Equine Gait Analysis, Body Part Tracking using DeepLabCut and Mask R-CNN and Biomechanical Parameter Extraction

by

Vinika Gupta

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August 7, 2021

Keywords: Biomechanics, Deep learning, Equine, Gait analysis

Approved by

Bo Liu, Chair, Assistant Professor Computer Science and Software Engineering
Yin Bao, Co-chair, Assistant Professor Biosystems Engineering
Elizabeth Staiger, Visiting Assistant Professor Animal Sciences

Abstract

Gait analysis plays a pivotal role in quantitatively defining equine biomechanical parameters for lameness detection and performance evaluation. Currently, equine gait analysis requires attaching markers or IMU sensors to a horse for motion capture, which is low-throughput and may impact the quality of the horse's movement. In this study, evaluation of the feasibility and utility of deep learning-based video processing as a marker-less motion capture method for equine gait analysis was performed. For that purpose, evaluation consisted of an annotated video dataset of horses performing their natural locomotion patterns, accounting for 4075 images, each with 21 tagged landmarks on the body. Detection of landmarks utilized DeepLabCut and Mask R-CNN models for each video. A performance comparison between both methods evaluated which landmark detection paradigm achieved higher detection accuracy. A fine-tuned Mask R-CNN model had a lower overall RMSE = 30.6 compared to DeepLabCut = 128.4, whereas DeepLabCut had a lower RMSE = 20.6 for the x coordinates of kepoints of each landmark compared to Mask R-CNN (RMSE = 26.1). Based on the x-axis body landmark tracking results of DeepLabCut, algorithms were developed to extract the two biomechanical parameters of stride length and stance time. The proposed post-processing pipeline correctly detected 92% of the strides and 95% of the stances. Subsequently, interpretive accuracies were found for the automatically extracted stride length (R^2 = 0.80, RMSE = 0.31) and stance time (R^2 = 0.81, RMSE = 0.03). The developed video processing pipeline has promising potential to become a convenient and efficient analytics tool for animal scientists and veterinarians to study the genetic control of equine locomotion and diagnose musculoskeletal problems, respectively.

Acknowledgments

I would like to express my heartfelt gratitude to Dr. Yin Bao, my graduate advisor, for his encouragement and guidance during my time at Auburn University. His encouragement and guidance paved the way for my successful research projects and thesis completion. I would like to express my gratitude to Dr. Samantha Brooks from University of Florida, for her experience and insight into equine gait analysis. Also, a special thanks to my committee member of the thesis, Dr. Elizabeth Staiger for her counsel to this research. I want to extend my gratitude to the academic advisor Dr. Bo Liu for his time, support, and advice towards my research and thesis. Thank you to all my lab mates and colleagues for the valuable information I acquired during my course and study work. The knowledge sharing during our lab meetings taught me a lot about my field of study. Finally, I would like to thank my parents and friends for their unwavering support during my academic career.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

DLC	DeepLabCut
GT	Ground Truth
RCNN	Region Based Convolution Neural Network
RPN	Region Proposal Network
FPN	Feature Pyramid Network
COCO	Common object in Context
CVAT	Computer Vision Annotation Tool
AP	Average Precision
IoU	Intersection over Union
RMSE	Root Mean Squared Error
R^2	R-squared, coefficient of determination
ROI	Region of Interest
MAPE	Mean Absolute Percentage Error
UFL	University of Florida

Chapter 1

Introduction

Horses possess wide variation in the pattern and timing of locomotion which is utilized in transportation, military, and sports activities. The performance of horses in sport, pleasure or race activities is determined by the quality of horse's locomotion (Serra Bragança et al., 2018) and thus, is of immense concern for the owners. Horse owners are interested to evaluate the soundness in locomotion through a lameness examination; lameness can lead to a loss of equine training or competition days causing financial losses for these owners and a possibility of an end to the horse's athletic career (Jeffcott et al., 1982; Murray et al., 2006; Dyson et al., 2008; Agneta Egenvall et al., 2008; A. Egenvall et al., 2013). Lameness is not a disease in itself but is a symptom of existing disturbance in locomotion. The goal of the lameness examination is to localize its root cause for possible veterinary therapy.

Locomotion in the horse can be defined into four major gait classifications based upon the speed and footfall pattern. The walk is a four-beat gait where two or three feet are in contact with ground at any instance and is the slowest gait. The trot is faster than the walk and is a two-beat diagonal gait with two moments of suspension (all four feet are off the ground simultaneously). In the trot, the opposite hind and fore limb travel and touch the ground simultaneously together (the left forelimb with the right hindlimb, and the right forelimb with the left hindlimb). The canter is a three-beat gait with a moment of suspension followed by one forelimb touching the ground, then the other forelimb and opposite hindlimb touching simultaneously, and the sequence ends with the remaining hindlimb touching the ground. The gallop is a four-beat gait with a moment of suspension where all four limbs are not in contact with the ground and is the fastest gait.

Analysis of equine gait patterns plays a pivotal role in the improvement of horse breeding, helps predict the performance potential of young horses, and reduces costs associated with training (Barrey, 1999). The horse gait patterns may be monitored regularly to assist in the early detection of an injury or lameness. Gait analysis is also used to assess the effectiveness of training or increase in fitness of the horse (Barrey et al., 1995; Rose et al., 2009). If the gait patterns are abnormal or drastically different from the baseline, it could be an indication that the horse requires physiotherapy or another treatment. In four year old or less Thoroughbred racehorses 53-68% horses could not race due to lameness in 1980 (Jeffcott et al., 1982). Even in 2016 in UK, 33% of the horses suffered due to lameness according to the National Equine Health Survey (Shrestha et al., 2017). Gait analysis is also used to identify the biomechanical parameters that are consistent among top performing sport and race horses (Echterhoff, Haladjian, & Brügge, 2018). Thus, equine gait analysis and its detection technique is of paramount importance.

1.1 Related Work

Current equine gait analysis techniques combine video recordings with commercial software, optoelectronic systems, as well as a variety of sensors such as electro-goniometers, force plates or shoes, strain gauges, and accelerometers (Shrestha et al., 2017). Stationary force-plate analysis techniques are repeatable as well as highly sensitive (Aviad, 1988; Merkens & Schamhardt, 1988). They are considered the "gold-standard" to date for kinetic gait analysis and lameness detection as they are precise and accurate instruments, but the trade-off is the laborious and time-consuming data collection process (Serra Bragança et al., 2018). Force-measuring shoes and treadmills have also been used in an attempt to refine the data collection process and

measure contiguous strides (Weishaupt et al., 2002), but these require training the horse to become accustomed to the treadmill (Bächi et al., 2018).

Wireless sensor-based systems are extensively used for clinical lameness investigations because they were claimed to be noninvasive, to not affect the horse's natural movement, and are easily attachable to the body (Marshall et al., 2012). These systems have advantages in terms of data collection as they measure lameness in real-time on a continuous scale (Marshall et al., 2012). However, recent studies have shown that the position of the sensor on the horse affects the kinematic data obtained by using inertial sensor-based gait analysis systems (Moorman et al., 2017).

Radar sensors can provide significant contributions in detecting and assessing lameness of horses by extending the already established techniques to distinguish between radar signatures of healthy horses to those exhibiting fore or hind limb lameness (Shrestha et al., 2017). With regards to equine lameness assessment, radar sensors may be interesting for their contactless and non-invasive detecting abilities, with no device required to be attached to the horse's body (Shrestha et al., 2017). However, there is not much exploration on the sustainability of this technique (Shrestha et al., 2017).

Motion analysis has advanced considerably. Biomechanical devices have grown significantly from manual explanation of images to marker-based optical trackers, inertial sensor-based frameworks and markerless frameworks utilizing refined human body models, computer vision and AI algorithms (Colyer et al., 2018). Applying computer vision and artificial intelligence for assessing horse locomotion has however not been significantly explored. This markerless

technology could be leveraged to assess horse locomotion and lameness detection (Wang et al., 2021).

1.2 Deep Learning

Over the last few years, the popularity of innovative techniques based on Deep Learning has engendered many research groups to apply ROI techniques for object segmentation in images and in videos, tracking for different purposes. The approach of this form appears to be very interesting and powerful as the steps required for the features extraction from the segmented ROIs are achieved, owing to DL architectures that use deep classifiers, i.e., Convolutional Neural Networks (CNNs) (Chua & Roska, 1993).

