Bachelor thesis

Artificial Intelligence

Radboud University

Examining Human Walking Characteristics Using Video-Based Motion Tracking

Author:
Michiel Koenis
s1014229

First supervisor:
prof. dr. W.P. MEDENDORP
Donders Institute
p.medendorp@donders.ru.nl

Second supervisor:
C.M.J. WILLEMSEN, MSc
Donders Institute
c.willemsen@donders.ru.nl

Second reader:
dr. Y. GÜÇLÜTÜRK
Donders Institute
y.gucluturk@donders.ru.nl

June 18, 2021

Abstract

Markerless gait recognition is a fast-evolving field. In the study of human movement, it can be a great asset and be preferred over marker-based methods, because markers themselves are obtrusive and may be inaccurate. The fast development in this field is marked by innovations such as DeepLabCut, a markerless pose estimation method based on transfer learning with deep neural networks that approaches human-level labeling accuracy with minimal training data. In order to test the viability of DeepLabCut, several gait parameters were identified from videos of walking participants to reproduce known differences in these parameters between men and women. These parameters are walking velocity, step frequency, and step length, in which previous research has shown significant differences between men and women, namely that men walk with greater velocity, with greater step length, but with a lower frequency. Using a well-known dataset consisting of people walking perpendicular to the camera, it was found that men walk with greater velocity and greater step length, but it was not found that women walk with greater frequency. These results have various consequences to the implementations of gait-identifying software, and the development of systems such as DeepLabCut. In the improvement of our understanding of human walking, this research could be expanded to investigate the effects of these parameters on the energy consumption of walking participants.

Contents

Chapter 1

Introduction

1.1 Historical Perspective

Edweard Muybridge first developed techniques to capture image sequences of equine gait (Cappozzo, Marchetti, & Tosi, 1992). The techniques to do this have developed from manual annotation of images to marker-based optical trackers to state-of-the-art machine learning technologies (Colyer, Evans, Cosker, & Salo, 2018). These have revealed great insight in the workings of animal and human locomotion, and in some cases have even discovered previously unknown motor sequences such as the so-called 'tap-dancing' songbird (Ota, Gahr, & Soma, 2015).

Capturing walking sequences starts at creating a human model, which is essentially a simplification of reality. In such a model a certain number of body parts, such as the head, shoulder, knee, and toes are established, and these are implicitly connected by frictionless rotational joints, which do not necessarily reflect reality, but do approach it (Colyer et al., 2018).

A wide range of motion analysis systems allows movement to be captured in a variety of settings, which can broadly be categorised into direct and indirect techniques. In the direct techniques, devices are attached to the body, whereas in indirect techniques vision-based data are used.

1.2 Direct Methods

Direct methods usually consist of applying reflective markers or sensors on body parts of the participants to assist with computer-based tracking. While this is an effective and time-efficient method of analysis, it has several drawbacks.

One of the caveats of direct methods is sensor drift: low frequency change in measurements of a sensor over time (Irish, 2005). This causes the sensors to be inaccurate to a certain degree. While drift can be reduced during post-processing, a complete solution is yet to be found.

Further, the inclusion of sensors might influence the movement of the participants by themselves. Sensors can often be distracting, especially to animals, which leads to behavior that is less natural (A. Mathis et al., 2018).

Finally, quality of motion capture using direct methods suffers from wobble of markers placed on the skin, and therefore measurements become less accurate. (A. Mathis, Schneider, Lauer, & Mathis, 2020).

1.3 Indirect Methods

In contrast to these direct methods, there are indirect methods. One of these is manual digitisation. Manual digitisation requires the manual localisation of several body parts in each frame, usually from multiple camera perspectives. Using these various different angles, these labels can be reconstructed into three-dimensional coordinates of the locations of body parts.

This method immediately has obvious disadvantages, namely that the labelling of all these frames is an incredibly time-consuming task, and, unlike marker-based methods, the quality of the labelling is dependent on the performance of the labeler.

Merits to this method are that markers do not have to be applied to the participant, preventing the issues mentioned above with the application of markers or sensors, and the ability to analyse video footage without markers or sensors, allowing the use of footage from more naturalistic settings outside of laboratories.

1.4 AI-Techniques

In the past few years, the field of human pose estimation has made great progress due to two factors. These factors are deep learning-based architectures, and the increasing availability of large datasets such as the MPII Human Pose dataset (Insafutdinov, Pishchulin, Andres, Andriluka, & Schiele, 2016).

The development of AI-techniques that are able to place labels on the appropriate spots in video data instead of humans avoids the biggest drawback of manual digitisation, and thus the development of such techniques is of great interest in the human movement field.

