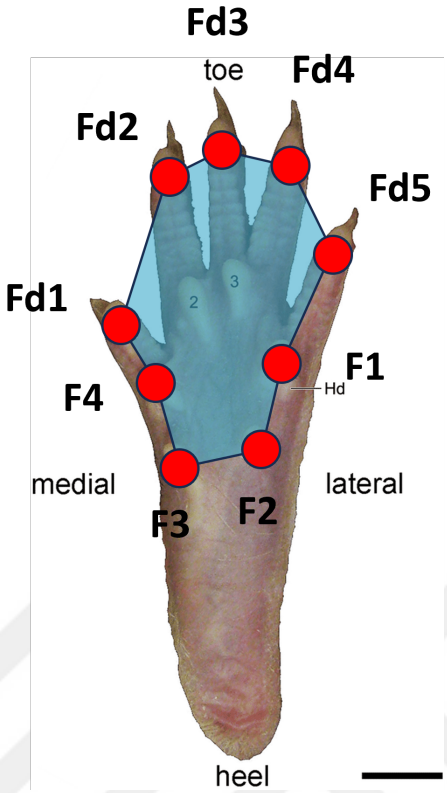


Figure 3.2 Paws labels as shown in the figure. Label name, which is denoted by F in the image, varies according to the paw.

data split as %95 for training and %5 for testing by default. We observed that our loss value plateaued after 60,000 iterations. Iteration and loss value plot is as shown at Figure :loss-iteration

3.2.1 Point-Oriented

Paw tracking model studies started with a single tag. Accordingly, each paw was associated with a tag. These tags were carefully followed during the labeling process of the frames. The tags we gave to each paw took the respective names of the paws. Accordingly: left front paw Left-Fore (LF), right front paw Right-Fore (RF), left back paw Left-Hind (LH), right back paw Right-Hind (RH).

At this stage, we gained useful information in the training of the paw tracking model. Thanks to this information, we were able to diagnose the deficiencies of our

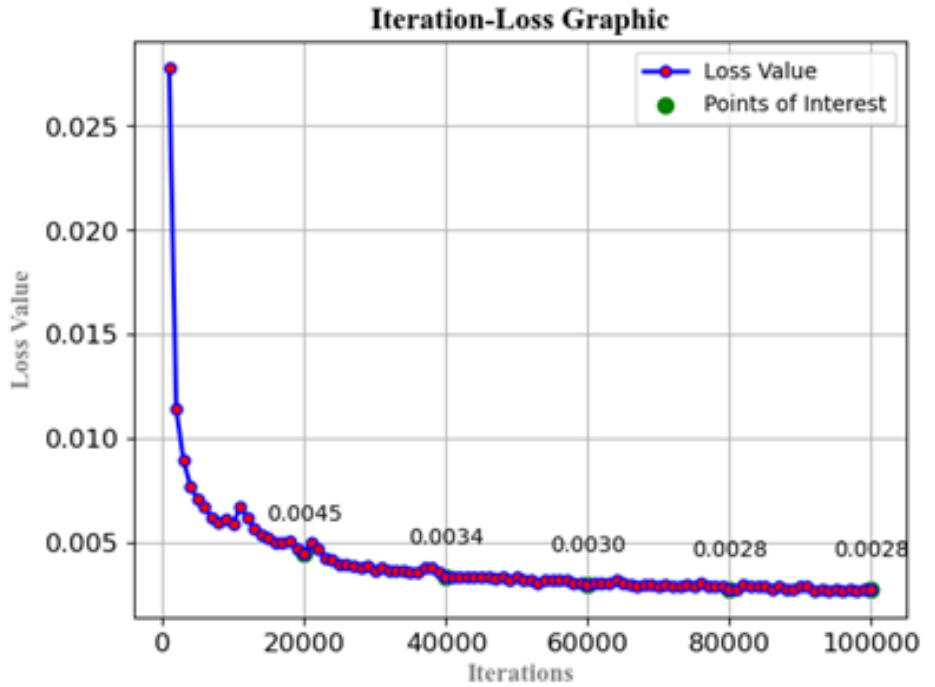


Figure 3.3 At the 100,000 iteration limit, the loss value up to the maximum iteration.

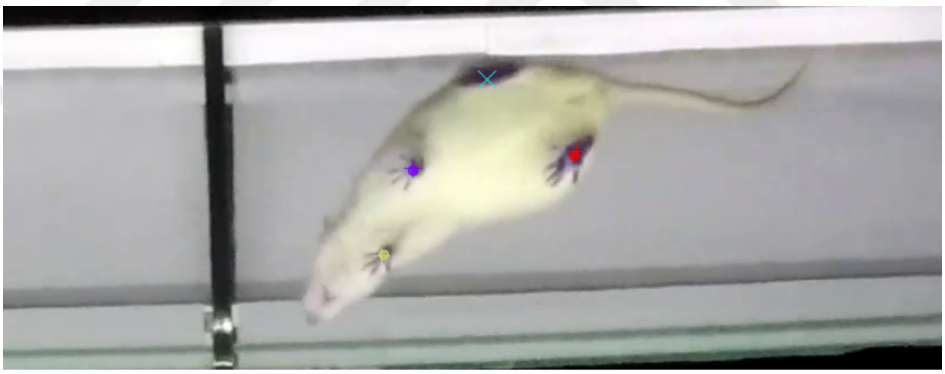


Figure 3.4 Point-Oriented labeling.

model and improve the paw tracking performance.

1. All extracted frames within the dataset must be labeled.
2. The size of the dataset should remain consistent in all videos.
3. Videos included in the dataset must share a uniform orientation.
4. The dataset must incorporate the lowermost frame of each video.
5. Frames included in the dataset should portray the animal with all four legs visible.

6. Video analysis must follow to a frame rate of 120 FPS.

Following these criteria, we separated left-right-oriented videos and increased the number of frames we extracted to avoid frame skipping in videos with high FPS values. We only used the bottom view perspectives of the videos so that the model can perform stable labeling. Following these rules, we trained our models and reported their performance. The models we trained in this process provided useful information for process optimization.

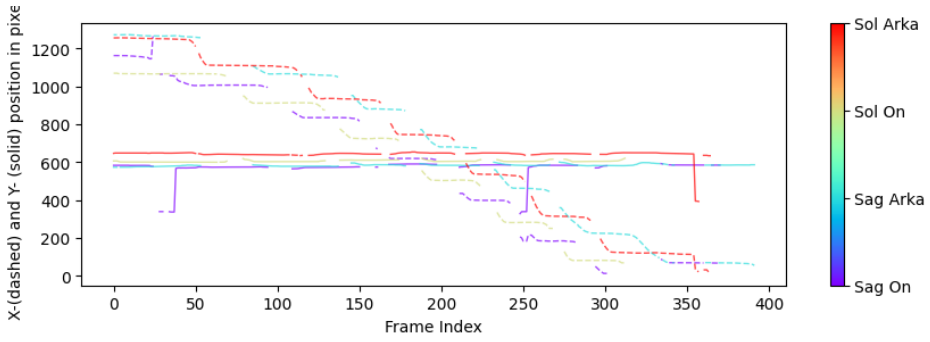


Figure 3.5 Figure showing flow of X and Y coordinates as the color-coded paws moving along the track. Since it moves on the horizontal axis, the Y coordinates remain constant while the X coordinates change with each stride. Stance, stride duration and stride distance can be observed. We used discrete color code for each paw but the DLC, in continous manner shows the color scale. As can be seen on the right legend, it is indicates which paw corresponds to which color code. In this study, Turkish equivalents were used for the labels.

We performed contact analysis to obtain contact data of the paw from a lateral perspective. However model performance was below our expectations. In moments when the paws were not in contact, it could label as if there was contact. To avoid this, we aimed to integrate the image in the lateral plane into our system. For this, we labeled the contact moments in the rat image in the lateral plane with the label respective paw. In addition, we analyzed the paw contact with image processing using low-pass, high-pass, and filters following different methods. However, we abandoned this analysis due to paws that could be hidden under the track or behind the paw.

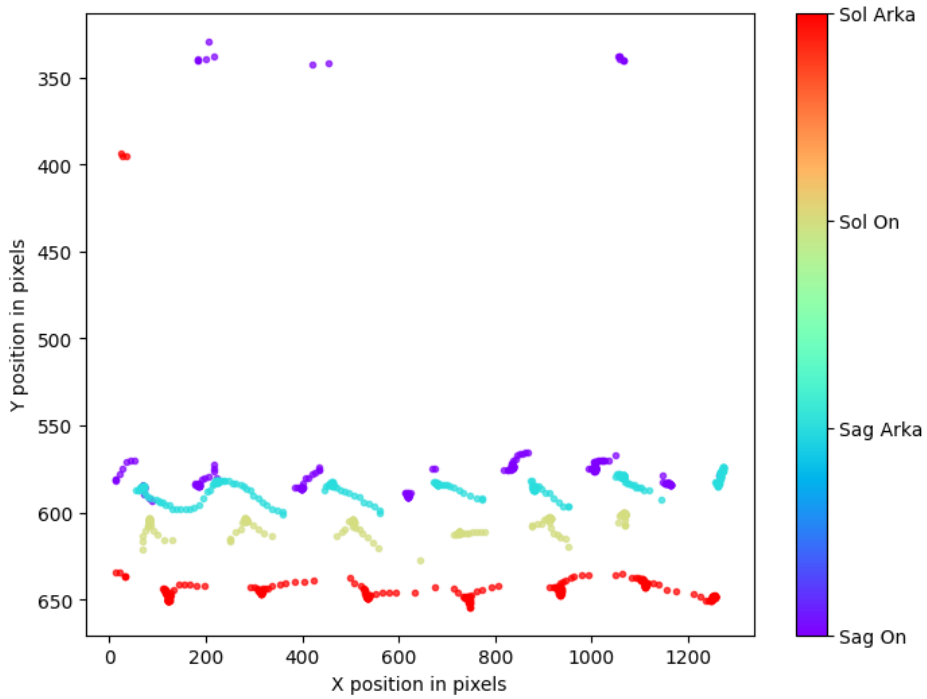


Figure 3.6 Trajectory plot showing the placement of time-independent paw labels in the coordinate system along the gait. In an ideal Trajectory Graph we would not expect to see following lines. This shows that the model follows the paw even in the absence of contact. Legend on the right is same as Figure 3.5.



Figure 3.7 A frame showing Laplacian of Gaussian (LoG) method being used.

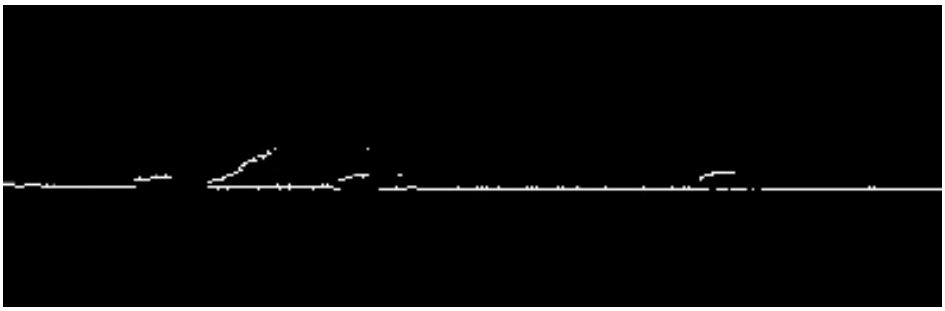


Figure 3.8 A frame showing Prewitt Operator (Prewitt) method being used.

3.2.2 Area-Oriented

We calculated the paw area of the mouse paw to include the paw area along the gait in our calculations and to use it as a feature. For this, we increased the number of

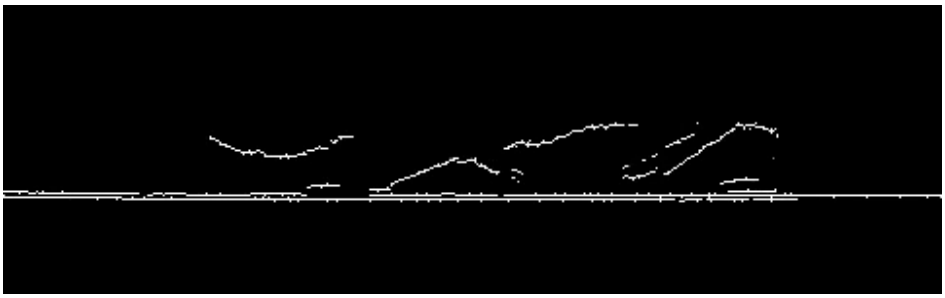


Figure 3.9 A frame showing Sobel Operator (Sobel) method being used. While we can detect edges along the rats gait, it does not provide any useful information about contact. Uneven ground and non-uniform lightning factors make it difficult.

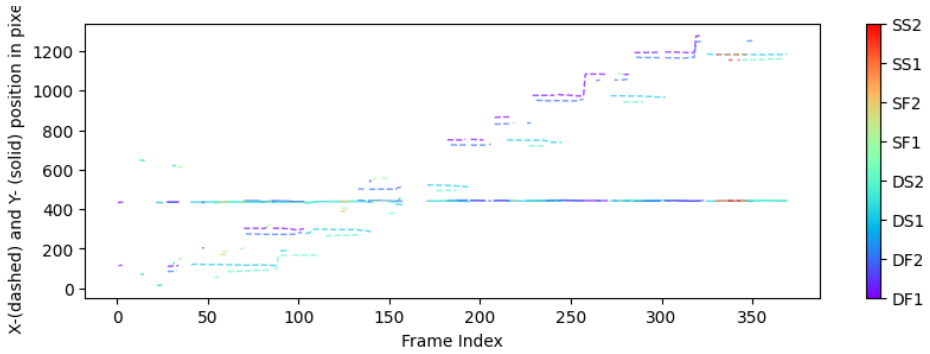


Figure 3.10 The scatter in the first frames shows the model's difficulty in following the paw. In an ideal figure we would expect to see no jumping lines. Mind the legend on the right, is changed. Instead of Turkish equivalents of paw labels, to ease the process of operators, Romanian equivalents, from laboratory interns, are used 2 for each paw; one for digits and one fore heel. SS: stanga spate (left hindlimb), SF: stanga fata (left forelimb), DS: dreapta spate (right hindlimb), DF: dreapta fata (right forelimb).

labels we used on each paw from 1 to 7. Accordingly, 4 labels were placed on the paw foot and 3 labels were placed on the paw digits. Since the DLC does not provide area information directly through its calculation, we integrated the points on the paws into the calculations externally with geometric area transformations. In order to make the area calculation more accurate and inclusive, we increased the number of labels from 7 to 9. We placed 2 new labels in the missing paw digits.

3.3 Preprocessing for Classification

For this study, we used version 2.3.5 of the DLC, which is the most recent version available at the time of this study. In this version, the tags in the processed videos were output as pixel coordinates in a CSV file. The records of the rats used in the

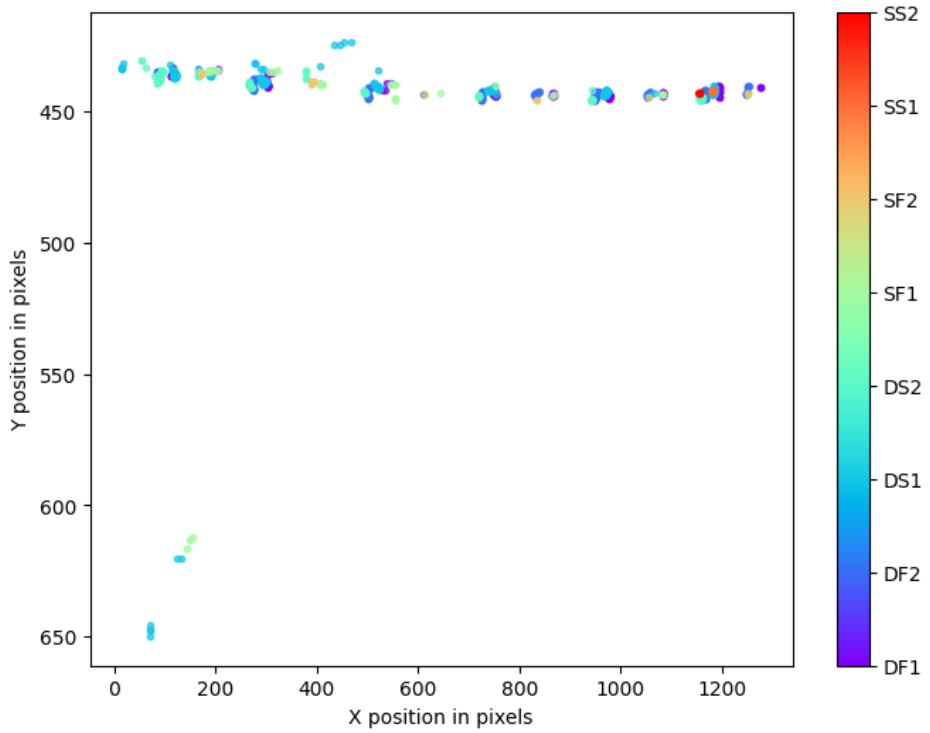


Figure 3.11 The jumps in the Figure 3.10 are clearly visible on the Y axis in here. Legend on the right is same as Figure 3.10.

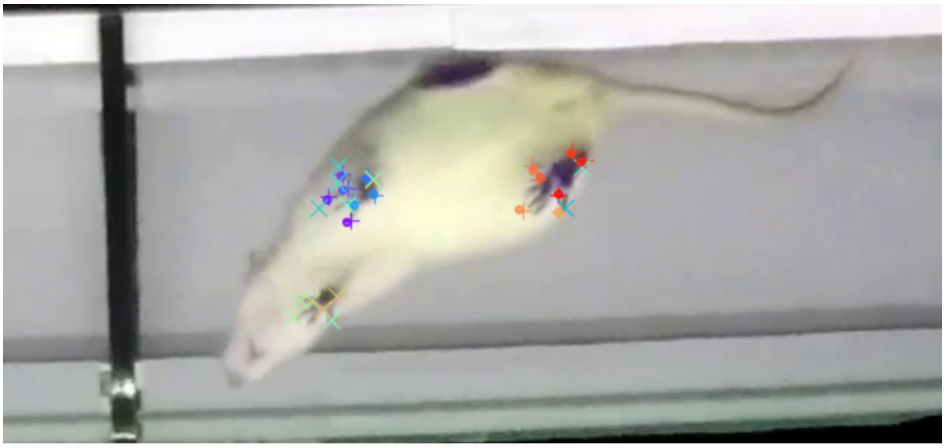


Figure 3.12 "+" signs are human labels. "X" and "O" signs are model labels. "O" for high likelihood, "X" for low likelihood markings. A frame from the 7-label model on a video. It can be seen that the right rear paw (at the top) was not in contact and the model did not recognize it. Both high likelihood and low likelihood markings can be seen on the human labels on the right front paw (top left). On the left front paw (bottom), it can be seen that the paw was recognized close to contact, but not yet with high enough likelihood to be marked.

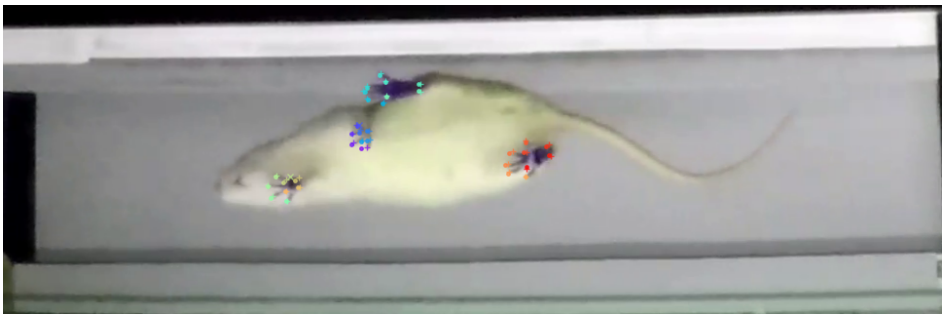


Figure 3.13 A frame from the 9-label model on a video. Almost perfect labeling with high likelihood values at the moment all paws are in contact. A label on the left front paw (far left) has a low likelihood for some reason. There may have been insufficient, incorrect labeling, or simply the model was not trained well enough.

