

Accurate Horse Gait Event Estimation Using an Inertial Sensor Mounted on Different Body Locations

Hamed Darbandi
Pervasive Systems Group
University of Twente
Enschede, The Netherlands
h.darbandi@utwente.nl

Filipe Serra Bragança
Dept. of Clinical Sciences
Utrecht University
Utrecht, The Netherlands
f.m.serrabraganca@uu.nl

Berend Jan van der Zwaag
Pervasive Systems Group
University of Twente
Enschede, The Netherlands
b.j.vanderzwaag@utwente.nl

Paul Havinga
Pervasive Systems Group
University of Twente
Enschede, The Netherlands
p.j.m.havinga@utwente.nl

Abstract— Accurate calculation of temporal stride parameters is essential in horse gait analysis. A prerequisite for calculating these parameters is identifying the exact timings of gait events, i.e., hoof-on and hoof-off moments. A hoof-mounted inertial measurement unit (IMU) can be used to identify these moments accurately, yet this approach is often impractical due to the vulnerability of IMU to the impacts during locomotion. In this study, we investigated the possibility of accurately estimating the gait events using the signals of an IMU mounted on a less vulnerable location, such as a limb or upper body. To achieve the goal, we equipped IMUs on horses limbs, withers, and sacrum and measured them during different gaits. Then, we estimated the gait events timings by training recurrent neural networks models on the output signals of each IMU. Finally, we evaluated the models by comparing their results to the gait events timings labeled from hoof-mounted IMUs. The best performing model represented the best location (between the limbs, withers, and sacrum) for gait event estimation. Compared to the previous studies, our models yielded higher accuracy and were more generic by supporting more gaits. In conclusion, accurate calculation of temporal stride parameters is feasible by estimating gait event timings using an IMU mounted on less vulnerable body locations.

Keywords— *Inertial sensors, Deep learning, Gait, Horse*

I. INTRODUCTION

The importance of horse gait analysis for health and performance assessment has been proliferating since the introduction of the horse motion picture in the late nineteenth century. Since then, researchers and biomechanical experts have defined various indicators for gait analysis and locomotion patterns comparison between horses. Temporal stride parameters are essential indicators used for analyzing the gait and evaluating the quality of locomotion for horses [1]. A horse stride can be defined as a repeated hoof placement pattern during locomotion [2]. This hoof placement pattern differs between gaits, making the temporal stride parameters comparable and meaningful for studying only within a single gait, e.g., walk, trot, and canter as natural gaits, and passage and piaffe as artificial gaits (i.e., horses learn to perform).

Temporal stride parameters were used as essential measures in different studies. For instance, in the training of sport horses, the long term changes in stride duration might indicate the fitness and performance level [3], while in the short-term, the changes in stride frequency, stride length, and speed might represent high internal workload and fatigue [4]. Also, as a practical example and according to dressage rules of the International Equestrian Federation, judges award scores during performance considering the temporal stride parameters values during competitions [5].

The accuracy of calculating the temporal stride parameters depends on identifying the exact moments of gait events: hoof impact on the ground (hoof-on) and hoof lift off from the ground (hoof-off). One of the approaches for estimating the hoof-on/off moments is using a force plate, which is the gold standard for identifying the moments in equine studies [6], [7]. In addition to the force plate, several studies estimated the moments kinematically using optical motion capture (OMC) [8], [9]. However, the usage of force plates and OMC is limited to the laboratories environment. Therefore, it is almost impossible to use them on the field during training or competition.

In contrast to force plate and OMC, an inertial measurement unit (IMU) can be used as a portable device. This device outputs three-dimensional acceleration and angular velocity signals. Several studies used the spikes in the output signals of hoof-mounted IMU as a reference for hoof-on/off moment estimation [6], [10], [11], [12]. However, mounting an IMU on the hoof is prone to damage during exercise or competition. Moreover, it is time-consuming and challenging, and the IMU can easily fall off during locomotion. On the other hand, attaching IMU to the limb using elastic straps or on the upper body using double-sided adhesive tapes is not as challenging as attaching IMU to the hoof. However, the limb or upper body IMU is farther from the hoof-on/off location than the hoof IMU, generating more complicated signal patterns compared to the clear spikes in hoof IMU signals. In addition, the patterns of IMU signals, including the moments of hoof-on/off, vary between different gaits. As a consequence, the complexity of signal patterns affects the accurate estimation of hoof-on/off moments and therefore, the accuracy of calculating temporal stride parameters.

This study aims to estimate the hoof-on and hoof-off moments during different gaits by using the signals generated from a single IMU mounted on limbs or upper body. This goal can be approached in three steps. The first step is to detect (label) the hoof-on/off moments in hoof-mounted IMU signals. The next step is to develop models using the output signals of a limb- or an upper body-mounted IMU to estimate the labeled hoof-on/off moments during different gaits. Finally, evaluating and optimizing the performance of the models based on the placement of IMU and comparing the outcome to state-of-the-art methods. This study approached the hoof-on/off moments estimation literature by developing a model to improve accuracy and to support more gaits (canter, piaffe, passage) and more IMU placement on the body (upper body locations).