In recent years, with the animal pose estimation tools like DeepLabCut (Nath et al., 2019), DeepPoseKit (Graving et al., 2019), and OpenPose (Yunus et al., 2020), it is easy to track the motion of an animal without the use of markers or IMU sensors. Most of these tools are open-source and therefore, inexpensive.

Biomechanical analysis using computer vision and deep learning has been widely used to evaluate and optimize performance of human athletes as mentioned above (Fang et al., 2018; Toshev & Szegedy, 2014). Now this innovative technology is also applied to measuring equine movement and performance (Kil et al., 2020). This technology allows objectively measuring of the horse's gaits patterns, providing an analysis of how the horse is moving. It also highlights any potential issues that might be adding to injury, behavioral concerns or restricting performance of horses in the pleasure, races or other sports. Little research has been done to measure the accuracy of these machine learning based methods.

Machine learning when integrated with biomechanics can provide a distinct advantage over current methods in animal research. A deep learning approach based on raw IMU sensor data that used a long-short term memory (LSTM) network were able to achieve high accuracy in gait classification (Serra Bragança et al., 2020). Deep learning has been utilized to classify animals through pictures with >93% accuracy (Norouzzadeh et al., 2018). Deep convolutional neural networks were explored to track horses (Kil et al., 2020). Around 34 horses kept in their stalls were followed by automated prediction of three horse body parts – nose, wither and tail. It was guaranteed that horse tracking was attained with a sensitivity of around 80% (Kil et al., 2020). This automated video tracking process was first of its kind for stabled horses. The limitation of this horse tracking automation is that it only considers the horses in stables, which means small area with restricted and slow movement (Kil et al., 2020). The proposed method in this thesis overcame that limitation by tracking horses moving outdoors, giving room for more space.

1.3 Research Objective

In this thesis, a video-based markerless motion capture, deep learning-based body part tracking and computer vision-based biomechanical parameter information extraction for equine gait analysis are proposed. The system developed consists of DeepLabCut based tracking of equine body parts and a fully automated pipeline that calculates the equine biomechanical parameters. The specific objectives of this research were to 1) Compare equine body part tracking of Mask R-CNN and DeepLabCut, and 2) Develop and evaluate a robust pipeline to process body part tracking data using computer vision to extract the trajectory of various equine body parts to compute biomechanical parameters, primarily stride length and stance time.

The dataset consists of tracking 21 different equine body parts while the horses are in motion from 12 videos and over 300 frames extracted from each video. This study aims to expand the horizon in the field of equine body parts tracking by processing and analyzing the videos with deep convolution neural networks. This could be used for multiple purposes – for example, to measure the horse's stride length, stance time, duty factor, and average forward limb extension for the detection of lameness or other pain related behaviors and quantitatively score athletic performance.

Chapter 2

Equine Body Part Tracking, Mask R-CNN vs DeepLabCut

Convolutional Neural Networks (CNN) have advanced significantly over the past years and been considered a practical detection technique for image recognition. Over the last decade, different deep convolutional neural networks like R-CNN (Girshick et al., 2016), Fast R-CNN (Girshick, 2015), Faster R-CNN (Ren et al., 2015) and Mask R-CNN (He et al., 2017) have shown great potential in object detection and classification of millions of worldwide pictures because of their ability to provide higher precision, accuracy and speed over the basic computer vision techniques.

Deep learning has also been evolving in the field of animal pose estimation. Animal pose estimation is based on tracking animal body parts. DeepLabCut (Mathis et al., 2018) is one of the technologies used in various animal pose estimation (Labuguen et al., 2019) projects owing to its qualities of being markerless, robust, open source and fast. DeepLabCut has already been used to estimate animal poses providing satisfactory results (Nath et al., 2019). Mask R-CNN is another technology that is used to track humans, could be extended for equine body part tracking. Both Mask R-CNN and DeepLabCut are based on CNN but they are two different paradigms within the CNN umbrella. The objective in this chapter was to compare the performance of equine body part (keypoint) detection using two CNN-based paradigms Mask R-CNN and DeepLabCut and assess their utility.

2.1 Mask R-CNN

Mask R-CNN is an improved version of Faster R-CNN (Ren et al., 2015) wherein an additional branch for segmentation mask prediction on each Region of Interest (ROI) is added

along with the classification and bounding box branches in parallel. It uses a two-stage technique, with RPN first stage. In the subsequent stage, besides predicting the class and box offset, Mask R-CNN additionally returns a binary mask for each ROI. Mask R-CNN follows the structure of Fast R-CNN that applies classification of bounding box and mask in parallel (He et al., 2017). It also uses an ROI alignment layer that increases the accuracy of the mask. For pose estimation or keypoint detection, a keypoint's location is modelled as one-hot mask, and Mask R-CNN is used to predict K masks (He et al., 2017), one for each of K keypoint types like shoulder, elbow, etc.

2.1.1 Detectron2

Facebook AI Research (FAIR) created Detectron2 to speed up the execution and assessment of novel computer vision research. Detectron2 is a ground-up modification of the already accessible Detectron form and is acquired from the Mask R-CNN benchmark (Wen et al., 2021). Detectron2 immerses top notch executions of cutting-edge object detection algorithms like DensePose (Wu et al., 2019).

2.2 DeepLabCut

DeepLabCut (Nath et al., 2019) is a deep convolution network formed by the combination of a pre-trained ResNet layer and a deconvolutional layer (Insafutdinov et al., 2016). The ResNet backbone was pre-trained on ImageNet dataset (He et al., 2016). The output of this layer is fed into a deconvolutional layer which is used to upsample the visual features and produce spatial probability densities of each pixel. When the model is trained with custom data, it learns from the labeled training images.

2.3 Materials and Methods

2.3.1 Data Collection

From a consumer quality high resolution camera, Sony Alpha a3600, videos of locomotion of different horse breeds were recorded. The camera was set up on the image plane parallel to the track where the horse was moving, with the right side of the horse facing the camera. The various horse breeds involved in the experiment included Quarter horses, Warmbloods, Thoroughbreds, a Percheron cross, Tennessee Walking horse, Kentucky Mountain Saddle horses, and a Haflinger. A total of 12 videos were collected by the University of Florida Equine Genetics lab in 2016. Each video records a horse moving from the left to right of the camera frame. These videos were recorded outdoors and processed for experimental results and analysis. The resolution of each video was 1280×720 and the frame rate 120 frames per second (FPS).

2.3.2 Data Preparation

Detectron2 model accepts annotated input data in COCO format. Thus, the data passed to the model should comply with the COCO annotation format. The CVAT (Pangal et al., 2021) tool was used to annotate the data for all 12 videos. The data labels consisted of the 21 different body parts of the horse (Figure 1) and a segmentation mask for the entire horse.

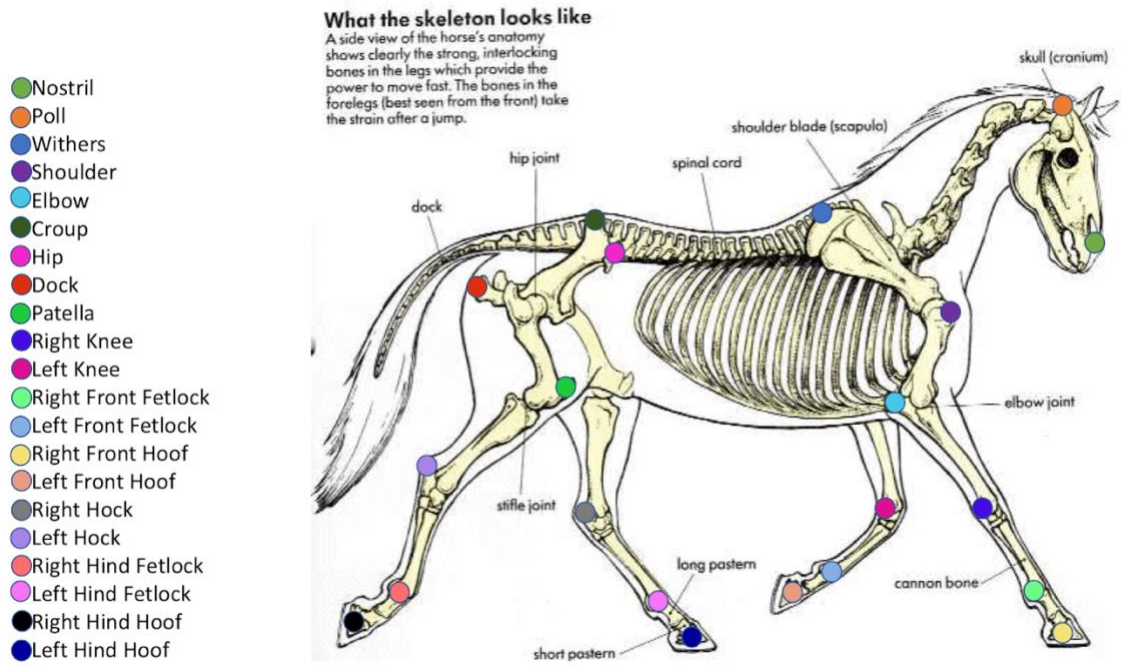

Figure 1: Illustration of 21 equine body parts

The CVAT tool allows exporting annotations in multiple formats. In the CVAT format, the annotations are in XML file and follow the CVAT formatting style which is different from the COCO format. This COCO format annotations are of JSON file type, but the format varies from the format used in Detectron2.

