autoWT: A Semi-Automated ML-Based Movement Tracking System for Performance Tracking and Analysis in Olympic Weightlifting

Gatis Jurkāns, Anikó Ekárt[a] and Ulysses Bernardet[b]

Aston University, Aston Triangle, Birmingham, U.K.

Keywords: Olympic Weightlifting, Human Action Recognition, AI Sport Coaching, Video Sports Analysis, Long Term Performance Tracking.

Abstract: As part of AI Coaching Assistant project research in sports performance tracking systems, we present autoWT, a novel semi-automated computer vision tracking system designed and developed for repeated long-term performance tracking of Olympic Weightlifting (OW) training. The system integrates multiple cameras and a heart rate sensor to capture, detect, and analyse OW movements, providing coaches and athletes with objective performance metrics. Key features include automated lift detection, clip extraction, and acquired performance metric visualisation based on markerless pose estimation data. The system architecture, consisting of a distributed system with multiple workers and a controller, enables efficient processing of high-bandwidth data streams. The paper provides an overall system architecture, operating principles and a detailed breakdown of action onset recognition and performance metric extraction system modules. We evaluate the system's lift detection accuracy and the repeatability of extracted performance metrics using data from Olympic lifts. Results demonstrate high accuracy in lift detection and consistent and repeatable metric extraction, indicating autoWT's potential as a valuable tool for conducting long-term Olympic weightlifting performance analysis studies and as an aid for coaching. The autoWT system can enhance the broader perspective and be an exemplary model for designing tracking systems in other sports.

1 INTRODUCTION

Significant progression in many sports disciplines, such as golf, shot put, or Olympic weightlifting, depends on gradual improvements of technical proficiency in a few complex dynamic movements.

Tracking and analysing these changes is traditionally the task of an experienced coach, who has spent years participating in, observing, and correcting these movements.

The AI Coaching Assistant (ACA) project aims to provide access to accurate tracking of long-term performance changes based on quantifiable and objective metrics using state-of-the-art computer vision and machine learning techniques.

Here, we present the auto Weightlifting Tracker (autoWT), a semi-automated long-term performance tracking system developed to support Olympic weightlifting training.

We have selected Olympic weightlifting (snatch and clean & jerk) movements because they are essential to strength and conditioning training in numerous sports disciplines such as judo, rugby, track and field, and CrossFit.

Olympic lifts are considered highly efficient for power development. However, these movements are often difficult to introduce into strength and conditioning programs due to their flat learning curve for acquiring movement efficiency. One challenge is that the core metric—weight on the barbell—can not directly inform about technical movement changes.

OW is well suited for long-term, high-resolution performance tracking, as the training process is spatially confined to being executed on weightlifting platforms. Multiple cameras can be permanently installed to capture the training. Our contribution proposes a novel system architecture and features that enable automated Olympic lift detection and annotated clip extraction for repeated data collection in long-term performance tracking studies.

To achieve the aforementioned functionality, the system needs to include the following features:

- To develop performance forecasting models integrating a combination of sensors, the system

[a] https://orcid.org/0000-0001-6967-5397
[b] https://orcid.org/0000-0003-4659-3035

Jurkāns, G., Ekárt, A. and Bernardet, U.
autoWT: A Semi-Automated ML-Based Movement Tracking System for Performance Tracking and Analysis in Olympic Weightlifting.
DOI: 10.5220/0012997400003828

In *Proceedings of the 12th International Conference on Sport Sciences Research and Technology Support (icSPORTS 2024)*, pages 60-71
ISBN: 978-989-758-719-1; ISSN: 2184-3201

requires continuous and synchronised recording from multiple sensors such as cameras, heart rate monitors, accelerometers, and positional transducers.

- We needed to develop a key action detection system to improve computational efficiency by reducing markerless pose model use and minimise coaches' need to curate individual Olympic lift clips.
- Ability to log supplementary data (exercise type, set/repetition count, used weight, additional commentary) – This data allows for easy annotation of the captured data of the extracted clips.
- System and metric data visualisation – Provides easy access to the system status and the ability to provide feedback using metrics extracted from Olympic lifts.
- System checks to ensure rigorous and repeatable data capture.

We will begin by reviewing related work in performance analysis technologies for barbell sports. Subsequently, we describe the system requirements, features, and architecture of autoWT. We will conclude with a description of the numeric evaluation of the system in terms of accuracy and repeatability.

2 RELATED WORK

2.1 Olympic Weightlifting

OW consists of two movements: snatch and clean & jerk. In both, a loaded barbell is lifted to the above-head position in one or two steps, and then the competition winner is awarded to the athlete with the highest combined weight. This study will refer to snatch and clean & jerk as Olympic lifts (see Figure 1) and any strength or bodybuilding movements trained using free weights or a barbell (barbell squats, bench press, clean pull, etc.) as assistance lifts.

2.2 Barpath Tracking in Olympic Weightlifting

The use of cameras for OW movement analysis goes back to the 1960s and 1970s when OW was a highly contested Olympic discipline during the Cold War era (Garhammer and Newton, 2013). This period established bar-path tracking as the preferred tool used by researchers interested in comparing the performance of Olympic lifts (AN, 1978). More recently improved bar tracking algorithms have been developed (Hsu et al., 2018), (Hsu et al., 2019), and there are now consumer-level tools such as BarSense, WL-Analysis, Dartfish and others that provide bar tracking. One of the first examples of a system to aid coaching in OW beyond bar-path tracking was implemented by (cha,) for extraction of 3D bar-path and performance metrics (barbell tilt and knee flexion angle) using depth data - Kinect cameras. While bar tracking is helpful, with the rapid advancement of computer vision and related computer science fields, many more tools and approaches have become available for extracting useful information from video data of human movement.

