applied sciences

MDPI

Article

Identifying Goalkeeper Movement Timing from Single-Camera Broadcast Footage through Pose Estimation: A Pilot Study

Chethan Reddy and Woohyoung Jeon *

Department of Kinesiology and Health Education, University of Texas, Tyler, TX 75799, USA; creddy@patriots.uttyler.edu
* Correspondence: wjeon@uttyler.edu

Featured Application: This research describes a methodology for identifying the moment when a goalkeeper initiates their save attempt in the frontal plane, providing a way to benchmark the timing component of a save attempt.

Abstract: This study explores how open-source pose estimation can be utilized to identify goalkeeper dive initiation during soccer penalty kicks. The purpose of this study is to provide an accessible, low-cost heuristic methodology for identifying goalkeeper dive initiation. This study uses single-camera broadcast footage (1080 p resolution, 50 frames per second) of all 41 penalty shootout kicks attempted during the 2022 FIFA Men's World Cup. We isolated each penalty kick and recorded the frames of goalkeeper dive initiation and flight. We then identified goalposts to create a homography matrix to account for camera movement and identified the goalkeeper's skeletal keypoints through pose estimation. From these keypoints, we derived frontal plane kinematics for the torso and legs. We identified local extrema for each kinematic variable and isolated the last observed extrema prior to goalkeeper flight for each variable. Using OLS regression, we found that the last local extremum of the goalkeeper centroid's y-value was the strongest predictor of labeled commitment to the dive side, with an R2 of 0.998 and a *p*-value of 0.00. The results of this research are preliminary but demonstrate the promise of pose estimation in identifying sport-specific action timing during live game play using a single camera.

Keywords: pose estimation; sports analytics; movement timing; penalty kicks

check for updates

Citation: Reddy, C.; Jeon, W. Identifying Goalkeeper Movement Timing from Single-Camera Broadcast Footage through Pose Estimation: A Pilot Study. *Appl. Sci.* **2024**, *14*, 5961. https://doi.org/10.3390/app14135961

Academic Editor: Arkady Voloshin

Received: 3 June 2024
Revised: 24 June 2024
Accepted: 2 July 2024
Published: 8 July 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Markerless human pose estimation has the potential to revolutionize sports analytics by providing detailed insights into athlete movement. Pose estimation provides biomechanical data on athlete movement by accurately detecting and tracking the spatial positions of key body joints in video frames [1]. This serves as a valuable tool for training, performance optimization, injury prevention, and rehabilitation [1–3]. Coaches can analyze an athlete's posture and movement in real-time or recorded footage, identify areas for improvement in technique, and tailor training strategies to individual athletes [2,4].

Currently, marker-based motion capture and force plate analysis are the most common methods used for athlete movement analysis [5–7]. While these conventional methods provide valuable data on athlete movement and biomechanics, they require specialized equipment and controlled environments, making them impractical for use during live game play. This creates a gap in information between what occurs in training sessions and what happens in live game play, where conditions are dynamic and unpredictable, resulting in low ecological validity. Pose estimation offers a practical alternative for the capture and analysis of player movement in live game play [8].

In recent years, markerless motion capture systems using pose estimation have gained popularity in professional sports due to their ability to accurately track and analyze player movement without the need for physical markers placed on the body [9]. For example, during the 2022 FIFA World Cup, Hawk-Eye Innovations unveiled a new Video Assistant Referee (VAR) system to assist referees in making accurate offsides calls [10]. Pose estimation has also found applications among MLB teams for pitcher and hitter biomechanical analyses [9] and among NBA teams for the automation of goaltending and out-of-bounds calls [11]. Despite its benefits, the widespread adoption of markerless motion capture is limited by its high costs. Estimates for Hawk-Eye markerless motion capture systems range from EUR 7000–EUR 8000 per match for Gaelic football [12], USD 60,000 for tennis courts, and GBP 250,000 for soccer stadiums [13]. In addition, these systems do not allow for the retroactive collection of data [14].

As pose estimation becomes more widespread, there has been growing interest in exploring more affordable and less invasive alternatives to motion capture systems [9]. Any sports organization or individual with access to footage, a computer for executing code, and open-source libraries such as MediaPipe and YOLO can utilize and benefit from markerless pose estimation technology [15].

In this study, we apply markerless human pose estimation to 2022 FIFA World Cup broadcast footage to identify the initiation of goalkeeper movement during penalty shootout kicks. Goalkeepers play a critical role in determining match outcomes, and the timing of dive initiation is a critical component of goalkeeper performance [16]. However, accurately identifying the initiation of goalkeeper movement during live game play remains a challenging task.

This study aims to advance sports analytics by leveraging markerless human pose estimation technology to provide valuable insights into athlete movement, particularly the initiation of goalkeeper movement during penalty shootout kicks in live soccer matches. Despite substantial research on goalkeeper performance during penalty kicks, several significant gaps exist. First, existing research predominantly focuses on the spatial and perceptual aspects of goalkeeping, leaving the temporal aspects of goalkeeper movement relatively understudied [16]. Second, most studies overlook the environmental factors present in professional matches, as they typically occur in controlled settings [5,6,17]. Third, frequently cited studies suffer from inappropriate reaction criteria (i.e., keyboard responses) [18], the use of stationary targets [5,6], or irrelevant stimulus presentation (i.e., light flashes) [5,19]. Fourth, the lack of research focused on elite goalkeepers limits the generalizability of findings to elite populations, with many studies focusing on amateur-level subjects or subjects with no soccer experience [7,18].

Traditional methods of determining goalkeeper movement initiation during penalties often rely on subjective observations. For example, in a study by Noël et al., a soccer coach with 10 years of coaching experience was asked to review footage of 395 penalty kicks and identify the frame at which he believed the goalkeeper initiated the dive [20]. However, the coach was not instructed on how to identify dive initiation [20]. Though no criteria were provided for how to identify dive initiation, one author independently labeled dive initiation and found inter-rater reliability of labels to be satisfactory [20]. Pinheiro et al. developed an observational analysis system for penalty kicks based on a questionnaire focused on variables that were likely to distinguish the characteristics of successful or unsuccessful penalty kicks [21,22]. The observational system was validated by experts, but the authors noted that observer perception could be influenced by the viewing angle and that camera angles behind either the penalty taker or goalkeeper were the most appropriate for assessing penalty kicks [22].