II. BACKGROUND AND PRIOR WORK

There are several studies on horse hoof-on/off moments estimation using body-mounted IMU. Ryan et al. [13] calculated

the resultant acceleration magnitude (Euclidean norm) from a three-dimensional accelerometer glued to the dorsal hoof wall and detected the hoof impact and break-over according to the peaks of the diagram. Tijssen et al. [6] calculated the variance of Euclidean norm of acceleration and gyroscope signals extracted from front and hind hoof-mounted IMUs and then identified the moments considering the peaks of the calculated signals.

Besides sensor mounting on the hoof, Starke et al. [11] attached an IMU on the pelvis for hoof-on/off moments determination during walk and trot. This study estimated the hoof-on timing using the events extracted from pelvis vertical velocity and roll angle signals. Bragança et al. [12] proposed an algorithm to detect the events from specific peaks and troughs of accelerometer and gyroscope signals extracted from a limb-mounted IMU to estimate the front and hind hoof moments during walk and trot. In one of the latest IMU studies, Sapone et al. [14] detected the front hoof-on and -off moments during trot by selecting the specific events from a discrete wavelet analysis on the gyroscope signals derived from a limb-mounted IMU.

All the mentioned studies were focused on detecting the signal peaks, troughs, and zero-crossings and then defining the hoof-on/off moments. However, these methods are highly dependent on the signal patterns, and the signal patterns can vary relative to the gaits [6], [8], [14].

In this study, we dedicated the task of signal pattern recognition to a machine learning algorithm, which is required for evaluating equine gait patterns [15]. Furthermore, in addition to the limb-mounted IMUs, we used upper body (sacrum and withers) mounted IMUs to estimate the hoof events. Finally, the mentioned studies focused only on walk and trot, while we trained models to estimate hoof-on/off moments during walk, trot, canter, passage, and piaffe.

III. METHODS

A. Data

The data were collected from twenty-one Warmblood horses participating in dressage competitions. The horses were measured while ridden at walk, trot, canter, piaffe, and passage on a soft surface (sand-fiber). Horses were equipped with six ProMove-mini IMUs [16] on withers, sacrum, the lateral aspect of the right front and hind limbs (cannon bones), and the lateral wall of the right front and hind hooves. Each IMU contained a tri-axial accelerometer and a tri-axial gyroscope and was set to collect data at a sampling rate of 200 Hz, an acceleration range of $\pm 16g$, and an angular velocity of ± 2000 deg/s. Fig. 1 demonstrates the IMU locations and orientations on the body.

The three axes of rotation for the sacrum, withers, and poll IMUs were x, y, and z (Fig. 1), which were defined in the order as longitudinal axis, mediolateral axis, and vertical axis. For limbs and hooves, x, y, and z-axis were aligned to the cannon bone and external hoof wall, (longitudinally), retraction/protraction angle axis, and abduction/adduction angle axis, respectively [16].

B. Labeling the data by gaits

Using the video footage of the measurements, the gaits (walk, trot, canter, piaffe, and passage) were labeled per stride. It should be noted that all the measurements were not gait-

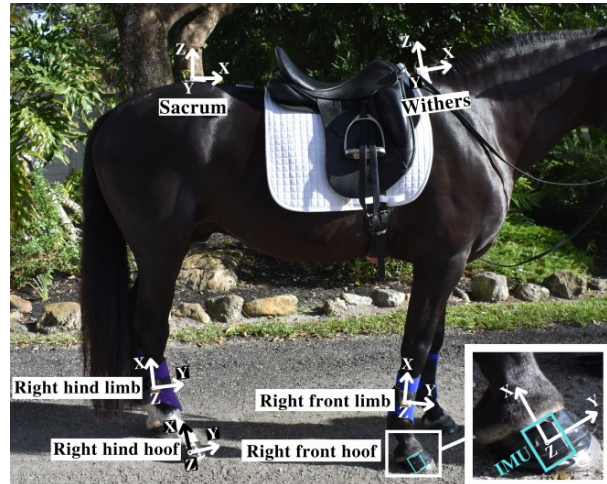


Fig. 1. IMUs locations and orientations on horse body. On the right bottom, the placement of IMU on right front hoof is magnified for detailed view.

labeled. To study the estimation model performance independent of gait and to increase the size of the data, the not labeled part of the measurements was used for model development and evaluation. These measurements are comprised of strides during walk, trot, canter, piaffe, and passage. Hereafter, the not labeled part is designated as "unlabeled" in this paper.

C. Labeling the data by hoof on/off moments

In the literature, various methods labeled the hoof on/off moments by using the signals extracted from a hoof-mounted IMU. In a number of studies, the moment that the hoof-IMU acceleration signal (vertical acceleration) changes abruptly was labeled as "hoof-on moment", and the final abrupt peak before silence on the signal pattern was labeled as "hoof-off moment" [6], [10], [11]. In this study, we used the methods for detecting and labeling the hoof on/off moments.