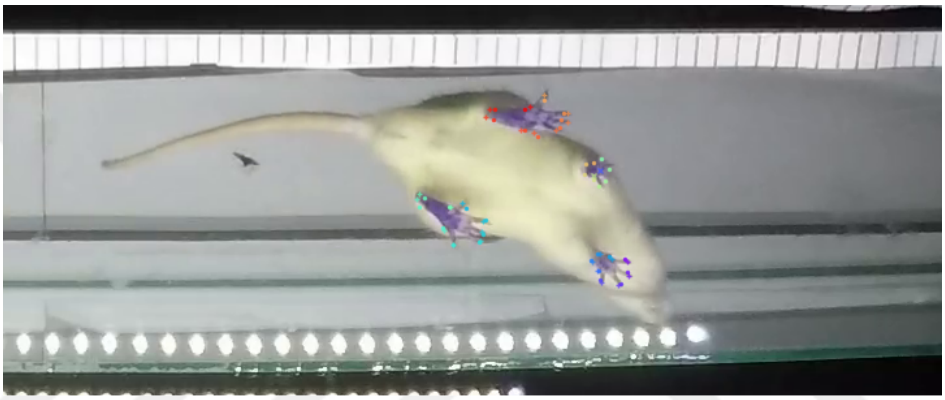


Figure 3.14 A frame from the 9-label model on a video. Almost perfect labeling with high likelihood values at the moment all paws are in contact. A label from the right front paw (top right) jumps from the left front paw (bottom) and an another label receives a low likelihood value. Figure 3.13 reasons apply here too.

experiment mentioned in the "3.1 Experiment and Data" section were analyzed with the DLC model in terms of their DPOs and CSV outputs were obtained. These CSV files contains coordinates of labels. Coordinates had to be converted into meaningful data. For this, MATLAB was used, the latest version of which was R2023a Update 4 (9.14.0.2306882) at the time of the study. The steps of reading the CSV files in MATLAB, processing the coordinates, and calculating them with appropriate gait parametric calculations were performed. Area, stance duration (stance_dur), stride distance (stride_dis), swing duration (swing_dur), limb duty factor (limb_duty) and speed parameters of a rat on each video were calculated in MATLAB. These parameters were formulated in MATLAB, taking into account the definitions in the publications [5],[11],[29]. According to this;

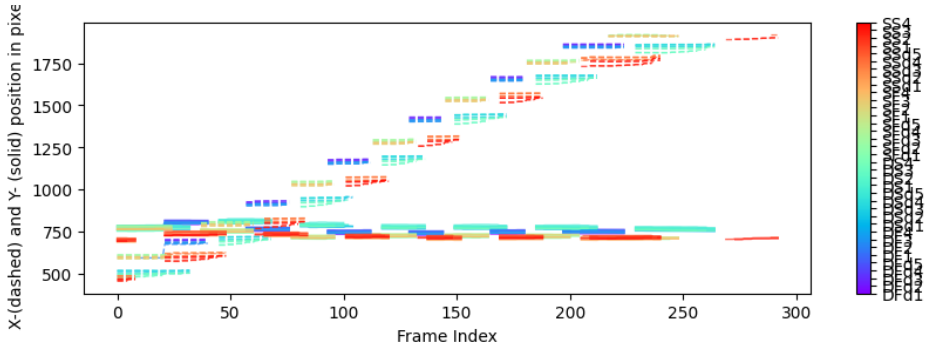


Figure 3.15 Figure showing a successful model. X and Y coordinates did not jump. The start and end of the Stride along the Gait are successfully recognized. Romanian equivalents are still in use as can be seen on legend at right. However as label size increased, for label count to each paw is also increased, legend grows. As digits and corners of paw started to label, order is given as Figure 3.2.

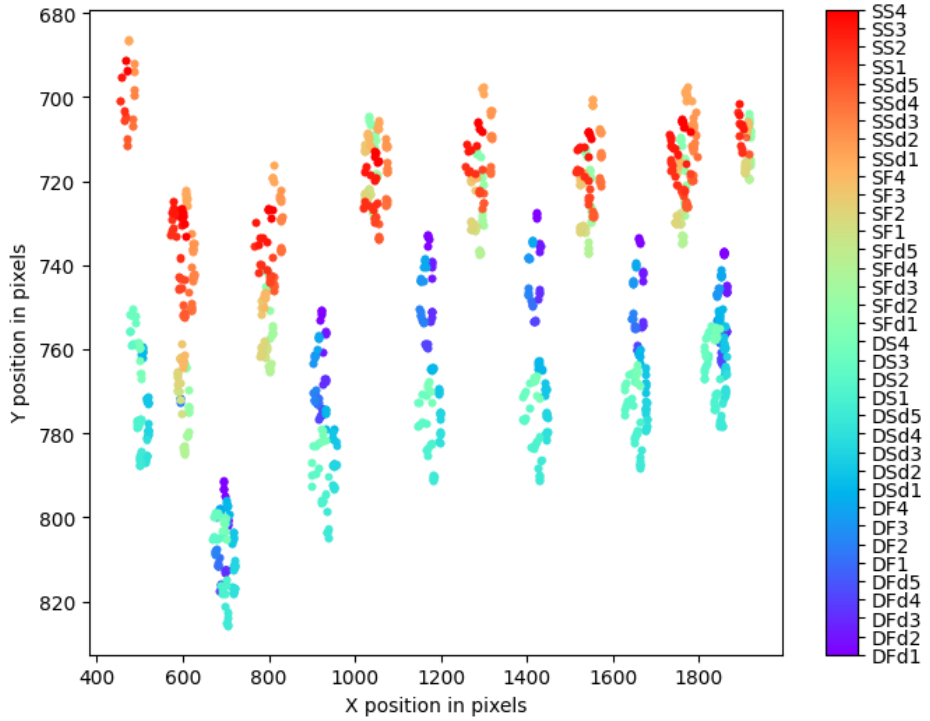


Figure 3.16 Paws and labels on paws are successfully separated from each other. Legend on the right is same as Figure 3.15.

- Area (area): Area, in square centimeters (cm^2), represents the surface area of a paw in contact with the ground during stance phase,
- Stance Duration (stance_dur): Stance Duration, in seconds (s), the duration from initial paw contact (heel-strike) to paw lift-off (toe-off) during a single stance phase,
- Stride Distance (stride_dis): Stride Distance, in centimeters (cm), the linear distance covered between two consecutive heel strikes of the same paw,
- Stride Duration (stride_dur): Stride Duration, in seconds (s), the total duration of a single stride cycle, from one heel strike to the next of the same paw,
- Swing Duration (swing_dur): Swing Duration, in seconds (s), the time during which a paw is in the air, from toe-off to the next heel strike,
- Limb Duty Factor (limb_duty): Limb Duty Factor, calculated as the stance duration divided by the stride duration of a paw, indicates gait symmetry and limb usage during locomotion,
- Speed (speed): Speed, in centimeters per second (cm/s), represents the rodent's forward velocity relative to the ground.

DLC provides a CSV file as analysis output. This CSV file contains the X and Y coordinates of each tag in a video, at each frame throughout the video, and "Likelihood" data, which indicates how confident the model is of contact. This CSV file needs to be processed and transformed from raw data to meaningful data. The CSV file we used to calculate the gait parameters contains the positions and confidence values of all the labels placed by the model throughout the video. This file needed to be processed to make sense of the data.

For this processing, I developed RatSuite, a fully customized computational framework designed to extract and process gait parameters from markerless motion tracking data in MATLAB. RatSuite was meticulously built from the ground up to address the unique challenges of SCI research, providing a cost-effective, comprehensive,

and open-source approach to gait parameter analysis. Unlike traditional tools, RatSuite is specifically tailored to work seamlessly with DLC-generated raw data, converting it into meaningful biomechanical metrics such as stride distance, swing duration, stance duration, and paw area. RatSuite is designed to handle data obtained from the bottom-view images of rats, enabling precise tracking and analysis using a robust labeling system with nine distinct labels per paw. This labeling approach captures detailed biomechanical information, ensuring accuracy and reliability across diverse datasets. RatSuite prepares the data in a structured and standardized format, making it ready for advanced machine learning models to classify SCI progression or treatment effects. By automating and streamlining the data extraction process, RatSuite empowers researchers to conduct high-quality analyses without the need for expensive proprietary tools. Its modular and open-source design allows for flexibility and adaptability, catering to various experimental setups and research objectives, ultimately bridging the gap between accessibility and scientific rigor in gait analysis.

3.4 RatSuite Code Group

3.4.1 Data Wizard

Converts CSV output to XLSX format. Discard tag records with a Likelihood value below 0.6. Divides the remaining tag records into X and Y coordinate values for each paw and saves them. In XLSX tables, each row corresponds to 1 frame.

3.4.2 Range Finder

Opens an environment to draw a line between marks on videos. Using the pixel size of this line, it gives a conversion ratio by proportioning the actual distance between consecutive lines. Saves this conversion ratio and stores it for later calculations.

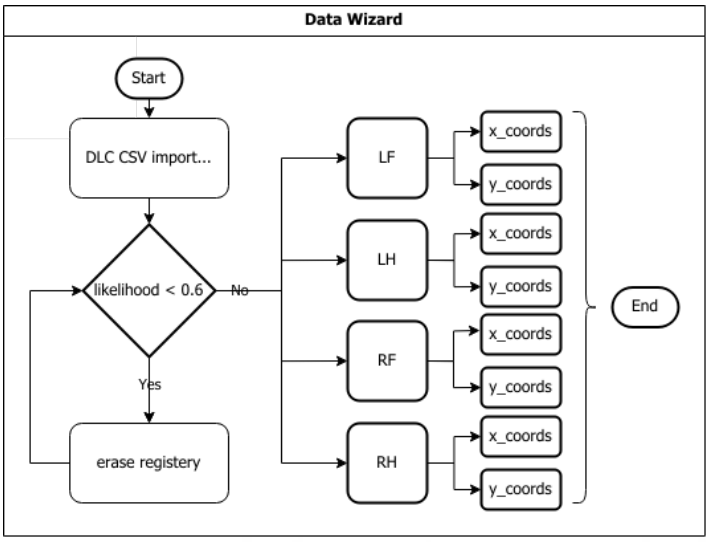


Figure 3.17 Flow chart of Data Wizard.

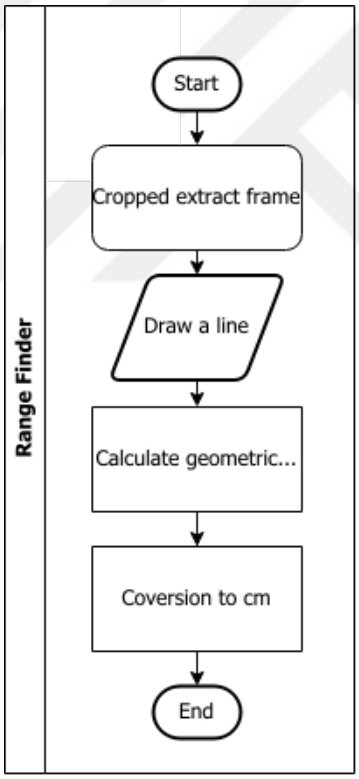


Figure 3.18 Flow chart of Range Finder.

3.4.3 Area Master

Records the contact area of the mouse during the stance duration during the video depending on the frame index. It performs the calculation using the X and Y coordinate files of the paws. The contact surfaces of each paw in the frames from heel-

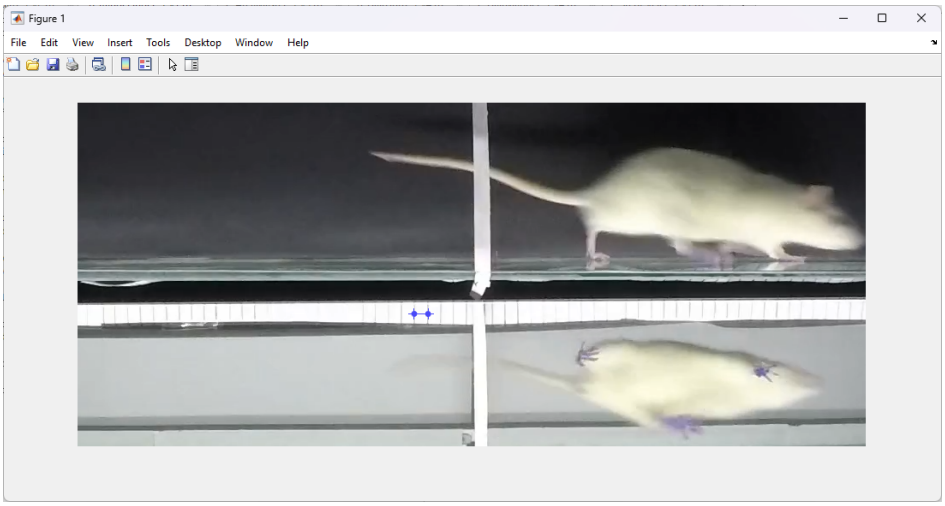


Figure 3.19 The user is instructed by the terminal to draw a line between the 2 markers. After the operation, the coversion ratio is saved in the workspace as a variable as it appears on the terminal.

strike to toe-off at each step throughout the video were calculated. These calculated areas were saved in an external file according to the corresponding frame sequence index.

3.4.4 Gait Guru

It uses all the files collected up to this stage. Calculates gait parameters for each paw. Gait parameter calculations are described in section "3.3 Preprocessing for Classification".

3.4.5 Gait Mapper

Keeps track of the frames where contact was made. Marks the frame areas where each paw made contact. Creates a gait map.

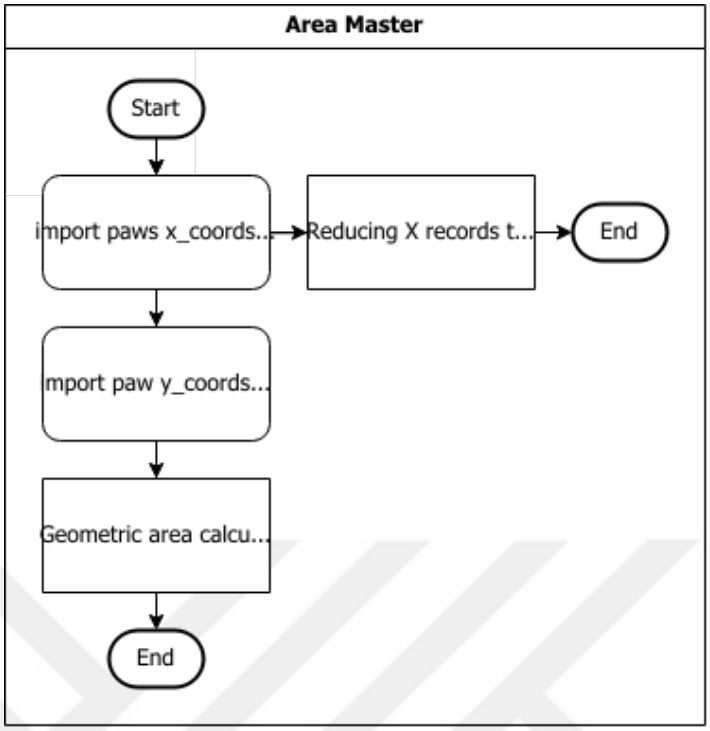


Figure 3.20 Flow chart of Area Master.

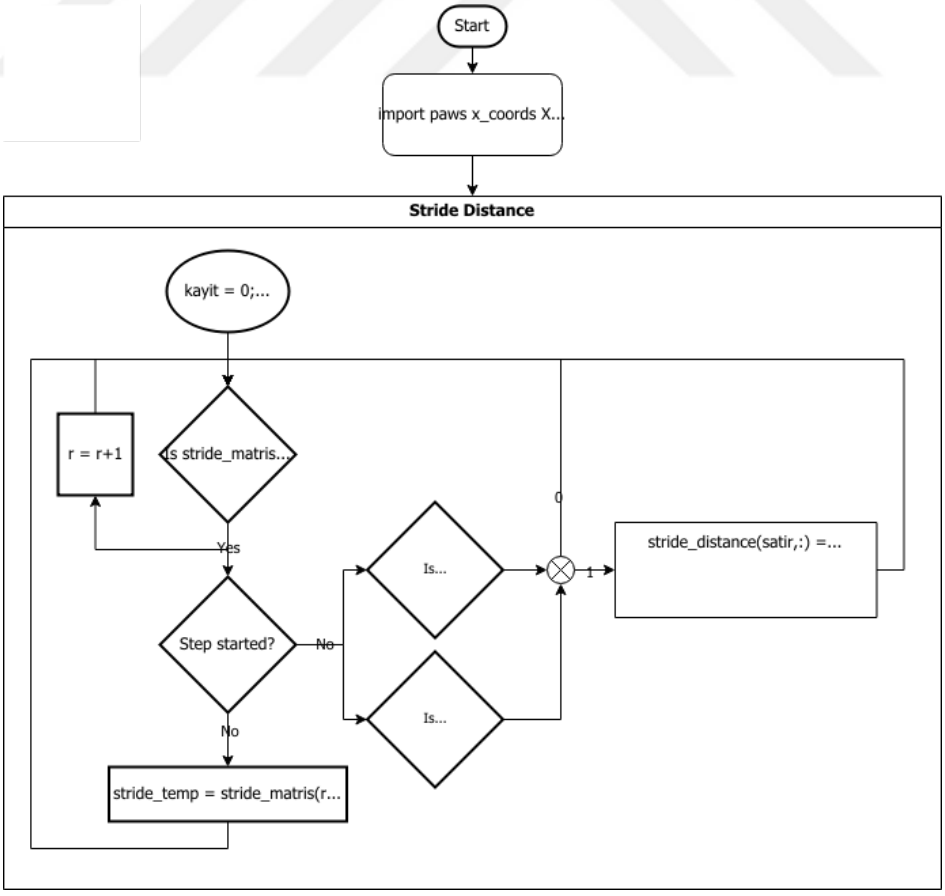


Figure 3.21 Flow chart of Stride Distance. At this stage Gait Guru code group started to work.

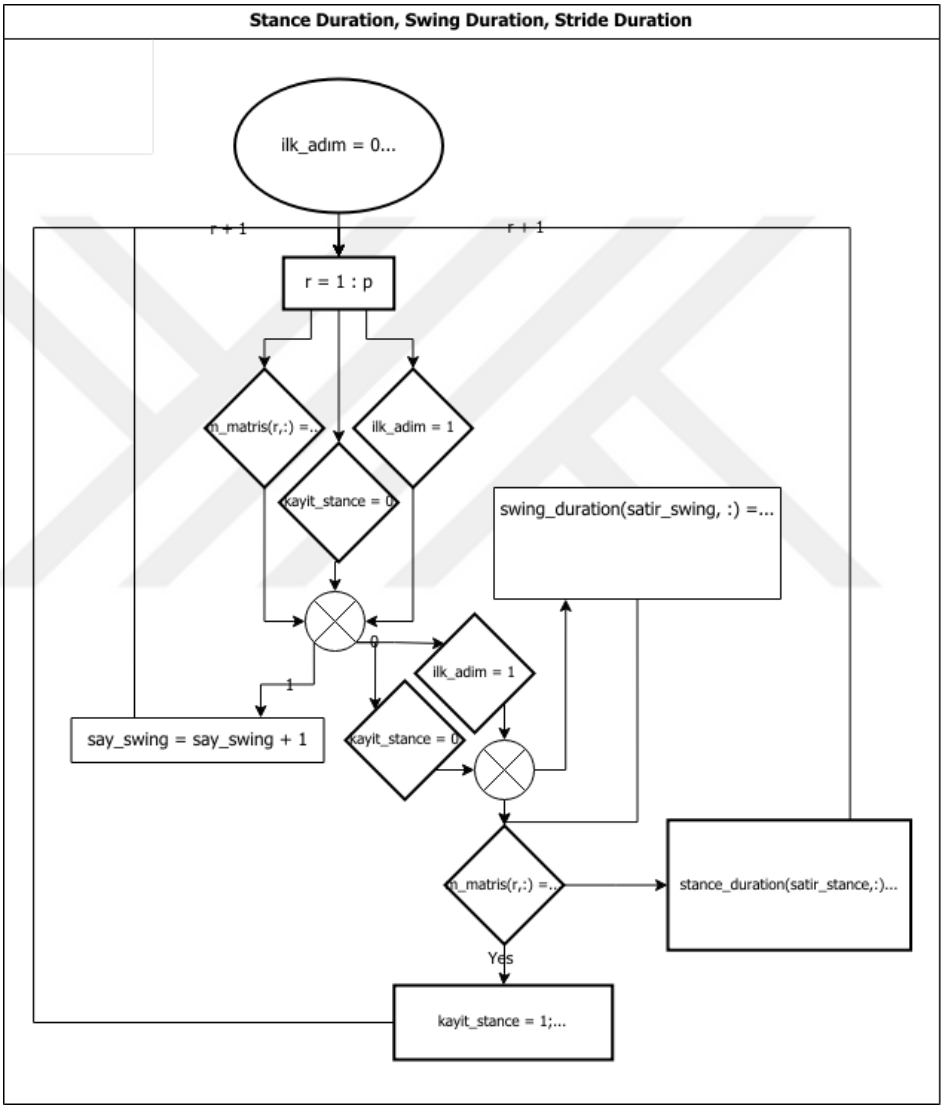


Figure 3.22 Flow chart of calculation the stance, swing and stride durations.