A template was created of the COCO format, JSON file type which was in accordance with the Detectron2 dataset type. A python script combined the segmentation mask from the CVAT extracted COCO file and the keypoint annotated values from the CVAT format XML file into our template file. Thus, a COCO annotation file was created which had the information about the keypoints and segmentation mask and complied to the Detectron2 syntax.

In order to have a fair comparison of Mask R-CNN based model performance to that of DeepLabCut, the training and test set used in both models should be identical. Thus, the same set of labeled images from each video were used for training and testing DeepLabCut as well as a

Mask R-CNN model. The validation set consisted of 10 images selected at random from the videos, exclusive of the frames used in the training set.

2.3.3 Ground truth data

The CVAT annotated data served as the ground truth data. This data consisted of 12 videos with each of its frame distinctly having annotations for all the 21 keypoints, the target body parts. Each keypoint had x and y coordinate values.

2.3.4 Evaluation Metrics

The 21 keypoints were predicted by both the Mask R-CNN model and the DeepLabCut model, respectively. The experiment used the root mean squared error (RMSE), the linear regression coefficients and coefficient of determination as the evaluation parameters. The metrics was calculated for the ground truth versus Mask R-CNN model predicted data and ground truth versus DeepLabCut model predicted data. The data consisted of x and y coordinates of every keypoint across the test dataset frames for all videos.

2.3.5 Experiments

The dataset used in our Mask R-CNN model was consistent with that used in DeepLabCut body part tracking. We used 12 videos and processed them to achieve body part tracking of 21 keypoints realized by the Mask R-CNN model as well as the DeepLabCut model. The training dataset consisted of 392 images randomly selected from all the videos and the validation dataset 10 random images from each video exclusive of the training images. For each body part, the Euclidean distance was obtained between the ground truth data and each model

output. This was obtained for 21 keypoints in every frame of each of the 12 videos. For the performance evaluation of both the models, we considered the final predicted output of 21 keypoints for testing dataset images. These images consisted of all the image frames excluding those in training and validation datasets. Table 1 shows the training, validation and test split of each video in the dataset. The variations in the numbers of training images per video were a result of using the DLC framework. Their linear regression equation and coefficient of determination were calculated. The RMSE values were calculated for the overall Euclidean distance between the ground truth and the predicted keypoints as well as for the x and y coordinates separately.

Table 1: Training, Validation and Test split of each video in the dataset

Video Number	Number of Training Images	Number of Validation Images	Number of Testing Images
1	33	10	281
2	39	10	287
3	33	10	235
4	42	10	229
5	38	10	195
6	48	10	206
7	27	10	225
8	37	10	311
9	35	10	371
10	20	10	441
11	1	10	450
12	38	10	333

2.3.5.1 Detectron2 Mask R-CNN model

Detectron2 was utilized to assess body part tracking of the horses with Mask R-CNN and induce from the outcomes, a comparison between the output of DeepLabCut following to that of Mask R-CNN. A pre-trained model zoo of Detectron2 was used and fine-tuned by customizing the hyperparameters suitable for our dataset. The parameters that can be customized are the

hyperparameters. They may differ based on the variation and requirement of the dataset. Thus, we fine-tuned a pretrained model zoo in detectron2 based on our dataset. For the purpose of fine tuning, the Detectron2 default trainer was used. The learning rate was set to 0.025. The ROIs batch size per image was set to 128. The maximum number of iterations were set at 400. The model weights configuration was model weights keypoint_rcnn_R_50_FPN_3x.

2.3.5.2 DeepLabCut Model

DeepLabCut was used to track 21 horse body parts as labeled in the Figure 2. The DeepLabCut training and body part tracking for all the 12 videos were performed by equine researchers at UFL. For each video, this software tracked each body feature of a horse and creates a digital trajectory of that part. The coordinates, in terms of pixels of image frame, of that feature per frame were recorded in a CSV format file. The CSV file along with the training dataset used to train a DeepLabCut model, were provided by the UFL team. This CSV file was used as input for the subsequent processing.

Figure 2 : A DeepLabCut output image horse with 21 labeled body parts and human feet.

The DeepLabCut model was trained by manually labelling the training dataset frames with the 21 keypoints. The model was run on each video to obtain the equine body part tracking for that video through DeepLabCut.

2.4 Results and Discussion

Figures 3 and 4 show qualitative results of the Mask R-CNN model predictions. The bounding box shows a confidence value of 100%. This means that the horse is detected with a 100% confidence by the model. In Figure 3, most of the keypoints were correctly detected. In Figure 4, there is an error in the detection of Left Hind Fetlock.

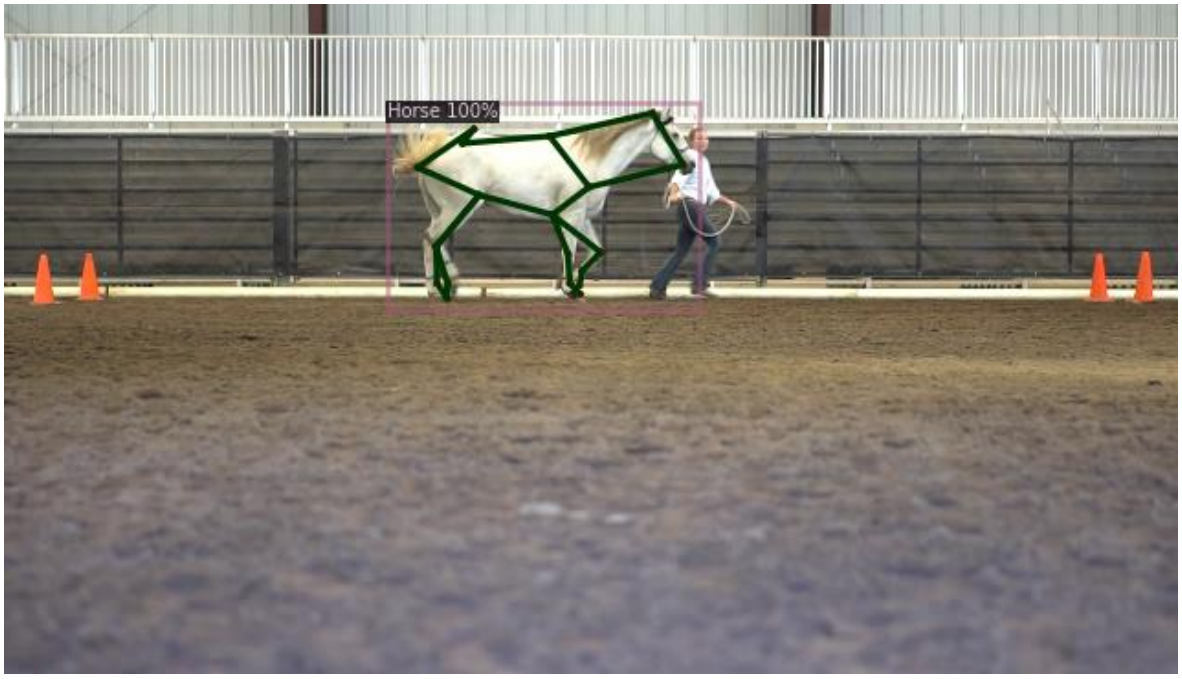

Figure 3: Mask R-CNN body part tracking output 1.