2.3 Action Recognition of Olympic Lifts

To perform long-term tracking of Olympic lifts using camera sensors, they must first be identified within the video stream data; this aim fits in the Human Action Recognition (HAR) sub-field of computer science. (Host and Ivašić-Kos, 2022) define categories for HAR in sports, according to which Olympic lifts are Individual Complex Actions - a combination of simple actions and interactions with an object. A popular method for detecting periodic activities, including exercise, is RepNet (Destro, 2024), which could be used for the repetition segmentation of assistance lifts like the squat and overhead press, yet for the detection of Olympic lifts, this is not applicable as the lifts are usually performed in sets ranging from 1 to 3 repetitions with each repetition usually intermittent by a pause between lifts. An Olympic lift onset detection method was developed by (Yoshikawa et al., 2010) using Cubic Higher-Order Local Auto-Correlation, which is a feature extraction method that captures complex spatial relationships computing third-order auto-correlation values of pixel values within a local neighbourhood. A strong feature of this method is the ability to detect the onset regardless of the different capture angles of the lifts. There are also convolutional neural network (CNN) based approaches using deep key-frame detection - detecting and extracting key positions of Olympic lifts from lift recordings (Jian et al., 2019), (Pan, 2022), (He et al., 2023). These methods are not applied for action recognition or lift onset detection but could be suited for these applications. The limiting factors for pure CNN-based systems are the high computational system requirements for data processing and the probabilistic nature of the key frame detection - delivering several frames associated with the same key position at high frame rates.

Figure 1: Examples of Olympic lifts. A - Snatch lift. B - Clean & Jerk lift.

2.4 Modern Computer Vision Use for Barbell Sports Performance Analysis

Olympic lifts are challenging to analyse due to their fast pace and highly technical nature. Currently, only two publications suggest technique adjustment and analysis methods for Olympic lifts using computer vision (Rethinam et al., 2023),(Bolarinwa et al., 2023). Markerless pose detection (MPD) is a core technology used in both papers. MPD enables the extraction of landmark locations of the body pose directly from image data. Several models are available; we use Google's - BlazePose (full-heavy) model implementation due to its reduced computational demand as being designed for application in mobile devices while delivering high pose estimation accuracy (Bazarevsky et al., 2020). MPD has shown reliable retrieval of body landmark data, the caveat being that there is a potential landmark offset compared to marker-based solutions and errors due to occlusion of body parts (Needham et al., 2021) (Mroz et al., 2021). (Rethinam et al., 2023) propose MPD data for an algorithm that calculates the athletes' centre of gravity extracted from their body proportions. This approach suggests that the algorithm can be used to determine the stability of the base (foot placement) for the athlete executing clean & jerk movement. (Bolarinwa et al., 2023) developed a system improving the refereeing process to reduce human bias in judging successful and failed attempts in OW. First, by recognising recovery parts of snatch and clean & jerk movements, a neural network classifies lifts as complete or incomplete. Then, MPD is used for lift analysis to determine common technique breakdowns, such as the press-out rule.

Several studies suggest MPD use to correct the technique of assistance lifts outside of OW applications, as they are technically much simpler and periodic. A good example is (Lin and Jian, 2022), where an algorithm was developed for assisting deadlift form correction.

(Arandjelović, 2017) did not use MPD, but developed an entire "monitoring-assessment-adjustment" loop for powerlifting exercises - squat and bench press, which allows for analysis and performance forecasting based on simulation and personalised athlete profiles.

Olympic weightlifting has a long history of using camera-based data analysis for performance tracking, which had historically been accessible only to sports scientists exploring the biomechanics of Olympic lifts. The last two decades of rapid developments in computer vision and machine learning have built algorithms that now can give greater access to pose data estimation directly from video footage. Some work has been shown to utilize these tools for real-time adjustment of simple periodic assistance lifts and for judging technical aspects of individual executed Olympic lifts. Based on our literature review, we have identified a gap in research for building an automated system for repeated long-term performance tracking in OW, integrating lift detection, performance metric extraction and feedback.

3 ARCHITECTURE

In this section, we will highlight the key features and introduce the overall selected system architecture. We will then describe in depth the two core modules, the Lift Detector and Metric Extractor.

3.1 autoWT Features

The features selected and implemented as part of autoWT aim to serve the goals of the AI Coaching Assistant project - to improve the coaching experience through long-term tracking of athletes using computer vision and machine learning. autoWT enables long-term data collection studies to develop performance forecasting models using markerless pose estimation data. Such research studies require long-term data collection, where athletes need to use the system to record their training sessions repeatedly throughout whole training blocks. This is achieved by easing the integration of autoWT as part of existing training practices; we selected features that enable data collection and provide benefits to the training process, encouraging the system's use.

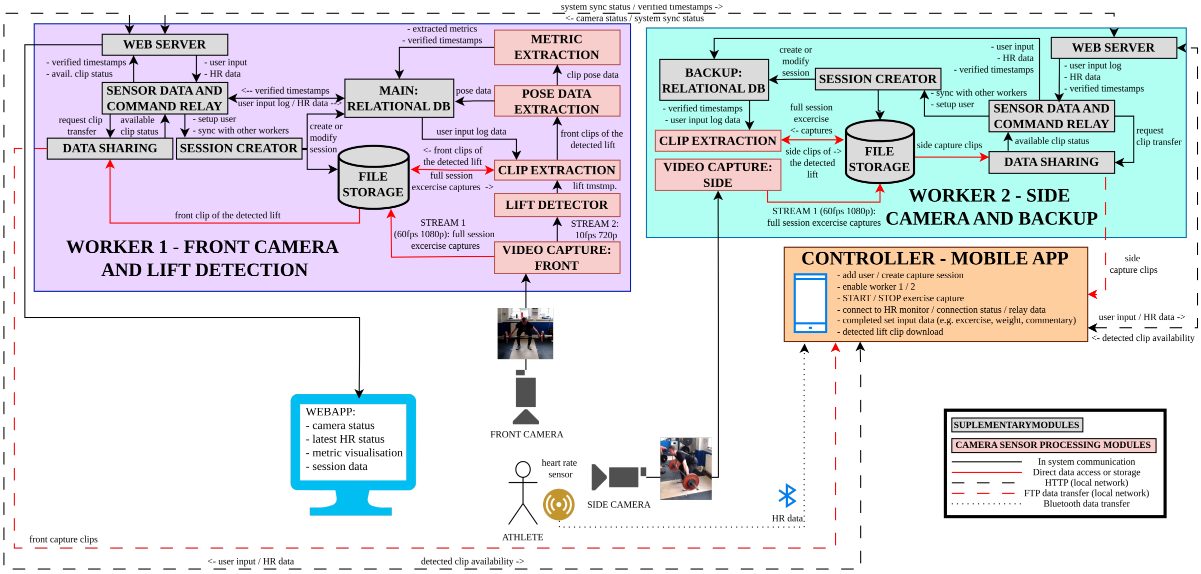


Figure 2: Auto weightlifting tracker architecture.