Most markerless pose estimation methods are based on deep neural networks (DNNs). One of these DNN implementations is called DeeperCut (Insafutdinov et al., 2016). Like many of its competitors, DeeperCut formulates ‘body part hypotheses’, which are estimations of the possible locations of body parts. DeeperCut made three main improvements to earlier versions of this model.

For one, it has improved body part detectors that generate accurate proposals for body part locations.

Secondly, it uses novel image conditioned pairwise body part locations; after first coming up with possible locations for all body parts, for each possible location it gives a score to the likelihood that some body part is at some location x supposing that another body part is at location y. This results in an estimate of the body part locations that are consistent with one another, improving accuracy. Such a measure is especially useful in multi-person tasks, where it is essential that the body parts of all people are placed at their appropriate owners.

These calculations of pairwise terms greatly increase time and model complexity, which is why many competing models do not include such a measure.

However, it is possible in DeeperCut due to the third improvement it makes: a more time-efficient algorithm, leading to better performance and, despite the implementation of pairwise body part locations, significant speedup.

The result is an accurate labeling of all available body postures in an image. DeepLabCut (A. Mathis et al., 2018; M. W. Mathis & Mathis, 2020) is a toolbox utilizing features from DeeperCut, which allows for accuracy comparable to human labellers in pose estimation from videos after training the network on only a few hundred training images.

The feature detectors that both DeeperCut and DeepLabCut use consist of variations of deep residual neural networks (ResNets), with readout layers that predict the location of body parts. In DeepLabCut these ResNets have been pretrained. Instead of immediately having the output layer after the ResNets, DeepLabCut employs deconvolutional layers to up-sample the visual information and produce spatial probability densities. For each body part, this represents the confidence of the network in the location of a body part.

The body parts that DeepLabCut can recognize are not fixed; users can choose which parts they might want to label, and what they consider a correct marking (for example, whether a marker named 'head' is placed on the top of the head or in the middle). In order to establish these markers, the user has to label individual images with the correct locations for the markers. During training, the weights of the network are adjusted in order to accurately label these custom labels.

DeepLabCut does not have a temporal component; videos are labelled frame by frame, and thus no temporal information is transmitted between these frames, for example to help in differentiating which leg is which in a pose where the legs are right next to each other in a walking motion.

1.5 Sex Differences

Ji and Pachi have studied human stepping frequency and walking velocity (Ji & Pachi, 2005). This was done by measuring 800 people walking over two footbridges and two shopping floors. The measurements of walking frequency, walking velocity and step length were processed using statistical methods, after which the step frequency and velocity of the walking were determined.

Men were found to walk with 1.35 m/s on the footbridges, and women with 1.25 m/s. On the shopping floors both men and women walked slightly faster; men walked with 1.46 m/s, and women walked with 1.37 m/s.

On the footbridges, men were found to walk with a step frequency of 1.80 Hz, while women walked slightly faster at a step frequency of 1.86 Hz. On shopping floors, men walked at a frequency of 1.97 Hz, and women walked with a frequency of 2.03 Hz.

Significant differences between shopping floors and footbridges were not found for the step lengths; men walked with an average step length of 0.75 meters on the footbridges, and 0.74 meters on the shopping floors, where women walked with average step length of 0.67 meters and 0.68 meters, respectively.

It was also found that there is a linear relationship between walking velocity and step frequency which is different for men and women. For this research, the differences in floors will be disregarded, but the findings across floors will be used, which are:

- men walk with a higher velocity than women,
- women walk with a higher step frequency than men,
- men walk with greater step lengths.

Wheeler found similar results for these parameters, although he did not separate between men and women. He found a mean velocity of 1.5 m/s, an average step frequency of 2 Hz, and an average step length of 0.75 meters (Wheeler, 1982).

Aoyagi et al. measured preferred walking speed over a 5-meter distance in a group of 23 healthy participants aged between 65 and 74 (Aoyagi et al., 2004). Of these, 10 were male and 13 were female. Preferred walking velocity for men was 1.36 m/s, and for women was 1.24 m/s, which are results similar to Ji and Pachi.

However, Hangland and Cimbalo reported some contrasting research (Hangland & Cimbalo, 1997). While control group participants in all age-groups walking along a track gave very similar results to the research above, different results were found in a shopping mall environment similar to that of Ji and Pachi. In three estimated age groups (15 to 35, 36 to 55, and over 55 years) it was found that women walked faster than men, with the other two walking parameters staying the same. This shows that walking differences in men and women may be situation-specific.

1.6 Aims

This project aims to further develop DeepLabCut to automatically extract gait parameters from the video-based data of walking men and women. These gait parameters will be movement velocity, step frequency, and step length. Using a dataset containing data from both men and women, this project will attempt to find known differences between men and women in these parameters along the lines of previous research as described above.