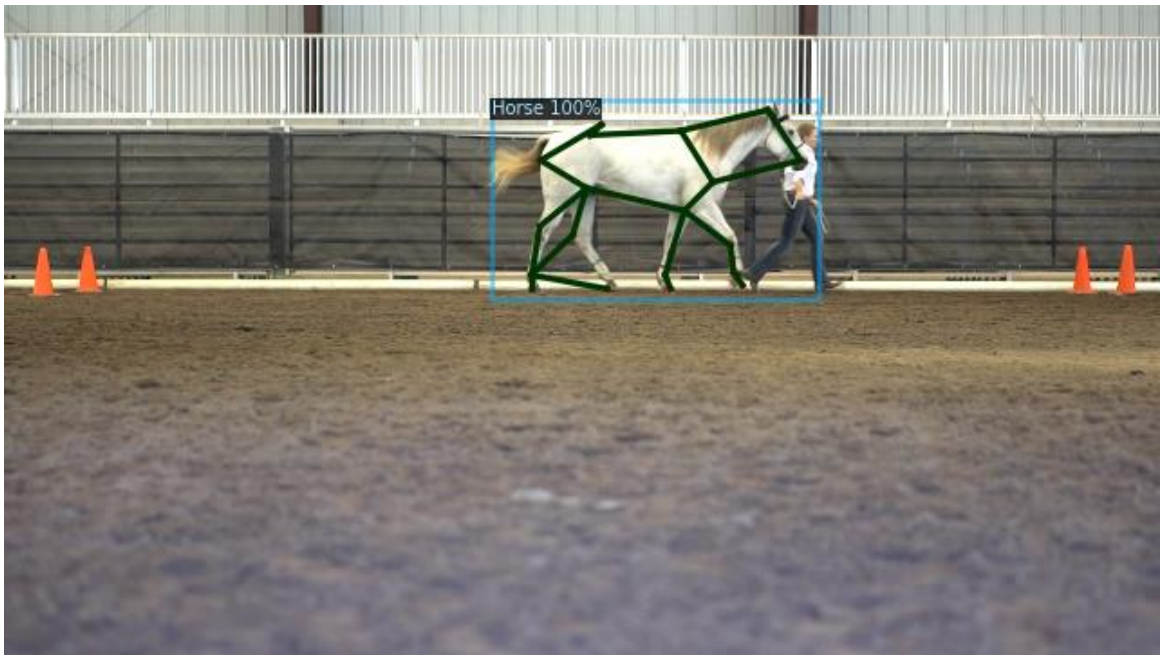

Figure 4: Mask R-CNN body part tracking output 2.

The overall RSME for the Euclidean distance between the ground truth and predicted values was high for Mask R-CNN for a video is 20.86. Figure 5 shows the Mask R-CNN vs Ground Truth plot of x and y coordinates respectively. The coefficient of determination is greater than 0.99 which means that most of the data was explained by the model. For this video, the Mask R-CNN was able to predict the y coordinates of the keypoints with an RMSE of 3.73 which is lesser than the RMSE of x coordinate, 20.52.

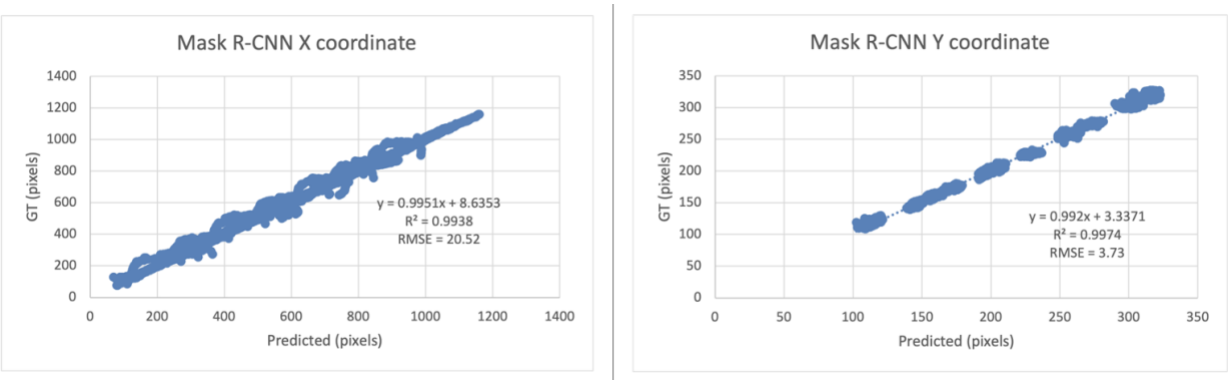

Figure 5: Linear regression, coefficient of determination, RMSE, Mask R-CNN versus Ground Truth of a video

The Figure 6 shows the output of DeepLabCut on an image in one of the videos. It has all 23 keypoints, including the ones on the operator's feet, represented by different colors.

Figure 6: The DeepLabCut labelled output image. Different colors represent different keypoints on the horse's body and the handler's feet.

The RSME for the Euclidean distance between the ground truth and predicted values for DeepLabCut for a video was 145.36 which was very high.

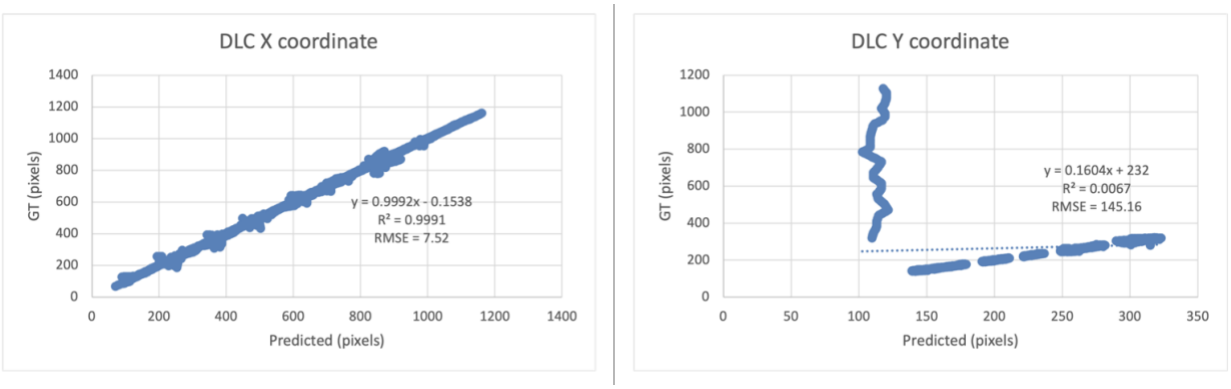

Figure 7: Linear regression, coefficient of determination (R^2), and RMSE of DeepLabCut (DLC) vs Ground truth (GT) of a single video

The overall RMSE for Mask R-CNN model for all the videos is 100.7. The overall RMSE for DeepLabCut model for all the videos is 138.9. Table 1 displays the RMSE for every coordinate for each model. Table 2 displays the overall RMSE of the Euclidean distance of each keypoint from the ground truth value to the predicted value for each model. The RMSE value for the Euclidean distance is less for the Mask R-CNN model as compared to that of the DeepLabCut model. However, considering the x coordinate alone, the RMSE values are smaller for DeepLabCut model for almost all the videos.

Figure 8: Incorrectly detected keypoints by DeepLabCut for Left Hock and Left Hind Hoof

Figure 7 for DeepLabCut realized body part tracking of a video suggests some extreme outliers in the DeepLabCut vs GT plot for the y coordinates of that video. In this case it failed to locate and track the data points in some cases and hence, there are outliers in the y-coordinate plot of DeepLabCut as seen in Figure 7. In some cases, the failure was caused by two overlapping front or rear legs from the camera viewing angle. This case is illustrated in Figure 8, where the left image shows correctly labelled parts. In the image on the right side of Figure 8, the points colored yellow(Left Hock) and blue (Left Hind Hoof) are incorrectly identified. It also happens when there is occlusion due to another object in the track. It is worth noting that in scenarios where there is no noise due to occlusion, DeepLabCut may provide better results compared to Mask R-CNN model.

