



HAL
open science

Preliminary results of YOLO-Based Recognition of Surface Markers in Aquatic Environments for Equine Motion Analysis Applications

Frédéric Marin, Chloé Giraudet, Claire Moiroud, Pauline Gaulmin, Audrey Beaumont, Emeline de Azevedo, Fabrice Audigié, Henry Chateau

► **To cite this version:**

Frédéric Marin, Chloé Giraudet, Claire Moiroud, Pauline Gaulmin, Audrey Beaumont, et al.. Preliminary results of YOLO-Based Recognition of Surface Markers in Aquatic Environments for Equine Motion Analysis Applications. 2024. hal-04737937

HAL Id: hal-04737937

<https://hal.science/hal-04737937v1>

Preprint submitted on 15 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Preliminary results of YOLO-Based Recognition of Surface Markers in Aquatic Environments for Equine Motion Analysis Applications

Frédéric Marin *

Université de technologie de Compiègne, France

Chloé Giraudet

Université de technologie de Compiègne, France

Claire Moiroud

Ecole Nationale Vétérinaire d'Alfort, France

Pauline Gaulmin

Ecole Nationale Vétérinaire d'Alfort, France

Audrey Beaumont

Ecole Nationale Vétérinaire d'Alfort, France

Emeline De Azevedo

Ecole Nationale Vétérinaire d'Alfort, France

Fabrice Audigié

Ecole Nationale Vétérinaire d'Alfort, France

Henry Chateau

Ecole Nationale Vétérinaire d'Alfort, France

***Corresponding author.**

Email : frederic.marin@utc.fr

Keywords: Motion Capture; Horse; Swim; Deep learning.

1. Introduction

Although markerless motion capture systems are currently under development and promising, the technique involving landmarks remains the gold standard either for obtaining labeled data for motion capture reference databases or conducting motion analysis in a conventional manner (Chaumeil et al. 2024). Although marker recognition techniques using retro-reflective spherical markers are well-established (Robertson et al. 2014), they prove impractical in aquatic environments. Consequently, the use of surface markers in natural light is preferred. Specifically, in the study of horse swimming for rehabilitation or muscle reinforcement, the use of color-differentiated markers corresponding to the horse's color and variations in light, which create shadows influenced by the animal's size in the swimming lane, pose challenges for conventional marker detection techniques (Giraudet, Moiroud, et al., 2023).

The purpose of this study was to assess the feasibility and reliability of using the deep convolutional neural networks (CNN) model of the YOLO-v2 type which allows the detection of objects in an image for the tracking of surface markers of various colors in the aquatic environment on the horse.

2. Methods

One horse was recorded swimming in the U-shaped pool at the Equine Clinic (CIRALE) of the National Veterinary College of Alfort, using six GoPro cameras (GoPro Hero 8 Black, GoPro, San Mateo, USA) with a resolution of 2.7K and a frame rate of 120 Hz. Anatomical landmarks on the horse were identified using surface markers applied with animal marking sticks (Raidex, Dettingen, Germany) (Giraudet, Moiroud, et al., 2023). This study was approved by the clinical research and ethics committee of the Alfort Veterinary School (protocol code: 2022-09-19). Each video underwent specific post-processing to correct distortions related to aquatic environments, and the surface markers were identified semi-automatically with manual corrections (Giraudet et al. 2023). A total of 2780 images with 17616 surface markers labeled were used. Each frame of the video was divided into 96 sub-images of a resolution of 190x208 px. Subsequently, a CNN YOLO v2 model based on pre-trained features extraction resnet-50 was trained (Redmon et al. 2016). A total of 29302 sub-images were used as training data set, 4884 as validation dataset and 14651 as test dataset. The training algorithm utilized the “adam” optimizer and a mini batch size of 20 over 5 epochs. The average precision of all marker detections was evaluated using the test dataset by comparing the markers detected by YOLO with the originally labeled markers. Computations were carried out using a DELL Precision 3660 computer equipped with an Intel Core i9 processor and an NVIDIA GeForce RTX 3080 graphics card. Calculations were performed with Matlab R2021b (The MathWorks Inc., Natick, MA, USA).

3. Results and discussion

The total training computation time with our setup was 18 minutes and 35 seconds, yielding a mean accuracy of 0.7 for the all detections. Initially, this accuracy may seem low, but it can be attributed to YOLO's capability to detect markers that were not originally labeled (Fig 1-D), including phantom markers specular reflected on the water surface (Fig 1-B), intentionally left unlabeled. Nonetheless, our observations reveal that YOLO detection outperforms semi-automatic method with manual corrections (Fig 1), by detecting a greater number of markers. Now, focusing only on markers originally labeled in the test set, we obtained a sensitivity of 0.97 and a specificity of 0.95 for marker detection by YOLO. The accuracy of the marker center

location by YOLO, compared to the originally labeled markers, has a median value of 2 pixels and is within 5 pixels in 95% of cases.