Chapter 2

Methods

2.1 Dataset

The project uses the GPJATK Dataset (Kwolek et al., 2019). This dataset contains videos of people walking straight forward while monitored using cameras from various angles. These movements were recorded using 10 mocap cameras, and 4 regular cameras. Each data sequence consists of four videos with RGB images with a resolution of 960×540 pixels at 25 Hz, along with motion capture data measured at 100 Hz using 10 MX-T40 cameras. The dataset consists of 166 data sequences that recorded 32 participants (22 men and 10 women). In 128 data sequences the participants wore their own clothes, in 24 data sequences some of the participants changed clothes, and in the final 14 data sequences some of the participants wore backpacks.

For this project only participants walking perpendicular to the camera were used, which are available from the one side when the participant walks to the left, and the other side when the participant walks to the right. All participants except one wore black pants, which made their legs easily distinguishable from the background.

There are in total 64 videos, consisting of 32 videos from the perspective of the front camera seeing participants walk right to left, and 32 videos from the perspective of the back camera of participants walking left to right. Of these 64 videos, 3 were excluded; in one the participant stood still in the middle of the recording, the two others were not perpendicular at all due to an error in the original recordings of the dataset. Of these 61 videos, 43 are of men, and 18 are of women.

2.2 Data Processing Using DeepLabCut

First, the locations of the labels in DeepLabCut were chosen. As previously mentioned, these body parts can be freely chosen and defined in any way the researcher chooses. The chosen body parts were the neck, the right ankle, and the left ankle. These three labels make it possible to obtain the average position of the body through the label of the neck, and to obtain accurate positions of both legs and feet, without the added distortion that would occur if the position of the toe were labelled, which would lag behind relative to the rest of the foot during a toe-off.

After labelling the 61 videos with the aforementioned labels, the network was trained for approximately 210.000 iterations. This results in an output of 61 labeled videos and 61 files containing pixel coordinates of the labeled body parts per frame. Using these pixel coordinates, the gait parameters can be established using the following procedure.

2.3 Gait Analysis

In all videos, a triangular shape of tape on the ground is visible. These have been placed to enable conversions from pixels to meters. According to previous research using this dataset, the distance between the two outmost points of the tape is 6.30 meters (Stenum, Rossi, & Roemmich, 2021).

This measure of tape is not equally large in each video, due to small changes in camera position and the fact that two different points of view were used. Thus, the distances of all 61 videos were measured in pixels p. With this, a scaling factor s = 6.30 / p consistent with the previous study was computed.

Next, methods of establishing the three desired gait parameters, walking velocity, step frequency, and step length, were determined as follows. For the velocity, the coordinates of the neck can be used, since the neck moves consistent with the center of mass of a person, which is usually the point from which the location of a person is measured (Faraji, Wu, & Ijspeert, 2018). Thus, the velocity of the neck coordinate was assumed to be the velocity of the body.

For determining step lengths and frequencies, it is imperative to establish what exactly a step constitutes. For this research, it is assumed that a step lasts from the toe-off to the heel-strike of the same foot. The toe-off is the moment in the step that the toe leaves the ground, and the heel-strike is the moment the heel touches the ground again.

Previous research has found that toe-offs and heel-strikes occur at specific relative velocities within a step (Huitema, Hof, & Postema, 2002). Toe-offs occur the moment in the step the velocity exceeds 30% of the maximal velocity within the step, and heel-strikes occur the moment the velocity drops below 35% of the maximal velocity within the step.

A step can be considered to have been performed when the foot is on the ground for a prolonged period of time. Thus, the ankle coordinate of that foot must not have moved more than a small amount over a certain period of time. When this occurs, a marker is placed at the time that this happens. Within two such markers, a step is made, and the moment of the toe-off and heel-strike can be calculated using the aforementioned metrics from Huitema et al. Thus, this measure gives the total number of steps.

For the calculation of step frequency two variables must be known; the number of steps and the time in which these steps were made. The time in which all these steps happened is from the first toe-off until the last heel-strike, which is used as the measure of total time. Since some participants start off standing still, using this measure as opposed to taking the total time of the video is essential in establishing an accurate measure of frequency. Thus, by dividing the total number of steps with this measurement of time, an accurate estimate of the step frequency of the participants was made.

Step length consists of the maximal distance between the two feet in the step cycle. Following other research in this field, this occurs when a heel strike occurs (Stenum et al., 2021). Thus the distance between the feet at the moment of a heel-strike was taken as the measure for step length.