Ibrahim et al. analyzed 10 elite Dutch goalkeepers diving off force plates towards suspended balls in response to an LED flash [5]. The dive initiation time was determined using ground reaction forces from the force plate data [5]. Similarly, Spratford et al. placed 37 reflective markers on six elite goalkeepers diving off a force plate towards stationary balls in response to a life-sized image of an outfield player projected on a screen [6]. Dive initiation time was identified by an exponential increase in vertical ground reaction forces [6]. In a similar study, Di Paolo et al. studied 19 adolescent goalkeepers with 17 sensors performing dives in response to a whistle [7]. The dive initiation time was marked by the contralateral foot toe-off, identified through a custom Matlab script (v2022a, The MathWorks, Natick, MA, USA) and video footage review. The stimuli presented in each of these studies is not reflective of real game play, and the use of markers and force plates is impractical for live game play.

Some studies have applied pose estimation to broadcast footage to evaluate goalkeeper strategies [21,23]. However, these studies only analyze selected key frames, rather than analyzing a continuous video. For example, Pinheiro et al. utilized Open Pose to classify goalkeeper strategy, a component of their validated observational analysis system [21]. This study only looked at two key frames and found that orientation and movement between the two key frames could accurately classify goalkeeper strategy [21]. However, this study did not attempt to understand the timing component of anticipation, instead opting to use the distance moved as a measure of goalkeeper anticipation [21]. Wear et al. assembled a dataset of 590 1v1 and penalty saves from broadcast footage, extracted a single frame at the moment of the kicker's contact with the ball, and used an unsupervised classifier to cluster goalkeeper pose data in the extracted frames [23]. However, it is important to note that this study assumed the goalkeeper was in a ready position and did not examine the development of goalkeeper positioning [23].

Markerless human pose estimation applied to broadcast footage presents a promising avenue to advance sports analytics, offering a non-invasive and cost-effective solution to capture and analyze athlete movements and bridging the gap between training and live game play.

This study provides a robust, safer, and more standardized methodology for determining when a goalkeeper initiates a save attempt. The use of broadcast footage eliminates the need for markers and platforms, making the process of capturing dives less invasive, safer, and more affordable. Additionally, this study provides a systematic alternative to manually reviewing footage to identify the initiation of movement. Furthermore, while pose estimation has been applied to broadcast footage of goalkeepers in prior studies, it has not been used to assess the timing of goalkeeper dive initiation [21,23]. By applying pose estimation to broadcast video, this study provides a basis for practitioners to connect game results to practice data. With this, coaches and player development professionals can use match outcomes to inform how a given goalkeeper's performance in live game play can inform potential training adjustments and vice versa. Coaches and training staff can review footage of unsuccessful save attempts in which goalkeepers initiated movement too early or too late. By providing targeted feedback, they can help players make better decisions in similar situations in the future.

This study shows that pose estimation can be applied to single-camera broadcast footage, and the resultant data can aid in the detection and analysis of goalkeeper movement initiation during penalty kicks by using frontal plane kinematics.

2. Materials and Methods

To create a heuristic methodology for identifying goalkeeper movement initiation, this study relies on broadcast footage of all penalty kicks attempted during the 2022 FIFA World Cup. Using this footage, the following framework was utilized:

1. Collect a dataset of penalty kicks from the broadcast footage.
2. Annotate the dataset with ground truth labels for goalkeeper movement and save outcomes.

3. Train and validate a pose estimation model using the annotated dataset.
4. Evaluate the accuracy of the pose estimation model for detecting goalkeeper movement during penalty kicks.

The primary data source for this study was broadcast video footage of all penalty shootouts from the 2022 FIFA World Cup. Typical soccer matches can end in a draw [24]. However, elimination games, such as the post-group stage matches at the World Cup, require a winner [24]. Penalty shootouts, which consist of a series of penalty kicks by both teams, occur when the match remains tied after regulation and extra time periods have expired [24]. We identified a total of 41 penalty kicks and collected broadcast footage of each kick from publicly available sources. All footage in the dataset has a resolution of 1080 p and a frame rate of 50 frames per second.

We cut the footage from each match to isolate the footage of each penalty kick attempt, starting from the kicker's run-up and ending with the kick outcome. For each penalty kick attempt, we identified two frames of interest: the frame of movement initiation and the frame of flight. These frames will be referenced as f0 and f1, respectively. Due to the interplay of goalkeeper and kicker strategies, there is no single definition of movement initiation. Movement initiation, in the context of this problem, is defined as the frame in which the goalkeeper initiates their save attempt, disregarding any extraneous motion prior to committing to the dive side. For example, if a goalkeeper jumps or shuffles on the line prior to the dive attempt, this is ignored when labeling. This serves to validate and understand the accuracy of automated movement detection. We marked f0 using Noël et al.'s subjective methodology of manually marking the frame and based our subjective observation on Ibrahim et al.'s findings that goalkeepers initiate a dive by pushing off with their contralateral leg, or the leg opposite of the dive side [5,20].

Ibrahim et al. described three strategies that goalkeepers use to start their dives: "(1) Exerting horizontal forces for horizontal displacement towards the ball, (2) Exerting vertical forces for a pre-push-off jump and (3) Exerting vertical forces with the contralateral leg for stepping sideward with the ipsilateral leg towards the ball" [5]. Ibrahim et al. identified "pushing off" using ground reaction force data collected from force plates [5]. This approach provided precise kinetic and kinematic data that allowed for an objective determination of dive initiation. However, given the constraints of our study, specifically the use of broadcast footage without access to ground reaction force data, we adapted this methodology for visual observation.