The position of IMU on the hoof was different between the studies. For example, in one study, the IMU was mounted on the dorsal surface of the hoof, and in another study, the IMU was attached to the lateral wall of the hoof, while both studies used the vertical acceleration signal to detect the hoof-on/off moments. In Fig. 2, it is demonstrated that x-axis acceleration (globally vertical) can present abrupt decelerations and accelerations. However, the x-axis of hoof-IMU was not aligned precisely vertical, as shown in the right bottom of Fig. 1. To cancel out the effect of different IMU positions and orientations on the hoof-mounted IMU alignment, the Euclidean norm of preprocessed tri-axial acceleration signals ($\sqrt{a_x^2 + a_y^2 + a_z^2}$) was calculated [6]. Euclidean norm resulted in a unidirectional acceleration signal, which was plotted in time for identifying and labeling the moments observationally (bottom plot of Fig. 2).

Developing the hoof-off moment estimation models are the same as hoof-on moment, presented in Fig. 3 as a summary. To avoid repetition in this paper, only the hoof-on moment estimation procedure is explained in the following. Each timestep was labeled (according to the Euclidean acceleration signal obtained from hoof-IMU) as "1" if it was a hoof-on

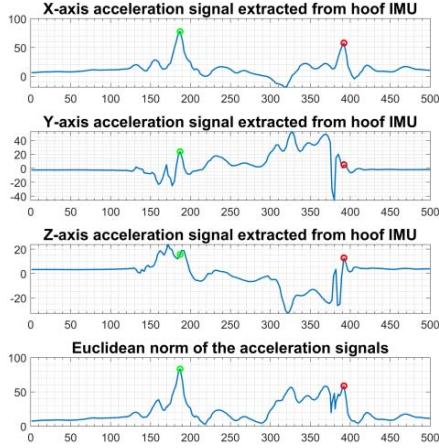


Fig. 2. Acceleration signals extracted from right front hoof-mounted IMU. The signals were aligned on global coordinate system. The green circles are the hoof-off moments and the red circles are the hoof-on moments. The plots were time-synchronized. The vertical and horizontal axis of the plots are in acceleration (m/s^2) and timestep (each timestep = 5 milliseconds), respectively.

moment or "0" if it was not a hoof-on moment. Therefore, a square wave signal as the label was produced for each hoof.

D. Feature selection

Eight signals were extracted from each IMU, mounted on the sacrum, withers, right front limb, and right hind limb. The extracted signals consisted of three acceleration signals, three angular velocity signals, one Euclidean norm of the three acceleration signals, and one Euclidean norm of the three angular velocity signals. All the eight signals and the label signal were time-synchronized with an accuracy of $< 100\text{ns}$. Therefore, the goal of each model is to estimate the label signal using the eight-dimensional input derived from a single IMU (sacrum, withers, front limb, or hind limb).

E. Dataset preparation

The models were trained using ten-fold cross-validation to prevent overfitting. The dataset from all subjects and gaits on each fold was split into approximately 90 percent training and 10 percent testing datasets. More specifically, on each split, the data from at least two horses were on the testing dataset, and the remaining were on the training dataset. After each split, the training dataset was normalized using z-score. Using the parameters from the z-score, the testing set was normalized afterward. The z-score parameters were calculated using the training dataset instead of the whole dataset (training and testing datasets) to remove the normalization bias from the testing data.

For each "1" in the label signal, 127 time-synchronized windows (256-timestep wide) were created from each input signal and label signal. For clarification, let assume there is a hoof-on moment ("1" in label signal) at timestep n . If $\{t \in \mathbb{N}; 1 < t < 127\}$, the windows for each n were in the range of $(n - 2 \times t, n + 2 \times (127 - t) + 1)$, which results in 127 windows. This sliding window for each "1" was applied for adapting the model to different locations of the moment within a window. It should be noted that all the windows used for training and testing datasets contained at least one hoof-on or hoof-off moment.

F. Loss function

Because of the scarcity of hoof-on/off moments compared to the non-moments, a loss function was required to focus the model on moments estimation since the goal of an estimation model is to minimize the loss. In this case, the loss function should be designed with the purpose of higher value returns from the model when it estimates the moment as "0" instead of "1", known as penalization of the model's incorrect estimation [17], particularly during events. A weighted binary cross-entropy loss function has been defined as follows [17],

$$L(y, \hat{y}) = -(W \cdot y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})) \quad (1)$$

where y and \hat{y} are the true and predicted output, respectively. W is the weight, representing the extent of penalizing the incorrect model estimation. In binary cross-entropy loss function, a logarithmic function is used instead of the linear form $(y \hat{y} + (1 - y)(1 - \hat{y}))$, to heavily penalize the model estimations that are confidently wrong. For instance, if $y = 1$ and $\hat{y} \approx 0$, the first part of binary cross-entropy function $y \log(\hat{y})$ outputs a very high value, while the second part $(1 - y) \log(1 - \hat{y})$ yields a value of zero (or near zero). Using a weighted loss function, the first part is multiplied by W , which yields a bigger value for the loss function when there is a wrong prediction for events ($y = 1$). Hence, the model optimizes the loss function, particularly during events. In contrast, if we do not use a logarithmic function in the loss function, the value of the loss function would become zero (or near zero), where the model had already been optimized. In the current study, we considered W as a hyperparameter and searched for the best model performance with $W = 50, 100$, and 200 .