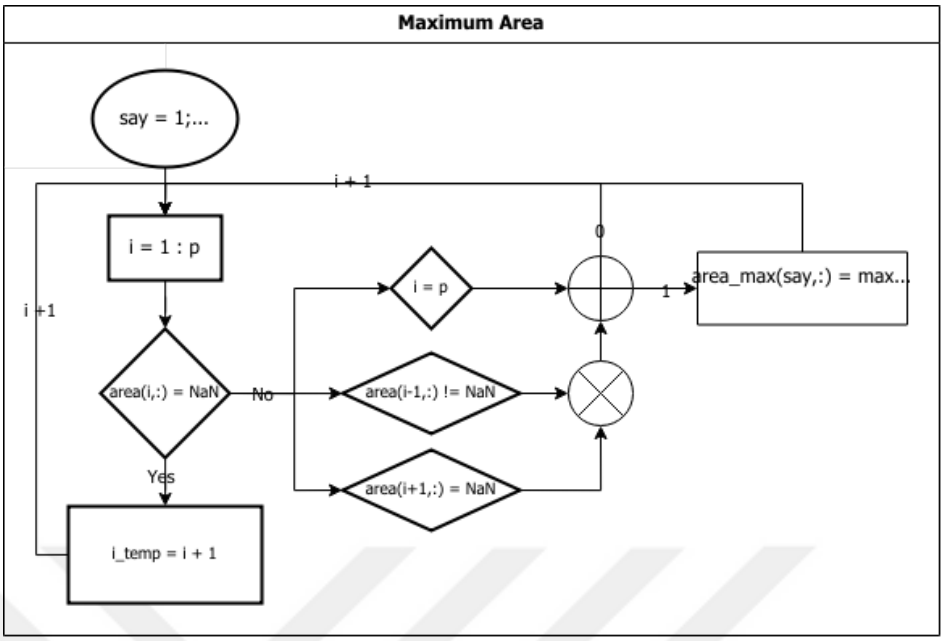


Figure 3.23 Flow chart of calculating the maximum area of each paw while rats gait.

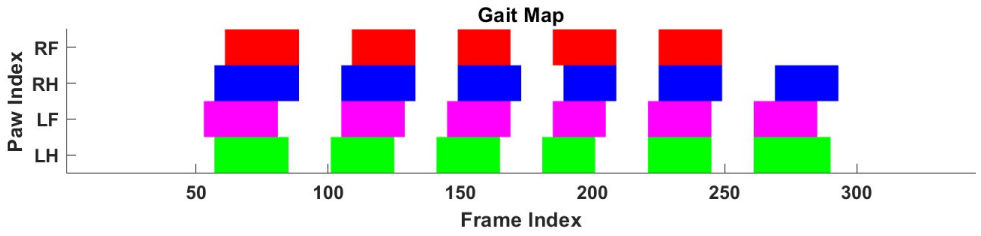


Figure 3.24 Gait map, showing in which frames the rat was in stance phase.

3.4.6 Step Cycler

Visualizes the mouse step sequence. We followed the sequence of steps for the Regularity Index, visualized this sequence and showed the sequence of steps. By examining this visual, the Regularity Index was extracted. Regularity Index was not used as a parameter in this study.

At the end of the whole computational process, the parameters were collected. It was compiled to form a dataset for the model for ML. This compilation was done as follows: rat rank index (rat), DPO time (dpo) at which the parameter was recorded, paw label (paw), mean paw area (area), mean stance duration (stance_dur), mean stride distance (stride_dis), mean stride duration (stride_dur), mean swing duration

(swing_dur), mean limb duty factor (limb_duty), mean speed (speed), BBB score at the time of DPO (bbb), vonFrey score at the time of DPO (vonFrey), injury class of the mouse (injury) and treatment class of the mouse (neuromod). This compilation formed the dataset for ML.

- "bbb": Basso, Beattie and Bresnahan, in BBB scoring (0-21), a standard neurological assessment tool for evaluating locomotor recovery post SCI, refers to EndPoint Classes,
- "vonFrey": Von Frey, in Von Frey scoring (0-100), tactile sensory threshold measurement, measures the force (in grams) required to elicit a withdrawal reflex in the rodent's paw, indicative of sensory sensitivity.

3.5 Machine Learning for Classification

The Classification Learner application in MATLAB was used for ML. To obtain this application, version 23.2 Statistics and Machine Learning Toolbox was obtained at the time of the study. The data sets created with the collected parameters, was divided into 2 different formats. Accordingly, 2 data sets were created; the common parameters of the rat were organized as one feature (Normal) and the common parameters of the rat were organized as one feature per DPO (Extended). In order to monitor the performance of the model in different paw configurations, the data sets were divided into 2 according to the paw configurations as well. In total, each data set was divided again according to the condition of the mice according to the experimental plan.

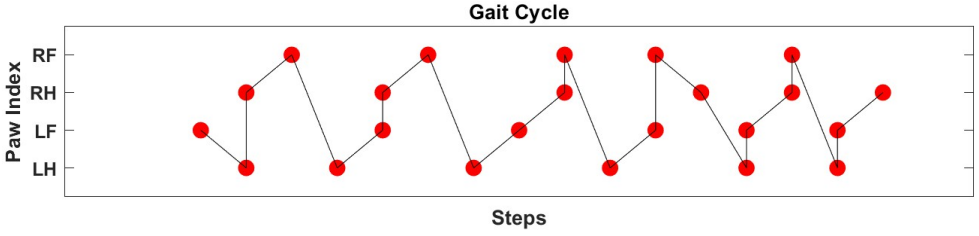


Figure 3.25 Gait cycle, showing the order in which the rat takes steps.

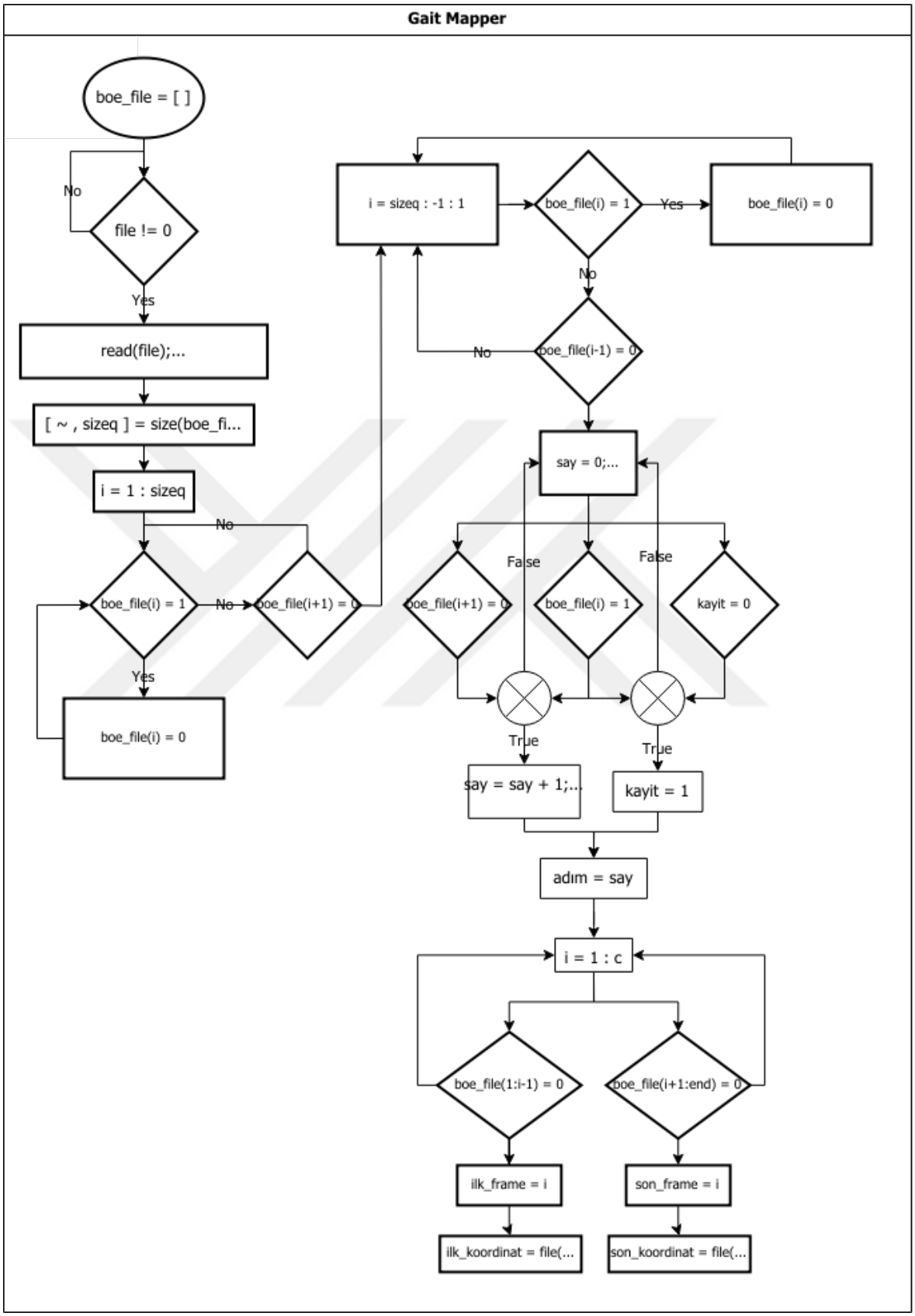


Figure 3.26 Flow chart of Gait Mapper and Gait Cycler.

Table 3.2
Number of rows of the data sets used for model training. Each row corresponds to one record.

Normal							Extended						
with Pre-Op	F+H			H			with Pre-Op	F+H			H		
	Sham+SCI	Sham	SCI	Sham+SCI	Sham	SCI		Sham+SCI	Sham	SCI	Sham+SCI	Sham	SCI
867	613	207	406	211	103	108	253	194	63	131	67	31	36

Pre-Op records were dropped from the model training data sets. Each data set was used for model training.

The records in the dataset show a heterogeneous distribution in terms of SCI severity, with mild, moderate and severe groups all present. ML for classification was performed using the mentioned toolbox. All model-related settings and applications were customized and executed using the GUI of the toolbox.

3.5.1 Features and Response

In the datasets, the row index of the mouse, the time of the DPO in which the parameters were recorded and the paw of the mouse were categorical classes; the remaining gait parameters, locomotor performance parameters ("BBB" and "vonFrey" scores) and injury classes (in binary format) and treatment classes (in binary format) were designated as numerical classes. As mentioned at 3.1 and plotted at 4.3 "dpo": Day-post-Operation (DPO), 0: Pre-Op (before operation), 1: D1 (1 day after), 2: D7 (7 day after), 3: D14 (14 day after), 4: M1 (30 day after), 5: M1,5 (45 day after), 6: M2 (60 day after), color coded from bright to dark, corresponding to Pre-Op to M2, refers to End-point Classes; "neuromod": Neuromodulation, in binary, 1: neuromodulated, 0: non-neuromodulated, neuromodulation refers to transcutaneous electrical nerve stimulation, refers to Treatment Classes; "injury": Injury, in binary, 1: SCI model, 0: sham, injury refers to SCI, refers to Injury Classes. The independent variables "dpo" and "neuromod" were set as response and the remaining variables were set as feature. "rat", "paw", "dpo", "injury" and "neuromod" variables that were non-

rat	dpi	paw	area	stance_dur	stride_dis	stride_dur	swing_dur	limb_duty	speed	bbb	vonFrey	injury	neuromod
RAT79	Pre-Op	LF	146,73850	0,25805	13,7347	0,44682	0,16323	61,27860	28,53620	21	60	1	1
RAT79	Pre-Op	LH	344,75880	0,26491	13,4566	0,41154	0,13161	63,18920	28,61810	21	60	1	1
RAT79	Pre-Op	RF	155,13290	0,23448	13,9793	0,39771	0,16323	61,99050	29,12770	21	60	1	1
RAT79	Pre-Op	RH	312,46490	0,26906	13,5931	0,42028	0,13280	64,25660	29,07310	21	60	1	1
RAT79	D7	LF	141,56190	0,35257	9,3497	0,51916	0,10393	76,92360	15,93230	7	180	1	1
RAT79	D7	RF	133,76640	0,43105	9,3381	0,57989	0,13114	75,98170	15,04970	7	180	1	1
RAT79	D14	LF	130,37000	0,25137	10,1808	0,37239	0,13446	64,46780	25,19860	14	100	1	1
RAT79	D14	LH	301,56940	0,26432	11,0558	0,36935	0,12450	70,86280	25,71460	14	100	1	1
RAT79	D14	RF	156,08250	0,25219	11,3602	0,36839	0,12386	68,12190	25,99880	14	100	1	1
RAT79	D14	RH	291,12520	0,28042	10,8733	0,39425	0,12003	69,97540	26,49860	14	100	1	1
RAT79	D1	LF	146,58990	0,63379	5,8432	0,65639	0,08300	87,61960	7,50430	1	180	1	1
RAT79	D1	RF	146,57950	0,72072	7,2050	0,90998	0,12616	86,30130	7,78230	0	180	1	1
RAT79	M15	LF	134,5606	0,3635	10,2933	0,5984	0,2015	65,1117	13,2471	19	100	1	1
RAT79	M15	LH	269,8564	0,3977	9,8664	0,5713	0,1552	69,7607	15,0424	19	100	1	1
RAT79	M15	RF	146,1673	0,3793	10,0286	0,5872	0,1916	68,7895	14,0402	19	180	1	1
RAT79	M15	RH	244,0879	0,3588	9,3593	0,5652	0,1600	66,4246	15,2012	19	180	1	1
RAT79	M1	LF	176,0020	0,2701	10,0485	0,3886	0,1185	70,1768	25,9130	14	100	1	1
RAT79	M1	LH	290,1929	0,2781	9,7835	0,4009	0,1228	69,0489	22,5160	14	100	1	1
RAT79	M1	RF	181,9953	0,2714	9,7903	0,3943	0,1124	67,9991	23,9367	14	180	1	1
RAT78	Pre-Op	LF	146,27340	0,31333	11,5098	0,46933	0,17128	62,88310	19,99730	21	60	1	0
RAT78	Pre-Op	LH	341,14510	0,37004	11,1550	0,45139	0,10918	77,00800	19,59000	21	60	1	0
RAT78	Pre-Op	RF	147,40850	0,24825	12,2027	0,45152	0,17505	59,70830	22,83910	21	60	1	0
RAT78	Pre-Op	RH	325,33910	0,35856	11,7402	0,49468	0,14110	72,58970	19,94620	21	60	1	0
RAT78	D7	LF	150,52130	0,33136	7,6081	0,47159	0,08466	80,96030	11,44810	3	60	1	0
RAT78	D7	RF	165,49890	0,35275	7,4934	0,38789	0,06640	88,32340	10,93520	0	100	1	0
RAT78	D14	LF	147,33700	0,24663	7,9689	0,34785	0,08669	73,08580	21,48930	1	100	1	0

Figure 3.27 Figure showing gait parameters, BBB, and vF scores extracted from the videos of rats 79 and 78 at the designated DPO dates in Normal Format. In the normal format, each column represents a feature, while DPOs and their corresponding data are listed as rows. This presentation highlights the straightforward alignment of data features across multiple time points.

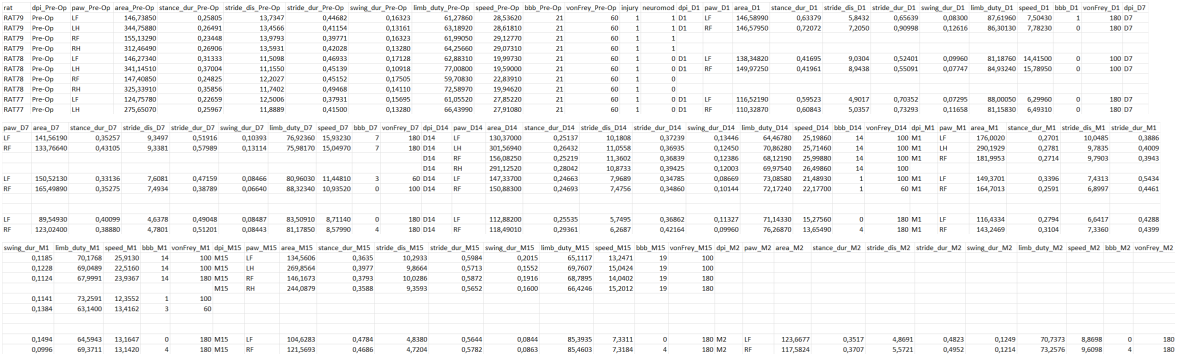


Figure 3.28 Figure showing gait parameters, BBB, and vF scores extracted from the videos of rats 79, 78, and 77 at the designated DPO dates in Extended Format. In the extended format, features at each DPO are represented as columns, which allows for proportional scaling of the dataset. Despite including more rats, the dataset is visually expanded. Injury and treatment classifications are presented in their respective columns, and the figure is displayed in three consecutive parts for clarity.

dependent, during the training of the model were dropped from the dataset. Records with Pre-Op in the "dpo" variable were dropped. Only dependent variables trained the model.

3.5.2 Data Splitting

Different training percentages were examined for data split. It was determined as %85-%15, %85 for training, %15 for testing.

3.5.3 Validation Scheme

For our small datasets that are out of the norm, we chose the Cross-Validation method [39]. For the selected validation scheme, we evaluated the validation performance of the model by experimenting with different k-fold values. We settled on 25 k-fold.

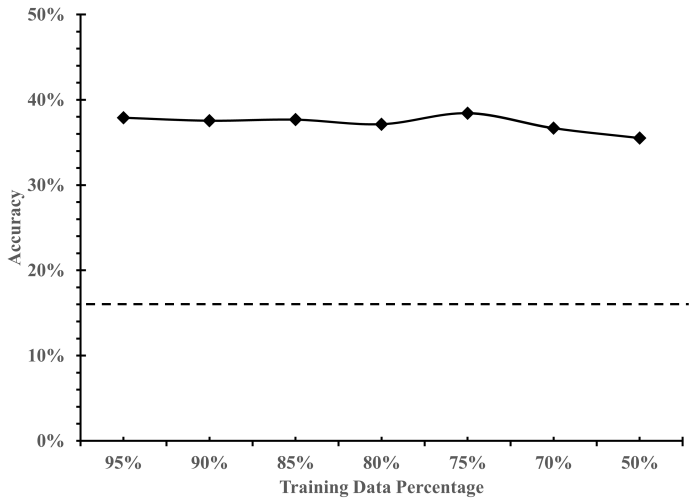


Figure 3.29 The validation accuracy of the model was evaluated under varying proportions of training data during the model-building phase. As the proportion of training data decreased, validation accuracy tended to decline. This decline can be attributed to the reduced volume of data available for the model to learn patterns effectively, as a larger percentage of the dataset is allocated to testing. These findings highlight the critical role of sufficient training data in optimizing model performance and ensuring robust generalization.