To mark f0 accurately, the first author visually observed the contralateral knee and ankle movement of the goalkeeper. This observation aligns with Ibrahim et al.'s findings, as the contralateral leg's motion is a critical indicator of the initiation of the dive. Specifically, the first author observed contralateral knee abduction and adduction, as well as contralateral ankle inversion and eversion in the mediolateral plane. By focusing on the contralateral knee and ankle, one can identify the moment when the goalkeeper committed to the dive. This method, while subjective, provides a practical solution for the limitations posed by the use of single-camera pose estimation applied to broadcast footage.

All data processing and analysis was conducted using Python version 3.9. The following dependencies were used: OpenCV 4.9.0.80 for image processing and feature extraction; YOLOv7 for pose estimation; and NumPy 1.23.5, pandas 2.2.0, and Statsmodels 0.14.1 for statistical analysis. All code was run on a Mac Studio with an M2 Ultra chipset and 192 GB of RAM.

The collected footage was filmed from a consistent vantage point, but the camera was not static. Therefore, we had to account for the pan, tilt, and zoom present in each video clip to properly track the pose estimation data's real-world coordinates from frame to frame. Because the goalposts are fixed dimensions, we were able to identify the corners of the goalposts and use these pixel coordinates to scale goalkeeper pose appropriately in each frame. Although there are "off-the-shelf" computer vision solutions to identify soccer field markings from broadcast footage, these models were not trained on the angles used in this research [25]. Therefore, we opted to identify the goalposts through a more traditional

color isolation method by identifying green pixels to mask the grass and white pixels to identify the goal.

The color of each pixel in the collected footage was in RGB format, which identifies the red, green, and blue values for each pixel. We chose the RGB values (100, 150, 50) and (220, 220, 220) to represent the grass and goalpost colors, respectively [26]. The RGB format presents problems in isolating colors because the time of day and weather can influence the perceived color of a given object. To account for this, we converted each frame to CIELUV color space to ensure that the perceived color of objects remained consistent across varying lighting conditions, thereby enhancing the accuracy of color-based object identification [27]. We subtracted the green from the frame and isolated the goalposts. The frame was then blurred and OpenCV's contour function was used to identify large shapes within the frame. We assumed that the largest contour represents the goal area and used that assumption to create a bounding box, as shown in Figure 1. We then used a Hough line transformation to identify all lines within the bounding box. These lines were then filtered into horizontal and vertical lists based on their angles [26]. The endpoints of these lines were then clustered, and the centroid of the identified cluster was used to create raw values for the goal corners.

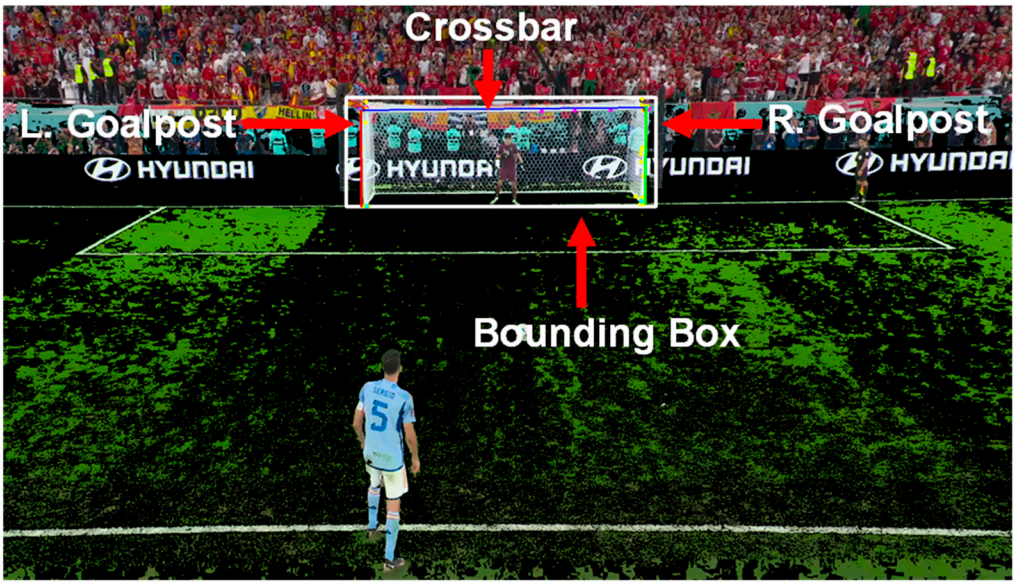


Figure 1. Estimates of goalpost bounding box.

Because the broadcast footage contained camera movement, the estimated goal corners needed to be mapped in each frame. We did so by using homography. Homography is a transformation between two planes whereby a given point can be mapped from one image to another [28]. To do so, a homography matrix must be calculated.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \quad (1)$$

Prior to creating the homography matrix, we smoothed the points using local polynomial fitting and overlayed them on the video to do a visual check, as shown in Figures 2 and 3. The FIFA rulebook stipulates that a regulation soccer goal must measure 732 cm in width and 244 cm in height. Using these dimensions, we defined a new image with a width of 732 px and a height of 244 px. The dimensions were chosen to have a 1:1 relationship between the goal dimensions in centimeters and pixels. The smoothed goalpost corners were then used to compute the homography matrix using OpenCV by mapping the upper goalpost corners in the original image to the upper corners of the new image and the bottom goalpost corners in the original image to the bottom corners of the new image. We

then added a buffer of 100 px to the new image in order to capture goalkeeper movement that occurs in front of the goal line. The homography matrix was then applied to warp the identified goal and buffer area to the dimensions of the new image. Examples of this transformation are shown in Figures 4 and 5.