G. Model training and performance evaluation

Single-layer 256 units RNN models with LSTM architecture (RNN-LSTM) were implemented on the training dataset extracted from each IMU (right front limb, right hind limb, sacrum, or withers) separately to optimize the customized loss function. Thus, eight RNN-LSTM models were trained and optimized, which were one hoof-on and one hoof-off moment estimation models for each IMU. RNN was chosen since we tried to model and estimate time-series data. In addition, we selected LSTM as the architecture since it solves the RNN issue of vanishing gradient during training time-series models. In total, the model received the eight-dimension input in 256-timestep windows (8×256), followed by a dropout layer, and a fully connected layer to the one-dimensional label signal as the output, as shown in Fig. 3.

Since the estimation approach in this study is regression and not classification, the predicted outputs would not be binary; instead, they were signals. Therefore, we detected the maximum peak of each window (global maxima) in the outputs signals and counted it as the predicted event. Then, for evaluating the models, the time difference (in milliseconds and timesteps) between the true moment from the label signal and the maximum peak in the model output was calculated as the model error. The estimations errors were presented as accuracy (mean) and precision (standard deviation) of the models.

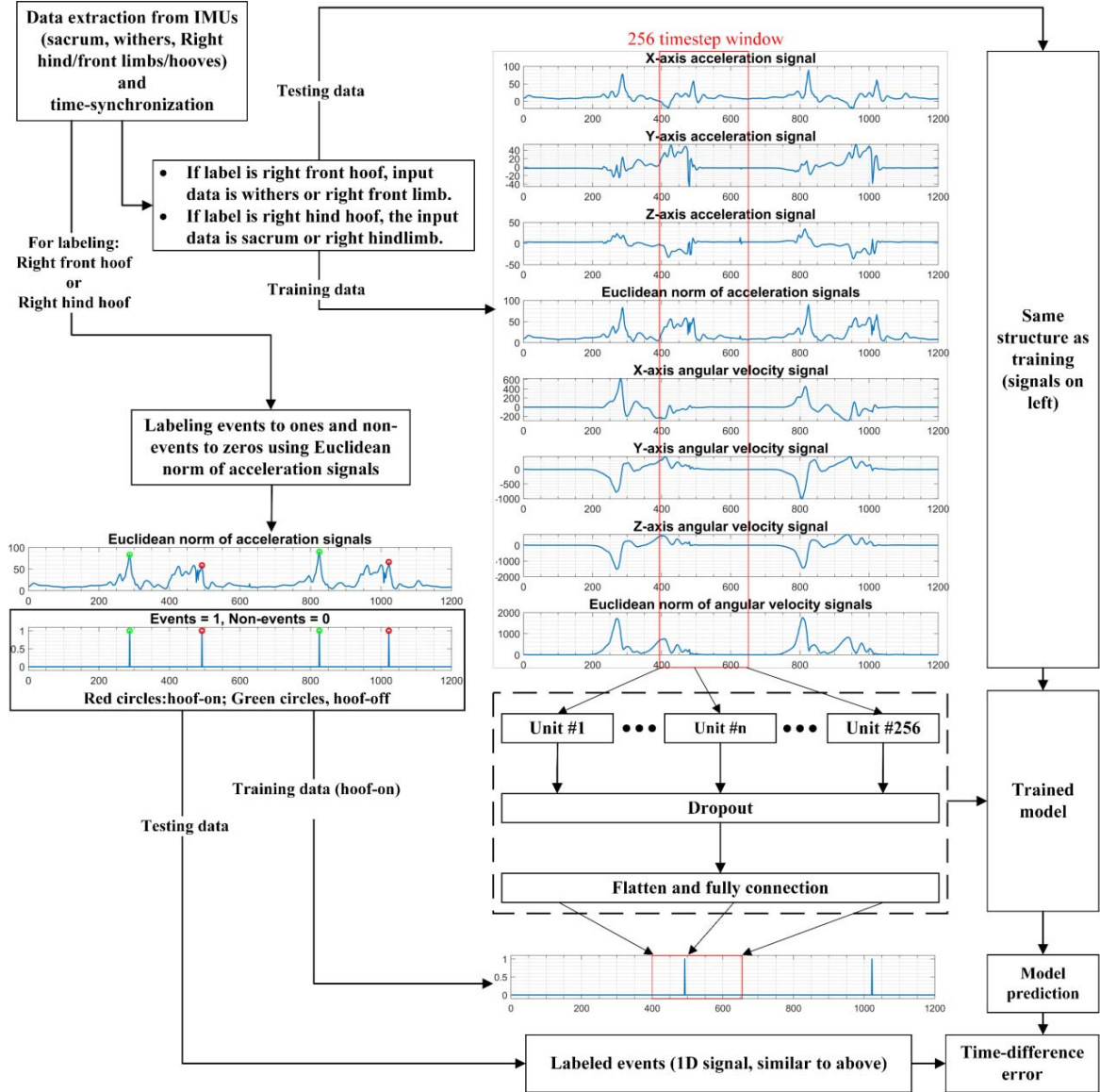


Fig. 3. A summary of the data processing, hoof-on model training, and hoof-on model evaluation (model training and evaluation procedure is same for hoof-off)

H. Hyperparameters optimization

In addition to W , batch size ($= 16, 32, 64, 128$), learning rate ($= 0.01, 0.05, 0.001, 0.005, 0.0001$), and dropout rate ($= 0.2, 0.4, 0.6$) were also tuned as hyperparameters with the purpose that the model achieve the best overall accuracy and precision.