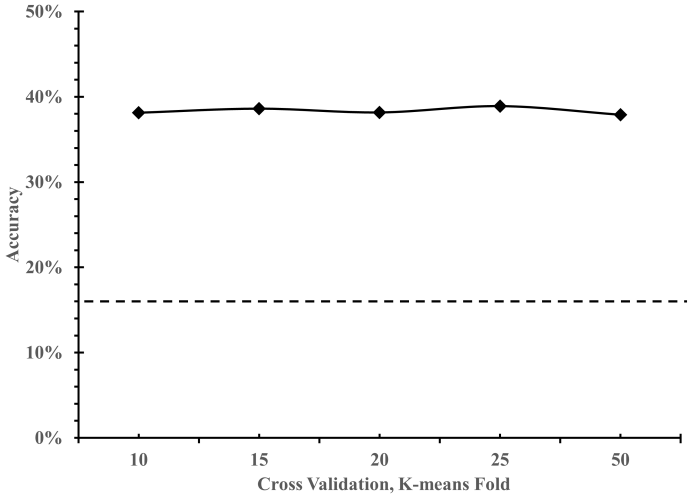


Figure 3.30 The relationship between model accuracy and the number of folds in k-means cross-validation was analyzed during the model-building phase. The general trend indicates that accuracy improves as the number of folds in k-means cross-validation increases. This is because increasing the number of folds reduces bias by allowing the model to train on a greater variety of data subsets. However, this also enhances generalization by providing a more robust evaluation of the model's performance across diverse data splits. These observations underscore the trade-off between model bias and generalization as the k-means cross-validation configuration is adjusted.

3.5.4 Feature Ranking

We used the the Toolbox for feature ranking. We examined the feature ranks using different methods in the Feature tool in this toolbox. Among these methods, we preferred ANOVA for feature ranking. Accordingly, we got the feature scores used in the data sets with the following ratios. Since we observed that feature ranking generally decreases accuracy, we did not use feature ranking in the final analysis.

3.5.5 Cost Matrix

For the cost matrix we used the relevant section in the toolbox interface. We tried 2 different cost matrixes.

But in the final analysis we concluded to not use it because we did not see a significant increase in prediction accuracy and there would be a confliction in rationale. Creating a cost matrix based solely on prediction outputs does not fit the biological nature of the study. For this reason, we looked for interconnecting relationships in the

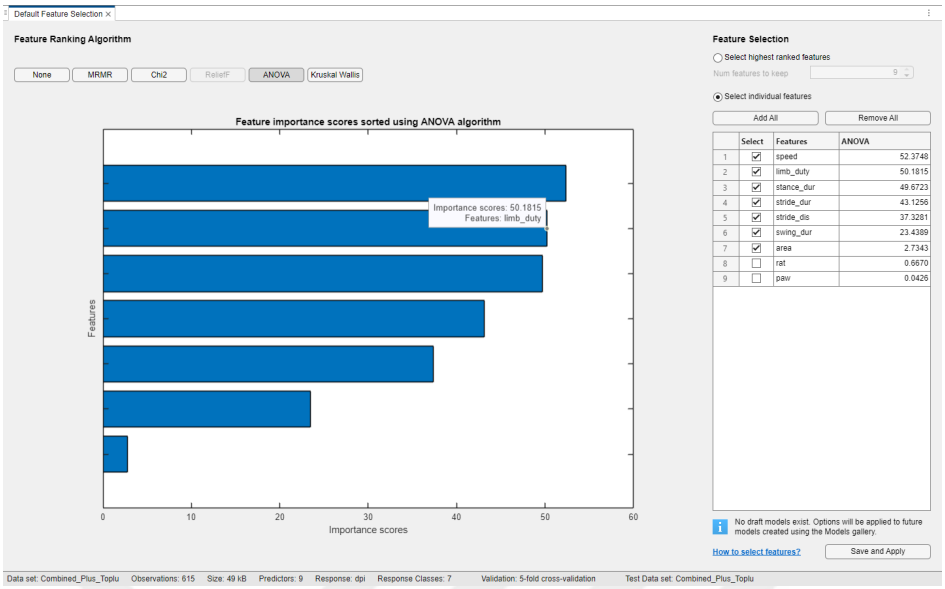


Figure 3.31 A screenshot of the window used for Feature Ranking. The results obtained with the ANOVA method can be seen.

Table 3.3
Default cost matrix. Tested to reinforce True Positive and False Negative predictions.

0	1	1	1	1	1	1
1	0	1	1	1	1	1
1	1	0	1	1	1	1
1	1	1	0	1	1	1
1	1	1	1	0	1	1
1	1	1	1	1	0	1
1	1	1	1	1	1	0

Table 3.4
Customized cost matrix. The confusion matrix output of the model was analyzed and weights were distributed in effort to correct False Positive predictions.

0	1.5	1	1	1	1	5.5
2.4	0	2	1	1	1	5
2.8	2	0	1	1	1	6.1
1.9	1.3	1.3	0	1	1	4.8
2.3	1.3	1	1	0	1	2.4
1.2	1	1	1	1	0	2.3
1	1	1	1	1	1	0

groups. We tried to distribute the cost matrix by dividing it into clusters. Since there is no detrending and the data sets are heterogeneously distributed, we could not create a raitonale since Clustering gives different results for each parameter.

3.5.6 Machine Learning Model

We employed Ensemble Learning, specifically the Bagging (Bootstrap Aggregating) method, for classification. Bagging enhances model stability and accuracy by reducing variance and overfitting, which is particularly beneficial for complex datasets like ours [40],[41],[42]. Our dataset comprises multiple gait parameters, each contributing uniquely to the classification of SCI levels. Bagging is well-suited for such multivariate data as it trains multiple models on different subsets of the data, thereby capturing diverse aspects of the feature space and improving overall predictive performance [43]. Additionally, the high-dimensional nature of our dataset poses challenges for single-model approaches, which may struggle to capture intricate relationships between features. Bagging addresses this by generating multiple subsets through bootstrapping, allowing each base model to focus on different feature interactions, leading to a more comprehensive understanding of the data [44]. Given the relatively small size of our dataset, there's an increased risk of overfitting. Bagging mitigates this risk by averaging predictions across multiple models, enhancing generalization to unseen data. This characteristic is crucial in biomedical studies, where data collection can be limited due to practical constraints [44]. Therefor, the Bagged ensemble model was selected for its ability to handle multivariate, high-dimensional datasets, its robustness against overfitting in small sample sizes, and its superior predictive accuracy, making it an appropriate choice for our classification tasks.

3.6 Performance Metrics

We used the performance metrics that Toolbox offers us: Validation and Test scores, ROC Curve and Confusion Matrix. Confusion Matrix also gives the rates to

calculate Precision, Recall (Sensitivity) and Specificity. Used equations are specified below:

$$Precision = \frac{TP}{TP + FP} \tag{3.1}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3.3}$$

These computationally derived performance metrics need to be linked to a rationale due to the biological nature of the study. These rationales are expressed as follows:

- Precision, when the model predicts that the rat is injured, how accurate that prediction is.
- Recall, the ability to detect real injured rats.
- Specificity, the ability to detect rats that are not truly injured.

4. RESULTS

4.1 Optimization of Data Collection

4.1.1 LtoR, RtoL Movement

We did not expect the orientation of the video, Left to Right (LtoR) or Right to Left (RtoL), to have a direct impact on the model's performance. Indeed, in the Figure 4.2, it can be seen that orientation has a slight positive effect. We believe that this is due to the operator factor in practical application rather than theoretical analysis.

4.1.2 Increasing Video Quality

Video quality has an indirect impact on the model's performance. However, at high video quality, at high resolution, paws are better distinguished from the background and digits are better distinguished from each other in close proximity, which is both useful for labeling frames and easier to distinguish for model during analysis.

4.1.3 Bottom View Perspective

Bottom view perspective was deliberately chosen. When lateral plane perspective or both lateral plane and bottom view perspectives used to train model, it negatively affected the labeling performance of the model. Especially as mentioned in the 3.2.1 subsection, the paws were not easily distinguishable due to uneven ground and non-uniform lightning conditions. In cases where perspectives are used together, since we are using a ML model that lacks reasoning, it cannot distinguish which perspective the marking area belongs to. This can lead to marking jumps or marking directly in the wrong area when paw images would invoke similar responses. For these reasons, focusing only on the bottom view perspective naturally improved the model's performance.

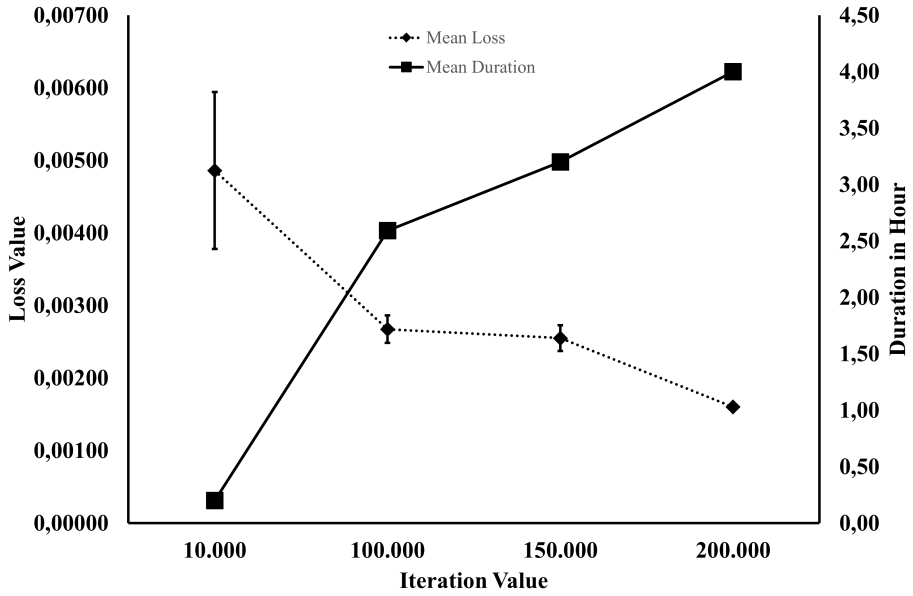


Figure 4.1 3 axis, showing Loss-Iteration-Training Duration correlation.

Indeed, bottom view perspective of the paw was not sufficient to supply enough information on paws contact state. For this reason, in addition to the normal half frame extraction where labeling takes place, a full frame extraction was also required, where both perspectives were visible in the whole video. While labeling the frames to be fed to the model, they were compared with full frames and in case of a frame match, the contacting paws were confirmed from the lateral plane and labeled then. It is critical for gait analysis that the model gives the coordinates of the correct points on the paws at the moment of contact. For this reason, we focused on the error values in pixels. Especially in the final model, the training error, determined with a P-cutoff, stood at 3.2 pixels, while the testing error was measured at 2.97 pixels. With our video setup, where 1 cm corresponds to 17 pixels, a 3-pixel error translates to 1.7 millimeters. Our loss value was 0.0028 after 100,000 iterations. In the Figure 4.2 on the error value in pixels, especially the negative correlation effect of the number of frames can be seen. A similar effect can be seen in the number of iterations, with higher iterations yielding lower error values. Although we attribute the high accuracy rate to the bottom view perspective, this perspective seems to have a relatively small impact on the error value of the model. In this context, the high accuracy rate can be attributed to the ease-of-use of the labeling operator.

We anticipated a negative impact of using a high number of tags on the paw. In particular, we expected that the increase in variance due to the increased number of labels would increase the false positive rate due to the narrow area of labeling. However, partially contrary to our expectation, the increased number of labels positively affected the performance of the model and resulted in low error values. Since the labels are point oriented and evaluate the limited frame in which they are placed, we interpret that the single label used in the point oriented approach performs poorly in contrast to the high number of labels due to insufficient discriminative features on the paw surface. With the same approach, we believe that a high number of labels leads to increased performance and lower errors, both because they are placed at the edges of the paw

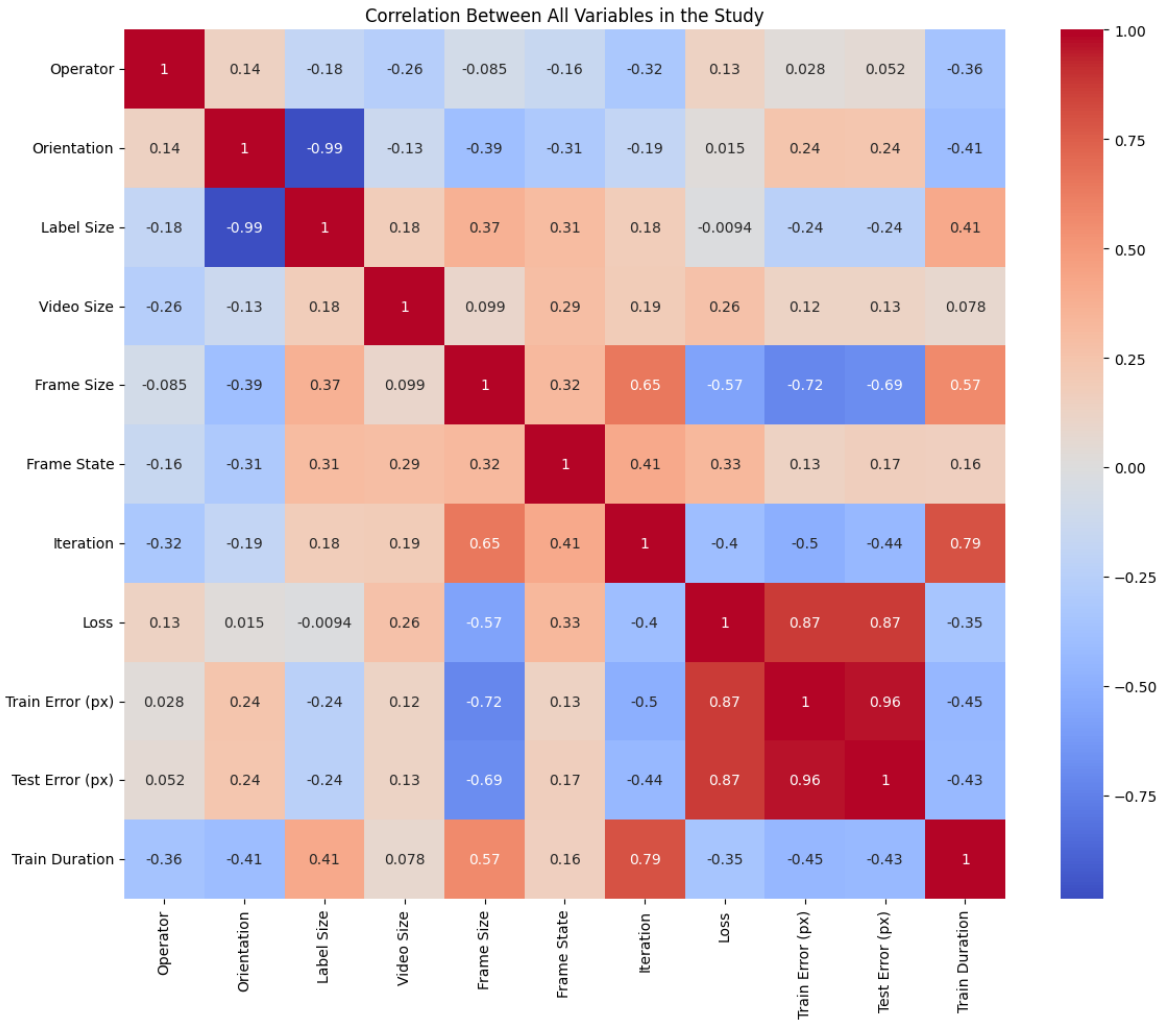


Figure 4.2 Correlation heatmap of parameters in training the paw tracking model.

and because there are more comparisons in a narrow space compared to a single label.

4.2 Optimization of Preprocessing

Data analysis was performed to feed the DLC outputs into the classification model. The gait parameter formulas mentioned in the 3.3 section were written in MATLAB. The code suite developed in MATLAB for the extraction of gait parameters accelerated the process.

4.3 Classification Data

Accuracy scores for all models decreased as the dataset coverage decreased. The highest accuracy scores were obtained in datasets with fore and hindlimb together and with both injury classes, sham and SCI model rats together. DPO and treatment prediction gave different accuracy results in the datasets. A direct reason for this is that treatment consisted of 2 classes while DPO consisted of 6 classes. This directly affected the percentage distribution of the dataset into classes.

In the area-stance duration plot, uninjured subjects exhibit a lower distribution of stance duration, while injured subjects display higher stance duration values, with the widest distribution particularly observed in 1 day post operation (DPO: D1 (1)) injured subjects. Stance duration decreases as the DPO (Days Post-Operation) progresses. Additionally, injured subjects demonstrate lower paw areas on D1, with paw area increasing as DPO advances. In the stance duration-stride distance plot, uninjured subjects are concentrated in the 11-15 cm range for stride distance, whereas injured subjects are more broadly distributed. For stance duration, uninjured subjects cluster around 0.25 seconds, while injured subjects show a wider distribution. Among injured subjects, those on D1 exhibit shorter stride distances and a broader stance duration range, while both parameters tend to converge toward the uninjured subjects as DPO

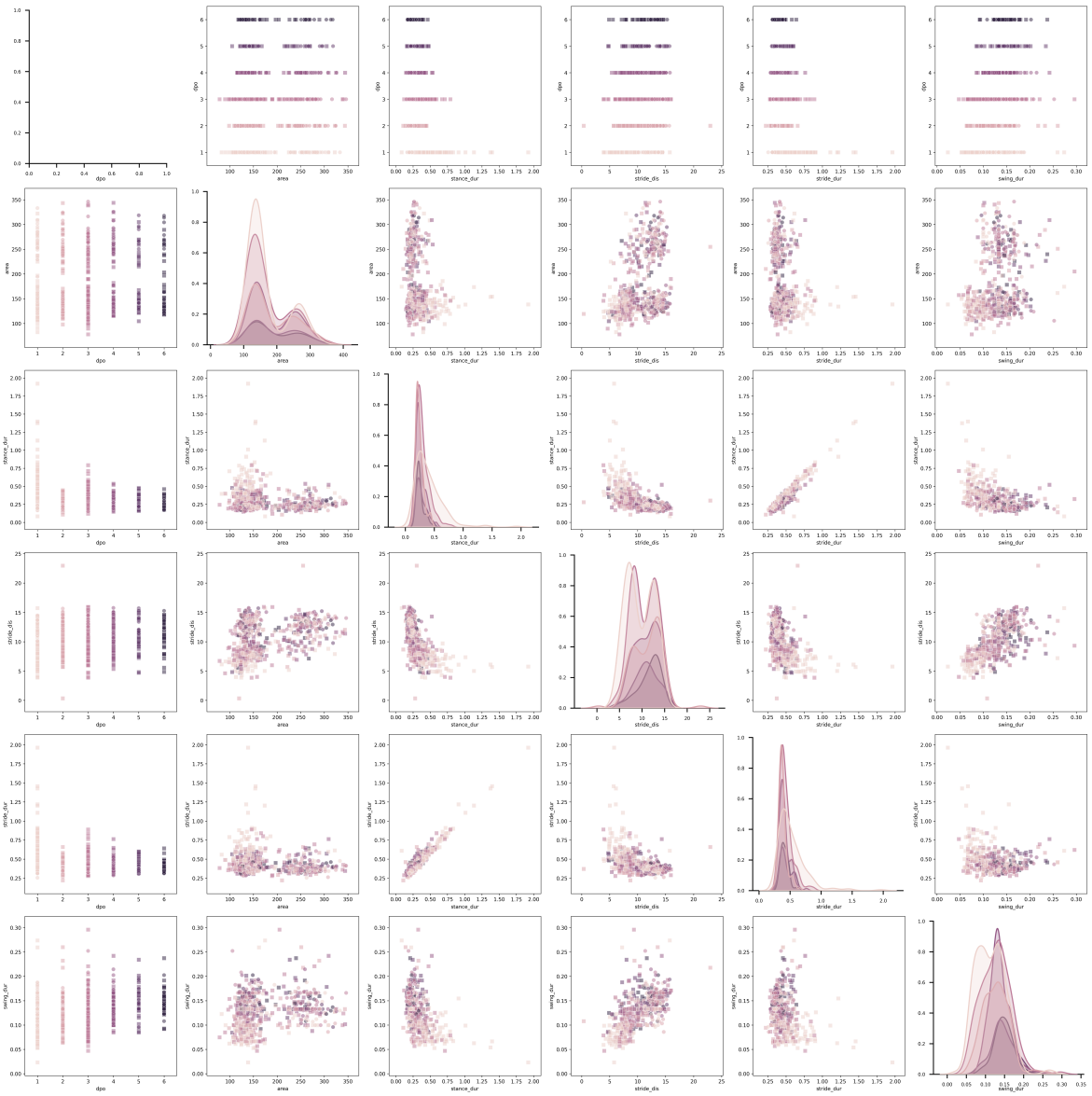


Figure 4.3 Distributions of all variables in the dataset containing Sham and SCI rats, colored according to End-point Classes, in the study. Both of the classes contains both neuromodulated and non-neuromodulated rats. Styled according to Injury Classes. Circles (•): sham, squares (■): SCI models. "dpo": Day post Operation, "area": Area, "stance_dur": Stance Duration, "stride_dis": Stride Distance, "stride_dur": Stride Duration, "swing_dur": Swing Duration, "limb_duty": Limb Duty Factor, "speed": Speed, "bbb": Basso, Beattie and Bresnahan scoring, "vonFrey": Von Frey, "neuromod": Neuromodulation, "injury": Injury.

progresses. In the stride distance-speed plot, uninjured subjects exhibit speed values that are not distinctly concentrated across DPOs. Conversely, injured subjects display a wider distribution in these parameters. Among injured subjects, those on D1 exhibit lower speeds and shorter stride distances, with both parameters increasing as DPO progresses. A positive correlation is observed between stride distance and speed for both injury classes. In the neuromodulation-BBB score plot, both treatment classes include subjects from both injury classes. Injured subjects without treatment show limited recovery, with only minor improvements in BBB scores. In contrast, injured subjects receiving treatment exhibit broader improvements in recovery. The uninjured classes, for both treatment and non-treatment groups, are concentrated at higher BBB scores.