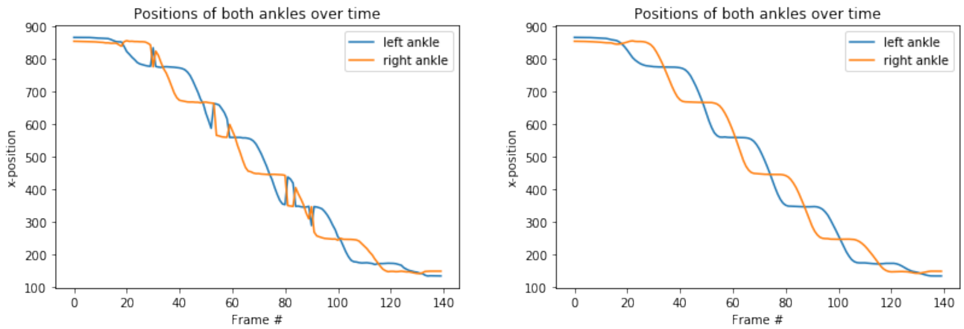
In order to test the linear relationship between velocity and step frequency that was found by Ji and Pachi, linear regression analysis has been performed.

2.4 Data Exclusion and Error Correction

When implementing these measures, some problems were encountered. The output of DeepLabCut can contain a variety of errors.

First, some labels are not placed or are placed at the far end of the screen due to DeepLabCut being unable to find an accurate placing for that body part. Such a problem can be resolved by removing values that are too different from their surroundings. A threshold of 25 pixels was established as a maximum 'acceptable' change between frames, and values surpassing that threshold are made more in line with the values around them, by changing to an average of those values.

Furthermore, sometimes DeepLabCut is incapable of accurately distinguishing which foot is which, and thus swaps them around. This has severe consequences for the accuracy of the measurements of step frequency and step length. Such 'swaps' were identified. When the x-coordinate of the left ankle at time i is almost equal to the x-coordinate of the right ankle at time i+1, and the x-coordinate of the right ankle at time i is almost equal to the x-coordinate of the left ankle at time i+1, it was assumed that the legs are swapped and thus incorrectly labeled. In consequence, one can 'unswap' all coordinates of the legs, as can be seen in Figure 1, where an example is shown of a particularly distorted signal, where using this solution, the original signal can be reconstructed.



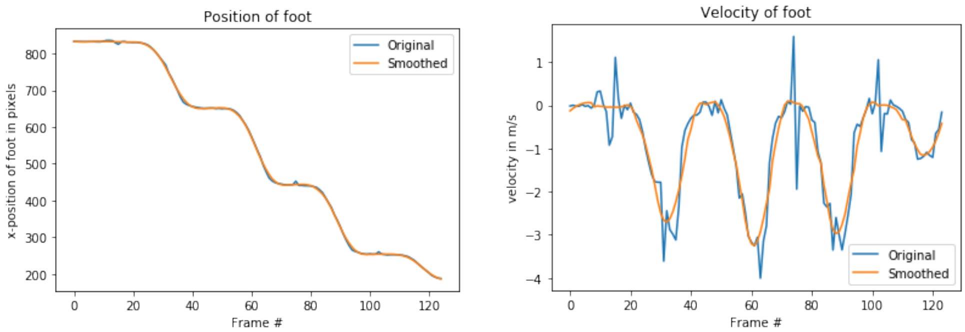
a) b)

Figure 1: a) An example of a distorted signal, b) the restored signal

Finally, there exists variability in the precise locations of the labels. For example, an ankle is not always labelled at the exact same physical part of the body; sometimes the label is placed slightly above the malleolus, other times it is placed on it. This variability most likely happens due to the fact that the training labels are not perfectly consistent, due to small errors in labelling.

This variability in combination with the fact that the coordinates of the labels are discrete measures of the continuous process of walking lead to large momentary changes in ankle velocity. This means the locations of the toe-offs and heel-strikes are less accurately established, such that in some cases the ankle does not fit the constraint of 'not moving for a certain period' required for the calculation of the steps anymore. Thus, some steps are not recognized.

A solution to this problem was found in the smoothing of such values with a Savitzky-Golay filter. This is a least linear squares approach which fits low-order polynomials to successive data points (Schafer, 2011a). As can be seen in Figure 2, such a filter has a seemingly small effect on the observed positions of the foot, only differing from the original signal in a noticeable fashion at some points. The difference is more observable in the graph for the velocity, where the original signal was not smooth at all, but the corrected signal is.



a) b)

Figure 2: An example of a signal being smoothed, and the resulting changes in a) the horizontal position of the ankle, and b) the velocity of that ankle

The Savitzky-Golay filter has two parameters, the window length and the polynomial order. The filter works by considering the middle value of the window, and selecting a new value for it as a weighed sum of the data points in the window. These weights are decided using a polynomial of some given order.

Attempts at developing a standard way of deciding these parameters have been made (Sadeghi & Behnia, 2018). However, no widely accepted way currently exists, and for this research, values were picked that looked to give good results on the data. These values were a window length of 9 and a polynomial order of 3.