I. Comparison of results with state of the art

The best performing method (algorithm number 3) of [12] was selected to compare estimation performance results between the current study and state-of-the-art methods. The selected method was based on the inputs from limb-mounted IMU signals. Therefore, it was only implemented on the dataset (combined training and testing datasets) extracted from right front and right hind limb IMUs. Then, the errors were calculated as the same as discussed for our models and were compared with our results.

Matlab R2020a (MathWorks Inc., Natick, MA, USA) was used for data preprocessing, and LSTM models were, trained, and tested with Python 3.7 (and Tensorflow 2.4.0). The described method for developing the gait event estimation models was summarized in Fig. 3.

IV. RESULTS

In total, 41000 strides (82000 hoof-on/off moments) were extracted from the data. The number of strides per gait are presented in Table I. The estimation errors of the trained models in all gaits are presented in Fig. 4 (top plot). In addition, the performances of the best performing models for front and hind hoof-on/off (front limb IMU and hind limb IMU models) within all gaits are also displayed in Fig. 4 (the four bottom plots).

TABLE I. NUMBER OF STRIDES PER GAIT

Gait	Walk	Trot	Canter	Passage	Piaffe	Unlabeled	Total
Number of strides	4600	7800	10400	1900	1500	14800	41000

It can be inferred from the top plot of Fig. 4 that the model based on front limb IMU yielded the best accuracy and precision for both hoof-on and hoof-off, with -0.2 and -0.1 ms accuracy and 9.0 and 6.0 ms precision, respectively. Between upper-body mounted IMUs, the hoof-on and hoof-off models based on withers IMU presented lower errors (-0.2 ± 21.0 and -2.8 ± 17.0 ms) than sacrum IMU for hind hoof-on (-8.7 ± 29.0 ms) and -off (7.9 ± 25.0 ms). A positive and negative value can be defined as an estimation delay and an early moment estimation, respectively.

All the models yielded the best performance when W and batch size were 100 and 32 , respectively. The tuned values of other hyperparameters, learning rate and dropout rate, differed among the models. The learning rate varied between 0.0001 and 0.0005 , while the dropout rate values were 0.2 , 0.4 , and 0.6 . The values of hyperparameters that resulted in the highest accuracy for each model were reported in Table II.

According to the results, the precision increased and decreased respectively as the models (except the right front limb model) estimated the hoof-on and hoof-off during faster natural gaits, from walk to canter ($speed_{canter} > speed_{trot} > speed_{walk}$). Between artificial gaits, both accuracy and precision were lower in piaffe compared to passage. Therefore, we achieved the lowest and highest error during walk and passage, respectively.

Fig. 5 presented the models performances in this study compared to the results of two state-of-the-art methods [21],[25]. Although the methods were only based on front and hind limbs IMUs (as the input), models based on withers and sacrum IMUs were also presented for comparison (based on the output IMUs) of models performances in terms of IMU placement on the body.

V. DISCUSSION

In this study, four outcomes have been achieved for the first time in equine literature. First, the models outperformed the previous studies in terms of estimation accuracy. Second, using the signals from upper body IMUs, we accurately estimated hoof-on/off moments. Third, by considering the different patterns between gaits, hoof-on/off moments during the five gaits were accurately estimated by developing single models that

TABLE II. THE HYPERPARAMETERS VALUES OF MOST ACCURATE MODELS (TESTING THE MODEL TRAINED WITH THE VALUES). IN ALL MODELS, $W = 100$ AND BATCH SIZE = 32 .

Input IMU	Output IMU	hoof-on/off	Learning rate	Dropout rate
Right front limb	Front hoof	On	0.0001	0.6
		Off	0.0005	0.4
Withers		On	0.0001	0.4
		Off	0.0005	0.6
Right hind limb	Hind hoof	On	0.0001	0.4
		Off	0.0005	0.4
Sacrum		On	0.0005	0.2
		Off	0.0005	0.4