Table 2: RMSE values for x and y coordinates from Mask R-CNN and DeepLabCut model outputs

Video No.	Mask R-CNN RMSE	DeepLabCut RMSE

	X coordinate	Y coordinate	X coordinate	Y coordinate
1	150.92	26.09	19.43	76.58
2	136.36	44.69	47.65	154.2
3	20.52	3.73	7.52	145.16
4	80.85	36.56	5.73	133.06
5	129.17	75.68	53.93	144.43
6	133.5	72	10.15	142.96
7	81.09	37.9	12.81	57.77
8	98.71	48.40	6.03	95.85
9	63.82	15.19	4.69	101.74
10	100.01	48.19	5.56	124.33
11	66.34	57.47	63.89	125.09
12	34.05	14.62	36.37	93.27

Table 3: RMSE value for Euclidean distance of keypoints from Mask R-CNN and DeepLabCut model outputs

Video Number	Mask R-CNN RMSE	DeepLabCut RMSE
1	153.16	230.37
2	143.49	161.39
3	20.86	145.36
4	88.73	133.178
5	149.71	154.16
6	151.68	143.32
7	89.51	135.93

8	109.94	96.04
9	65.60	101.85
10	111.01	124.45
11	87.77	140.47
12	37.06	100.11

2.5 Conclusions

The coefficient of determination is greater than 0.9 for the x and y coordinates, respectively, for Mask R-CNN body part tracking. But the RMSE values of the Mask R-CNN model were much higher for x coordinates than y coordinates. Based on our results, it can be clearly interpreted that the Mask R-CNN model could be utilized for body part tracking, in the case of horses. It is important to note that our data is completely based on horses and the configurations and weights used in the pre-trained detectron2 model that we have used in our experiment, had a number of classes and were not oriented towards animals or horses alone. Better results could be expected if the model were trained from scratch with a horse dataset which is beyond our scope.

When compared to Mask R-CNN model, DeepLabCut has more acceptable tracking results for x coordinates but does not perform very well in tracking the y coordinates. The extremely high overall RMSE values for DeepLabCut-based body part tracking are a result of the outliers present in the data. It can be concluded that Mask R-CNN has overall better performance when the data is noisy. However, if the data is smooth, DeepLabCut outperforms Mask R-CNN in body part tracking. To add further, based on our experimental results, if valuable information lies in x coordinate values, the DeepLabCut model should be used over the Mask R-CNN model.

Chapter 3

Biomechanical Parameters Extraction

In this Chapter, a fully automated pipeline was developed that calculates the equine biomechanical parameters based on DeepLabCut-based body part tracking. From the previous chapter, it could be inferred that although Mask R-CNN had overall better performance, DeepLabCut gave better results for the x coordinates alone. For stride length and stance time extraction, the proposed pipeline utilizes x coordinates results. Thus, the DeepLabCut tracking results were used for our pipeline. There are multiple biomechanical parameters that can be extracted from horse body part tracking like stride length, stance time, flexion, and maximum forward extension. However, this study focuses on extracting stride length and stance time as they have multiple gait analysis applications like lameness detection (Peham et al., 2001), training impact analysis (Rose et al., 2009), and behavior assessment (Ashley et al., 2005).

3.1 Ground Truth Data

The ground truth values of stride length and stance time were obtained in each video. First, the x coordinates of the center of each hoof for each stance were manually measured to calculate the stride length in pixels. Second, the frame numbers at the beginning and end of each stance were distinguished to identify the stance duration as the number of frames. The head length was measured as the Euclidean distance between the poll and the nostril of the horse at the start and end of each stride. An average value of this distance served as a normalization factor to compensate for unknown parameters which include the distance of between the horse and the camera and the size of the horse. The stride length was then normalized by dividing the

normalization factor. In a similar fashion, the stance duration in number of frames was divided by the frame rate, in frames per second, to obtain stance time in seconds.

3.2 Post-Processing of Hoof Trajectories

Figure 9: Hoof trajectory from the beginning to the end of video in an image

3.2.1 Trajectory Smoothing

The trajectory of a horse hoof generated by DeepLabCut contained noise and errors as seen in Figure 9. A Kalman filter (Kalman, 1960) was used to remove noise and smooth the data. This filter consists of a 2-step process of prediction and estimation. In the first step it produces an estimate of the next state based on current state variables and in the second step, it uses a weighted average to re-estimate the next state. The pykalman python library (Metz et al., 2017) was used with a pre-defined transition matrix, a pre-defined observation matrix and an initial state mean as parameter to the Kalman filter. We gave our x-y coordinates to the filter as initial measurement vector, z.

$$z = [x,y]$$

where

x = array of x coordinates of a hoof

y = array of y coordinated of a hoof

We used an observation matrix:

$$\boldsymbol{z} = \boldsymbol{H}o = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} o$$

where $\boldsymbol{H}$ represents the sensor matrix and o is the new best estimate matrix (mean)

and a dynamics or weighted average matrix:

$$\boldsymbol{o}(k) = \boldsymbol{F}\boldsymbol{o}(k-1) => \boldsymbol{F} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where $\boldsymbol{F}$ represents the transition matrix and k is the current state, and initial state mean:

$$initial\_state\_mean = [x, 0, y, 0]$$

When the x and y coordinate values are passed to the filter, it returns the smoothed values. To smooth minutely crooked regions as observed in Figure 10, these values are further passed to another Kalman Filter where the observation coefficients are 10 times those of the initial filter. After the second Kalman filter, the peak and valley in the hoof trajectory during each stride can be clearly identified as shown in Figure 11.

Figure 10: Hoof trajectory smoothed once from the beginning to the end of video in an image

Figure 11: Hoof trajectory smoothed twice from the beginning to the end of video in an image

3.2.2 Stride Length and Stance Time Extraction

For the required parameters, stride length and stance time, only the x coordinates of the hoof is used due to the relatively small change in the y coordinates. Based on the change in gradient of the x coordinate trajectory, the x values were segmented into plateaus and slopes. This change is calculated by calculating the slope at every point and taking its second derivative. dy = xs(i) – xs(i-1) is the slope and dpy = dy(i) – dy(i-1) is the second derivative where xs is the array of smoothed x coordinate values and i is the current frame.

Based on the value of change in gradient, the smoothed x values are segmented into two groups. One group where the slope is negligible, in the range -2 to 2, is identified as the points where the hoof is still. The other group where the slope is beyond the above-mentioned range corresponds to the points where the hoof is in motion. To retain the frame-coordinate relation, the frames where the values are not present, are filled with zeroes. Plateaus are the regions where the x values are constant relative to video frame ID. They are the stances when the hoof is in contact with the ground. On the contrary, the slopes represent the hoof in motion during a stride.

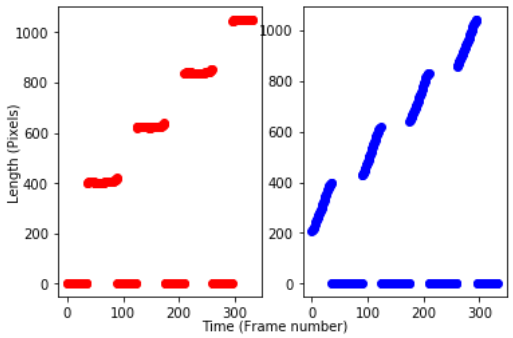

Figure 12: (Left) stances and (right) strides as separated list of points. The zeroes are fillers for maintaining frame-coordinate relation

Each slope in the trajectory corresponds to a stride. To make the pipeline more robust, incomplete slopes were eliminated. Due to occlusions or incorrect detection in the video, if a slope was split into two or more parts, the occluded part was removed and the fragments were combined into one slope. A threshold of 6 frames was used to measure the gap between two adjacent slopes. If the difference between two slopes is less than the threshold, the gap is filled, and the slopes are merged. Incomplete strides might occur in the beginning and at the end of a video. To remove these incomplete strides, the strides that do not fall in the range of median ± 50 of all strides, were removed from the list of strides. Lastly, the stride length in pixels were normalized by head length to minimize the effect of varying camera-to-horse distances and horse size.

Each plateau represents a stance. Similar to swing phase, the gaps between stances are filled and the incomplete stances are removed. The length of a continuous plateau is referred to as a stance in terms of number of frames of the video. All the values are normalized by dividing the stance time by the frame rate (120 FPS).

3.2.3 Evaluation

The strides and stances computed for each video were compared to the corresponding ground truth values. The stride length and stance time were evaluated for right front hoof, left front hoof, right hind hoof, and left hind hoof separately. Methods and metrics used to evaluate the results include linear regression, coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute percentage error (MAPE).

3.3 Results and Discussion

The proposed processing pipeline correctly identified 92% of the total 168 strides and 95% of the total 180 stances in the 12 videos. The estimated stride length and stance time were found strongly correlated with the ground truth for individual hoofs. The coefficient of determination ranges from 0.64 to 0.88 for stride length and from 0.75 to 0.86 for stance time. Figures 13 and 14 show the linear regression analysis and error metrics for the stride length and stance time estimations of individual hoofs, respectively. Because the normalized stride length is unitless, interpretation of the numeric values may be difficult. However, the distributions in Figure 13 suggests that large variations in stride length existed in the 12 videos.