3.1.1 Multi Camera Capture with Detected Olympic Lift Extraction

Camera use for performance analysis and review is a common practice in OW. Yet, it is often left out as the production of individual clips and their annotation is time-consuming and labour-intensive. AutoWT automates lift detection and clip extraction with easy access to the user. By providing this sought-out feature, we enable coaches and athletes to direct more focus on training. The extracted clips are also annotated with a unique identifier number and weight used on the barbell, extracted from the user-logged data.

3.1.2 User Data Logging and Visualisation

Besides lift data accessible through the controller application, autoWT worker systems serve a WebApp displayed on a large screen during system use. The WebApp gives easy access to system information: camera capture status, detected lift performance metric visualisations, latest heart rate sensor readings, and tabulated user-logged data. The ability to quickly see the progression of the current training session in a singular overview encourages the logging of the training session data.

3.1.3 Multi Sensor Data Capture

Currently, autoWT allows for camera and heart rate sensor data capture. The system enables the setting of heart rate warning thresholds. When the user's heart rate is above the threshold, visual cues flag this change by increasing the font and changing the colour of the displayed heart rate readings. This data can enable a simple method for tracking athletes' exertion levels and inform rest time selection. This data collection supports future research exploring athlete fatigue level forecasting based on combined heart rate tracking and performance metric data.

3.2 autoWT System Procedure

Before the system's first use, the controller system and the workers are set up on the same local network, the cameras are positioned in permanent static locations, and markers are set on the weightlifting platform to indicate the setup starting position when lifting.

The worker systems are first powered on when using the system, which starts a series of automated scripts enabling autoWT software. Then, when the controller app is opened, it waits to receive synchronising checks from the worker systems. Once workers are available, the controller app enables the initialisation of a user using the Session Creator module to create new storage locations and database entries associated with a new training session in all worker systems.

After initialisation, athletes can start using cameras and the heart rate sensor - these sensors can be used simultaneously or independently. As soon as the user enables the heart rate sensor on the controller app, the sensor data starts streaming to workers and appears for display on the WebApp.

The software is designed to automate the exercise capture and analysis, so when capturing an exercise, the athlete only needs to start and stop capturing when changing to a new exercise. To start camera capture, the user selects the intended exercise from a list on the controller app. Depending on the exercise selection, different autoWT features will be enabled. If Olympic lifts are selected, worker modules associated with lift detection and metric extraction will be enabled. After exercise selection and pressing the “start capture” button, worker systems are notified to start recording using the front and side cameras simultaneously.

When recording Olympic lifts, detected and extracted clips populate the available clips table on the controller app, which, upon selection, can be downloaded to the device. Additionally, if the Snatch movement is being recorded, the metric extraction system generates the Snatch Pull Height (SPH) metric, visualised on the WebApp.

During the capture, the athlete can log information about completed sets - set/repetitions count and additional commentary on execution. Each time a new set is added, the information appears on the WebApp in a tabulated format. Every 5 seconds, the WebApp updates newly added sets or calculated metrics available from the previously executed sets. Once the athlete has completed the exercise, they press the “stop capture” button, notifying termination of exercise capture; once the lift processing modules complete any outstanding tasks, they stop until a new exercise recording commences, repeating the process.

3.3 System Design Considerations and Architecture Selection

To implement the functionality described, the system requirements include:

- Concurrently process multiple sensor data streams, such as capturing and encoding multiple high-bandwidth video streams and timestamped physiology data, including heart rate sensor data.
- The system needs to be enabled by single-board computers rather than GPU-enabled, high-powered PC systems.
- System needs to be accessible via low-specification devices.
- Function effectively with unreliable external network access.

Given these requirements, a distributed architecture was selected because it can handle high-bandwidth data processing through parallel processing, support low-specification control devices by offloading intensive tasks to distributed workers, and operate effectively without consistent network access. autoWT comprises two worker sub-systems that handle the processing for camera sensor data combined with a single controller system (MobileApp) that coordinates activities and relays data for a heart rate monitor sensor. Each worker sub-system comprises modules covering different system functions divided into core and sensor processing modules. The overall system architecture and the interactions between sub-systems and their specific modules can be seen in Figure 2.

3.3.1 Worker Sub-System - Core Modules

Core modules are crucial for integrating the worker sub-systems into the autoWT system. The current autoWT core modules include the Web Server, which enables communication between systems through the local network using the HTTP protocol and can host a WebApp to display system status information. Another module is the Sensor/Command Relay, which processes command packets received from the Web Server and forwards them to the relevant sub-system modules. Additionally, there is the Data Sharing module, which controls an FTP server to transmit extracted Olympic lift video clips to the MobileApp (controller). Finally, the Session Creator module manages the creation and administration of the ongoing session, including file storage and database setup.

3.3.2 Worker Sub-System - Sensor Processing Modules

The sensor processing modules - capture, process, and analyze sensor data, specifically camera sensor data. The video capture module manages the capture and storage of multiple video streams. Meanwhile, the Lift Detector module (Worker 1) uses a low-framerate video stream to detect Olympic lift start and end onset timestamps, with a more detailed description in the Lift Detector section. Following this, the Clip Extraction module (Worker 1) generates individual annotated clips from the stored high-framerate stream data using detected onsets and user-logged data. Then, the Pose Data Extraction module (Worker 1) applies the markerless pose model to extracted clips and stores the pose in the relational database. Finally, the Metric Extraction module (Worker 1) extracts performance metrics for visualisation by the WebApp - more details in the Metric Extractor section.