In the DPO-von Frey plot, of the Figure 4.7, the group receiving neuromodulation demonstrates a response to lower gram forces of stimulation up to Day 14 (DPO: D14 (3)). Swing duration values are observed to cluster at 0.1328 and 0.166 seconds. This pattern results from an imputation process applied during calculations.

The most notable distinction in the DPO-BBB plot in the Figure 4.11 is observed at point 6, corresponding to M2 in the DPO timeline. In rats with SCI, the treatment group demonstrates improved BBB scores at advanced DPO stages. In plots comparing speed along the horizontal axis, it is evident that the groups receiving neuromodulation did not exceed a speed of 32 units.

Groups receiving neuromodulation in the Figure 4.15 showed a response to lower gram-level stimulation by day 14 post-operation (DPO: D14 (3)). In graphs comparing swing duration along the horizontal axis, recordings are concentrated at 0.1328 and 0.166. This pattern is attributed to an imputation process performed during the analysis. Groups receiving neuromodulation did not exceed a speed of 38 units.

We tested various feature combinations to evaluate their impact on accuracy. In some cases, we excluded certain features to observe how this arrangement affected the model's performance. For instance, we analyzed the results by excluding or including



Figure 4.4 Continuation of the Figure 4.3.



Figure 4.5 Continuation of the Figure 4.4.

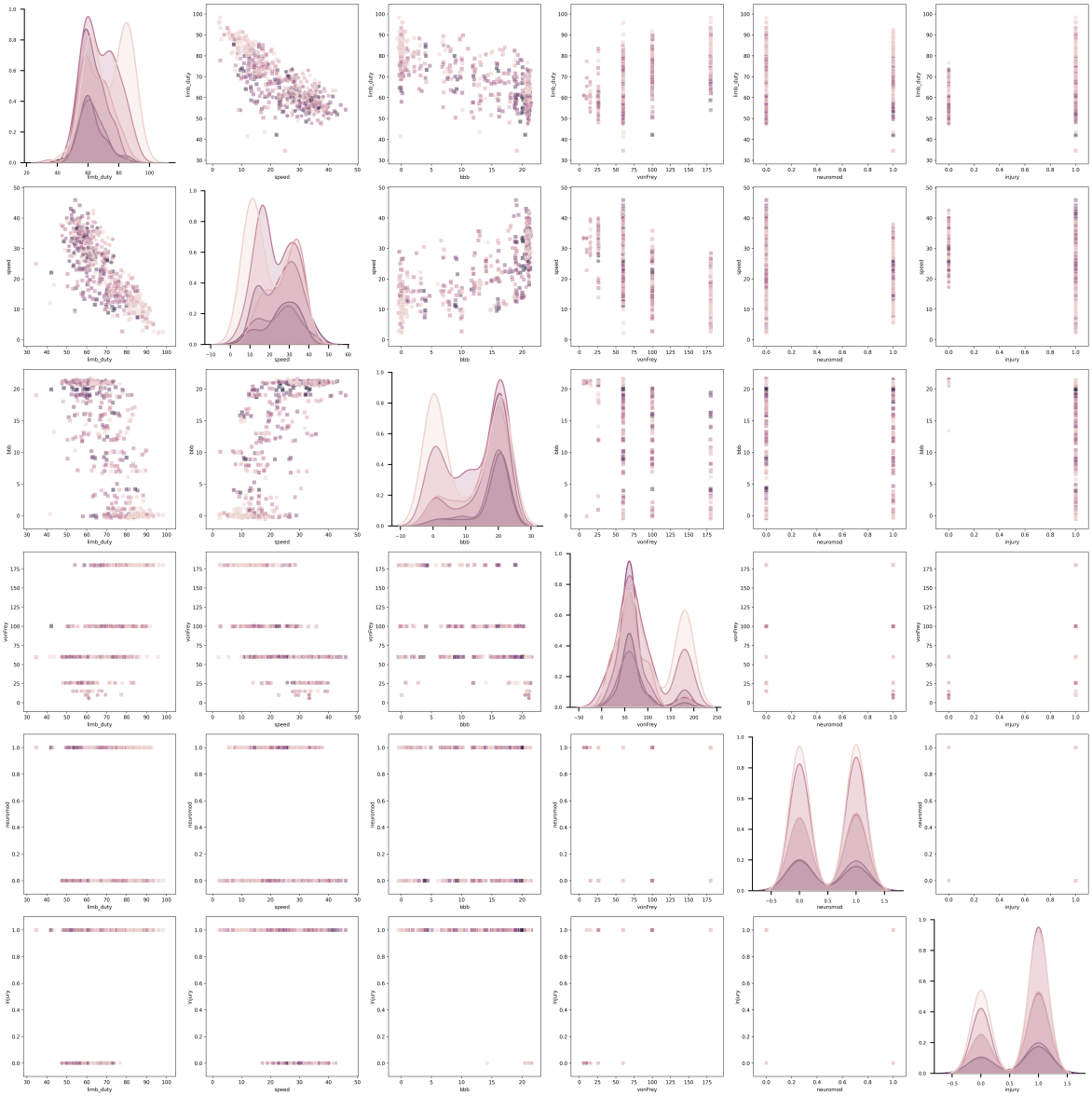


Figure 4.6 Continuation of the Figure 4.5.



Figure 4.7 Distributions of all variables in the dataset containing Sham and SCI rats, colored according to Treatment Classes, in the study. Sized according to End-point Classes. Orange colored (•): neuromodulated, Blue colored (•): non-neuromodulated. Parameters are the same as Figure 4.3.

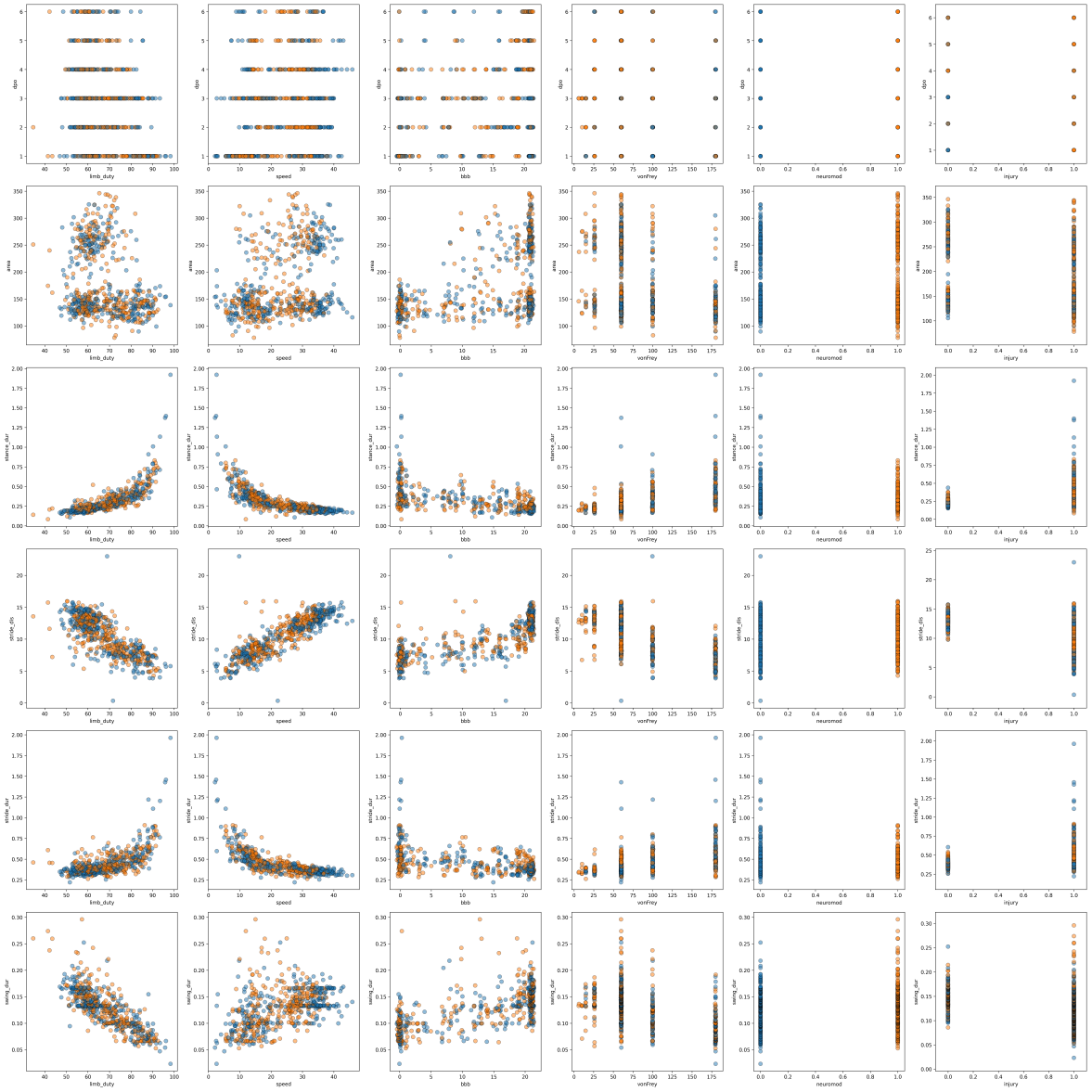


Figure 4.8 Continuation of the Figure 4.7.

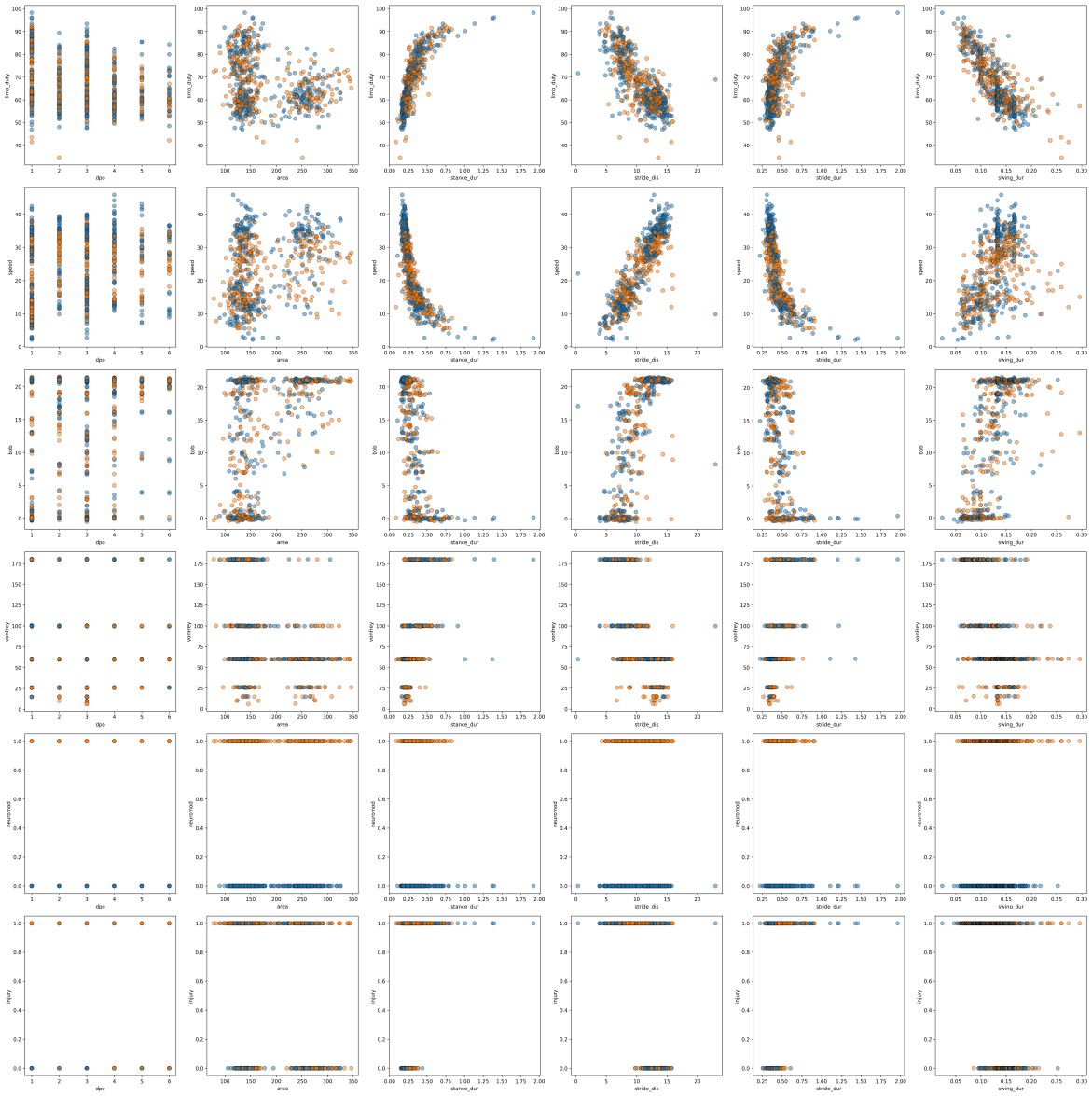


Figure 4.9 Continuation of the Figure 4.8.

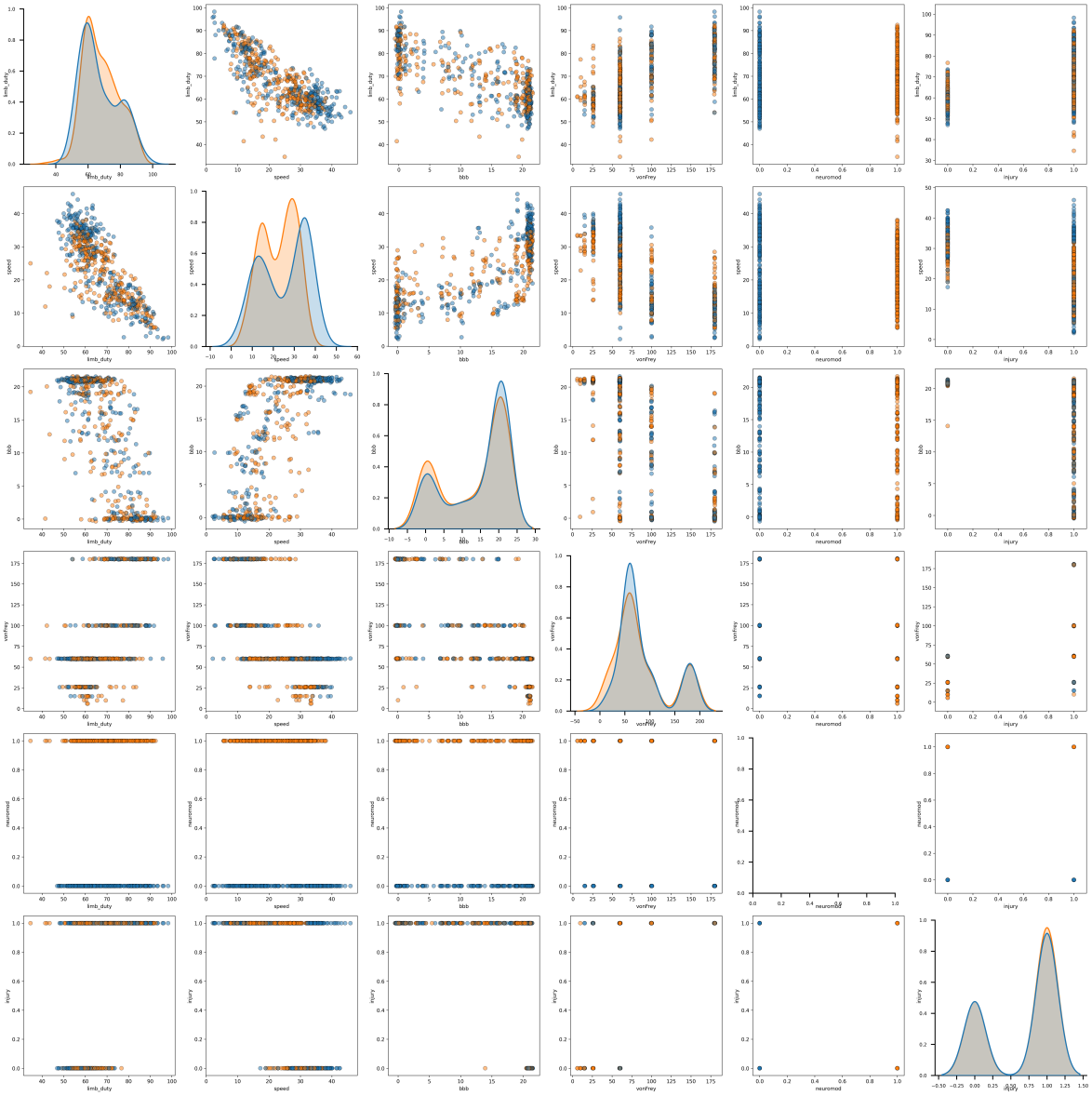


Figure 4.10 Continuation of the Figure 4.9.