It is worth noting that the stride length estimations are prone to error due to occlusions. Since most of the videos cover trotting gaits, the diagonally opposite hoofs tend to show similar strides and stances. The right front hoof and the left hind hoofs share similar behavior, and the behavior of left front hoof is closely related to that of right hind hoofs. It is only a matter of coincidence that the right hind hoof trajectory in these 12 videos was least affected by occlusions and thus, right hind hoof stride and stance values have overall least Mean Absolute Percentage Errors (MAPE). It was expected that the left-side hoofs would be more affected by occlusion

caused by the right-side hoofs (all horses moves from left to right in the videos). But the pipeline was able to predict both the left and right hoof strides and stances with satisfactory coefficients of determination.

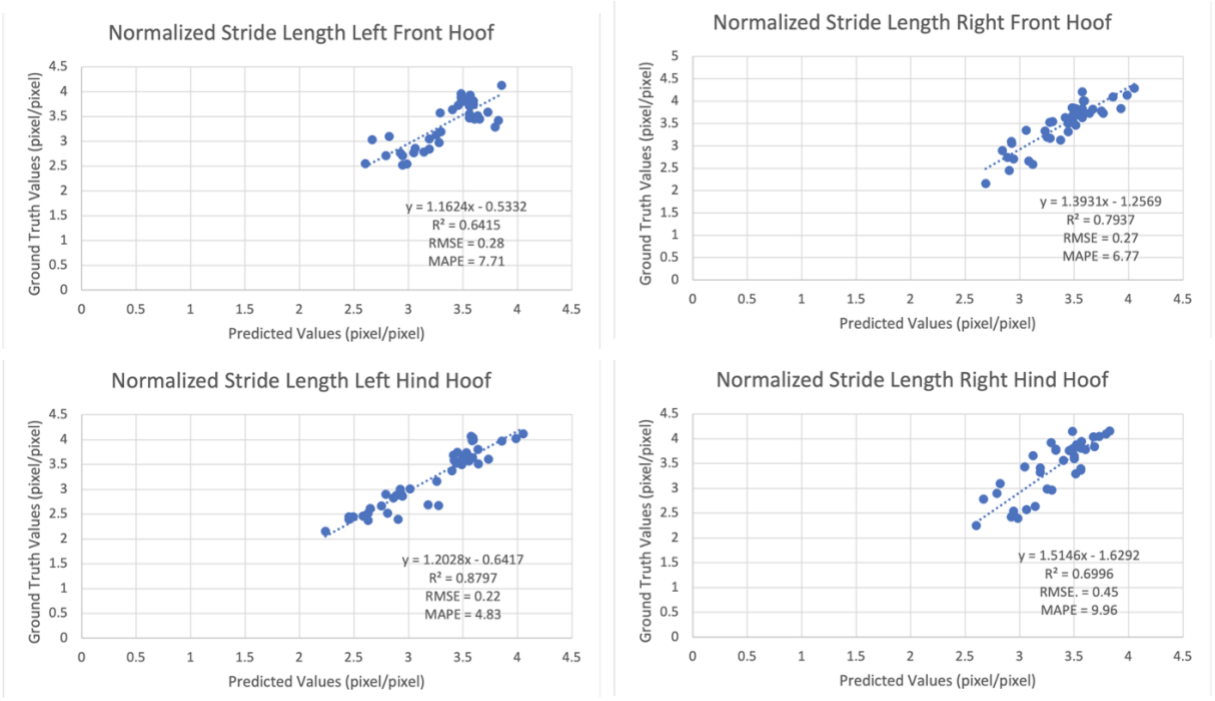

Figure 13: Normalized stride lengths of each hoof with the equation of regression line, coefficient of determination (R^2), root mean square error (RMSE)(pixel/pixel) and mean absolute percentage error (MAPE)(pixel/pixel)

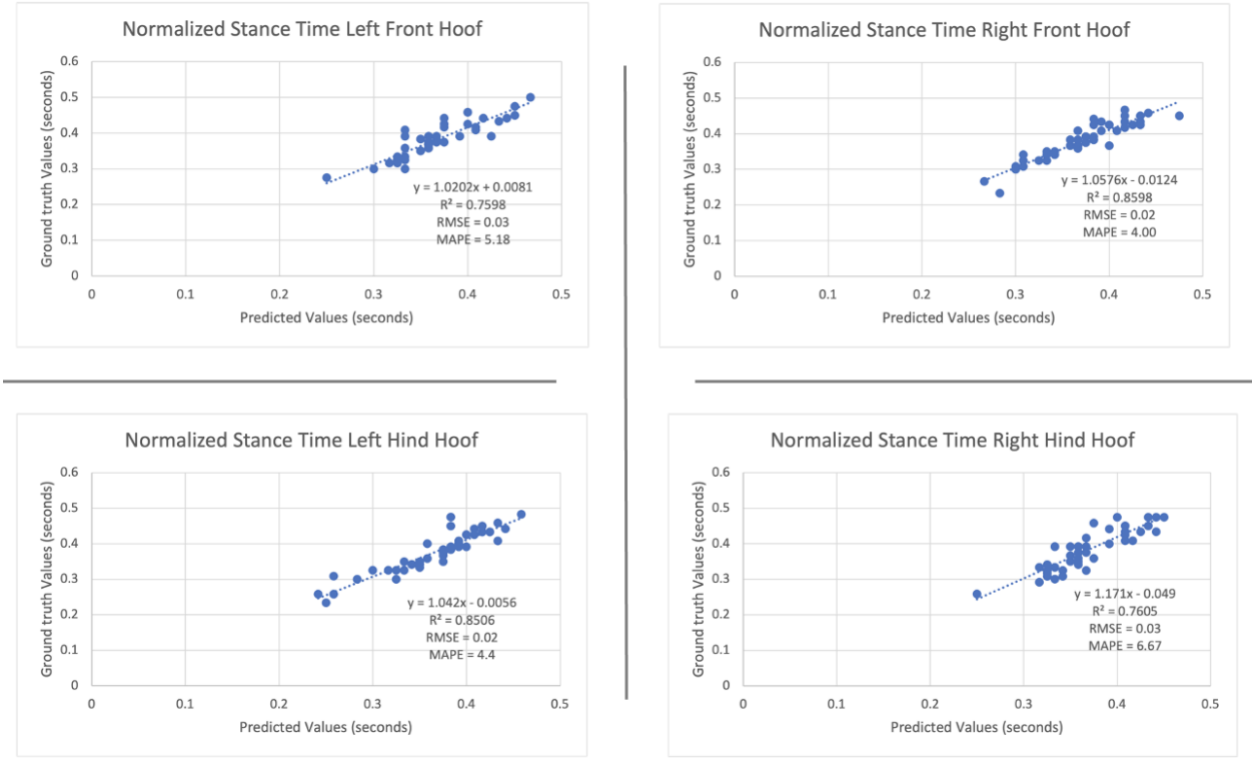

Figure 14: Normalized stance time of each hoof with the equation of regression line, coefficient of determination (R^2), root mean square error (RMSE) and mean absolute percentage error (MAPE)

Figure 15 shows the linear regression results of overall stance time and stride length estimations. The coefficient of determination, root mean squared error and mean absolute percentage error are comparable to those of the individual hoofs. Therefore, our biomechanical parameter extraction pipeline can accurately measure stride length and stance time, at least in the 2D image space.

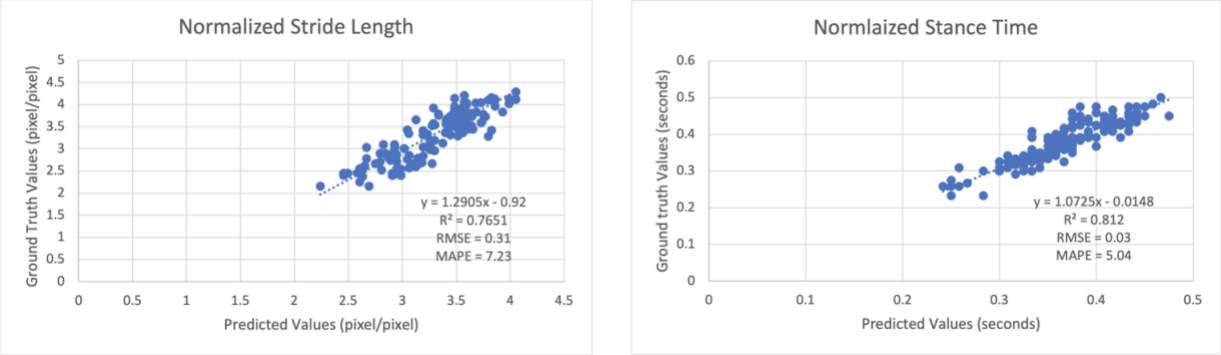

Figure 15: Normalized Stride length and stance time of all hoofs with the equation of regression line, coefficient of determination (R^2), root mean square error (RMSE) and mean absolute percentage error (MAPE)

3.4 Conclusions

The research in this paper was oriented to tracking the horse body parts and extract their biomechanical traits. We used DeepLabCut to process the horse locomotion videos and track horse body parts. We developed a pipeline to interpret the DeepLabCut data and process it to output useable information on the biomechanical parameters of the horses.