3.4 Lift Detector Module

The Lift Detector is the module that detects the start and end of Olympic lift movements from a low-framerate video stream. It can currently detect snatch

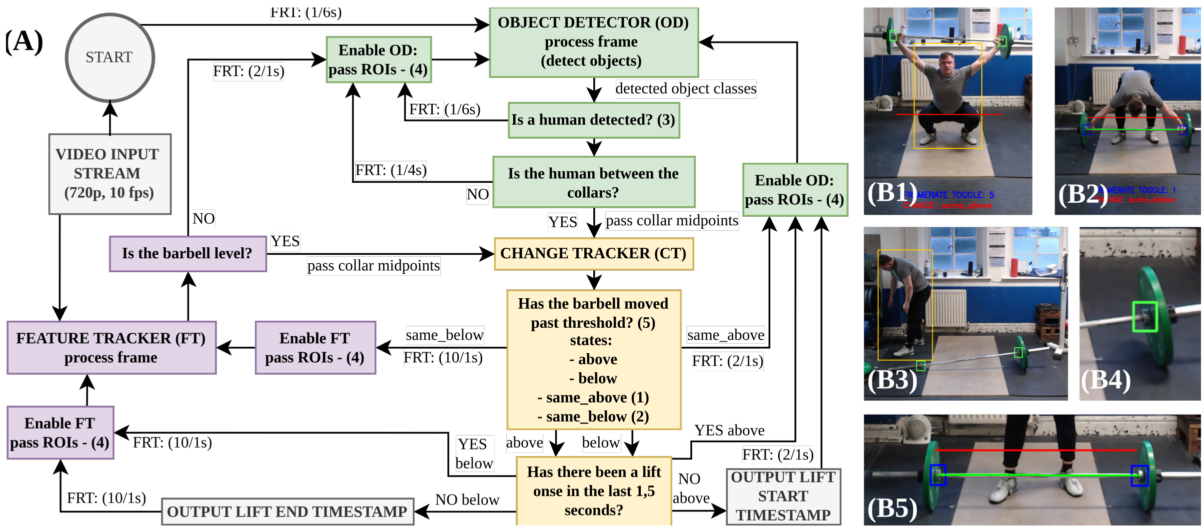


Figure 3: Lift Detector module state diagram. (A) The state diagram shows the interaction between three parts: the Feature Tracker (FT), Object Detector (OD), and Change Tracker (CT). FRT - Framerate toggle indicates an adjustment to the input framerate for the OD and the FT, e.g. (1/6s) (1 frame per 6 seconds). (B) A series of images detailing parts of the state diagram. (B1) Lift detector output during the "same_above" state. (B2) Output during the "same_below" state; (B3) is an example of output when a human is detected in the frame but not between the barbell collars; (B4) Example of a barbell collar Region of interest (ROI), identified by the object detector. (B5) ROIs tracked by the feature tracker; the red line is the threshold line, and the green line is the barbell identified by the midpoints of the barbell collar ROIs.

and clean & jerk movement onsets with high accuracy (see Evaluation - Lift Detector Testing). The "Change Tracker” component of the Lift Detector tracks the barbell’s position around a threshold line, set based on readings of the barbell’s location while stationary on the floor (Figure 3). Depending on whether the barbell is detected above or below the threshold and whether a person is detected in the frame or between the barbell collars, the system switches between the “Object Detector” and the "Feature Tracker”. This design enables efficient lift onset detection with optimised system resource use.

3.4.1 Change Tracker

The Change Tracker is the go-between the Object Detector and the Feature Tracker, keeping track of the state changes of the overall lift detector module. There are four states in the system:

- above —- the barbell just moved past the threshold line, triggering the lift start onset detection;
- same_above —- the barbell continues to be above the threshold;
- below—the barbell just moved below the threshold line, triggering the lift end onset detection;
- same_below —- the barbell continues to be below the threshold line.

In addition to tracking the lift detector state, the change tracker employs additional checks for improved performance. First, to reduce false positive consecutive onset detections due to the barbell being thrown on the ground and bouncing off the floor past the threshold line, the change tracker checks if an onset has been detected in the last 1.5 seconds. Second, the change tracker adjusts the height of the threshold line above the barbell based on the last 400 readings while the barbell is level on the floor.

3.4.2 Object Detector

The Object Detector is a fine-tuned Region-based Convolutional Neural Network (R-CNN) trained on a set of 1400 images of Olympic lifts. The images were manually annotated with two classes of objects: Collars and Humans. Collars are the inner part of the weightlifting barbell (see Figure 3(B4) - example for region of interest of the barbell inner collar). When both humans and collars are detected in the frame, additional checks are performed to ensure the correct object detection, such as checking if the person is detected between the collars at the same height in the image frame. Then, if the barbell is below the change threshold line and the additional checks are passed, the system switches to the Feature Tracker. Depending on the Lift Detector’s overall state, the Object Detector will read the input frames at a rate between 1 frame every 6 seconds to 2 frames every second. This low frame rate is due to the high system performance needs of the R-CNN, which is why we use it in combination with a more traditional Feature Tracker system.

3.4.3 Feature Tracker

Once the Object Detector passes the barbell collar ROIs, the Feature Tracker acquires the stream, tracking each ROI at an increased frame rate of 10 frames per second. The Channel and Spatial Reliability Tracking feature tracker allows for efficient barbell tracking but is prone to drift when a single side of the barbell is moved, creating partial occlusion, or when the barbell is accelerating very rapidly, creating motion blur and, therefore, losing features being tracked. To avoid drift, the feature tracker checks that the barbell is level and switches to the object detector when the barbell is uneven. To prevent the feature tracker from losing collars due to barbell acceleration, we only use the feature tracker up to the threshold line, past which the object detector is used. This works because the mechanics of the Olympic lifts have a slow initial acceleration pattern from the floor; the barbell never moves fast enough in this region for the feature tracker to lose tracking.

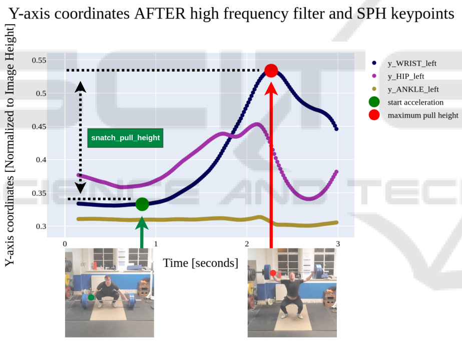


Figure 4: Snatch Pull height metric. Y-axis markerless pose estimate data coordinates after filtering with SPH metric key points added.