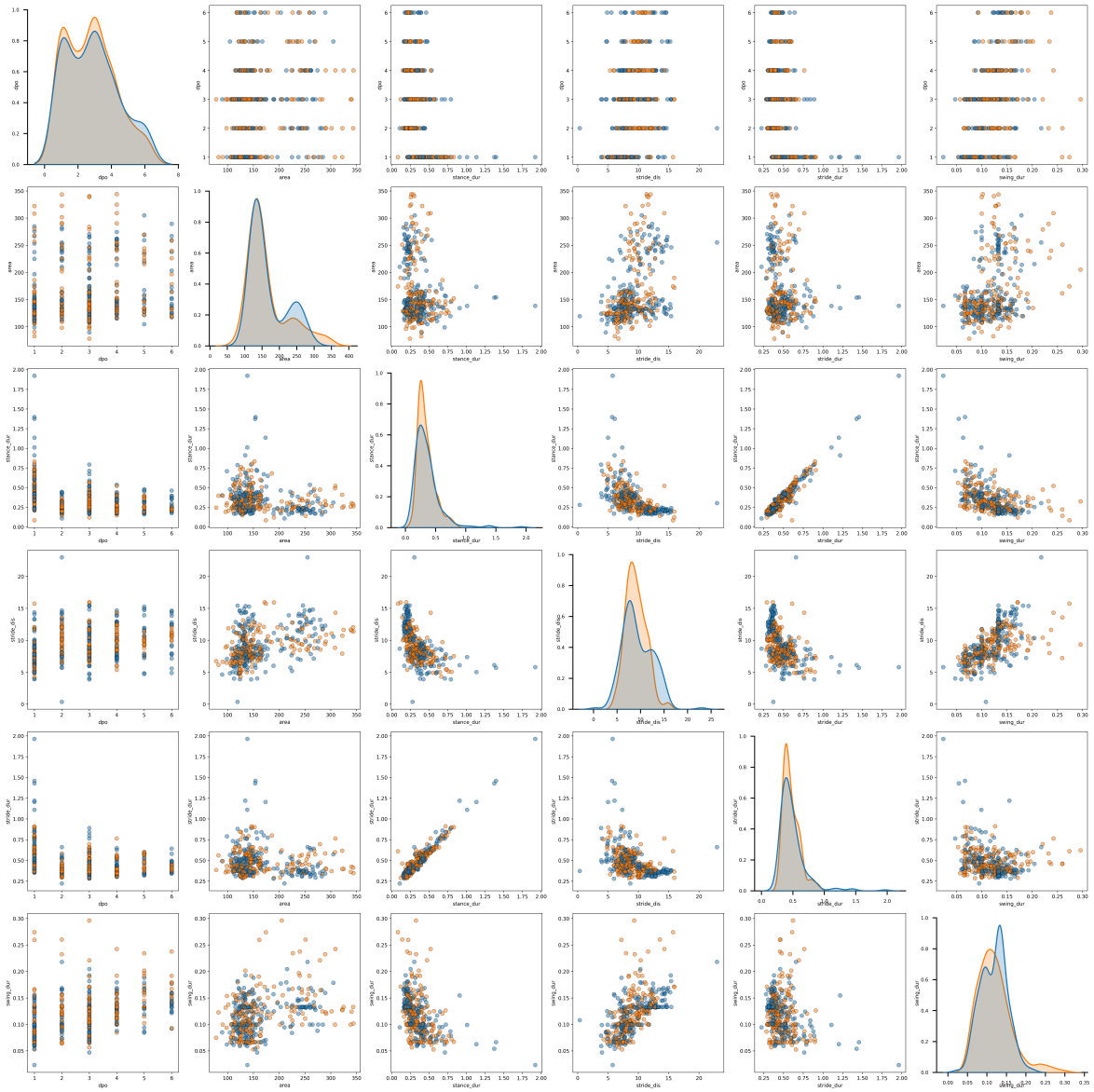


Figure 4.11 Distributions of all variables in the dataset containing only SCI model rats, colored according to Treatment Classes, in the study. Orange colored (•): neuromodulated, Blue colored (•): non-neuromodulated. Parameters are the same as Figure 4.3.

specific features to determine their significance in prediction accuracy. Contrary to our expectations, area was not an effective feature. When ranked using feature importance methods, area consistently appeared at the bottom with the lowest significance scores as in Figure 3.31. BBB was also not directly correlated with treatment prediction. The treatment classes within each injury classes did not show significant differences in locomotor performance, and large variations in BBB scores were only noticeable in cases of substantial differences in locomotor abilities, making it less discriminative.



Figure 4.12 Continuation of the Figure 4.11.

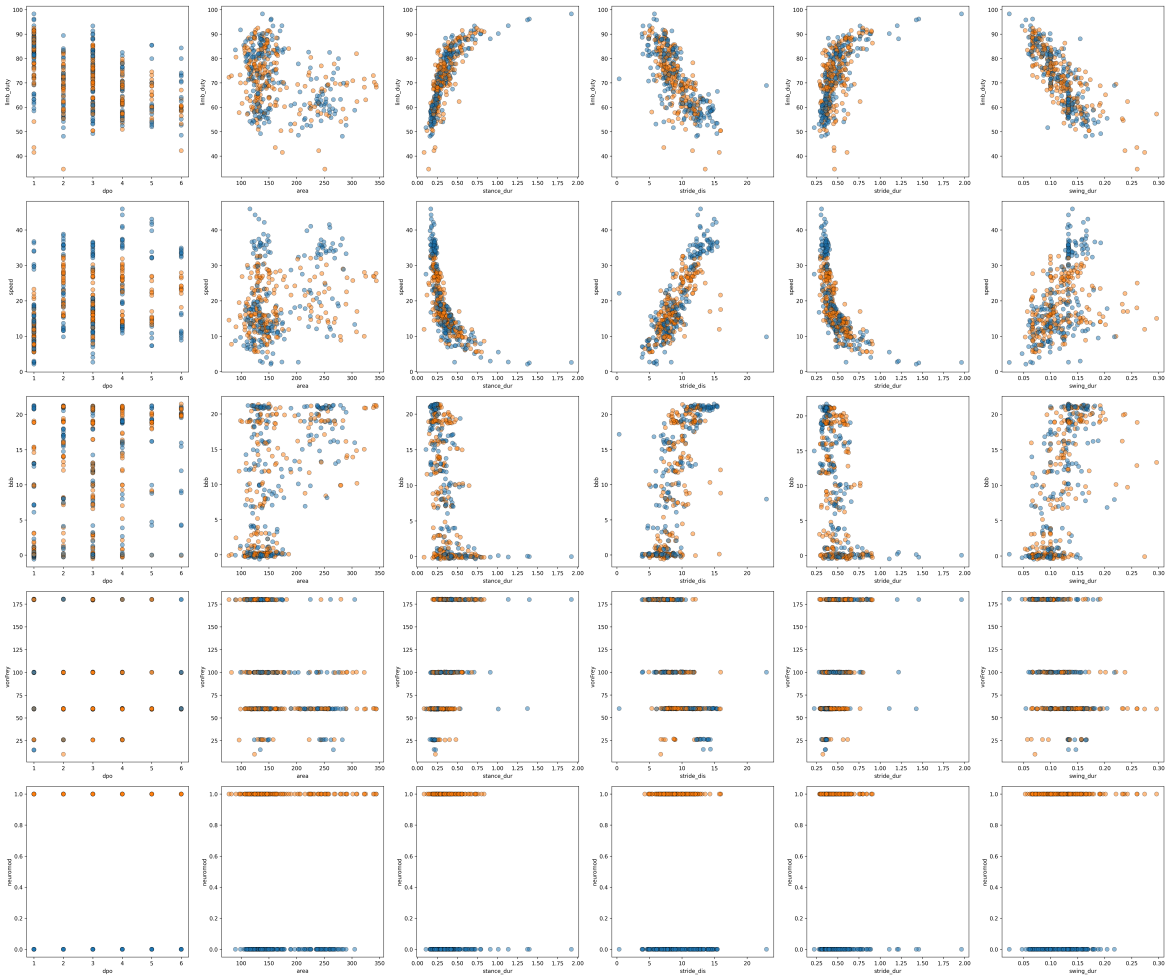


Figure 4.13 Continuation of the Figure 4.12.

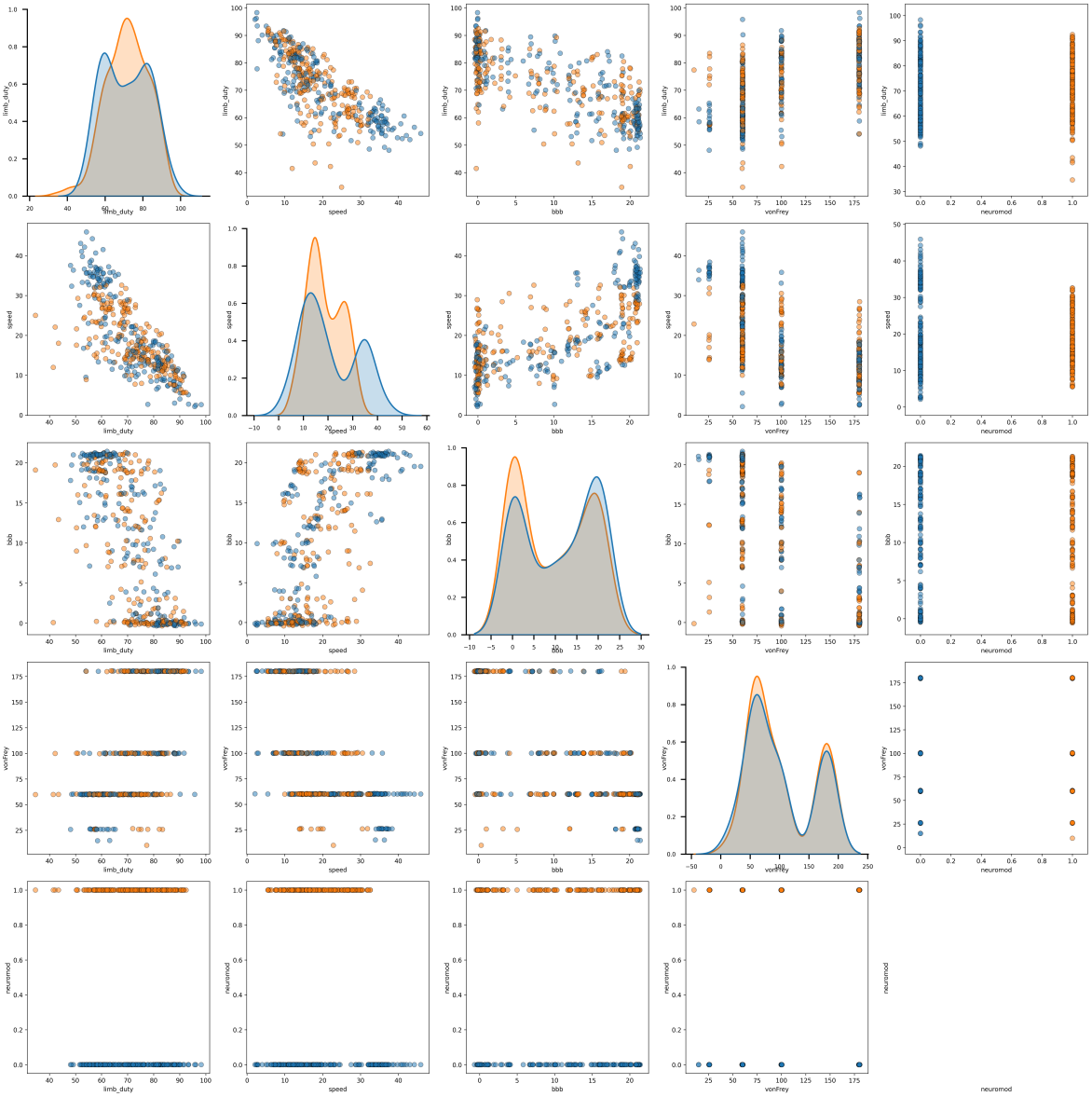


Figure 4.14 Continuation of the Figure 4.13.



Figure 4.15 Distributions of all variables in the dataset containing only Sham rats, colored according to End-point Classes, in the study. Orange colored (•): neuromodulated, Blue colored (•): non-neuromodulated. Parameters are the same as Figure 4.3.

In contrast, vF proved to be the most effective feature for treatment prediction. The physical stimulus effect of vF played a key role in distinguishing treatment classes. Additionally, vF showed high discriminative power within the Sham group for different DPO stages. Similarly, gait parameters also contributed significantly to treatment prediction. These parameters were able to capture subtle differences that BBB alone could not detect.

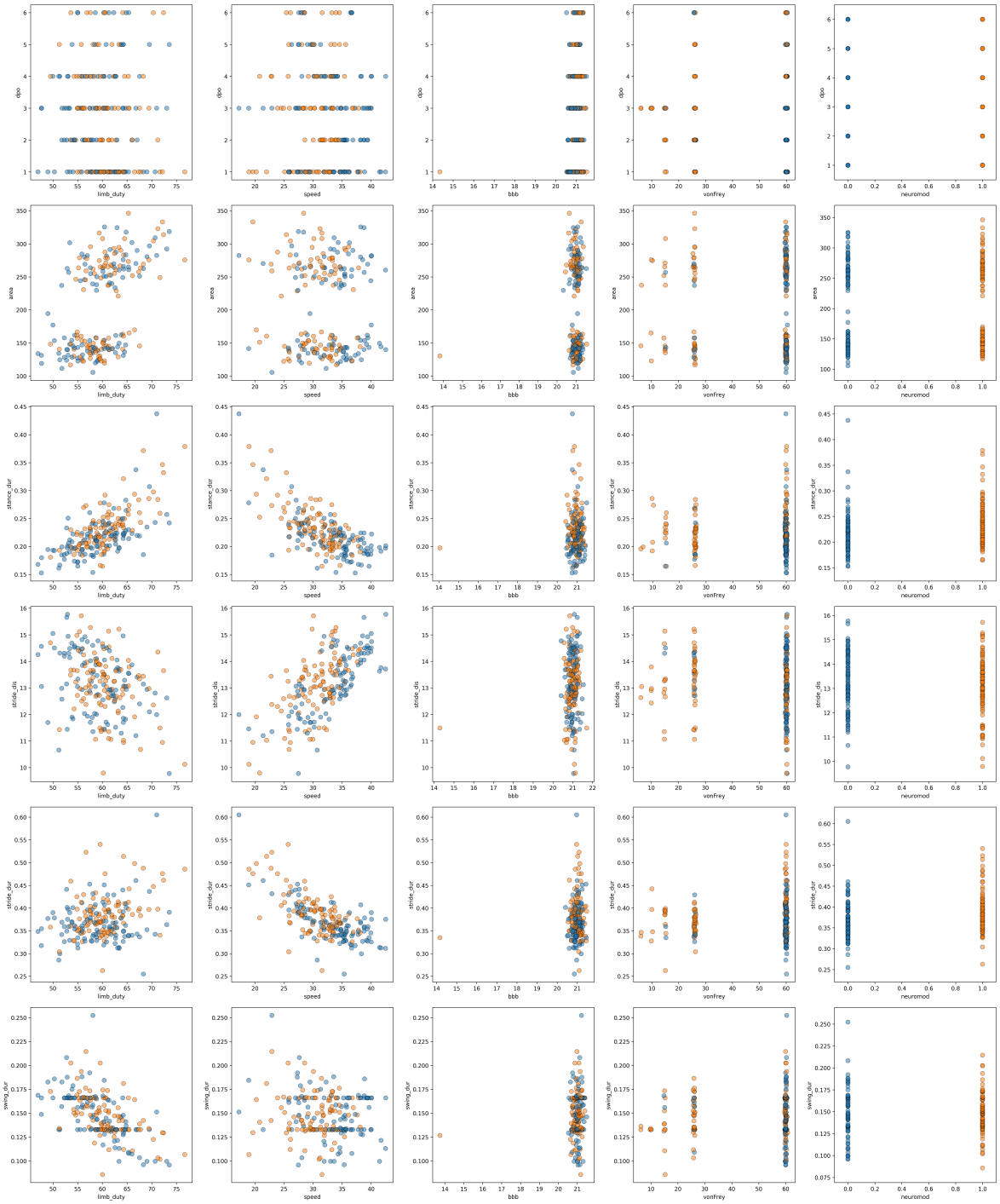


Figure 4.16 Continuation of the Figure 4.15.

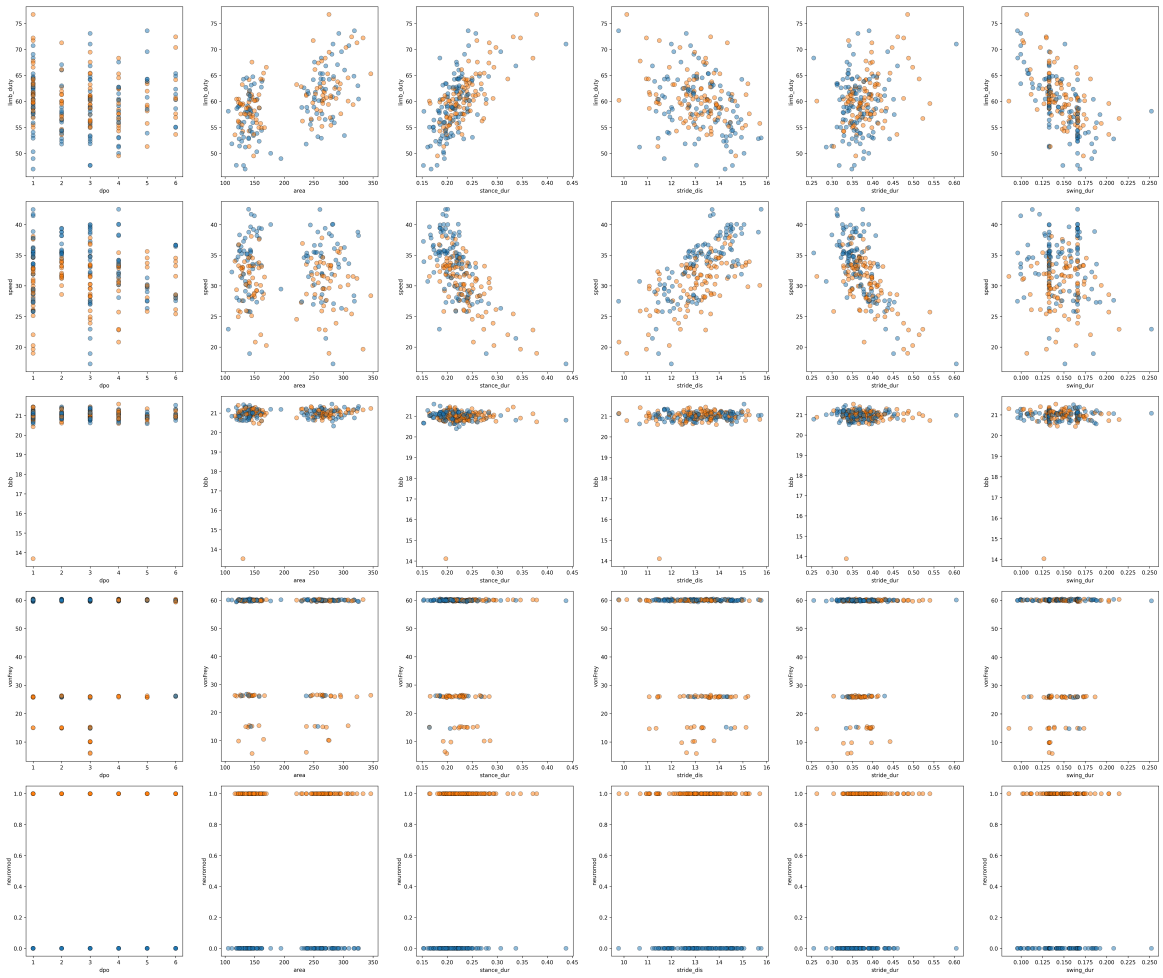


Figure 4.17 Continuation of the Figure 4.16.

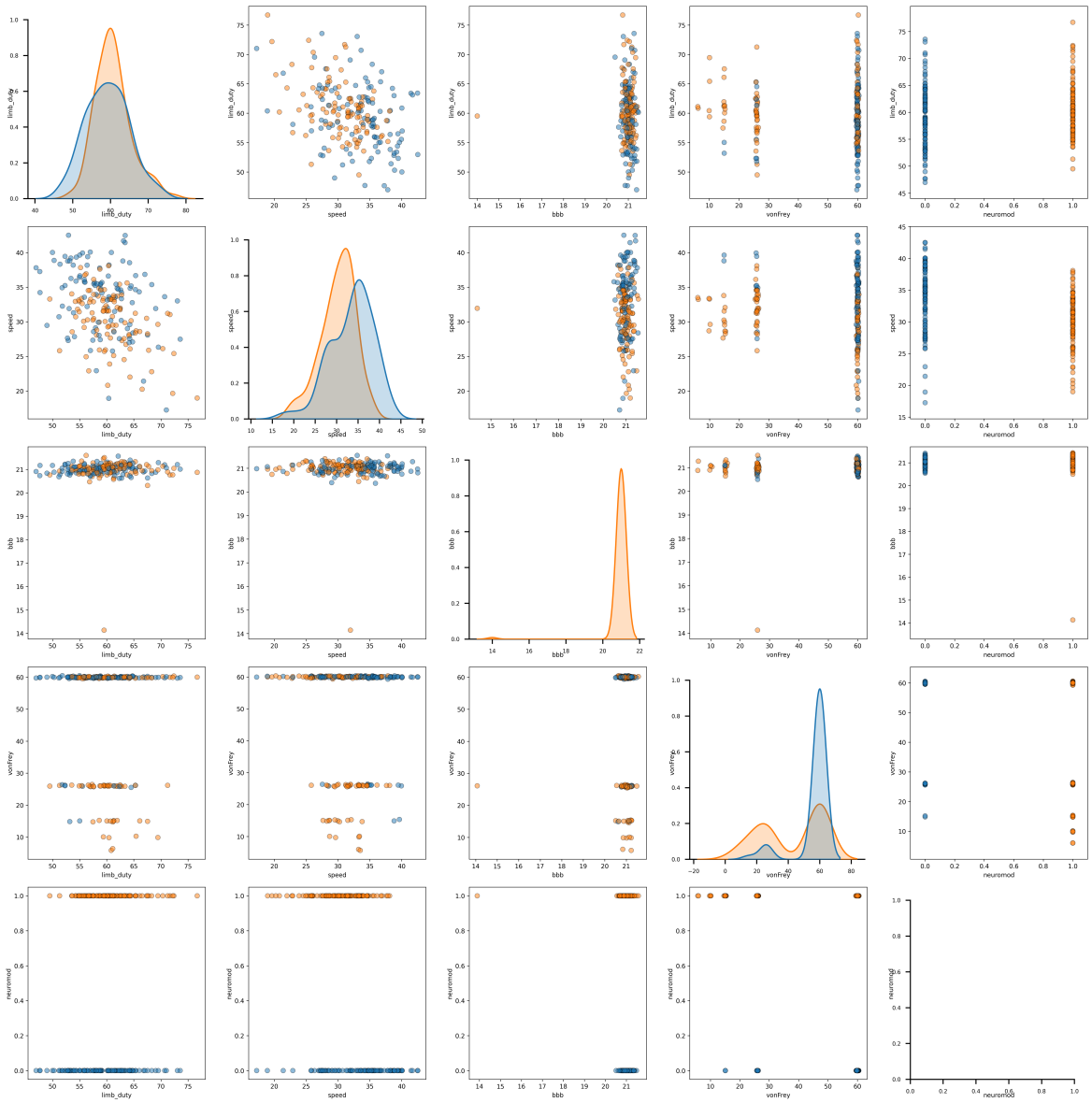


Figure 4.18 Continuation of the Figure 4.17.