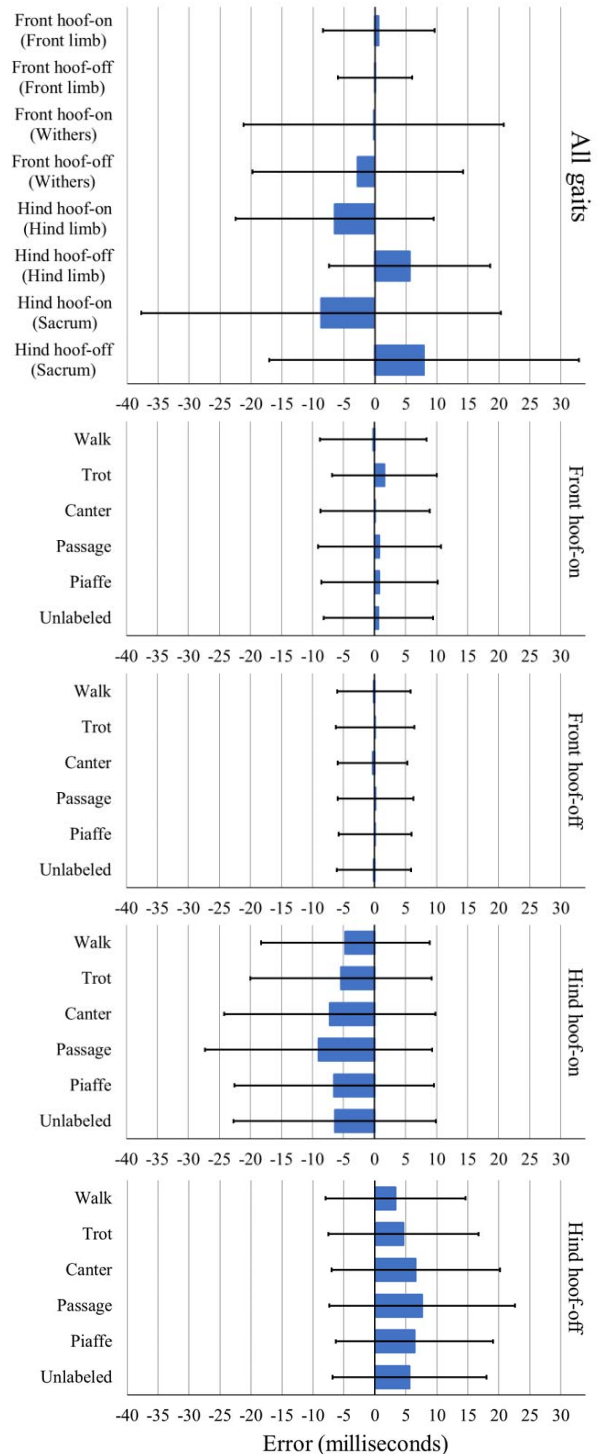


Fig. 4. The Performances of the estimation models. Top: Performance of the models within all gaits per IMU. Second: Performance (per gait) of the front hoof-on model trained on front-limb IMU. Third: Performance (per gait) of the front hoof-off model trained on front-limb IMU. Fourth: Performance (per gait) of the hind hoof-on model trained on hind-limb IMU. Fifth: Performance (per gait) of the hind hoof-off model trained on hind-limb IMU. The bars and the error bars indicate the accuracy (mean) and precision (standard deviation) of estimation errors, respectively.

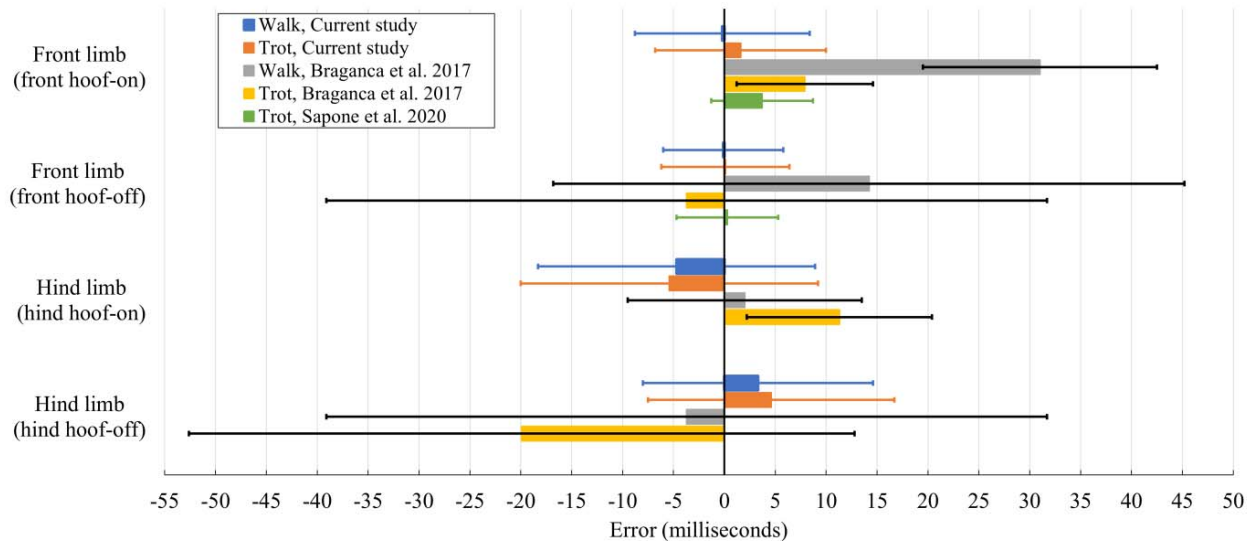


Fig. 5. The comparison between estimation models of the current study and the literature. The bars and the error bars indicate the accuracy (mean) and precision (standard deviation) of the estimation models, respectively.

support all gaits. Lastly, a deep learning approach was implemented for hoof moment estimation to ease the way for developing automatic and real-time applications.

A. Presentation of the models accuracy and precision

Since the models were fed with 200 Hz data, the output was also 200 Hz. Therefore, the lowest error value is one timestep, which equals five milliseconds. For models based on front limb and withers IMUs, we achieved mean accuracies lower than a full timestep (< 5 ms), indicating the model's capability to estimate many testing samples without error.