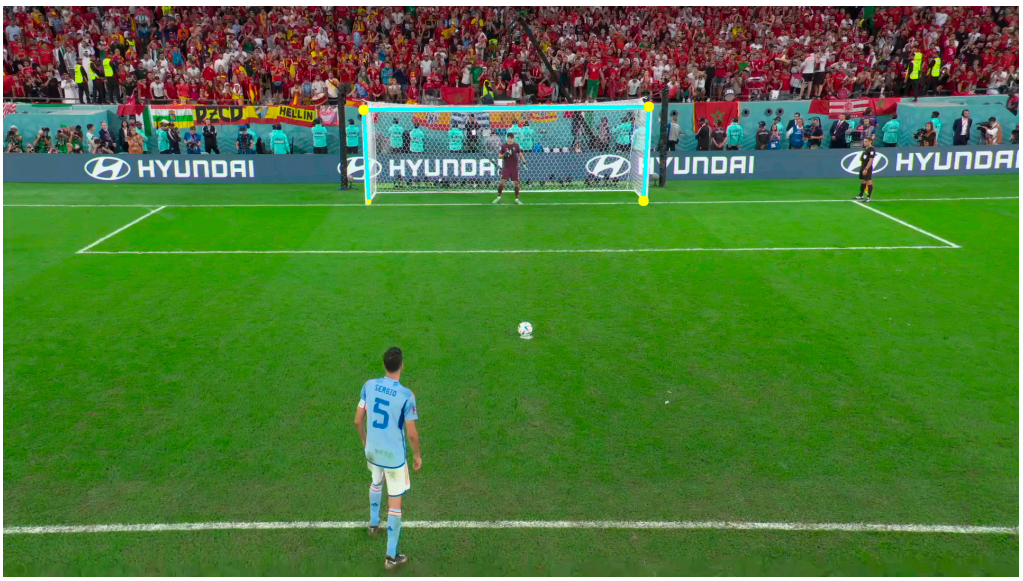


Figure 2. Smoothed goalposts and corners at start of run-up.



Figure 3. Smoothed goalposts and corners at goalkeeper flight.



Figure 4. Isolated area of interest at first frame.



Figure 5. Isolated area of interest at goalkeeper flight.

After creating the isolated goal area videos, we applied the YOLOv7 pose estimation algorithm to each video [29]. We chose YOLOv7 because it is open-source, general-purpose, and does not have licensing restrictions. Pose estimation must be applied at the frame level. YOLOv7 identifies all possible people in a given image, but identity across images is not recorded. Therefore, we applied a SORT object tracking algorithm to track the bounding box of the goalkeeper across frames. The SORT tracker allowed us to assign an integer ID label to all detected persons in each frame and isolate the pose data of interest. Examples of the pose estimation person labels are shown in Figures 6 and 7. YOLOv7′s pose estimation model provides the estimated locations of 17 keypoints (joints), listed in Table 1, as well as confidence scores for the bounding box of the human and all keypoints [30]. The confidence score is the probability that a given person or joint in the image has been correctly identified. This probability is provided on a 0–1 scale, with a higher score implying higher certainty of identification. By averaging the confidence scores for each joint over a set of frames, we can interpret the stability of the pose estimation algorithm in that set.

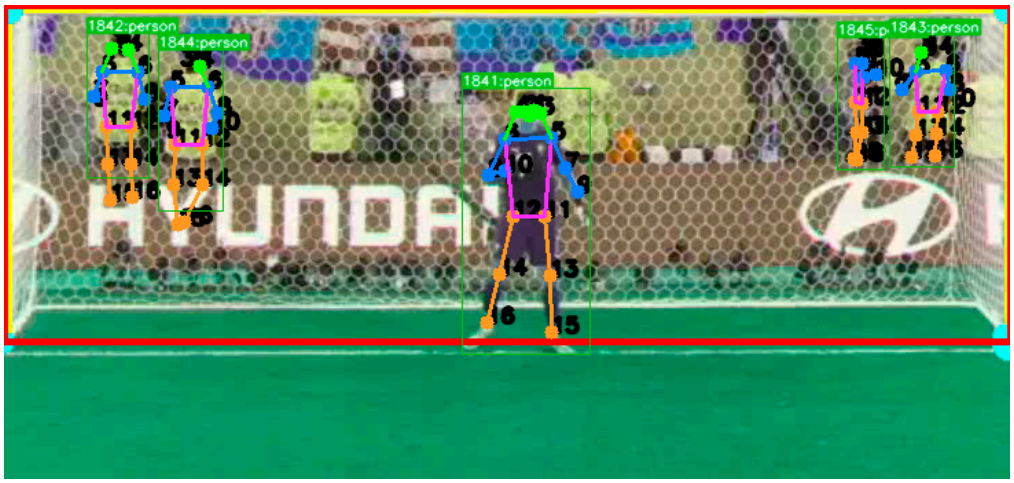


Figure 6. Pose estimation at start of run-up.



Figure 7. Pose estimation at goalkeeper flight.

Table 1. YOLOv7 joint numbers and labels.

Joint #	Joint Name
0	Nose
1	Left Eye
2	Right Eye
3	Left Ear
4	Right Ear
5	Left Shoulder
6	Right Shoulder
7	Left Elbow
8	Right Elbow
9	Left Wrist
10	Right Wrist
11	Left Hip
12	Right Hip
13	Left Knee
14	Right Knee
15	Left Ankle
16	Right Ankle

After applying the pose estimation algorithm, we calculated the centroid of the torso, as well as the angle in the frontal plane from the hip to the ankle, the hip to the knee,

and the knee to theankle of each leg. The centroid of the torso (C) was calculated as the mathematical center of the shoulder (LS—Left Shoulder, RS—Right Shoulder) and hip (LH—Left Hip, RH—Right Hip) coordinates as shown in Equation (2).

$$C(x,y) = (\frac{LSx + RSx + LHx + RHx}{4}, \frac{LSy + RSy + LHy + RHy}{4}), \quad (2)$$

The raw keypoints, centroid, and angles were then smoothed using local polynomial fitting over a 5-frame (100 ms) window [31]. This window was used in order to reflect the complexity of human movement. Whereas the smoothing of the goalpost corners was performed over the span of each video because the camera movement was not sudden, goalkeeper movement can be sudden. Accordingly, we chose a small window in order to preserve movement that may have otherwise been smoothed out if applying polynomial fitting over an entire kick.