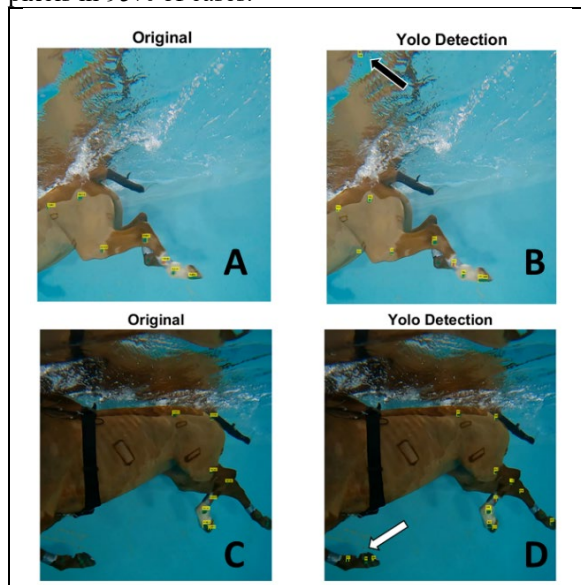


Figure 1. Sample marker detection with YOLO. The black arrow shows a detection of a phantom marker linked to the specular reflection of an existing marker. The white arrow indicates markers that were not labeled on the original image.

However, within the context of the study conducted here, even the semi-automatic detection of surface markers required a significant amount of manual correction work. The utilization of the YOLO algorithm for detection has demonstrated its relevance and reliability. Our approach's novelty lies in analyzing not the entire image, but rather sub-images, which allows for an adequate number of pixels related to the surface markers while reducing calculation time. Furthermore, employing a YOLO-v2 type facilitates reasonable computation times, even with a mid-range computer equipped with a moderately performing graphics card, all at an affordable cost. Training conducted here involved data augmentation and could be considered a potential area for improvement in the future.

Our preliminary study has limitations. Here we used videos made on a single horse. We limited the risk of overfitting, using the images of the 6 cameras that had 6 different viewpoints with variable lighting and contrasts. This choice was motivated by the fact that already in this configuration, we were limited in memory capacity on our workstation. More computational capacity is needed for a larger database.

4. Conclusions

Based on our preliminary observations, the use of CNN-YOLO-v2 appears to be well suited for surface landmarks detection even in an unconventional

environment such as underwater. This technique opens new perspectives on automating the laborious task of labeling markers on video images. The following work is to increase the size of databases with videos with other horses while using augmentation techniques. This implies computing power that is not in the mid-range computer capacity.

Funding

This research was funded by the Agence Nationale de la Recherche (ANR-20-CE19-0016).

Conflict of Interest Statement

none

References

- Chaumeil, A., Lahkar, B. K., Dumas, R., Muller, A., & Robert, T. (2024). Agreement between a markerless and a marker-based motion capture systems for balance related quantities. *Journal of Biomechanics*, *165*, 112018. doi: 10.1016/j.jbiomech.2024.112018
- Giraudet, C., Hattrisse, C., Macaire, C., Gaulmin, P., Moiroud, C., Ben Mansour, K., ... Marin, F. (2023). Development of a pipeline of 3D underwater motion capture : Application to the analysis of horse swimming. *48ème congrès de la Société de Biomécanique*, Grenoble, France: <https://hal.science/hal-04269709>
- Giraudet, C., Moiroud, C., Beaumont, A., Gaulmin, P., Hattrisse, C., Azevedo, E., ... Marin, F. (2023). Development of a Methodology for Low-Cost 3D Underwater Motion Capture : Application to the Biomechanics of Horse Swimming. *Sensors*, *23*(21), 8832. doi: 10.3390/s23218832
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once : Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779-788. Las Vegas, NV, USA: IEEE. doi: 10.1109/CVPR.2016.91
- Robertson, D. G. E., Caldwell, G. E., Hamill, J., Kamen, G., & Whittlesey, S. N. (2014). *Research methods in biomechanics* (Second edition). Champaign, Ill.: Human Kinetics.

Received date:04/04/2024

Accepted date: 28/06/2024

Published date: XX/XX/2024

Volume: 1

Publication year: 2024