Following research by Schafer, a cutoff frequency can be approached when these two parameters are known (Schafer, 2011b). Schafer posits that the cutoff frequency f_c is linearly dependent on the polynomial order N and inversely linearly dependent on window length M according to the following formula: $f_c = \frac{N+1}{3.2M-4.6}$. Thus, for this research with N = 3 and M = 9, $f_c = 0.165$.

2.5 Statistical Analysis

The data consists of two independently sampled groups, namely the results for the men and for the women. However, these results are not normally distributed and group sizes are not equal, with 43 men compared to 18 women (See Figure 3). This means that common methods of significance testing such as the parametric t-test can not be used, and instead only non-parametric testing can be used.

The non-parametric test that was used was the Mann-Whitney U test, also known as the Wilcoxon rank sum test, which could be considered to be the non-parametric version of the parametric t-test (McKnight & Najab, 2010). This test tests for differences between two groups on a single, ordinal variable with no specific distribution, irrespective of group size.

Chapter 3

Results

Figure 3 presents the results that were obtained from the analysis on the data of DeepLabCut with the aforementioned adjustments.

As can be observed in Figure 3a, 3b, and 3c, the data are not normally distributed. The data consists of 43 men and 18 women.

As seen in Figure 3a, men walk with significantly greater velocity compared to women, with 1.18 m/s compared to 1.03 m/s (statistic = 182.0, p=6.1 x 10^{-4}, Mann-Whitney U).

Regarding step frequency (Figure 3b), men walk, on average with 1.84 steps per second, as compared to women, who walk with 1.78 steps per second. This difference is not statistically significant (statistic = 313.0, p = 0.12, Mann-Whitney U). Thus the result from Ji and Pachi is not reproduced (Ji & Pachi, 2005).

Figure 3c illustrates that men walk with significantly greater step length compared to women, with 0.51 meters as opposed to 0.46 meters (statistic=205.0, p=2.1 x 10^{-3})

Finally, Ji and Pachi found a linear relationship between walking velocity and step frequency, albeit different for men and women. As can be seen in Figure 3d, such a relationship is found here as well.

Relationships in linear regression are described as $v = af + b$, where v is the velocity, f is the frequency, and a and b denote the slope and the intercept, respectfully.

For men, a slope of $a = 0.67$ and an intercept of $b = -0.056$ were found, with $R^2 = 0.68$.

For women, a slope of $a = 0.59$ and an intercept of $b = -0.017$ were found, with $R^2 = 0.62$.

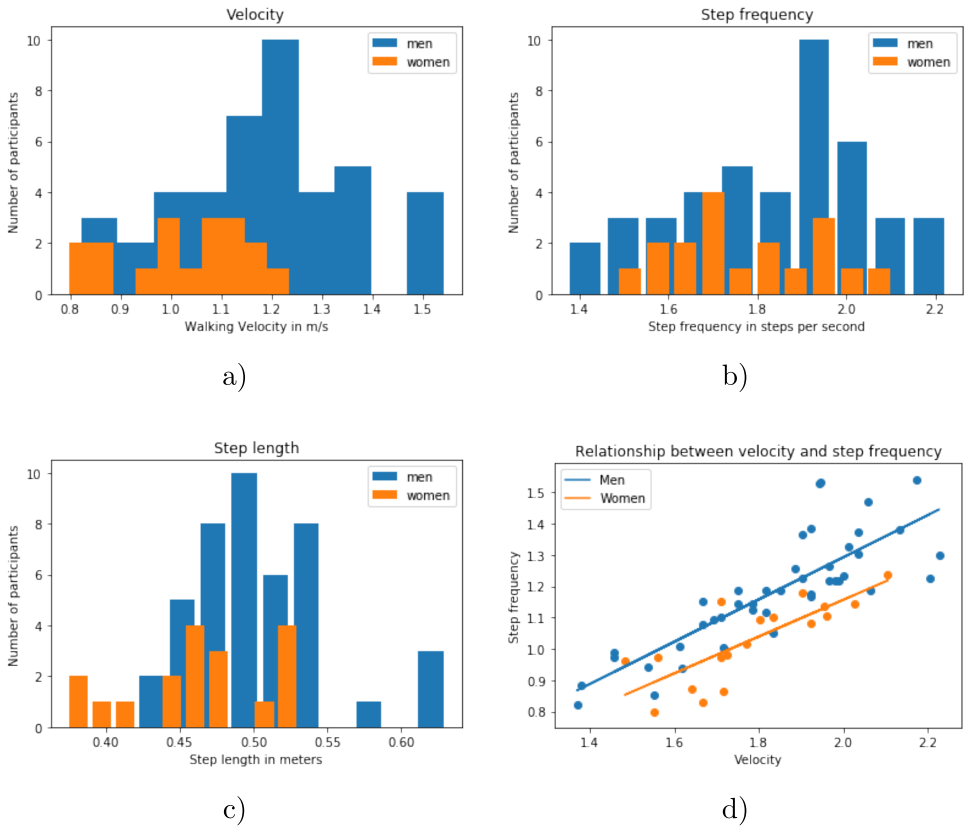


Figure 3: a) Velocity distributions of men and women. b) Step frequency distributions of men and women. c) Step length distributions of men and women. d) Linear relationship between velocity and step frequency for both men and women