After cleaning the pose data, we plotted each variable as a function of frame number along with vertical lines at f0 and f1. This gave us an isolated visual of how each variable changed over the course of a given penalty kick in relation to goalkeeper actions. Subsequently, we identified local extrema for each variable using a sliding window. This simply found the minimum and maximum value in a rolling 5-frame window. We identified the last extrema prior to f1 for each variable. We then used Ordinary Least Squares (OLS) regression to model f0 as a function of the last extrema of each variable in order to determine which variables were significant predictors of goalkeeper dive initiation. Using a significance level of 0.05, we identified the variables with the highest significance level and conducted a single regression to demonstrate the most parsimonious model.

$$f_0 \sim x_0 + \ldots + x_n$$

Equation: *Initial Linear Regression*

$$f_0 \sim x_i$$

Equation: *Parsimonious Linear Regression*

3. Results

3.1. Pose Detection

The goalkeeper pose was detected in 88% of frames. YOLOv7's pose estimation algorithm assigned an 85% probability of the goalkeeper being a person in each frame. The confidence scores for shoulder, hip, knee, and ankle joints remained relatively stable. The confidence scores for the shoulder, knee, and ankle joints were all 90% or higher. The confidence scores were significantly lower in the goalkeeper's flight phase compared to other segments of the video. This was accompanied by lower confidence scores for each joint (Table 2). Taken together, these results demonstrate that pose estimation can be successfully applied to soccer penalty kicks prior to goalkeeper flight, with the torso and lower limbs providing stable points for analysis.

Table 2. YOLOv7 confidence scores.

		Video Segment			
Joint	**Total Sample**	**Standing**	**Crouching**	**Committed**	**Flight**
Body	0.85	0.87	0.87	0.82	0.67
Nose	0.85	0.90	0.86	0.70	0.58
Left Eye	0.80	0.86	0.81	0.61	0.40
Right Eye	0.77	0.83	0.78	0.61	0.46
Left Ear	0.61	0.66	0.61	0.50	0.34
Right Ear	0.58	0.60	0.59	0.55	0.51
Left Shoulder	0.97	0.97	0.97	0.96	0.92

Table 2. *Cont.*

	Video Segment				
Joint	**Total Sample**	**Standing**	**Crouching**	**Committed**	**Flight**
Right Shoulder	0.97	0.97	0.97	0.96	0.94
Left Elbow	0.90	0.92	0.92	0.88	0.79
Right Elbow	0.90	0.92	0.90	0.88	0.82
Left Wrist	0.87	0.89	0.89	0.82	0.71
Right Wrist	0.87	0.89	0.87	0.82	0.73
Left Hip	0.97	0.98	0.97	0.96	0.92
Right Hip	0.97	0.98	0.97	0.96	0.92
Left Knee	0.95	0.96	0.95	0.92	0.86
Right Knee	0.95	0.96	0.95	0.92	0.86
Left Ankle	0.90	0.92	0.91	0.88	0.81
Right Ankle	0.90	0.92	0.91	0.88	0.82

3.2. Timing Detection

The initial regression (Table 3) showed that the last extrema for the centroid x component, centroid y component, and contralateral knee-to-ankle (lower leg) angle were the most significant predictors of goalkeeper movement initiation. Figure 8 shows the correlation between each of these variables and the labeled dive frame. Of these two, the centroid y component was more significant. Because the centroid is agnostic to the dive direction and goalkeeper strategy, we chose to create a parsimonious model using the centroid y component. The parsimonious model results indicated that the centroid y ($\beta = -0.96$, $p < 0.001$) component accounts for 99.8% of the variance in the labeled initiation of movement. Figure 9 provides a comparison of predicted and labeled dive initiation for a selected video.

Table 3. Full regression results. Values with ** denote significant variables.

Variable	**β**	**t**	***p***
Centroid X	0.3646	3.339	0.002 **
Centroid Y	0.5132	4.785	0.000 **
Contralateral Hip-to-Knee Angle	−0.0256	−1.103	0.277
Contralateral Knee-to-Ankle Angle	0.1350	3.211	0.003 **

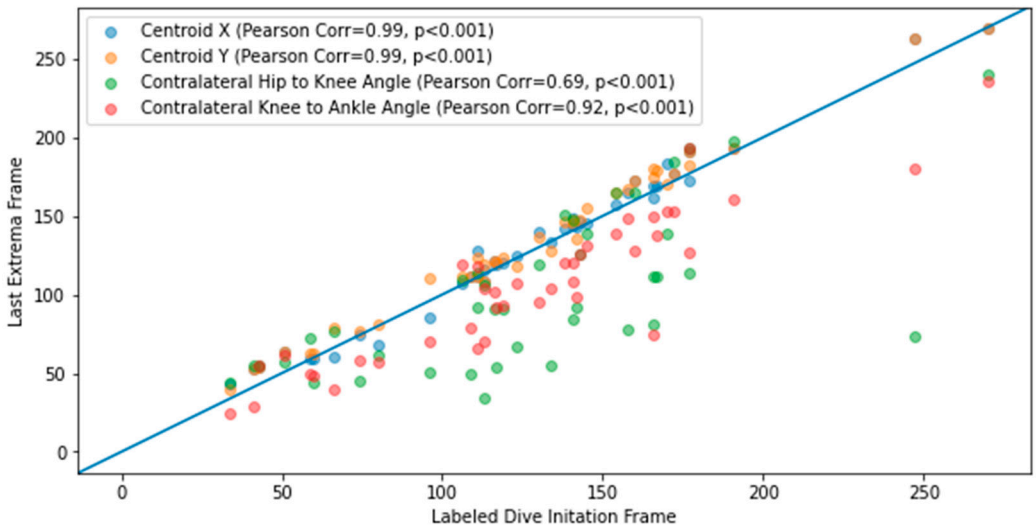


Figure 8. Labeled dive initiation vs. last extrema for regression variables.