3.5 Metric Extraction Module

The Lift Detector identified onsets are used to extract high-framerate clips. Then, the Metric Extraction module uses markerless pose data extracted from the clips to generate defined metrics. Following discussions with Olympic weightlifting coaching staff supporting the AI Coaching Assistant project, several metrics were identified as potentially helpful in tracking long-term performance. One such metric is snatch pull height SPH, which is the distance that the barbell travels between the start of the lift and the maximum height the barbell reaches following the last pull of the snatch (see Figure 4); the metric can identify snatch lift performance efficiency when examined for maximal lifts. Additionally, we propose that SPH can help effectively analyse athletes' ability to adjust to weight increases during multiple repetition sets. The following is an example of an analytical approach for extracting the key points of interest from the markerless pose data necessary to generate the snatch pull height metric.

3.5.1 Extracting Key Points for Snatch Pull Height Metric

To calculate the metric, we primarily use the y-axis data for the wrist movement. First, the pose data is filtered to remove high-frequency noise using a rolling average filter with a window size of 10, which is 1/6th of a second for a capture frame rate of 60fps. We find two key points to calculate SPH: the start of the snatch movement and the maximum pull height. The differences between these two points result in the SPH metric (see Figure 4).

3.5.2 Finding Maximum Pull Height Point

To determine the maximum point where the wrist reaches its peak height during a pull, we locate the first local maximum peak in the upward movement of the wrist along the y-axis that is greater than the average height of the wrist when the barbell is on the floor, plus an additional 20%.

3.5.3 Finding Start Acceleration Point

The onsets provided by the lift detector give an approximate start and end position for the lift movements. However, our identified key points must be exact, as the start acceleration point is essential for calculating many potential performance metrics. This task is not trivial. We cannot simply use a single signal, e.g., y-axis points, for the wrist movement, as the athlete's body is not fully rigid; there is a slight offset for different body landmarks when the body starts moving upwards before the barbell leaves the ground. At the same time, just before the barbell starts accelerating from the floor, the athlete remains steady for a fraction of a second as the force transfers into the barbell before the combined system starts moving upwards. By comparing the y-axis movement across multiple landmarks, we can find this steady period and define the start acceleration point at the start of this period using Algorithm 1. The algorithm initially filters the pose data only to include the specified y-axis signals from the beginning of the lift recording to the maximum pull index. Then, for each signal, standard deviations over a rolling window are calculated, and the mean of these deviations is used to calculate a threshold value. Finally, to determine the start acceleration point, all the signals within the defined window are compared to the threshold; if the start and end values are below the set threshold, a steady period is identified. If a steady period is not found, the window size is reduced, and the process is repeated. As the steady period is a fraction of a second, we selected the window size as 15 or below for our system. Given that our system's pose data is extracted from recordings at 60fps, the window equals 15 frames or less than a quarter of a second.

Input: Pose data D (all 33 y points),
Maximum pull height index M
Output: start_acc_point index
Initialize Parameters:
Define signals_of_interest: $\{s_1, s_2, \ldots, s_m\} \subseteq$ {"wrist_y", "hip_y", "ankle_y", ...};
Filter D to D' with signals of interest;
Truncate D' at M, resulting in D'';
Let n be the number of data points in D'';
Iterate over window sizes:
Define $W = \{15, 14, \ldots, 5\}$;
foreach *window_size* $w \in W$ **do**
    Calculate Tolerance:
    foreach *signal* $s \in \{s_1, s_2, \ldots, s_m\}$ **do**
        Compute $\sigma_{t,s}$ over a rolling window of size w for $t \in \{1, 2, \ldots, n-w+1\}$ in D'';
        Calculate mean standard deviation:

$$\sigma_{w,s} = \frac{1}{n-w+1} \sum_{t=1}^{n-w+1} \sigma_{t,s}$$

    end
    Calculate overall mean:

$$\sigma_w = \frac{1}{m} \sum_{s \in \{s_1, s_2, \ldots, s_m\}} \sigma_{w,s}$$

    Define tolerance $\tau = \frac{\sigma_w}{2}$;
    Check Moving Averages:
    for $i = n$ ***down to*** w **do**
        if $\forall s \in \{s_1, s_2, \ldots, s_m\}, \left| D''_{i,s} - D''_{i-w,s} \right| < \tau$ **then**
            Set start_acc_point to $i - w$;
            break;
        end
    end
    if *start_acc_point is set* **then**
        break;
    end
end

Algorithm 1: Find start acceleration algorithm.

4 EVALUATION

A data collection study was conducted to provide data, first, to develop and test the Lift Detector module, and second, to test the extracted metric repeatability across and within capture sessions using the integrated markerless pose model in the Metric Extraction module.

We captured full training sessions using the base autoWT system of two cameras and a heart rate monitor, capturing Olympic and assistance lifts. Following each capture session, data was annotated manually, with individual start and end timestamps of each executed Olympic lift and start and end timestamps for executed set for assistance movements. Additionally, user data was logged, including set, repetition numbers, weight on the barbell, and additional coach commentary. Data was collected using the following hardware:

- Controller - MobileApp: Samsung Galaxy Tab 6A (Android 8.1)
- Workers: 2 * (single board computers) - Jetson Xavier NX (Ubuntu 22.04 JetPack)
- Cameras: 2 * (web cameras) - Logitech StreamcCam
- Heart Rate Monitor: Polar H10

Eight athletes (7m,1f) participated in a total of 21 training sessions; each participant conducted between 1 to 6 full sessions, providing us with 381 snatch and 190 clean & jerk recordings. The following sections explore insights gained from this data.

4.1 Lift Detector Testing

Apart from the data used to train the Object Detector for the Lift Detector, out of the twenty-one sessions, four full capture sessions, each featuring a different athlete performing Olympic lifts, were set aside for testing. This dataset comprises 4 hours and 23 minutes of complete recordings, with 268 manually annotated timestamps marking the start and end of each lift.