Chapter 4

Discussion

In this research, the aim was to convert estimates from DeepLabCut into data from which gait parameters could be inferred. In doing so, it would give insight into the applicability of DeepLabCut, and the use of the data that it produces. By comparing these results against previous research on these parameters, a test of viability could be done. There are however, several considerations to be made when interpreting the results.

In the GPJATK dataset, participants walk a very short distance, from approximately 4 to 6 meters. Results of the inferred gait parameters might not be the same for people walking longer distances. In addition, participants were not instructed to walk at their normal walking velocity, leading some people to walk slower than they usually would. These results influence the data.

Furthermore, calculation of the toe-offs and heel-strikes, which is used in the calculation of all three parameters, is done in an indirect fashion by calculating the maximal velocity within a step and from there calculating the moments toe-offs and heel-strikes occur by taking a standard fraction of the maximal velocity (30% and 35%, respectively). While this method is accurate, it is not as accurate as the data from Ji and Pachi, and other similar research, where these moments were calculated manually.

First, the reason as to why these sex differences exist in the first place must be considered. Why is it that men walk with greater step size and with faster velocity than women?

One explanation would be that men have, on average, longer legs than women, and thus, assuming the motion of a step to be exactly the same, would be expected to make greater steps. These greater step sizes would, assuming all else being equal, lead to greater velocity. When normalising for leg length, Debi et al. found that differences in walking velocity and step length between men and women disappear, although they did not find a difference in step frequency between men and women at all (Debi et al., 2009).

While these differences in walking velocity and step length disappeared, significant differences in the gait between men and women remained, such as the difference in both single limb support (one foot on the ground) and double limb support (both feet on the ground) as a percentage of a gait cycle, and foot placement angle. Thus, differences between men and women are not solely due to the fact that men have, on average, longer legs.

Ji and Pachi found velocities of 1.35 m/s for men on footbridges, and 1.25 m/s for women. On the shopping floors they found velocities of 1.46 m/s for men and 1.37 m/s for women. Aoyagi found velocities of 1.36 m/s for men, and 1.24 m/s for women.

In this research, velocities of 1.18 m/s and 1.03 m/s were found for men and women, respectively. This is slower than the results found in the other papers, which is likely due to the shorter distance that the participants walked, leading to some opting to walk slower than they would over longer distances.

As for walking frequency, on the footbridges Ji and Pachi found step frequencies of 1.80 Hz for men and 1.86 Hz for women, whereas on the shopping floors these frequencies were 1.97 Hz for men, and 2.03 Hz for women. In this research, this finding was not replicated. In contrast, men actually walked with higher frequency, with 1.84 Hz as compared to 1.78 Hz for women. However, this difference was not significant.

Women might simply have chosen not to walk as fast as they normally would since they did not evaluate it as worthy of the extra expense in energy to use their normal acceleration, realizing that they would soon have to slow down again. While this also holds for men, they might make a different evaluation, thus leading to men walking with greater frequency.

The reason for this difference might lie in differences in acceleration between men and women. It has been reported that it takes women longer than men to reach their target velocity (Brown, Whitehurst, Gilbert, & Buchalter, 1995).

Further, it has been found that inherent sex differences such as lower muscle strength, greater fat mass, shorter limb length, and greater joint instability cause women to be less able to rely on their ability to produce force during walking, and thus they are more reliant on their acceleration capacity (Osawa, Studenski, & Ferrucci, 2018). Since the distance walked here was quite short, the factor of acceleration was not as relevant, which in turn might have led to women walking with lower frequency.

As to step length, results by Ji and Pachi on footbridges and shopping floors were almost the same, with 0.75 meters and 0.74 meters for men, respectively, and 0.67 meters and 0.68 meters for women, respectively.

In this research, it was found that men walk with an average step length of 0.51 meters, as opposed to women with an average step length of 0.46 meters.

These differences are large, although the sex difference point in the same direction. The results have been verified by manual measuring of the data, and the result has been confirmed to be accurate.