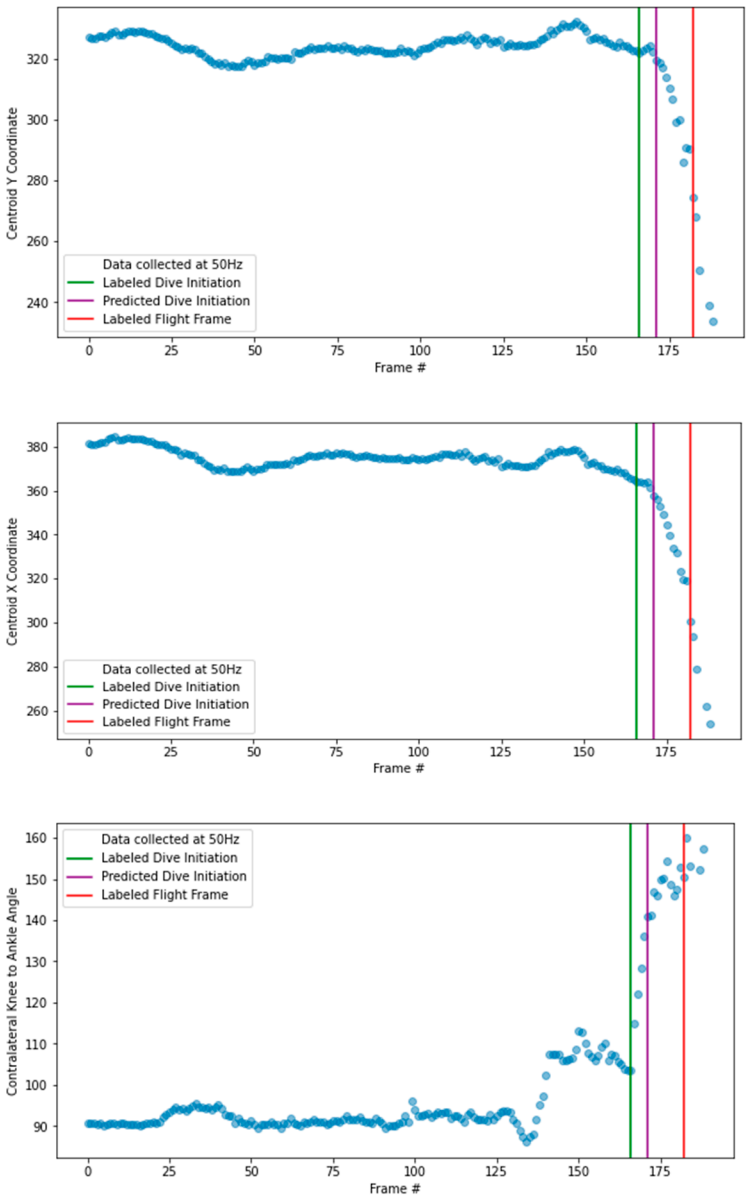


Figure 9. Significant variables by frame for selected video.

4. Discussion

This study created a heuristic methodology to identify goalkeeper movement initiation during penalty kicks by applying pose estimation to a single camera angle of broadcast footage. This is a novel approach that provides a robust and safe methodology to conduct analyses of live game performance in elite populations.

From the results of the OLS regressions, it appears that we can estimate goalkeeper movement initiation by using the centroid's y value. This demonstrates the potential applications of pose estimation to identify action timing in sport. It highlights the feasibility of conducting a descriptive analysis of athlete movement and timing, even outside of the constraints of a controlled environment, solely by using standard camera footage.

The methodology used in this study provides a way to derive absolute goalkeeper movement timing, which has the potential to inform more contextual measures such as timing relative to ball kick or kicker visual cues. The significance of this study should be viewed through its potential to fill in the gaps that other studies leave. Because this is a novel approach to measure timing in live game play, there are no direct comparisons to compare results to. It provides a way to standardize action timing and enables further

exploration of the timing aspect of a skill that otherwise cannot be studied outside of a laboratory environment.

Our labeling methodology used labels that sought to assign visual cues from previous studies. Namely, it used Ibrahim et al.'s description of goalkeeper movement during the dive and focused on the contralateral leg as a visual cue for dive initiation [5]. We demonstrated that mediolateral movement in the knee and ankle can be considered a significant predictor of contralateral leg push-off as defined by Ibrahim et al. [5]. Additionally, the torso centroid coordinates were significant predictors of goalkeeper movement initiation. Spratford et al. used center of mass (COM) displacement to estimate timing [6]. The centroid and COM are fundamentally different, but the centroid of the torso in this study could be a viable proxy for COM when understanding movement timing. Together, the results of this study should not be mistaken for measuring ground reaction forces but rather interpreted as being able to approximate the visual manifestation of these forces. Moreover, this study provides a more structured methodology for assigning timing labels to soccer penalty kicks than Noël et al. and allows for a more repeatable and less time-intensive process [20].

Further, this study introduces the use of pose estimation over continuous frames to analyze goalkeeper movement. While previous studies by Pinheiro et al. and Wear et al. used pose estimation to analyze goalkeeper save attempts, they looked at specific key frames rather than continuous movement [21,23]. This study expands upon their strategies and allows for a more granular assessment of goalkeeper strategy and performance. In particular, it opens the possibility of integrating timing into Pinheiro et al.'s evaluation network [21]. Pinheiro et al. measured anticipation by total movement between two frames [21]. Because anticipation is a measure of timing, this study could provide supplemental or alternative measures to assess anticipation. Pinheiro et al. and Noël et al. both assess goalkeeper strategies, with Noël et al. incorporating timing into their assessment of strategy [20,21]. Assigning labels using the methods in this study could add more value to evaluation of strategy and outcomes.