4.1.1 Lift Detector Test Procedure

The ground truth timestamps were annotated to include the whole lift, with the start timestamp being in an interval of ±2 seconds of the actual start of the lift and the end being recorded as ±2 seconds of the

Table 1: Lift Detector Test Results. LD detections - total number of all detections by the Lift Detector. AP - Actual Positives, TP - True Positives, TN - True Negatives, FP - False Positives, FN - False Negatives, MCC - Matthew's Correlation Coefficient, JCC - Jaccard Coefficient.

Participant	Exercise	LD detections	AP	TP	TN	FP	FN	Precision	Recall	F1	MCC	JCC
p1	cnj	27	28	27	1783	0	1	1.000	0.964	0.982	0.982	0.964
p1	snatch	70	68	60	1918	10	8	0.857	0.882	0.870	0.865	0.769
p14	cnj	24	22	22	2275	2	0	0.917	1.000	0.957	0.957	0.917
p14	snatch	42	40	37	2468	5	3	0.881	0.925	0.902	0.901	0.822
p6	cnj	21	26	21	1825	0	5	1.000	0.808	0.894	0.897	0.808
p6	snatch	33	40	30	2040	3	10	0.909	0.750	0.822	0.823	0.698
p9	cnj	23	20	17	1720	6	3	0.739	0.850	0.791	0.790	0.654
p9	snatch	24	24	23	1432	1	1	0.958	0.958	0.958	0.958	0.920
TOTAL_SNATCH	**snatch**	**169**	**172**	**150**	**7858**	**19**	**22**	**0.888**	**0.872**	**0.880**	**0.877**	**0.785**
TOTAL_CNJ	**cnj**	**95**	**96**	**87**	**7603**	**8**	**9**	**0.916**	**0.906**	**0.911**	**0.910**	**0.837**
TOTAL	**ALL**	**264**	**268**	**237**	**15461**	**27**	**31**	**0.898**	**0.884**	**0.891**	**0.889**	**0.803**

barbell being dropped back on the floor. Yet the Lift Detector detects onsets once the barbell moves past a threshold above or below the barbell. We consider that the detector detects an onset correctly if the correct type of onset (start or end) is detected within ± 2 seconds of the annotated onset. We treat each recording as a time series of length equal to the number of seconds (assuming one frame/second). This means that a recording of length N with an onset at moment t and a time window around the onset from $t-n$ to $t+n$ will have N data points;

- Time window within which we accept one positive data example: $t-n....t+n$
- Negative data examples: $0...t-n-1$ and $t+n+1...N$

For each recording and the lift detector identifying an onset at some moment d, we will have N data points in time, as seen in Figure 5.

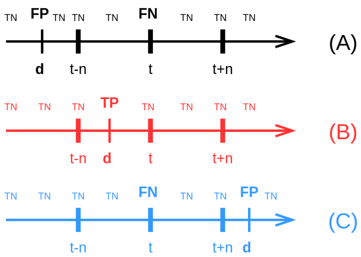


Figure 5: Lift detector onset detection. A - $(d < t-n)$; B - $(t-n \leq d \leq t+n)$; C - $(d > t+n)$.

The counting of the different categories for each recording can be simplified:

$$ift - n < d < t + n:$$

$$TN = N-1, TP = 1, FP = 0, FN = 0$$

$$Otherwise:$$

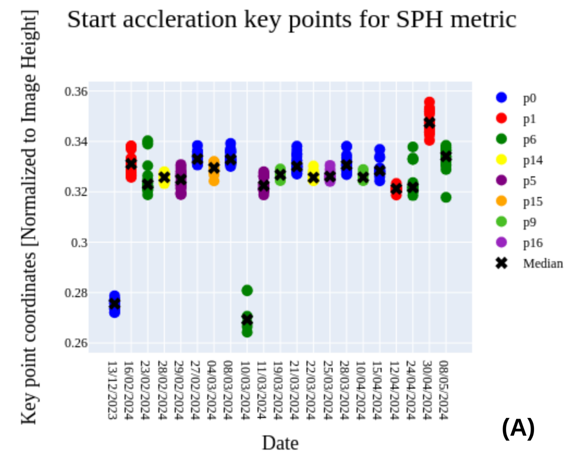
$$TN = N-2, TP = 0, FP = 1, FN = 1$$

4.1.2 Lift Detector Test Findings

We used five metrics to measure performance: precision, recall, F1 score, Matthew's correlation coefficient, and the Jaccard Coefficient. These metrics mitigated the misleading effects of dataset imbalances caused by the high number of True Negatives. It provided a more accurate assessment of the test's performance by focusing on both positive and negative outcomes. Table 1 shows the breakdown of the results across individual participants in either snatch or clean & jerk movements. The Lift Detector demonstrates overall solid performance with total test values across both Olympic lifts of 0.898 and 0.884 for precision and recall, 0.891 for the F1 score, 0.889 for the MCC, and 0.803 for the Jaccard coefficient. The values indicate that the implemented detector system has balanced precision and recall, a strong correlation between predicted and actual classifications, and a significant overlap between predicted positives and actual positives. This suggests that the Lift Detector is highly effective at correctly identifying true positive cases, indicating lift onsets while maintaining a low rate of false positives and negatives.

4.2 Metric Extractor Testing

The AutoWT system relies on integrating markerless pose as a key technology enabling performance data collection. The system is intended for long-term data collection through repeated capture sessions. The underlying measurement system—the markerless pose model—must deliver repeatable measurements given the same conditions, i.e., the same results are obtained repeatedly under unchanged conditions. Additionally, based on this repeatability assumption, the autoWT system should be able to identify deviations in the extracted key point data if the camera setup has been altered.