Further, the results can be compared with research by Stenum et al., who also used the GPJATK dataset in their research (Stenum et al., 2021). This research used motion capture to find an average step length over all participants of 0.598 meters. However, it only uses full steps in this research. At the start or the end the participants often make a small extra step, in order to align their feet. Stenum et al. filtered these results, which is something this research opted not to do, since they are actual steps according to the requirements that were set for a step. However, this affects the average step size in this research.

It was also found that there is a linear relationship between walking velocity and step frequency which is different for men and women. This linear relationship is explained with the formula $v = L_s f$, where v is the velocity, L_s is the step length, and f is the step frequency.

For Ji and Pachi, this different linear relationship is explained by the difference of L_s, which was 0.75 for men and 0.67 for women. The same principle works with filling in the values for L_s found in this research, thus 0.51 for men and 0.46 for women, but this is less accurate than performing linear regression.

In the linear regression that was performed, different values were found for this linear relationship, where the formula was adapted to $v = L_s f + b$, where b is a constant. For men, the best fit was L_s = 0.67, and b = -0.056, with R^2 = 0.68, while for women the best fit was L_s = 0.59, and b = -0.017, with R^2 = 0.62. Thus a linear relationship between walking velocity and frequency does indeed exist in the data of this research.

The constant b that is added in the linear regression is relatively small, meaning that the original formula is still approached. The difference in the values for L_s found by the linear regression and found as the average of all values can, just like step length, be explained with the small steps at the start and end. If these would be removed, the value for the step length would presumably be closer to the values found by the linear regression.

Chapter 5

Further Research

Participants only had to walk a very short distance, which likely has affected their gait pattern compared to normal walking.

Further research could consider under which circumstances different decisions about gait pattern are made, and what the influence of different factors is in the decision to walk in a certain manner.

One factor to consider in this process is energy consumption; walking costs energy, and people want to keep as much energy as possible. Manners of gait such as walking or running have different velocities at which their energetic cost is minimized.

It has been found that not only energetic cost is important when deciding on walking velocity, but also reward, implying that when reward is higher, it is worthwhile to be energetically inefficient (Shadmehr, Huang, & Ahmed, 2016). One example of differing rewards would be when you are at the airport awaiting a passenger. Surely you would walk with greater speed towards the arriving passenger if they were your child, as opposed to a thesis student!

Shadmehr et al. created a framework involving a utility function which depends on both the reward and the energetic cost of an action.

In future research, a more detailed dataset should be used, in which participants walk for a longer distance. Since energy consumption is calculated per unit of mass, this should record the mass of the participants. In addition, it should include rewards for completing the task.

Then, following the framework proposed by Shadmehr et al., an estimate of the energetic cost of walking could be made, from which should then follow that there exists a speed of movement that minimizes energetic cost.

Using the rewards in this future dataset, an estimate of the influence of those rewards on walking velocity could be made. This would give more insight into the reasons for why people walk at different speeds in different circumstances.

Other possible research making use of DeepLabCut could be the identification of other gait parameters for the purpose of estimating metabolic cost. It has been found that the metabolic cost of walking can largely be explained by four factors (Faraji et al., 2018). These are:

- Weight Support: the required muscular force to keep the leg from collapsing during walking. An estimation of this cost can be made using the angles of the knees. Thus this factor can be identified by placing markers on the top of the leg, the knee, and the foot, in order to get this angle.
- Ground Clearance: the cost in energy to lift the leg. The height of this lift relative to the rest of the body can be identified by placing markers on the feet.
- Center of Mass Velocity Redirection: the work required to change centre of mass velocity from step to step. The magnitude of this velocity redirection depends on the angle between the two legs, and thus can be identified by placing markers on both legs.
- 3LP Dynamics: the balancing cost of the sagittal and frontal dynamics; the velocity and mass of individual bodyparts, such as the toes, knees, and torso. All these velocities can be identified using DeepLabCut, as this research has already shown. The masses can be estimated as a fraction of the total body weight.

Total metabolic cost can largely be explained as a linear combination of these four factors, all of which can be identified using DeepLabCut. The development of an automatic method to identify metabolic cost would be usable in many fields such as rehabilitative care, professional sports, and the study of movement in general, as it would allow any video to be analysed in the context of metabolic cost.

Chapter 6

Conclusions

The results show that DeepLabCut is a useful tool for the analysis of gait parameters such as walking velocity, step frequency, and step length. It is a tool which is fast and easy to use; with the labelling of a couple hundred frames it can give accurate results due to the use of transfer learning.

While there are still inaccuracies in the identifications at some points, where labels are erroneously swapped or otherwise misplaced, lightweight programmatic solutions can be employed to correct this, such as identifiers for swapped labels, and filters which smooth out the data.