More globally, human action recognition has largely relied on unsupervised clustering methods [21,23,32]. The focus of such technologies is sometimes to identify human actions in real time, but this study demonstrates a simple, descriptive solution for identifying when a specific action occurs and demonstrates that Wu's choice to classify actions over five-frame intervals can be extended to action timing without the use of clustering [32]. This study could be used in conjunction with clustering/classifying algorithms to enhance the accuracy of each method.

Perhaps the most interesting application of this study is the possibility of evaluating affordance-based performance as outlined by van der Kamp et al. [16]. In their study, they created a theoretical framework for assessing the performance of goalkeeper dives within the context of their physical capabilities. One of the main components of this framework was the timing of the goalkeeper's dive relative to the timing of the ball kick. This study is a step forward in being able to make that framework actionable, as it allows for an absolute measurement of the dive time and thereby a comparison to the time of the ball kick. Van der Kamp et al.'s framework is promising in its approach to benchmark a goalkeeper's dive time to their physical capabilities and further evaluate their decision making [16].

The two main limitations in this study are the sample size and the information that can be captured by the pose estimation framework. We limited our sample to the 2022 World Cup to ensure as consistent footage as possible. However, the small sample size of both attempts and goalkeepers likely limit the statistical takeaways from this study. While the dataset includes repeated instances of the same goalkeepers, this does not necessarily introduce bias. The diversity in direction, approach, and strategy within the sample ensures a wide variety of scenarios, as goalkeepers face different opponents and situations during each penalty kick. In this study, we used two-dimensional pose estimation and therefore could only observe motion in the frontal plane. Due to this limitation, we lose information in other planes. More robust coordinate mapping (i.e., 3D instead of 2D) could help provide

more accurate pose estimation and enable the analysis of velocities and accelerations. Additionally, there is currently no feasible way to capture ground reaction forces from broadcast footage, potentially leading to a lag in the identification of goalkeeper dive initiation compared to methods that directly measure ground reaction forces. Further, the methodology used here relies on processed videos and qualitative labeling. The results cannot be applied to a continuous broadcast without prior annotation or labeling, and the criteria for labeling are loose. Although we describe movements of interest, the exact criteria are ambiguous and subject to human evaluation, potentially impacting the accuracy of a trained model.

The research presented here only presents a methodology to detect when a goalkeeper initiates their movement but does not explore the context of this timing. This research provides a stepping stone to better understand timing outside of binary save outcomes. Future studies may consider analyzing the kinematics of the penalty taker as part of a more holistic evaluation of the penalty kick. It may be possible to use pose estimation data to create a probabilistic model based on kicker kinematics that attempts to assign probabilities of kick direction at each frame leading up to foot-to-ball contact and benchmarks goalkeeper timing and dive direction to the model's confidence in kick direction. Future studies may also look to combine pose estimation predictions of timing with force plate data to understand how goalkeeper strength and ability interplay with the timing of movement.

5. Conclusions

This study has demonstrated the effectiveness of pose estimation in identifying the initiation of goalkeeper movement during live game play using single-camera broadcast footage. Through pose estimation and a heuristic methodology, it is possible to estimate goalkeeper movement initiation from footage. While pose estimation can reliably be applied to analyze lower-velocity segments of movement, its accuracy diminishes with motion blur, particularly during goalkeeper flight. This research presents a reliable methodology to estimate timing using 2D pose estimation. Future studies may explore how detected timing can be combined with performance data for a more comprehensive analysis of goalkeeper movement.

Author Contributions: Conceptualization, C.R. and W.J.; Methodology, C.R. and W.J.; Software, C.R.; Validation, W.J.; Formal analysis, C.R.; Investigation, C.R.; Writing—original draft, C.R.; Writing—review & editing, W.J.; Visualization, W.J.; Supervision, W.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Due to copyright laws, the authors cannot redistribute the isolated video clips used in this study, however matches are available at www.fifa.com.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Stenum, J.; Cherry-Allen, K.M.; Pyles, C.O.; Reetzke, R.D.; Vignos, M.F.; Roemmich, R.T. Applications of Pose Estimation in Human Health and Performance across the Lifespan. *Sensors* **2021**, *21*, 7315. [CrossRef] [PubMed]
2. Sharma, D. Review of application of gesture and poses for reducing injury in sports. *Int. J. Converg. Health* **2024**, *4*, 31–35. [CrossRef]
3. Blanchard, N.; Skinner, K.; Kemp, A.; Scheirer, W.; Flynn, P. "Keep Me In, Coach!": A Computer Vision Perspective on Assessing ACL Injury Risk in Female Athletes. In Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 7–11 January 2019; pp. 1366–1374. [CrossRef]
4. Einfalt, M.; Zecha, D.; Lienhart, R. Activity-conditioned continuous human pose estimation for performance analysis of athletes using the example of swimming. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 446–455. [CrossRef]