This study primarily focused on the horse hoof body part, and the stride length and stance time biomechanical parameters. There are other traits such as flexion angle, withers range, poll range, forward head extension and so on. These traits can also be obtained by tracking withers, poll, fetlock angle, nostril and other equine body parts. The most significant capability of horses lies in their locomotion. The stride lengths and other information that comes from hoof tracking could be used to assess horse locomotion. Regular monitoring of horse gait patterns can help detect diseases like lameness at an early stage. These trends may also assist in improving horse performances in sports activities and races. Horse gait analysis, thus, have numerous clinical, scientific and physical performance-based applications.

References List

Ashley, F. H., Waterman-Pearson, A. E., & Whay, H. R. (2005). Behavioural assessment of pain in horses and donkeys: Application to clinical practice and future studies. In *Equine Veterinary Journal* (Vol. 37, Issue 6, pp. 565–575). https://doi.org/10.2746/042516405775314826

Aviad, A. D. (1988). The use of the standing force plate as a quantitative measure of equine lameness. *Journal of Equine Veterinary Science*, *8*(6), 460–462. https://doi.org/10.1016/S0737-0806(88)80095-9

Bächi, B., Wiestner, T., Stoll, A., Waldern, N. M., Imboden, I., & Weishaupt, M. A. (2018). Changes of Ground Reaction Force and Timing Variables in the Course of Habituation of Horses to the Treadmill. *Journal of Equine Veterinary Science*, *63*, 13–23. https://doi.org/10.1016/j.jevs.2017.12.013

Barrey, E. (1999). Methods, applications and limitations of Gait analysis in horses. *Veterinary Journal*, *157*(1), 7–22. https://doi.org/10.1053/tvjl.1998.0297

Barrey, E., Auvlnett, B., & Couroucei, A. (1995). Gait evaluation of race trotters using an accelerometric device. In *EQUINE VETERlNARY JOURNAL Equine vet. J* (Vol. 18).

Chua, L. O., & Roska, T. (1993). The CNN Paradigm. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, *40*(3), 147–156. https://doi.org/10.1109/81.222795

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. T. (2018). A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. In *Sports Medicine - Open* (Vol. 4, Issue 1). https://doi.org/10.1186/s40798-018-0139-y

DYSON, P. K., JACKSON, B. F., PFEIFFER, D. U., & PRICE, J. S. (2008). Days lost from training by two- and three-year-old Thoroughbred horses: A survey of seven UK training yards. *Equine Veterinary Journal*, *40*(7). https://doi.org/10.2746/042516408X363242

Echterhoff, J. M., Haladjian, J., & Brugge, B. (2018). Gait and jump classification in modern equestrian sports. *Proceedings - International Symposium on Wearable Computers, ISWC*, 88–91. https://doi.org/10.1145/3267242.3267267

Echterhoff, J. M., Haladjian, J., & Brügge, B. (2018, December 4). Gait analysis in horse sports. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3295598.3295601

Egenvall, A., Tranquille, C. A., Lönnell, A. C., Bitschnau, C., Oomen, A., Hernlund, E., Montavon, S., Franko, M. A., Murray, R. C., Weishaupt, M. A., Weeren, van R., & Roepstorff, L. (2013). Days-lost to training and competition in relation to workload in 263 elite show-jumping horses in four European countries. *Preventive Veterinary Medicine*, *112*(3–4). https://doi.org/10.1016/j.prevetmed.2013.09.013

Egenvall, Agneta, Bonnett, B., Wattle, O., & Emanuelson, U. (2008). Veterinary-care events and costs over a 5-year follow-up period for warmblooded riding horses with or without previously recorded locomotor problems in Sweden. *Preventive Veterinary Medicine*, *83*(2). https://doi.org/10.1016/j.prevetmed.2007.06.008

Fang, H. S., Lu, G., Fang, X., Xie, J., Tai, Y. W., & Lu, C. (2018). Weakly and Semi Supervised Human Body Part Parsing via Pose-Guided Knowledge Transfer. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 70–78. https://doi.org/10.1109/CVPR.2018.00015

Girshick, R. (2015). Fast R-CNN. *Proceedings of the IEEE International Conference on*

Computer Vision, *2015 Inter*, 1440–1448. https://doi.org/10.1109/ICCV.2015.169

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2016). Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(1), 142–158. https://doi.org/10.1109/TPAMI.2015.2437384

Graving, J. M., Chae, D., Naik, H., Li, L., Koger, B., Costelloe, B. R., & Couzin, I. D. (2019). DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *ELife*, *8*, e47994. https://doi.org/10.7554/eLife.47994

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (n.d.). *Mask R-CNN*.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Insafutdinov, E., Pishchulin, L., Andres, B., Andriluka, M., & Schiele, B. (2016). Deepercut: A deeper, stronger, and faster multi-person pose estimation model. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 9910 LNCS* (pp. 34–50). Springer International Publishing. https://doi.org/10.1007/978-3-319-46466-4_3

JEFFCOTT, L. B., ROSSDALE, P. D., FREESTONE, J., FRANK, C. J., & TOWERS-CLARK, P. F. (1982). An assessment of wastage in Thoroughbred racing from conception to 4 years of age. *Equine Veterinary Journal*, *14*(3). https://doi.org/10.1111/j.2042-3306.1982.tb02389.x

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of*

Fluids Engineering, Transactions of the ASME, *82*(1), 35–45. https://doi.org/10.1115/1.3662552

Kil, N., Ertelt, K., & Auer, U. (2020). Development and validation of an automated video tracking model for stabled horses. *Animals*, *10*(12), 1–12. https://doi.org/10.3390/ani10122258

Labuguen, R., Bardeloza, D. K., Negrete, S. B., Matsumoto, J., Inoue, K., & Shibata, T. (2019). Primate markerless pose estimation and movement analysis using deeplabcut. *2019 Joint 8th International Conference on Informatics, Electronics and Vision, ICIEV 2019 and 3rd International Conference on Imaging, Vision and Pattern Recognition, IcIVPR 2019 with International Conference on Activity and Behavior Computing, ABC 2019*, 297–300. https://doi.org/10.1109/ICIEV.2019.8858533

Marshall, J. F., Lund, D. G., & Voute, L. C. (2012). *Use of a wireless, inertial sensor-based system to objectively evaluate flexion tests in the horse*. https://doi.org/10.1111/j.2042-3306.2012.00611.x

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289. https://doi.org/10.1038/s41593-018-0209-y

MERKENS, H. W., & SCHAMHARDT, H. C. (1988). Evaluation of equine locomotion during different degrees of experimentally induced lameness II: Distribution of ground reaction force patterns of the concurrently loaded limbs. *Equine Veterinary Journal*, *20*, 107–112. https://doi.org/10.1111/j.2042-3306.1988.tb04656.x

Metz, J., Castro, I., & Schrader, M. (2017). Peroxisome Motility Measurement and

Quantification Assay. *BIO-PROTOCOL*, *7*(17). https://doi.org/10.21769/bioprotoc.2536

Moorman, V. J., Frisbie, D. D., Kawcak, C. E., & McIlwraith, C. W. (2017). Effects of sensor position on kinematic data obtained with an inertial sensor system during gait analysis of trotting horses. *Journal of the American Veterinary Medical Association*, *250*(5), 548–553. https://doi.org/10.2460/javma.250.5.548

Murray, R. C., Dyson, S. J., Tranquille, C., & Adams, V. (2006). Association of type of sport and performance level with anatomical site of orthopaedic injury diagnosis. *Equine Veterinary Journal*, *38*(SUPPL.36), 411–416. https://doi.org/10.1111/j.2042-3306.2006.tb05578.x

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature Protocols*, *14*(7), 2152–2176. https://doi.org/10.1038/s41596-019-0176-0

Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(25), E5716–E5725. https://doi.org/10.1073/pnas.1719367115

Pangal, D. J., Kugener, G., Shahrestani, S., Attenello, F., Zada, G., & Donoho, D. A. (2021). A Guide to Annotation of Neurosurgical Intraoperative Video for Machine Learning Analysis and Computer Vision. *World Neurosurgery*, *150*, 26–30. https://doi.org/10.1016/j.wneu.2021.03.022

Peham, C., Licka, T., Girtler, D., & Scheidl, M. (2001). The Influence of Lameness on Equine Stride Length Consistency. *Veterinary Journal*, *162*(2), 153–157.

https://doi.org/10.1053/tvjl.2001.0593

Ren, S., He, K., Girshick, R. B., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*, 1137–1149.