B. Comparison between front and hind hoof on/off models

According to Fig. 4, the model based on limbs IMUs data yielded the highest precision (i.e., lowest standard deviation of errors), while the models based on the right front limb (0 ms) and withers IMUs (-0.2 ms) were the most accurate (i.e., lowest mean of errors) for hoof-off and hoof-on. This result indicates that calculating the temporal stride parameters using the front hoof-on/off estimated moments is more accurate than the hind hoof-on and hoof-off (6.5 and 5.6 ms for the model based on hind limb IMU), independent of IMU placement (on a limb or upper body). Even with lower accuracy of the hind hoof-on/off models, their errors for calculating the stride duration meet the criteria for distinguishing between different degrees of hind limb lameness [18] as well as for front limb lameness [19]. In addition, calculating stride duration from front hoof-on/off models is sufficient for detecting training impacts on racehorses (< 0.6 ms) [3].

C. Comparison of gaits and IMU placements

As shown in the four bottom plots of Fig. 4, the accuracy and precision of the models are correlated inversely with speed, and are lower in artificial gaits than in natural gaits. The reason might be more sudden lower limb movements (creating a more complex pattern) during higher speed gaits, especially piaffe and passage, where the dressage horses perform special movements focusing on the lower limbs.

As displayed in the top plot of Fig. 4, the models based on the right front and hind limb IMUs performed better than those based on withers and sacrum IMUs, respectively. From the vulnerability aspect, equipping a limb, withers, or sacrum for the measurement is more secure than attaching an IMU to the hoof. Moreover, from a practical standpoint, equipping a limb with IMU is more straightforward than mounting an IMU on top of the withers or sacrum. The limb IMU can be placed in a pocket, where the pocket is wrapped around a limb or a tendon boot (horse boot) using hook-and-loop straps. Fixing the pocket on the limb and putting the IMU inside the pocket is feasible with simple guidance for an average user. On the other hand, mounting an IMU on withers or sacrum is usually done by sticking a double adhesive tape; However, if the measurement duration gets prolonged, the adhesive tape becomes loose due to the horse's body sweat, and the sensor falls off. Moreover, finding the anatomical location of the sacrum requires equine anatomical knowledge. Therefore, the practical advantage of using limb IMU for estimation adds up to its better performance result compared to upper body IMU. The advantage of placing the IMU on a limb was also discussed in the machine learning studies within equine gait literature [20],[21].

D. Comparison with state-of-the-art

The models performances results were validated by applying a state-of-the-art method [12], where accelerometer and gyroscope signals were processed during walk and trot. Our models outperformed the study except for hind hoof-on during walk and trot (Fig. 5). In addition, our model based on withers IMU presented better accuracy and precision in both right front hoof-on and -off during walk and trot, compared to the results from state-of-the-art right front limb.

Apart from the above-mentioned study, another study developed an estimation model by implementing a discrete wavelet analysis on the longitudinal axis of the gyroscope mounted on the right front distal limb while the horses were trotting on a treadmill, which achieved accuracy and precision

of 3.7 ± 17.0 ms for hoof-on and 0.3 ± 15.4 ms for hoof-off (by considering the mean trotting stride duration was 680 ms in the study) [14]. The hoof estimation studies that used IMU have been focused on walk or trot [11], [12], [24], [14], whereas our model includes more gaits, especially passage and piaffe. In addition, a large number of strides were used for model development compared to the small number of samples used in previous studies. Furthermore, the data in this study was measured from a training field, while the mentioned studies collected data in laboratory settings or treadmills.

E. Considerations and limitations of the study

This study did not consider the left limbs' data since all quadrupedal vertebrates perform bilateral movement symmetry between front limbs and hind limbs [22].

The models were limited in terms of input, which must contain a hoof-on and/or hoof-off. The models will select hoof-on/off moments regardless of the availability of hoof-on/off moments in the window, and in the case of no moment in the window, the output would be a false positive. A possible solution can be detecting the window containing the hoof-on or -off moment using the Euclidean signal peaks [6] and then sending the detected window to the current model for accurate moment estimation.

Lastly, the signal pattern of hoof IMU might be different depending on the surface type. The data was collected on a soft surface (sand and fiber). The effect of hard surfaces on the model should be analyzed in further studies.

VI. CONCLUSION

This study demonstrated that by mounting only one IMU on less vulnerable body locations and using its generated signals during five different gaits, it is possible to calculate temporal stride parameters with higher than the required accuracy for detecting the differences between degrees of lameness and training impacts on the performance. The estimation models developed in this study can be implemented on the IMU to calculate the temporal stride parameters in real-time and automatically. The outcome of this study in practice can help researchers and equestrians to measure the gait parameters by safely attaching an IMU to the body of horses, preferably to a front limb for higher accuracy. The models can be developed further in future studies by adding data from different breeds and different surface types.

ACKNOWLEDGMENT

The authors wish to thank Sarah Jane Hobbs from University of Central Lancashire, Hilary Clayton from Sport Horse Science, Marie Rhodin and Elin Hernlund from Swedish University of Agricultural Sciences, Mick Peterson from University of Kentucky, Rosalie Bos from Utrecht University, and all the riders and caretakers that helped with the data collection. EFRO OP-Oost partly funded this study. The ethical permissions were given by University of Central Lancashire and University of Kentucky (ethics codes: STEMH961 and 2019-3150).

