From video to watts—predicting mechanical power in stationary cycling using computer vision

Pawel Kudzia, Peter Pham, Peter Cripton
School of Biomedical Engineering
Email: pawel.kudzia@ubc.ca

Summary

Cyclists rely on power output for training, but traditional power meters are expensive and require professional installation. To address this, we developed a video-based system that estimates mechanical power using computer vision and machine learning. We recorded synchronized video and cycling data from participants using frontal and sagittal cameras and a Smart Trainer. Using pose estimation, we extracted kinematic features used as input to train a neural network. We are experimenting with transformer-based architectures to optimize accuracy. Preliminary results from five participants suggest the model successfully tracks power fluctuations, with ongoing data collection (target n=30) expected to improve performance. This hardware-free system may provide a cost-effective alternative to power meters, making cycling performance metrics more accessible.

Introduction

Power output is a key metric for cycling performance, essential for training optimization and performance analysis. However, commercial power meters are expensive and require professional installation, limiting accessibility. Researchers have explored alternative low-cost methods using surrogate metrics like cadence and heart rate [1,2], but no existing system estimates power from video. We explore a computer vision-based system that estimates mechanical power output using pose estimation and neural networks to extract biomechanical features from video. By eliminating the need for dedicated hardware, our approach aims to democratize power-based training**,** making high-precision cycling analytics more affordable and accessible.

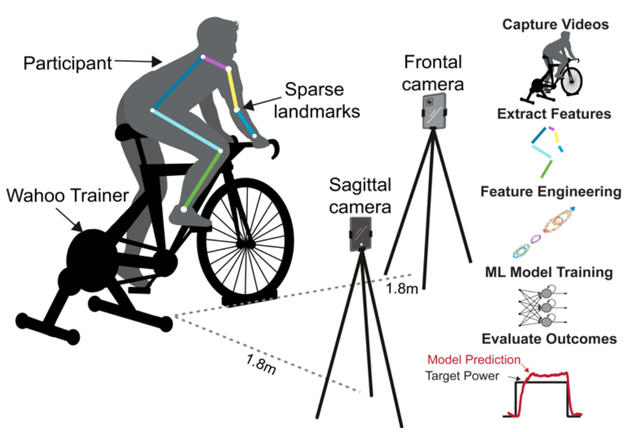

Figure 1: Our experimental setup and methods pipeline.

Methods

We developed an experimental setup to collect synchronized video and cycling data (**Fig. 1**). Two cameras positioned in frontal and sagittal planes recorded 60 Hz video of participants cycling on a Smart Trainer, while we simultaneously recorded power, cadence, and heart rate at 1 Hz. A software trigger synchronized all data streams. During each session, participants followed real-time power targets displayed on a screen. We randomized three power conditions—1 W/kg (low), 1.5 W/kg (medium), and 2 W/kg (high)—and kept gear settings constant to ensure consistency. We processed the video using MediaPose, a pose estimator, extracting sparse kinematic features. We then combined these features as an input to train a neural network for power prediction as an output. To develop the model, we tested transformer-based architectures, training these models using a leave-one-subject-out approach. We assessed model performance by comparing predicted power with empirical power using RMSE. At this point, we have collected data from five participants, with an ongoing target of n=30.

Results and Discussion

Our preliminary results suggest that the model effectively tracks power output trends, with predictions aligning well with trainer data (**Fig. 2**). Visual comparisons indicate that while the model captures power fluctuations, it may slightly underestimate higher intensities. Despite the small dataset (n=5), the model exhibits stable learning behaviour, demonstrating its potential for hardware-free power estimation. As we collect more data (target n=30), we expect to refine accuracy by incorporating additional biomechanical features such as leg extension velocity and pedal symmetry.

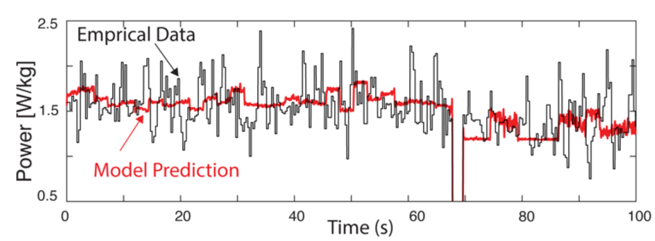

Figure 2: Comparing model to empirical data predictions.

Conclusions

This study presents a hardware-free, cost-effective alternative to traditional power meters, making high-precision cycling analytics more accessible. Beyond cycling, this technology could extend to sports science and real-time fitness tracking, offering camera-based biomechanical analysis without dedicated sensors.

References

[1] Townsend NE et al. (2018). *J. of Sports Sciences*, **36(9)**: 1012–1018.

[2] Karavirta L et al. (2019). *J of Sport Eng. and Tech*, **233(4):** 537–545.