Participant	Date	median	CV	Shapiro-Willk (reject H0, a=0.01)	p-val	sample size
p0	13/12/2023	0.2756	0.671	TRUE	8.54E-01	18
	27/02/2024	0.3330	0.570	TRUE	1.36E-01	24
	08/03/2024	0.3328	0.663	TRUE	7.52E-02	28
	21/03/2024	0.3301	0.860	TRUE	2.77E-02	25
	28/03/2024	0.3306	0.964	TRUE	9.00E-02	18
	15/04/2024	0.3284	1.061	TRUE	1.01E-01	13
p1	16/02/2024	0.3311	1.193	TRUE	9.07E-02	22
	12/04/2024	0.3212	0.361	TRUE	5.30E-01	11
	30/04/2024	0.3474	1.141	TRUE	6.36E-01	21
p14	28/02/2024	0.3257	0.368	TRUE	4.43E-01	17
	22/03/2024	0.3256	0.481	TRUE	2.72E-02	19
p15	04/03/2024	0.3295	0.608	TRUE	1.87E-01	17
p16	25/03/2024	0.3261	0.443	TRUE	1.86E-02	18
p5	29/02/2024	0.3249	1.176	TRUE	2.63E-01	20
	11/03/2024	0.3225	0.846	TRUE	4.13E-02	23
p6	23/02/2024	0.3230	2.198	FALSE	1.29E-03	16
	10/03/2024	0.2692	2.680	TRUE	1.52E-01	6
	24/04/2024	0.3217	1.735	FALSE	7.74E-05	19
	08/05/2024	0.3341	1.409	FALSE	8.46E-04	19
p9	19/03/2024	0.3268	0.420	TRUE	9.85E-01	12
	10/04/2024	0.3257	0.430	TRUE	1.43E-01	10

(B)

Figure 6: Start Acceleration key points across all capture sessions from the data collection study. A - graphic showing the distributions of the key points in each capture session. B - Table showing each capture session: Each session data point distribution median, CV - Coefficient of Variation, Shapiro-Wilk test results for distribution normality and sample size.

4.2.1 Start Acceleration Point Distribution and Repeatability

We will examine the start acceleration key points to test the data repeatability. Given that the front camera has not been moved, the key point we extract for the same participant across capture sessions should deliver the same y-axis landmark data within a normal distribution.

We use the Shapiro-Wilk test to determine the normality of the start acceleration point distributions for each capture session. Before testing, extreme outliers—values greater than three standard deviations—were removed (5/384 key points). The test results show that 18 of the 21 sessions tested were Gaussian. The test results can be seen in Figure 6 - B.

After testing that the distributions are normal, we can compare the distributions of points using central tendency measures, such as the median, and variance measures, such as the coefficient of variation.

To highlight other indications of the data repeatability, we draw attention to a few additional notes. As this data is based on the wrist point detections from the markerless pose model, they are slightly different for each athlete. Yet we can observe that the median of the data and the spread are similar for participants who have multiple sessions (e.g., p0, p5, p14). This can also be seen when looking at the coefficients of variation in Figure 6 - B and the shapes of the distributions in Figure 6 - A.

4.2.2 Detecting Camera Offset Using Start Acceleration Points

We know the front camera used for data collection was not moved during the captured sessions. Therefore, we expect the wrist points to lie within a normal distribution for each captured training session, which we have shown for most sessions.

When setting up the autoWT system for data collection, the weightlifting gym staff and regular members were instructed not to alter the setup. However, the cameras were adjusted during the data capture study period due to space restrictions - the front camera was near a weight stand and was moved by accident. As the camera mounts are static, they can only be moved up, down, or side to side. This should be noticeable in the data as an offset for the session following the alteration.

By reviewing the session key point distributions and their median values in Figure 6 - A, we can see that the camera was moved three times during the study period: following the first session on 13/12/23, then following the sessions on 8/3/24 and 24/04/24, respectively. Each time following a session, when the camera was tilted, it was subsequently readjusted, but the newly adjusted height did not match exactly with the previous viewing angle.

Even when accounting for slight variations in the athlete's setup position before the lift, we found that the captured key points are normally distributed and can be acquired with high repeatability across repeated capture sessions. Lastly, we observed that deviations for the median of the start acceleration point distributions can be used to determine if the camera

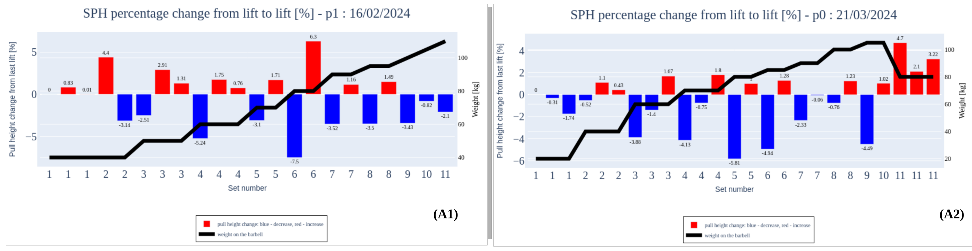


Figure 7: A1 and A2 - Examples of the Snatch Pull Height metric visualised as a percentage change from lift to lift across a full snatch capture session for participants p0 and p1. Figures display the snatch pull height change as a bar chart with pull height increase in blue and decrease in red; the weight on the barbell for each repetition and the set is overlaid as a black line with the y-axis markings on the right side.

position has been altered, therefore giving a method to flag setup changes and help improve the rigour of the data capture process.

5 CONCLUSIONS

This paper presents autoWT, a system for enhancing research and coaching practice in Olympic weightlifting training through automated, long-term performance tracking. The distributed architecture efficiently manages multiple high-bandwidth camera data streams and supports low-specification control devices for processing and analysing Olympic weightlifting movements. Key features such as multi-camera capture, user data logging, and heart rate monitoring have been carefully selected to build a versatile, powerful, yet easily deployable system.

Our Lift Detector module automates Olympic lift onset detection with high precision (0.898) and recall (0.884) across snatch and clean & jerk movements, providing reliable automated clip extraction for further analysis. The Metric Extraction module, exemplified by the snatch pull height (SPH) metric, showcases the system's ability to provide meaningful performance data.

Furthermore, our analysis of start acceleration key points across multiple sessions demonstrates the reliability of using markerless pose estimation for long-term performance tracking, with high repeatability and normally distributed data. We have also shown that deviations in data distribution can give the autoWT the ability to flag camera position changes, enhancing data capture rigour and ensuring consistent measurement conditions.