5. Ibrahim, R.; Kingma, I.; de Boode, V.A.; Faber, G.S.; van Dieën, J.H. Kinematic and kinetic analysis of the goalkeeper's diving save in football. *J. Sports Sci.* **2018**, *37*, 313–321. [CrossRef] [PubMed]
6. Spratford, W.; Mellifont, R.; Burkett, B. The influence of dive direction on the movement characteristics for elite football goalkeepers. *Sports Biomech.* **2009**, *8*, 235–244. [CrossRef] [PubMed]
7. Di Paolo, S.; Santillozzi, F.; Zinno, R.; Barone, G.; Bragonzoni, L. On-Field Biomechanical Assessment of High and Low Dive in Competitive 16-Year-Old Goalkeepers through Wearable Sensors and Principal Component Analysis. *Sensors* **2022**, *22*, 7519. [CrossRef] [PubMed]
8. Tuyls, K.; Omidshafiei, S.; Muller, P.; Wang, Z.; Connor, J.; Hennes, D.; Graham, I.; Spearman, W.; Waskett, T.; Steel, D.; et al. Game Plan: What AI can do for Football, and What Football can do for AI. *J. Artif. Intell. Res.* **2021**, *71*, 41–88. [CrossRef]
9. Sarris, E.; Lewis, A. The Future—And Present—Of Baseball is in Biomechanics. The Athletic. Available online: https://theathletic.com/3144548/2022/02/25/the-future-and-present-of-baseball-is-in-biomechanics/ (accessed on 25 February 2022).
10. Lemire, J. FIFA World Cup 2022 Will Use Semi-Automated Offside Technology for 'Accurate and Faster Decisions'. FIFA World Cup 2022 Will Use Semi-Automated Offside Technology. 1 July 2022. Available online: https://www.sportsbusinessjournal.com/Daily/Issues/2022/07/01/Technology/fifa-world-cup-2022-semi-automated-offside-technology (accessed on 21 March 2023).
11. Lemire, J. NBA Bringing in Hawk-Eye for Tracking Data in 2023-24. NBA to Start Working with Hawk-Eye Innovations. 9 March 2023. Available online: https://www.sportsbusinessjournal.com/Daily/Issues/2023/03/09/Technology/nba-hawkeye-player-ball-tracking-data.aspx (accessed on 21 March 2023).
12. Fogarty, J. HawkEye's Presence in Championship 2020 Hangs in Balance. Irish Examiner. Available online: https://www.irishexaminer.com/sport/gaa/arid-40052699.html (accessed on 22 September 2020).
13. Wilson, J. Arsène Wenger Calls for Extension of Technology in Football as Hawk-Eye Goal-Line System Introduced. Available online: https://www.telegraph.co.uk/sport/football/teams/arsenal/10231277/Arsene-Wenger-calls-for-extension-of-technology-in-football-as-Hawk-Eye-goal-line-system-introduced.html (accessed on 8 August 2013).
14. Sutherland, J. Goal-Line Technology 'Unaffordable' for Scottish Premiership. *BBC Sport*. 28 December 2017. Available online: https://www.bbc.com/sport/football/42504610 (accessed on 21 March 2023).
15. Lugaresi, C.; Tang, J.; Nash, H.; McClanahan, C.; Uboweja, E.; Hays, M.; Zhang, F.; Chang, C.-L.; Yong, M.G.; Lee, J.; et al. MediaPipe: A Framework for Building Perception Pipelines. *arXiv* **2019**, arXiv:1906.08172.
16. van der Kamp, J.; Dicks, M.; Navia, J.A.; Noël, B. Goalkeeping in the soccer penalty kick. *Ger. J. Exerc. Sport Res.* **2018**, *48*, 169–175. [CrossRef]
17. Higueras-Herbada, A.; Lopes, J.E.; Travieso, D.; Ibáñez-Gijón, J.; Araújo, D.; Jacobs, D.M. Height After Side: Goalkeepers Detect the Vertical Direction of Association-Football Penalty Kicks from the Ball Trajectory. *Front. Psychol.* **2020**, *11*, 311. [CrossRef] [PubMed]
18. Diaz, G.J.; Fajen, B.R.; Phillips, F. Anticipation from biological motion: The goalkeeper problem. *J. Exp. Psychol. Hum. Percept. Perform.* **2012**, *38*, 848–864. [CrossRef] [PubMed]
19. Yang, J.; Park, Y.; Kim, K.; Ryu, J.-K. Goalkeeper's position for defending short range shots. *Int. J. Sports Sci. Coach.* **2017**, *12*, 603–610. [CrossRef]
20. Noël, B.; van der Kamp, J.; Klatt, S. The Interplay of Goalkeepers and Penalty Takers Affects Their Chances of Success. *Front. Psychol.* **2021**, *12*, 645312. [CrossRef] [PubMed]
21. Pinheiro, G.d.S.; Jin, X.; Da Costa, V.T.; Lames, M. Body Pose Estimation Integrated with Notational Analysis: A New Approach to Analyze Penalty Kicks Strategy in Elite Football. *Front. Sports Act. Living* **2022**, *4*, 818556. [CrossRef] [PubMed]
22. Pinheiro, G.d.S.; Nascimento, V.B.; Dicks, M.; Costa, V.T.; Lames, M. Design and Validation of an Observational System for Penalty Kick Analysis in Football (OSPAF). *Front. Psychol.* **2021**, *12*, 661179. [CrossRef] [PubMed]
23. Wear, M.; Beal, R.; Matthews, T.; Norman, T.; Ramchurn, S. Learning from the Pros: Extracting Professional Goal-keeper Technique from Broadcast Footage. *arXiv* **2022**, arXiv:2202.12259.
24. The International Football Association Board (IFAB). *Laws of the Game 2022/23. Zurich*; The International Football Association Board: Zurich, Switzerland, 2022.
25. Chen, J.; Little, J.J. Sports Camera Calibration via Synthetic Data. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, USA, 16–17 June 2019; pp. 2497–2504. [CrossRef]
26. Te, I.C. Goalpost-Detector. GitHub. 2018. Available online: https://github.com/ianchute/goalpost-detector (accessed on 1 May 2024).
27. Xu, H.; Lu, J.; He, Q.; Gu, L.; Feng, C. Research on Detecting Moving Objects in Football Match Video. *J. Phys. Conf. Ser.* **2020**, *1651*, 012161. [CrossRef]
28. Claasen, P.J.; de Villiers, J.P. Video-Based Sequential Bayesian Homography Estimation for Soccer Field Registration. *arXiv* **2023**, arXiv:2311.10361. [CrossRef]
29. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. YOLOv7: Trainable BAG-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 17–24 June 2023; pp. 7464–7475. [CrossRef]
30. Wong, K.Y. Yolov7. GitHub. 2022. Available online: https://github.com/WongKinYiu/yolov7 (accessed on 1 May 2024).

31. Cleveland, W.S.; Loader, C.R. Smoothing by Local Regression: Principles and Methods. *Contrib. Stat.* **1996**, 10–49. [CrossRef]
32. Wu, Q.; Xu, G.; Chen, L.; Luo, A.; Zhang, S. Human action recognition based on kinematic similarity in real time. *PLoS ONE* **2017**, *12*, e0185719. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.