4.4 Classification Accuracies

4.4.1 Classification in Normal Format

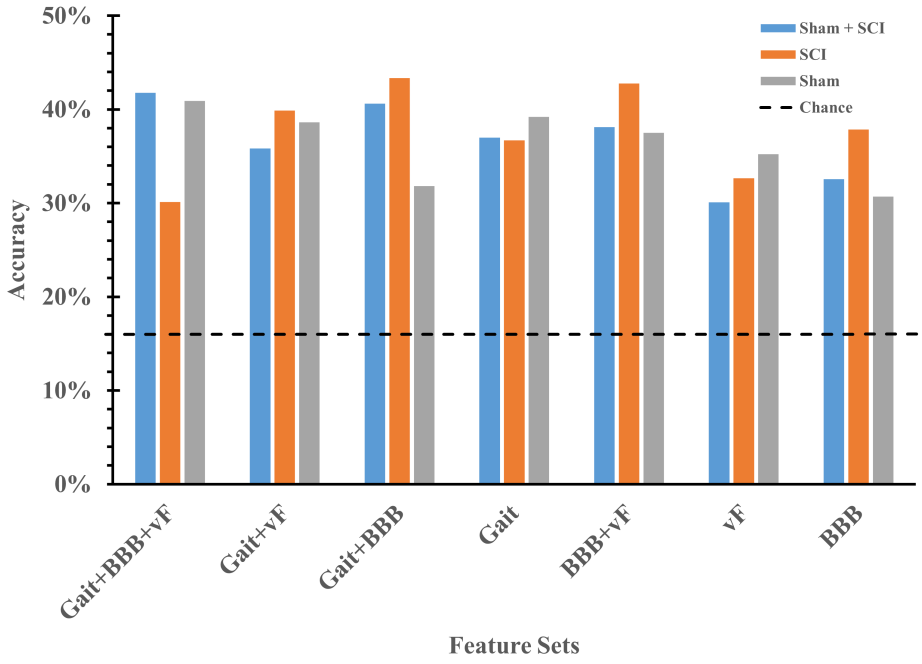


Figure 4.19 Prediction of DPO, using dataset with containing fore and hindlimb records. Dataset formatted as Normal. All groups include groups receiving and not receiving neuromodulation. "Gait+BBB+vF": Full feature set comprising gait parameters used in this study (area, stance duration, stride distance, stride duration, swing duration, limb duty factor, speed), BBB scoring, and Von Frey (vF) scoring, "Gait+vF": Gait parameters and Von Frey scoring, excluding BBB scoring, "Gait+BBB": Gait parameters and BBB scoring, excluding Von Frey scoring. "Gait": Gait parameters only, "BBB+vF": Combination of BBB and Von Frey scoring, excluding gait parameters, "vF": Von Frey scoring only, "BBB": BBB scoring only. For each feature set, results are categorized into three subject groups: “Sham+SCI”: combined injured and non-injured subjects, “SCI”: injured subjects, “Sham”: noninjured subjects. A chance level line is set at %16, representing the minimum accuracy required for random prediction across six endpoint classes (Pre-Op, D1, D7, D14, M1, M1.5, M2). This level serves as a baseline for evaluating model performance.

In the Figure 4.19 the dataset with the largest data, which included Sham and SCI model rats, scored the highest when all features were present. While BBB alone had a higher effect than vF, BBB and vF together had a more discriminative effect than gait parameters alone for the Sham+SCI group. The SCI group gave the highest score when only BBB was added to the gait parameters. When all features were combined, it gave the lowest score among all injury classes. This draws attention to the heterogeneous distribution in SCI model rats. In the Sham group, all BBBs are the same. Therefore, the lowest discriminative effect is observed in the BBB. Since the neuromodulated group is also in this class, the effect can be observed. The

discriminative effect of the vF with physical stimulation increased accordingly. The model has the highest prediction accuracy in the feature set with gait parameters and BBB in the SCI group. Only hindlimb used dataset results are in Figure 4.20.

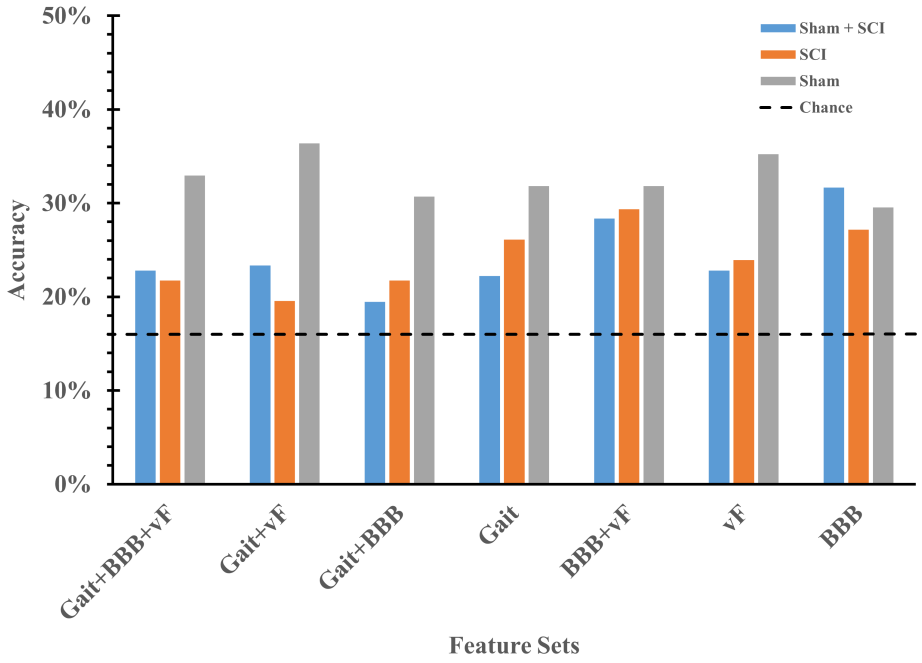


Figure 4.20 Prediction of DPO, using dataset with containing only hindlimb records. Dataset formatted as Normal. All groups include groups receiving and not receiving neuromodulation. Axis are same as Figure 4.19.

When forelimbs are removed from the dataset, it is clearly seen that prediction accuracy decreases across classes in the Figure 4.20. While BBB is the most discriminative feature in the Sham+SCI group, it also provides the highest prediction accuracy for this group when used alone. The SCI group follows the same trend. The fact that the highest prediction accuracy in the feature sets with gait parameters is obtained only in the feature set with gait parameters shows that there is a contradictory relationship between gait parameters and locomotor performance assessment scores. Among all feature sets, the highest prediction accuracy is obtained only in the feature set with BBB and vF together. The Sham group has the highest prediction accuracy among the other two groups, except for the feature set with only BBB. As in the dataset with both front and back paws, vF has the highest discriminative accuracy.

In scenarios where treatment is predicted in the Figure 4.21, the overall prediction accuracy increases. The primary reason for this is the smaller response class

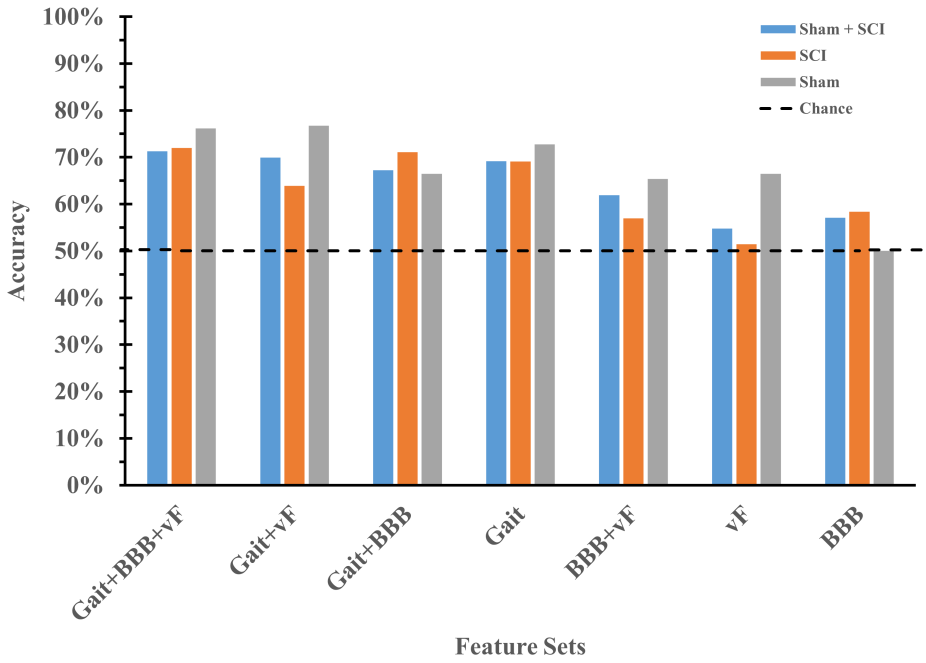


Figure 4.21 Prediction of treatment, using dataset with containing fore and hindlimb records. Dataset formatted as Normal. All groups include groups receiving and not receiving neuromodulation. A chance level line is set at %50, representing the minimum accuracy required for random prediction across two treatment classes (neuromodulated, non-neuromodulated). This level serves as a baseline for evaluating model performance. Axis are same as Figure 4.19.

space. Among all datasets, BBB and vF together and individually have the lowest prediction accuracy. BBB has the lowest discriminative effect for the Sham group and vF has the lowest discriminative effect for the Sham+SCI and SCI groups. In this context, we can make the observation that treatment does not increase BBB scoring noticeably, but small changes can be interpreted with gait parameters. vF, on the other hand, contains direct physical stimulation and has the highest discriminative effect in treatment, so it can be inferred that these two features have the highest discriminative effect in treatment prediction. The model predicts the Sham group with the highest accuracy in the set with Gait parameters and only the vF feature, and with the highest accuracy among the other prediction scores. Same analysis done with only hindlimbs in the Figure 4.22.

Here, in data sets with only the hind limb, the Figure 4.22 shows that the BBB has almost no effect, as in data sets with both front and back paws. For the Sham+SCI group, the accuracy drops when BBB and vF features are removed one by one from the set of all features, while it almost reaches the score in the set of all features

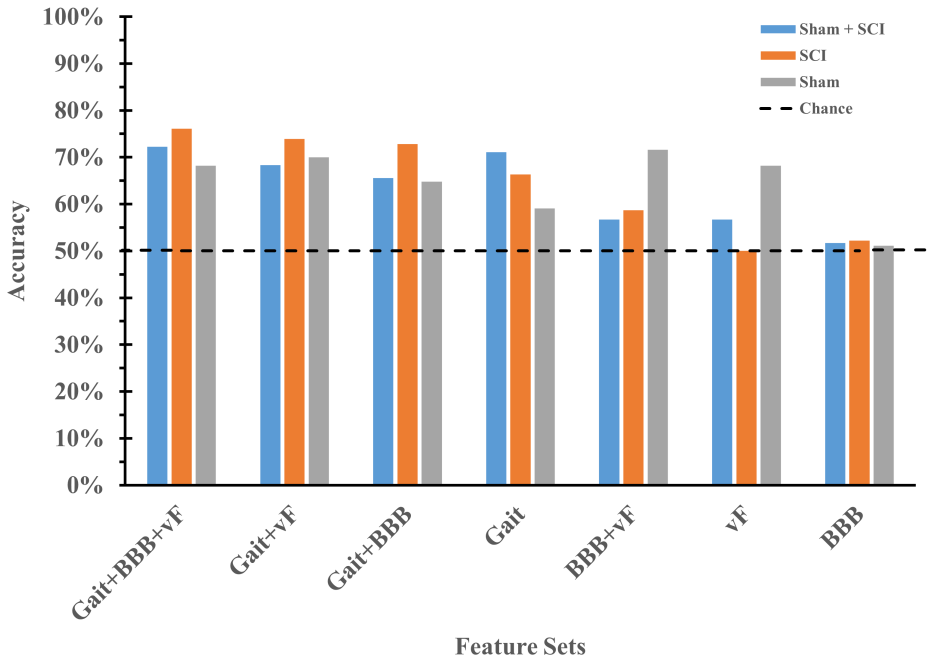


Figure 4.22 Prediction of treatment, using dataset with containing only hindlimb records. Dataset formatted as Normal. All groups include groups receiving and not receiving neuromodulation. Axis are same as Figure 4.19.

when both are removed. It can be inferred that the gait parameters have the highest discriminative information for the Sham+SCI group, when including only hindlimb in the treatment prediction. For the SCI group, there is a linear decrease from left to right in the feature sets. SCI group gives the highest prediction accuracy in the set with all features. The discrimination of vF for the Sham group is also seen here. While BBB has the lowest effect for the Sham group, the model gives the highest prediction accuracy in the feature set where BBB and vF are together.

The overall performance of the machine learning models supported by the toolbox in DPO prediction on this dataset of SCI model rats with fore and hindlimb paw recordings can be seen in Figure 4.23. The grouped bars represent different configurations of the algorithm.

In Figure 4.24, the same data set was used for treatment prediction.

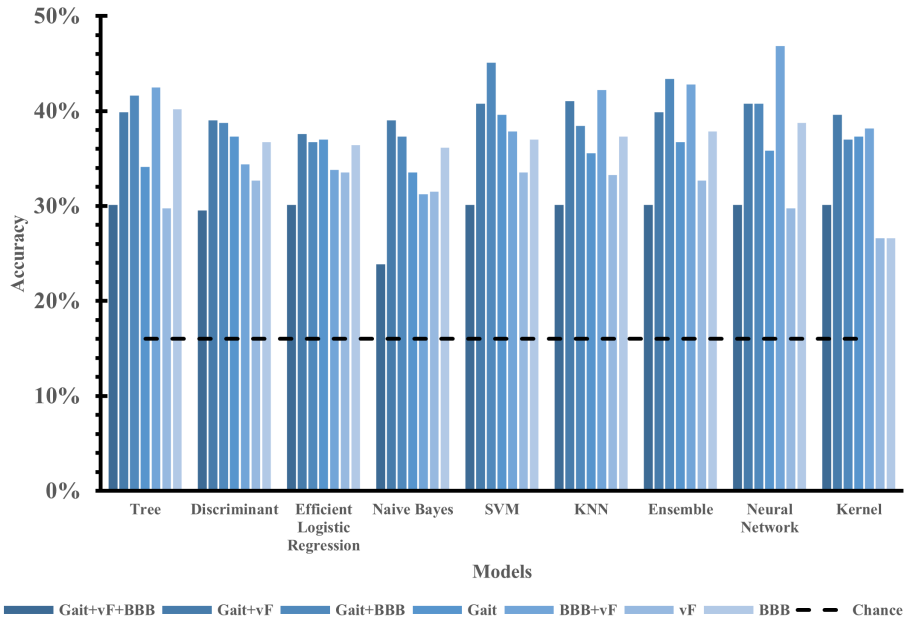


Figure 4.23 Prediction of DPO of models supported by the toolbox, using dataset with containing fore and hindlimb records. Dataset formatted as Normal and contains only SCI model rats. All groups include groups receiving and not receiving neuromodulation. Bars are colored according to feature sets as stated in Figure 4.19.

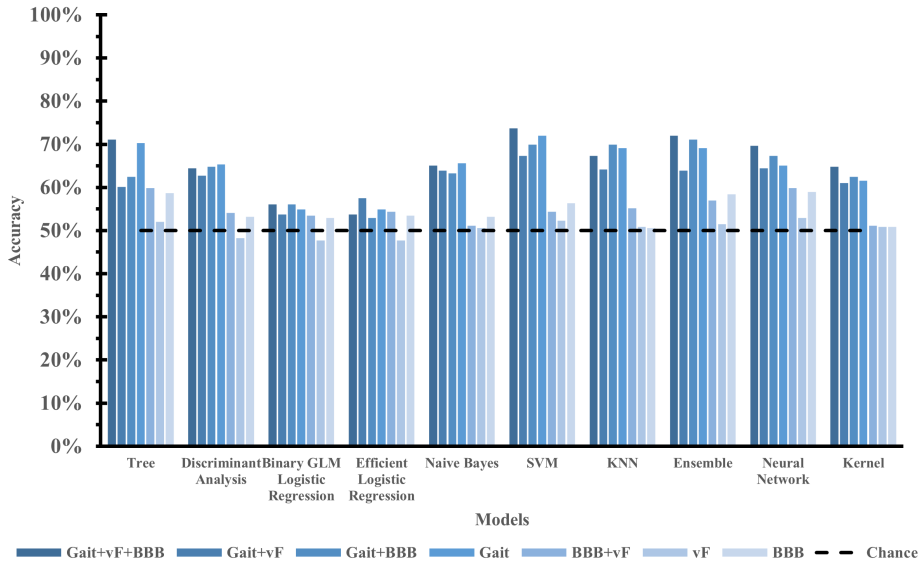


Figure 4.24 Prediction of treatment of models supported by the toolbox, using dataset with containing fore and hindlimb records. Dataset formatted as Normal and contains only SCI model rats. Bars are colored according to feature sets as stated in Figure 4.19.

4.4.2 Classification in Extended Format

The nature of formatting in the Extended format dataset means that only treatment prediction can be done for all rats. In this Figure 4.25, for the SCI group, the

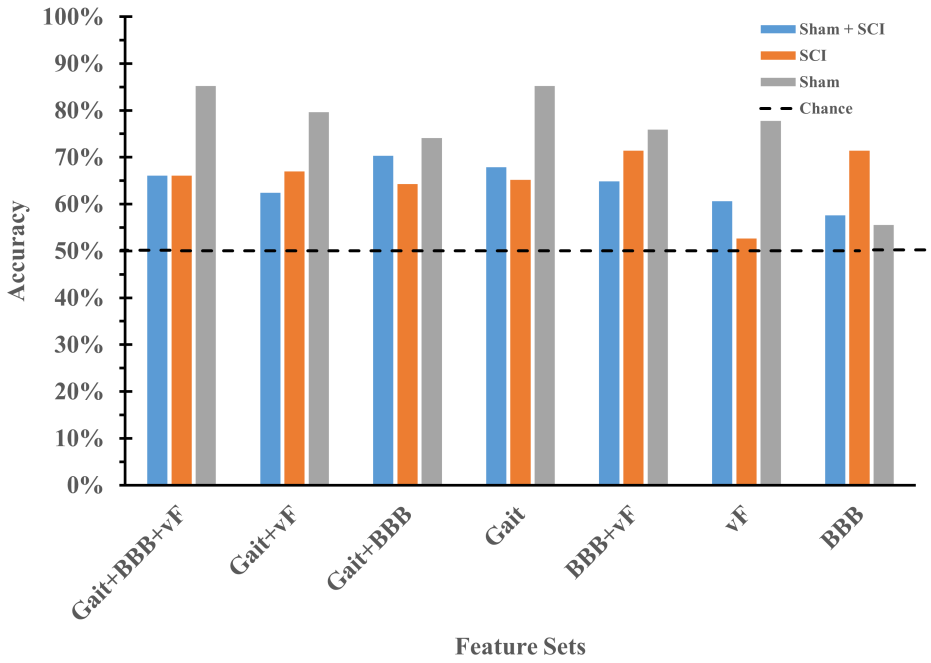


Figure 4.25 Prediction of treatment, using dataset with containing fore and hindlimb records. Dataset formatted as Extended. Axis are same as Figure 4.19.