Despite these advancements, the current study has limitations, as not all aspects of the OW training process have been integrated into the autoWT system. Future work includes improvements to the Lift Detector module to automate periodic assistance movement action recognition and individual repetition segmentation. A full comprehensive system validation will be needed using a larger more diverse dataset with more participants.

Currently, autoWT computes a single performance metric—snatch pull height. The range of extracted metrics will be expanded to cover clean & jerk, and assistance movements to enable performance tracking across the entire OW training process. The planned metrics will include the clean pull height, jerk displacement, and jerk velocity, with each metric informing us about lift efficiency changes.

Currently, the user interface of the controller app and the WebApp are mainly geared towards control of the system by an operator. In the future, the UI will be further developed not only to allow control by end users but also to provide users with instant performance feedback. Figure 7 shows a proposed version of the SPH metric feedback for observing performance adjustments to weight increases across multiple repetitions. This visualisation can effectively show how SPH changes during a training session and inform the coach of different performance patterns. In Figure 7, we draw attention to the changes from set to set. Both athletes use comparatively light weights for the first two sets, which do not give enough resistance to control the weight with significant precision. Yet from set 3 onwards, two different patterns for adjusting to increased weight from the previous set become apparent. Participant 0 consistently shows a more substantial drop when weight increases with a slight increase in pull height in subsequent repetitions. At the same time, Participant 1 also shows a decrease in pull height but with an increase in every consecutive repetition. The overall relative change range for the athletes is similar. However, participant 1's ability to increase the pull height for every following set aligns with the athlete's overall higher strength base than Participant 0. The autoWT system will be used to collect data for longitudinal studies to investigate the effect of providing short-term and long-term feedback to users using the identified metrics and data visualisations. Notably, the data from these studies will serve to develop performance forecasting models integrating data from multiple sensors, including the combination of heart rate data and performance metrics acquired from the camera sensors. By addressing these areas, the autoWT system can further contribute to performance optimization in Olympic weightlifting and serve as an example for long-term performance research of other complex sports movements.

In conclusion, the autoWT system offers a promising approach to objective, repeatable, long-term performance tracking. The system has the potential to transform research and coaching practice, opening new avenues for future performance optimization in Olympic weightlifting.

ACKNOWLEDGEMENTS

This research was funded and supported by the EPSRC's DTP, Grant EP/T518128/1 and the industrial partner - Gymshark. Additional thanks to Gian Singh Cheema and Warley Weightlifting Club.

REFERENCES

Three-dimensional monitoring of weightlifting for computer assisted training. In *VRIC*, Laval France.

AN, V. (1978). *Weightlifting*. International Weightlifting Federation, Budapest.

Arandjelović, O. (2017). Computer-Aided Parameter Selection for Resistance Exercise Using Machine Vision-Based Capability Profile Estimation. *Augmented Human Research*, 2(1):4.

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., and Grundmann, M. (2020). BlazePose: On-device Real-time Body Pose tracking. Number: arXiv:2006.10204 arXiv:2006.10204 [cs].

Bolarinwa, D., Qazi, N., and Ghazanfar, M. (2023). Shifting the Weight: Applications of AI in Olympic Weightlifting. In *2023 IEEE 28th(PRDC)*, pages 319–326. ISSN: 2473-3105.

Destro, M. (2024). materight/RepNet-pytorch. original-date: 2023-02-28T15:28:43Z.

Garhammer, J. and Newton, H. (2013). Applied Video Analysis for Coaches: Weightlifting Examples. *International Journal of Sports Science & Coaching*, 8(3):581–594.

He, Q., Li, W., Tang, W., and Xu, B. (2023). Recognition to weightlifting postures using convolutional neural networks with evaluation mechanism. *Measurement and Control*, page 00202940231215378. Publisher: SAGE Publications Ltd.

Host, K. and Ivašić-Kos, M. (2022). An overview of Human Action Recognition in sports based on Computer Vision. *Heliyon*, 8(6):e09633.

Hsu, C.-T., Ho, W.-H., and Chen, J.-S. (2019). High Efficient Weightlifting Barbell Tracking Algorithm Based on Diamond Search Strategy. In Arkusz, K., Będziński, R., Klekiel, T., and Piszczatowski, S., editors, *Biomechanics in Medicine and Biology*, pages 252–262, Cham. Springer International Publishing.

Hsu, C.-T., Ho, W.-H., Tsai, W.-I., and Lin, Y.-C. (2018). Realtime weightlifting barbell trajectory extraction and performance analysis. | Physical Education Journal | EBSCOhost. ISSN: 1024-7297 Issue: 1 Pages: 73 Volume: 51.

Jian, M., Zhang, S., Wu, L., Zhang, S., Wang, X., and He, Y. (2019). Deep key frame extraction for sport training. *Neurocomputing*, 328:147–156.

Lin, C.-Y. and Jian, K.-C. (2022). A real-time algorithm for weight training detection and correction. *Soft Computing*, 26(10):4727–4739.

Mroz, S., Baddour, N., McGuirk, C., Juneau, P., Tu, A., Cheung, K., and Lemaire, E. (2021). Comparing the Quality of Human Pose Estimation with BlazePose or OpenPose. In *2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART)*, pages 1–4.

Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., and Colyer, S. L. (2021). The accuracy of several pose estimation methods for 3D joint centre localisation. *Scientific Reports*, 11(1):20673. Number: 1 Publisher: Nature Publishing Group.

Pan, S. (2022). A Method of Key Posture Detection and Motion Recognition in Sports Based on Deep Learning. *Mobile Information Systems*, 2022:e5168898. Publisher: Hindawi.

Rethinam, P., Manoharan, S., Kirupakaran, A. M., Srinivasan, R., Hegde, R. S., and Srinivasan, B. (2023). Olympic Weightlifters' Performance Assessment Module Using Computer Vision. In *2023 IEEE International Workshop on Sport, Technology and Research (STAR)*, pages 8–12.

Yoshikawa, F., Kobayashi, T., Watanabe, K., Katsuyoshi, S., and Otsu, N. (2010). Start and End Point Detection of Weightlifting Motion using CHLAC and MRA:. In *B-Interface 2011*, pages 44–50, Valencia, Spain. SciTePress.