REFERENCES

1. S.J. Hobbs, et al., "A scoping review of determinants of performance in dressage," *PeerJ*, vol. 8, no. e9022, 2020.

2. D.H. Leach, K. Ormrod, and H.M. Clayton, "Standardised terminology for the description and analysis of equine locomotion," *Equine Vet J.*, vol. 16, no. 6, pp. 522-528, 1984.
3. R.S.V. Parkes, R. Weller, T. Pfau, and T.H. Witte, "The Effect of Training on Stride Duration in a Cohort of Two-Year-Old and Three-Year-Old Thoroughbred Racehorses," *Animals (Basel)*, vol. 9, no. 7, pp. 466, 2019.
4. S.J. Wickler, H.M. Greene, K. Egan, A. Astudillo, D.J. Dutto, and D.F. Hoyt, "Stride parameters and hindlimb length in horses fatigued on a treadmill and at an endurance ride," *Equine Vet J Suppl.*, vol. 36, pp. 60-4, 2006.
5. Fédération Equestre Internationale. Dressage Rules 2021. 2021-03-08; Available from: <https://inside.fei.org/>.
6. M. Tijssen, et al., "Automatic hoof-on and -off detection in horses using hoof-mounted inertial measurement unit sensors," *PLoS ONE*, vol. 15, no. 5, pp. e0233649, 2020.
7. N. Zahradka, K. Verma, A. Behboodi, B. Bodt, H. Wright, and S.C.K. Lee, "An Evaluation of Three Kinematic Methods for Gait Event Detection Compared to the Kinetic-Based 'Gold Standard'," *Sensors (Basel)*, vol. 20, no. 18, pp. 5272, 2020.
8. J.K. Boye, M.H. Thomsen, T. Pfau, and E. Olsen, "Accuracy and precision of gait events derived from motion capture in horses during walk and trot," *J Biomech.*, vol. 47, no. 5, pp. 1220-4, 2014.
9. C. Roepstorff, et al., "Reliable and clinically applicable gait event classification using upper body motion in walking and trotting horses," *J Biomech.*, vol. 114, pp. 110146, 2021.
10. T.H. Witte, K. Knill, and A.M. Wilson, "Determination of peak vertical ground reaction force from duty factor in the horse (*Equus caballus*)," *J Exp Biol.*, vol. 207, no. Pt 21, pp. 3639-3648, 2004.
11. S.D. Starke, T.H. Witte, S.A. May, and T. Pfau, "Accuracy and precision of hind limb foot contact timings of horses determined using a pelvis-mounted inertial measurement unit," *J Biomech.*, vol. 45, no. 8, pp. 1522-1528, 2012.
12. F.M. Braganca, et al., "Validation of distal limb mounted inertial-measurement-unit sensors for stride detection in Warmblood horses at walk and trot," *Equine Vet J.*, vol. 49, no. 4, pp. 545-551, 2017.
13. C.T. Ryan, B.L. Dallap Schaer, and D.M. Nunamaker, "A novel wireless data acquisition system for the measurement of hoof accelerations in the exercising horse," *Equine Vet J.*, vol. 38, no. 7, pp. 671-4, 2006.
14. M. Sapone, P. Martin, K. Ben Mansour, H. Château, and F. Marin, "Comparison of Trotting Stance Detection Methods from an Inertial Measurement Unit Mounted on the Horse's Limb," *Sensors (Basel)*, vol. 20, no. 10, pp. 2983-2983, 2020.
15. S. Mouloudi, H. Rahmanpanah, S. Gohari, C. Burvill, K.M. Tse, H. Davies, "What can artificial intelligence and machine learning tell us? A review of applications to equine biomechanical research," *J Mech Behav Biomed Mater.*, vol. 123, pp. 104728, 2021.
16. S. Bosch, et al., "EquiMoves: A Wireless Networked Inertial Measurement System for Objective Examination of Horse Gait," *Sensors (Basel)*, vol. 18, no. 3, pp. 850, 2018.
17. Y. Ho and S. Wookey, "The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling," *IEEE Access*, vol. 8, pp. 4806-4813, 2020.
18. M.A. Weishaupt, T. Wiestner, H.P. Hogg, P. Jordan, J.A. Auer "Compensatory load redistribution of horses with induced weightbearing hindlimb lameness trotting on a treadmill," *Equine Vet J.*, vol. 36, no. 8, pp. 727-33, 2004.
19. M.A. Weishaupt, T. Wiestner, H.P. Hogg, P. Jordan, J.A. Auer "Compensatory load redistribution of horses with induced weight-bearing forelimb lameness trotting on a treadmill," *Vet J.*, vol. 171, no. 1, pp. 135-146, 2006.
20. H. Darbandi, et al., "Using Different Combinations of Body-Mounted IMU Sensors to Estimate Speed of Horses—A Machine Learning Approach," *Sensors (Basel)*, vol. 21, no. 3, pp. 798, 2021.
21. H. Darbandi and P. Havinga, "A Machine Learning Approach to Analyze Rider's Effects on Horse Gait Using On-Body Inertial Sensors," 2022 IEEE International Conference on Pervasive Computing and Communications Workshops.
22. A. Abourachid, "A new way of analysing symmetrical and asymmetrical gaits in quadrupeds," *C R Biol.*, vol. 326, no. 7, pp. 625-630, 2003.