Rose, N. S., Northrop, A. J., Brigden, C. V., & Martin, J. H. (2009). Effects of a stretching regime on stride length and range of motion in equine trot. *The Veterinary Journal*, *181*(1), 53–55. https://doi.org/10.1016/J.TVJL.2009.03.010

Serra Bragança, F. M., Broomé, S., Rhodin, M., Björnsdóttir, S., Gunnarsson, V., Voskamp, J. P., Persson-Sjodin, E., Back, W., Lindgren, G., Novoa-Bravo, M., Roepstorff, C., van der Zwaag, B. J., Van Weeren, P. R., & Hernlund, E. (2020). Improving gait classification in horses by using inertial measurement unit (IMU) generated data and machine learning. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-73215-9

Serra Bragança, F. M., Rhodin, M., & van Weeren, P. R. (2018). On the brink of daily clinical application of objective gait analysis: What evidence do we have so far from studies using an induced lameness model? In *Veterinary Journal* (Vol. 234, pp. 11–23). W.B. Saunders. https://doi.org/10.1016/j.tvjl.2018.01.006

Shrestha, A., Kernec, J. Le, Fioranelli, F., Marshall, J. F., & Voute, L. (2017). Gait analysis of horses for lameness detection with radar sensors. *International Conference on Radar Systems (Radar 2017)*, 1–6. https://doi.org/10.1049/cp.2017.0427

Toshev, A., & Szegedy, C. (2014). DeepPose: Human pose estimation via deep neural networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1653–1660. https://doi.org/10.1109/CVPR.2014.214

Wang, Y., Li, J., Zhang, Y., & Sinnott, R. O. (2021, March 22). Identifying lameness in horses

through deep learning. *Proceedings of the 36th Annual ACM Symposium on Applied Computing*. https://doi.org/10.1145/3412841.3441973

Weishaupt, M. A., Hogg, H. P., Wiestner, T., Denoth, J., Stüssi, E., & Auer, J. A. (2002). Instrumented treadmill for measuring vertical ground reaction forces in horses. *American Journal of Veterinary Research*, *63*(4), 520–527. https://doi.org/10.2460/ajvr.2002.63.520

Wen, H., Huang, C., & Guo, S. (2021). The application of convolutional neural networks (CNNs) to recognize defects in 3D-printed parts. *Materials*, *14*(10). https://doi.org/10.3390/ma14102575

Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., & Girshick, R. (2019). *Detectron2*. https://github.com/facebookresearch/detectron2. https://github.com/facebookresearch/detectron2

YUNUS, A. P., SHIRAI, N. C., MORITA, K., & WAKABAYASHI, T. (2020). Time Series Human Motion Prediction Using RGB Camera and OpenPose. *International Symposium on Affective Science and Engineering*, *ISASE2020*, 1–4. https://doi.org/10.5057/isase.2020-C000037

Appendix List

Appendix 1

Mask R-CNN model, detectron2 configuration of hyperparameters

```
from detectron2.engine import DefaultTrainer
from detectron2.config import get_cfg
import os

cfg = get_cfg()
cfg.merge_from_file(model_zoo.get_config_file("COCO-
Keypoints/keypoint_rcnn_R_50_FPN_3x.yaml"))
cfg.DATASETS.TRAIN = ("data_horse101", "data_horse102", "data_horse103",
"data_horse104", "data_horse105", "data_horse106", "data_horse107",
"data_horse108","data_horse109","data_horse10210","data_horse10211","data_
horse10212",)
cfg.DATASETS.TEST = ("data_horse1011", "data_horse1021", "data_horse1031",
"data_horse1041", "data_horse1051", "data_horse1061", "data_horse1071",
"data_horse1081","data_horse1091", "data_horse102101", "data_horse102111",
"data_horse102121",)  # no metrics implemented for this dataset
"data_horse1011",
cfg.DATALOADER.NUM_WORKERS = 2 #Default 2
cfg.MODEL.WEIGHTS = model_zoo.get_checkpoint_url("COCO-
Keypoints/keypoint_rcnn_R_50_FPN_3x.yaml")

cfg.SOLVER.IMS_PER_BATCH = 1 #Default 2
cfg.SOLVER.BASE_LR = 0.025 #Default 0.2
# cfg.SOLVER.STEPS = []        # do not decay learning rate
cfg.SOLVER.MAX_ITER = (
    400
)  # 400 iterations seems good enough
cfg.MODEL.ROI_HEADS.BATCH_SIZE_PER_IMAGE = (
    128
)
cfg.MODEL.ROI_HEADS.NUM_CLASSES = 1  # 1 classes (horse)
cfg.MODEL.RETINANET.NUM_CLASSES = 1
cfg.MODEL.ROI_KEYPOINT_HEAD.NUM_KEYPOINTS = 21
os.makedirs(cfg.OUTPUT_DIR, exist_ok=True)
trainer = DefaultTrainer(cfg)
trainer.resume_or_load(resume=False)
torch.cuda.empty_cache()
trainer.train()
```

Appendix 2

DeepLabCut configurations:

```
   # Project definitions (do not edit)
Task: HTU_AW
scorer: Ariana West
date: Mar19
multianimalproject: false

   # Project path (change when moving around)
project_path: C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19

   # Annotation data set configuration (and individual video cropping parameters)
video_sets:
  C:\DeepLabCut_Trainings_Master_Folder\AW_HTU\C0009.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\AW_HTU\C0010.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\AW_HTU\C0018.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\AW_HTU\C0021.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\AW_HTU\C0025.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\AW_HTU\C0028.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0029.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0375.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0371.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0370.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0368.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0363.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0021.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0034.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0028.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0010.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0025.mp4:
```

```
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0018.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU\C0009.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU - 2\C0021.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU - 2\C0028.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU - 2\C0034.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU - 2\C0363.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU - 2\C0368.mp4:
    crop: 0, 1920, 0, 1080
  C:\DeepLabCut_Trainings_Master_Folder\HTU_AW-Ariana West-2021-03-19\AW_HTU - 2\C0375.mp4:
    crop: 0, 1920, 0, 1080
bodyparts:
- rightFhoof
- rightHhoof
- righthock
- rightFfetlock
- rightHfetlock
- rightknee
- handlerRfoot
- nostril
- poll
- withers
- shoulder
- elbow
- croup
- hip
- stifle
- pointofbuttock
- leftFhoof
- leftHhoof
- lefthock
- leftFfetlock
- leftHfetlock
- leftknee
- handlerLfoot
start: 0
stop: 1
numframes2pick: 20

    # Plotting configuration
skeleton:
- - nostril
```

- poll
- - poll
- withers
- - withers
- shoulder
- - shoulder
- elbow
- - withers
- croup
- - croup
- hip
- - croup
- dock
- - hip
- stifle
- - hip
- pointofbuttock
- - pointofbuttock
- croup
- - pointofbuttock
- stifle
- - hip
- dock
- - dock
- stifle
- - stifle
- righthock
- - righthock
- rightHfetlock
- - rightHfetlock
- rightHhoof
- - elbow
- rightknee
- - rightknee
- rightFfetlock
- - rightFfetlock
- rightFhoof
- - leftknee
- leftFfetlock
- - leftFfetlock
- leftFhoof
- - lefthock
- leftHfetlock
- - leftHfetlock
- leftHhoof

skeleton_color: blue
pcutoff: 0.6
dotsize: 12

```
alphavalue: 0.7
colormap: jet

    # Training,Evaluation and Analysis configuration
TrainingFraction:
- 0.95
iteration: 0
default_net_type: resnet_50
default_augmenter: default
snapshotindex: -1
batch_size: 8

    # Cropping Parameters (for analysis and outlier frame detection)
cropping: false
croppedtraining: false
    #if cropping is true for analysis, then set the values here:
x1: 0
x2: 640
y1: 277
y2: 624

    # Refinement configuration (parameters from annotation dataset configuration also relevant in this
  stage)
corner2move2:
- 50
- 50
move2corner: true
```