With these changes, all labels are correctly recognized, leading to the conclusion that the known gait parameters, walking velocity, step frequency, and step length, could be deduced from the coordinate data that was given by DeepLabCut.

Considering this, future research should be able to develop a full model of energetic cost on the basis of videos analysed in DeepLabCut.

References

Aoyagi, Y., Togo, F., Matsuki, S., Kumazaki, Y., Inoue, S., Takamiya, T., ... Shephard, R. J. (2004). Walking velocity measured over 5 m as a basis of exercise prescription for the elderly: preliminary data from the nakanojo study. *European journal of applied physiology*, *93*(1), 217–223.

Brown, L. E., Whitehurst, M., Gilbert, R., & Buchalter, D. N. (1995). The effect of velocity and gender on load range during knee extension and flexion exercise on an isokinetic device. *Journal of Orthopaedic & Sports Physical Therapy*, *21*(2), 107–112.

Cappozzo, A., Marchetti, M., & Tosi, V. (1992). *Biolocomotion: a century of research using moving pictures* (Vol. 1). Promograph.

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. (2018). A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports medicine-open*, *4*(1), 1–15.

Debi, R., Mor, A., Segal, O., Segal, G., Debbi, E., Agar, G., ... Elbaz, A. (2009). Differences in gait patterns, pain, function and quality of life between males and females with knee osteoarthritis: a clinical trial. *BMC Musculoskeletal Disorders*, *10*(1), 1–10.

Faraji, S., Wu, A. R., & Ijspeert, A. J. (2018). A simple model of mechanical effects to estimate metabolic cost of human walking. *Scientific reports*, *8*(1), 1–12.

Hangland, A., & Cimbalo, R. S. (1997). Human ethology: age and sex differences in mall walking. *Perceptual and Motor Skills*, *85*(3), 845–846.

Huitema, R. B., Hof, A. L., & Postema, K. (2002). Ultrasonic motion analysis system—measurement of temporal and spatial gait parameters. *Journal of biomechanics*, *35*(6), 837–842.

Insafutdinov, E., Pishchulin, L., Andres, B., Andriluka, M., & Schiele, B. (2016). DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model. *ECCV*.

Irish, J. (2005, Feb). *Ocean instrumentation, course 13.998, lecture on instrumentation specifications.* Retrieved from `https://ocw.mit.edu/courses/mechanical-engineering/`

2-693-principles-of-oceanographic-instrument-systems -sensors-and-measurements-13-998-spring-2004/readings/ lec2_irish.pdf

Ji, T., & Pachi, A. (2005). Frequency and velocity of people walking. *Structural Engineer*, *84*(3), 36–40.

Kwolek, B., Michalczuk, A., Krzeszowski, T., Switonski, A., Josinski, H., & Wojciechowski, K. (2019). Calibrated and Synchronized Multi-View Video and Motion Capture Dataset for Evaluation of Gait Recognition. *Multimedia Tools and Applications*.

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.

Mathis, A., Schneider, S., Lauer, J., & Mathis, M. W. (2020). A primer on motion capture with deep learning: principles, pitfalls, and perspectives. *Neuron*, *108*(1), 44–65.

Mathis, M. W., & Mathis, A. (2020). Deep learning tools for the measurement of animal behavior in neuroscience. *Current opinion in neurobiology*, *60*, 1–11.

McKnight, P. E., & Najab, J. (2010). Mann-whitney u test. *The Corsini encyclopedia of psychology*, 1–1.

Osawa, Y., Studenski, S. A., & Ferrucci, L. (2018). Knee extension rate of velocity development affects walking performance differently in men and women. *Experimental gerontology*, *112*, 63–67.

Ota, N., Gahr, M., & Soma, M. (2015). Tap dancing birds: the multimodal mutual courtship display of males and females in a socially monogamous songbird. *Scientific reports*, *5*(1), 1–6.

Sadeghi, M., & Behnia, F. (2018). Optimum window length of savitzky-golay filters with arbitrary order. *arXiv preprint arXiv:1808.10489*.

Schafer, R. W. (2011a). What is a savitzky-golay filter?[lecture notes]. *IEEE Signal processing magazine*, *28*(4), 111–117.

Schafer, R. W. (2011b). On the frequency-domain properties of savitzky-golay filters. In *2011 digital signal processing and signal processing education meeting (dsp/spe)* (pp. 54–59).

Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016). A representation of effort in decision-making and motor control. *Current biology*, *26*(14), 1929–1934.

Stenum, J., Rossi, C., & Roemmich, R. T. (2021). Two-dimensional video-based analysis of human gait using pose estimation. *PLoS computational biology*, *17*(4), e1008935.

Wheeler, J. E. (1982). Prediction and control of pedestrian-induced vibration in footbridges. *Journal of the structural division*, *108*(9), 2045–2065.