BBB score alone and in combination with vF provides the highest discrimination and the highest prediction accuracy score. The Sham group has the highest prediction scores in all feature sets except for the feature with only the BBB score. This is due to the fact that the Sham group has the highest density of sham rats and therefore has the richest data on prediction data. BBB has the lowest discriminability because all rats have the same BBB score.

In contrast to Figure 4.23 and Figure 4.24, the Figure 4.26 shows that the Ensemble method stands out together with Naive Bayes. In the scenario with forelimb and hindlimb paws, where treatment is predicted within the SCI group, Ensemble score higher relative to other models.

In the presented Figure 4.23, Figure 4.24 and Figure 4.26, the Ensemble method may not appear to stand out as the most distinctive classifier. However, analyses conducted in different data spaces and configurations revealed that Ensemble consistently outperformed other methods. This was particularly evident in datasets formatted in the extended structure, where Ensemble emerged as a highly effective approach. Ad-

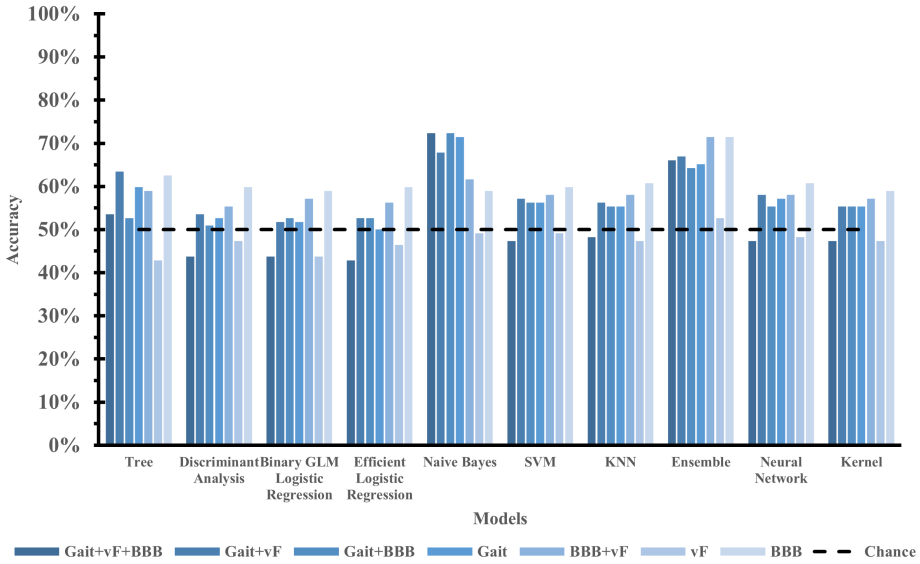


Figure 4.26 Prediction of treatment of models supported by the toolbox, using dataset with containing fore and hindlimb records. Dataset formatted as Extended and contains only SCI model rats. Bars are colored according to feature sets as stated in Figure 4.19.

ditionally, the choice of Ensemble was influenced by the nature of the dataset, which included a relatively low sample size but a high number of features. In such scenarios, Ensemble methods leverage the combined decisions of multiple weak learners, which helps to mitigate overfitting and enhances overall performance. This ability to integrate diverse algorithms into a single decision framework makes Ensemble methods robust and versatile. Furthermore, the strength of Ensemble lies in its capacity to balance the strengths of individual algorithms. For instance, while neural networks excel in capturing complex, non-linear relationships, and decision trees provide interpretability and speed, Ensemble combines these benefits into a unified and optimized model. This synergy helps in maintaining consistency across various feature combinations and experimental setups.

5. DISCUSSION

5.1 Main Conclusions

Here, in this thesis we presented a novel methodology utilizing DLC for gait parameter extraction and machine learning-based classification in SCI model rats. Using open-source tools makes the system flexible and cost-effective, allowing researchers to adjust and expand it for different experiments. This approach reduces the need for expensive proprietary systems and makes it more accessible. It also decreases human error by automating important steps in gait analysis, such as parameter extraction and classification, which improves consistency and reliability. By incorporating innovative approaches such as bottom view imaging and combining these extracted parameters with traditional metrics like BBB scores and vF thresholds, we demonstrated significant advancements in predicting the day after injury and treatment classifications. Specifically, in day after injury classification achieved an accuracy of %43, while neuromodulation therapy classifications reached %71. These findings underscore the feasibility and effectiveness of this accessible and cost-effective approach for SCI research. However, there are some limitations. The system relies heavily on having high-quality and diverse training data, which can limit its performance in new or less common situations. Additionally, while the method shows great potential, it needs further testing to ensure it works well with other animal models or experimental conditions.

The proposed methodology addresses challenges in gait parameter extraction and SCI classification by offering a simple, affordable, and effective solution. Unlike traditional tools like CatWalk or VICON, our approach reduces dependency on expensive equipment, making it accessible for a broader range of research facilities. Additionally, this method effectively classifies gait recovery phases as day after injury classes and differentiates neuromodulation therapies, offering practical applications for both research and therapeutic evaluations.

It is difficult to get a clear and precise indication of how much traditional tools like CatWalk and VICON will cost a lab. While these high-end systems remain the gold standard in gait analysis due to their precision and reliability, their prohibitive cost often limits their accessibility, especially for smaller research facilities or those with constrained funding. Ideally, such advanced systems would be utilized if sufficient resources were available. However, the methodology employed in this study demonstrates that robust and reproducible results can also be achieved with significantly lower-cost setups. By leveraging accessible tools such as DeepLabCut and general-purpose imaging equipment, we successfully circumvented the financial barriers commonly associated with traditional gait analysis systems, without compromising the scientific integrity of the results.

5.2 Previous Studies

We proposed a ML-based method to classify SCI in rats by using gait parameters from bottom view imaging. We showed that this approach can achieve classification accuracies similar to high-cost systems like CatWalk or VICON, while being simpler and cheaper. With DeepLabCut, we extracted detailed gait parameters, especially paw area, which helped us overcome the limitations of traditional lateral-plane methods. By using datasets containing both forelimb and hindlimb data, we successfully differentiated treatment and injury classes. We also showed that treated groups had distinct gait parameter changes, with von Frey scores improving over time. Previous studies often used tools like CatWalk and VICON for SCI analysis in rats. While these are reliable, their high costs and complexity limit their use. Low-cost tools like Agatha, GAITOR, and PrAnCER are alternatives, but our approach stands out by using bottom view imaging to calculate paw area, a parameter often overlooked. By using DeepLabCut, we also addressed challenges like variations in rat physiology and the need for consistent setups, making the method more robust.

5.3 Limitations

One major challenge was the relatively small amount of data available for training and validating. Although we achieved good results for treatment prediction, this was likely because the treatment class has fewer response space. ML requires a lot of data, and splitting it further into groups between forelimb, hindlimb and SCI, Sham reduces the available data even more. Also, the heterogenous distribution of rats across SCI severities and DPO classes affected our results. For example, we had fewer records for later DPO classes. Mixing rats with different severities in the same groups also created variability in gait parameters. While ML helps normalize trends, unclear patterns in small datasets become a bigger problem. Another issue is the reliance on video quality in the bottom view imaging method. If the video quality is poor, paw details may be hard to differentiate, affecting parameter extraction. Even with high-quality videos, if the frame rate is too high, the model may struggle to identify the exact paw contact moment, leading to errors. While this method works well for low-budget and small-scoped studies, more detailed research may need other more robust approaches. Lastly, while gait parameters, BBB scores, and von Frey measures are useful for assessing locomotor performance, they don't work equally well for all classes. For example, BBB scores were not adequate for distinguishing Sham rats, but von Frey scores diffirantiates well alone. Ensuring unbiased measurement is critical, and focusing on features that provide class-specific discriminative power could significantly enhance the success of future studies.

5.4 Future Work

With this thesis we believe that we opened the door to several future research directions in gait analysis and SCI classification. Firstly, predicting DPO can help track the progression of injuries and evaluate treatments more effectively. By linking gait patterns to specific days after injury, it becomes possible to better understand recovery processes and the impact of different therapies. This can also reveal how

treatments vary in their effectiveness over time. Secondly, the method could be used to predict secondary complications and related conditions that develop after injuries, such as inflammation or scar tissue. These predictions could help identify such problems earlier and provide insights into their progression without the need for post-mortem analysis. Thirdly, the approach could be applied to other neurological diseases like Parkinson's, ALS, or MS. These conditions also affect movement, and this method could help analyze changes in gait to track disease progression or evaluate treatments. Using this approach for other diseases would make it more versatile and useful. Additionally, combining the bottom view imaging method with 3D motion capture systems could provide a more complete understanding of movement. This would allow researchers to recreate detailed digital models of animal gait, making it easier to train machine learning models and analyze movements repeatedly. Finally, a fully automated system could be developed to handle video analysis, parameter extraction, and classification in one seamless process. This would make the method more user-friendly and accessible for a wider range of researchers.

5.5 Conclusion

In this study, we developed a simple, affordable, and effective method for classifying SCI in rats. Our framework worked well while avoiding the high costs and complexity of traditional systems. We addressed key challenges in SCI research and made these tools more accessible for labs with limited resources. However, while we achieved %43 accuracy in predicting days after injury and %71 accuracy in treatment classification, there is room for improvement. To enhance these outcomes, further work is needed to address the limitations outlined in this study. In the future, refining this method and extending its application to other conditions and experiments could make it even more valuable for researchers in the field.

REFERENCES

1. García-Alías, G., H. A. Petrosyan, Schnell, and et. al., "Chondroitinase ABC Combined with Neurotrophin NT-3 Secretion and NR2D Expression Promotes Axonal Plasticity and Functional Recovery in Rats with Lateral Hemisection of the Spinal Cord," *Journal of Neuroscience*, Vol. 31, pp. 17788–17799, Dec. 2011.

2. Timotius, I. K., L. Bieler, S. Couillard-Despres, and et. al., "Combination of Defined CatWalk Gait Parameters for Predictive Locomotion Recovery in Experimental Spinal Cord Injury Rat Models," *eneuro*, Vol. 8, pp. ENEURO.0497–20.2021, Mar. 2021.

3. Aljovic, A., S. Zhao, M. Chahin, and et. al., "A deep learning-based toolbox for Automated Limb Motion Analysis (ALMA) in murine models of neurological disorders," *Communications Biology*, Vol. 5, Dec. 2022. Publisher: Nature Research.

4. Verma, R., J. K. Virdi, N. Singh, and et. al., "Animals models of spinal cord contusion injury," Jan. 2019. ISSN: 20930569 Issue: 1 Pages: 12-21 Publication Title: Korean Journal of Pain Volume: 32.

5. Heinzel, J., N. Swiadek, M. Ashmwe, and et. al., "Automated gait analysis to assess functional recovery in rodents with peripheral nerve or spinal cord contusion injury," *Journal of Visualized Experiments*, Vol. 2020, pp. 1–22, Oct. 2020. Publisher: Journal of Visualized Experiments.

6. Kjell, J., and L. Olson, "Rat models of spinal cord injury: from pathology to potential therapies," *Disease Models & Mechanisms*, Vol. 9, pp. 1125–1137, Oct. 2016.

7. Basso, D. M., M. S. Beattie, and J. C. Bresnahan, "A Sensitive and Reliable Locomotor Rating Scale for Open Field Testing in Rats," *Journal of Neurotrauma*, Vol. 12, pp. 1–21, Feb. 1995.

8. Zheng, G., H. Zhang, M. Tail, H. Wang, and et. al., "Assessment of hindlimb motor recovery after severe thoracic spinal cord injury in rats: classification of CatWalk XT® gait analysis parameters," *Neural Regeneration Research*, Vol. 18, pp. 1084–1089, May 2023. Publisher: Wolters Kluwer Medknow Publications.

9. Koopmans, G. C., R. Deumens, W. M. Honig, and et. al., "The Assessment of Locomotor Function in Spinal Cord Injured Rats: The Importance of Objective Analysis of Coordination," *Journal of Neurotrauma*, Vol. 22, pp. 214–225, Feb. 2005.

10. Hamers, F. P., A. J. Lankhorst, T. J. Van Laar, and et. al., "Automated Quantitative Gait Analysis During Overground Locomotion in the Rat: Its Application to Spinal Cord Contusion and Transection Injuries," *Journal of Neurotrauma*, Vol. 18, pp. 187–201, Feb. 2001.

11. Isvoranu, G., E. Manole, and M. Neagu, "Gait Analysis Using Animal Models of Peripheral Nerve and Spinal Cord Injuries," *Biomedicines*, Vol. 9, p. 1050, Aug. 2021.

12. Sato, Y., T. Kondo, M. Shinozaki, and et. al., "Markerless analysis of hindlimb kinematics in spinal cord-injured mice through deep learning," *Neuroscience Research*, Vol. 176, pp. 49–56, Mar. 2022. Publisher: Elsevier Ireland Ltd.

13. S, P., A. S, D. A, and et. al., "Detection of Spinal Cord Injury using Deep Learning Algorithm," in *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, (Erode, India), pp. 270–275, IEEE, Apr. 2022.

14. García-Alías, G., A. Torres-Espín, C. Vallejo, and et. al., "Functional involvement of the lumbar spinal cord after contusion to T8 spinal segment of the rat," *Restorative Neurology and Neuroscience*, Vol. 28, no. 6, pp. 781–792, 2010.

15. Redondo-Castro, E., G. García-Alías, and X. Navarro, "Plastic changes in lumbar segments after thoracic spinal cord injuries in adult rats: An integrative view of spinal nociceptive dysfunctions," *Restorative Neurology and Neuroscience*, Vol. 31, no. 4, pp. 411–430, 2013.

16. Viswakumar, A., V. Rajagopalan, T. Ray, and et. al., "Development of a Robust, Simple, and Affordable Human Gait Analysis System Using Bottom-Up Pose Estimation With a Smartphone Camera," *Frontiers in Physiology*, Vol. 12, p. 784865, Jan. 2022.

17. Cicirelli, G., D. Impedovo, V. Dentamaro, and et. al., "Human Gait Analysis in Neurodegenerative Diseases: A Review," *IEEE Journal of Biomedical and Health Informatics*, Vol. 26, pp. 229–242, Jan. 2022.

18. Gillani, S. I., M. A. Azam, and M. Ehatisham-ul Haq, "Age Estimation and Gender Classification Based on Human Gait Analysis," in *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*, (Karachi, Pakistan), pp. 1–6, IEEE, Mar. 2020.

19. Niu, X., D. Varoqui, M. Kindig, and et. al., "Prediction of gait recovery in spinal cord injured individuals trained with robotic gait orthosis," *Journal of NeuroEngineering and Rehabilitation*, Vol. 11, no. 1, p. 42, 2014.

20. Arac, A., P. Zhao, B. H. Dobkin, and et. al., "DeepBehavior: A Deep Learning Toolbox for Automated Analysis of Animal and Human Behavior Imaging Data," *Frontiers in Systems Neuroscience*, Vol. 13, p. 20, May 2019.

21. Xu, D., H. Zhou, W. Quan, and et. al., "A new method proposed for realizing human gait pattern recognition: Inspirations for the application of sports and clinical gait analysis," *Gait & Posture*, Vol. 107, pp. 293–305, Jan. 2024.

22. Haulcomb, M., R. Meadows, W. Miller, and et. al., "Locomotor analysis identifies early compensatory changes during disease progression and subgroup classification in a mouse model of amyotrophic lateral sclerosis," *Neural Regeneration Research*, Vol. 12, no. 10, p. 1664, 2017.

23. Metz, G., V. Dietz, M. Schwab, and et. al., "The effects of unilateral pyramidal tract section on hindlimb motor performance in the rat," *Behavioural Brain Research*, Vol. 96, pp. 37–46, Nov. 1998.

24. Fouad, K., G. A. Metz, D. Merkler, and et. al., "Treadmill training in incomplete spinal cord injured rats," *Behavioural Brain Research*, Vol. 115, pp. 107–113, Oct. 2000.

25. Mathis, A., P. Mamidanna, K. M. Cury, and et. al., "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning," *Nature Neuroscience*, Vol. 21, pp. 1281–1289, Sept. 2018. Publisher: Nature Publishing Group.

26. Sperry, R. W., "The functional results of muscle transposition in the hind limb of the rat," *Journal of Comparative Neurology*, Vol. 73, pp. 379–404, Dec. 1940.

27. Sperry, R. W., "The effect of crossing nerves to antagonistic muscles in the hind limb of the rat," *Journal of Comparative Neurology*, Vol. 75, pp. 1–19, Aug. 1941.

28. Sperry, R. W., "Transplantation of motor nerves and muscles in the forelimb of the rat," *Journal of Comparative Neurology*, Vol. 76, pp. 283–321, Apr. 1942.

29. Bounds, H. A., D. L. Poeta, P. M. Klinge, and et. al., "Paw-Print Analysis of Contrast-Enhanced Recordings (PrAnCER): A Low-Cost, Open-Access Automated Gait Analysis System for Assessing Motor Deficits," *Journal of Visualized Experiments*, p. 59596, Aug. 2019.

30. Ueno, M., and T. Yamashita, "Kinematic analyses reveal impaired locomotion following injury of the motor cortex in mice," *Experimental Neurology*, Vol. 230, pp. 280–290, Aug. 2011.

31. Lauer, J., M. Zhou, S. Ye, and et. al., "Multi-animal pose estimation and tracking with DeepLabCut," Apr. 2021.

32. Klapdor, K., B. G. Dulfer, A. Hammann, and et. al., "A low-cost method to analyse footprint patterns," *Journal of Neuroscience Methods*, Vol. 75, pp. 49–54, July 1997.

33. Hamers, F. P., G. C. Koopmans, and E. A. Joosten, "CatWalk-Assisted Gait Analysis in the Assessment of Spinal Cord Injury," *Journal of Neurotrauma*, Vol. 23, pp. 537–548, Apr. 2006.

34. Kloefkorn, H. E., T. R. Pettengill, S. M. F. Turner, and et. al., "Automated Gait Analysis Through Hues and Areas (AGATHA): A Method to Characterize the Spatiotemporal Pattern of Rat Gait," *Annals of Biomedical Engineering*, Vol. 45, pp. 711–725, Mar. 2017.

35. Duan, H., Y. Pang, C. Zhao, and et. al., "A novel, minimally invasive technique to establish the animal model of spinal cord injury," *Annals of Translational Medicine*, Vol. 9, pp. 881–881, May 2021.

36. Jeffery, N. D., K. Brakel, M. Aceves, and et. al., "Variability in Open-Field Locomotor Scoring Following Force-Defined Spinal Cord Injury in Rats: Quantification and Implications," *Frontiers in Neurology*, Vol. 11, p. 650, July 2020.

37. Kappos, E. A., P. K. Sieber, P. E. Engels, and et. al., "Validity and reliability of the CatWalk system as a static and dynamic gait analysis tool for the assessment of functional nerve recovery in small animal models," *Brain and Behavior*, Vol. 7, p. e00723, July 2017.

38. Huang, R., A. A. Nikooyan, B. Xu, and et. al., "Machine learning classifies predictive kinematic features in a mouse model of neurodegeneration," *Scientific Reports*, Vol. 11, Dec. 2021. Publisher: Nature Research.

39. The MathWorks Inc., Natick, Massachusetts, *Statistics and Machine Learning Toolbox User's Guide*, Sept. 2023.

40. Salman, R., A. Alzaatreh, H. Sulieman, and et. al., "A Bootstrap Framework for Aggregating within and between Feature Selection Methods," *Entropy*, Vol. 23, p. 200, Feb. 2021.

41. Brodeur, Z. P., J. D. Herman, and S. Steinschneider, "Bootstrap Aggregation and Cross-Validation Methods to Reduce Overfitting in Reservoir Control Policy Search," *Water Resources Research*, Vol. 56, p. e2020WR027184, Aug. 2020.

42. Breiman, L., "Bagging predictors," *Machine Learning*, Vol. 24, pp. 123–140, Aug. 1996.

43. Polikar, R., "Ensemble learning," *Scholarpedia*, Vol. 4, no. 1, p. 2776, 2009.

44. DataCamp, "Ensemble learning in python," n.d. Accessed: 2024-12-30.