F-measure and Intra-operator Reliability of Aerial Imagery Tracking in Football

Khairul Imran Sainan[1,2], Ahmad Khushairy Makhtar[1], Zulkifli Mohamed[1,2]*

[1]School of Mechanical Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, MALAYSIA
[2]Sports Engineering Artificial Intelligence Centre, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, MALAYSIA
[*]imransainan@uitm.edu.my

ABSTRACT

Current player tracking methods using multiple fixed stadium-based cameras and wearable sensors have limitations. To address this, a new computer vision system using aerial drone imagery has been developed to track football players. This approach is less expensive, has a wider field of view, and captures data not accessible through sensors. However, the visual performance has not been extensively evaluated. In this study, we aimed to determine the system's tracking performance using an F-measure score, which is calculated based on the number of true positives, false positives, and false negatives identified during the tracking process. We also investigated the tracking reliability by comparing the intra-operator performance using the ICC for distance, speed, and time metrics by repeating the measurements five times. The aerial-imagery data were taken from a test match recorded using a drone that was hovered away from the touchline. Four players were tracked and measured simultaneously. The system demonstrates accuracy by performing admirably with average F-measure scores of 0.80, 0.80, 0.89, and 0.84 for player A0, player A1, player B0, and player B1, respectively. Meanwhile, the intra-operator reliability for distance and speed was deemed good to moderate with %MD < 10%. The findings suggest that the system is a capable and reliable computer vision tracking tool with potential applications in performance analysis, training feedback, and injury prevention.

Keywords: *F-measure; Reliability; Aerial Imagery; Tracking in Football*

ISSN 1823-5514, eISSN 2550-164X

https://doi.org/10.24191/jmeche.v13i1.2860

Received for review: 2024-06-14
Accepted for publication: 2024-09-03
Published: 2024-11-15

Introduction

The advancement of the Optical Tracking System (OTS) has allowed football teams and associations access to valuable kinematic data for football analysis. Substantial efforts focused on improving the capability of detection and tracking [1]-[2], validation of data accuracy [3]-[4], repeated tests for reliability [5]-[6], and between devices interchangeability [7]-[8]. However, limitations in cost [1], data acquisition ability and scalability persist [9].

Traditional OTS uses cameras that are permanently placed around the stadium. This approach requires a large number of cameras, from a minimum of three [10] to twenty-eight cameras [11], necessitating time-consuming and specialized calibration [12]-[13] that requires expertise. The system may also require the installation of tall temporary structures [10] specifically for small stadiums to achieve good viewing angles and avoid occlusion issues [14]-[15]. Additionally, combining multiple camera views (image stitching) adds complexity and processing power, leading to high overall costs. These limitations restrict OTS use to teams or associations with significant and stable financial resources. Mobile and portable OTS that are easy to set up, operate, and use might be a preference for small teams. Teams that are constantly on the move, using different venues for training or going for away matches may also benefit from such an OTS setup.

Recently, drones have emerged as potential alternatives due to their affordability, single but powerful camera, and ability to achieve a wider field of view from a higher position which is unbound by stadium structure. At present, existing drone research in football is limited even though there is a growing interest in using the system as part of a coaching tool [16]. Studies have discussed the impact of drone applications on football [17], explored player detection accuracy for indoor activities [18], or reported position validity [19] in non-standard settings [20], but have not assessed tracking performance or reliability in a real match on a standard field neglecting the challenges in handling smaller objects. Researchers have also delved into implementing drones with other sports i.e. ultimate frisbee [20]-[21], badminton [22], and tennis [20]. These studies often employed a bird's eye view [20]-[23] as opposed to a non-overhead drone position. Clearly, an assessment of visual performance and intra-operator reliability using aerial drone imagery for a real match from a standard-sized football field is justified.

Therefore, this study aims (1) to assess the visual tracking performance using an F-measure score and (2) to investigate the reliability of intra-operator performance for distance, speed, and time metrics using aerial drone imagery in real matches from non-overhead positions. We hypothesize that the new algorithm will achieve high accuracy (F-measure score exceeding 0.8) and good reliability (Intraclass Correlation Coefficients (ICC) values above 0.7). This assessment will contribute greatly to understanding the potential of computer vision and drone application in football coaching.

Methodology

Twenty-two male (n = 22) football players (age: 21 ± 3 yrs; height: 1.75 ± 0.10 m; body mass: 70 ± 10 kg) playing for the varsity team, participated in this study. The participants were actively involved in football tournaments and were free from any injury. Prior to participation, researchers were briefed on the purpose, procedures, and potential risks of this study to the participants and the match officials. Institutional board approval for the study was obtained from the Research Ethics Committee of Universiti Teknologi MARA (UiTM). The ethics approval guidelines comply with the Declaration of Helsinki (2013), Malaysian Good Clinical Practice (2018), UiTM Ethics Policy as well as relevant laws in Malaysia. All participants voluntarily gave informed consent to participate and allowed data collection of kinematics value through drone footage. All performance data were anonymized for confidentiality. In the 24 hours leading up to the test, participants were advised to consume a well-balanced diet and maintain adequate fluid intake.

The participants were divided into two teams (11 vs. 11) and played a football test match at Padang Pusat Sukan UiTM. The teams wore distinct colour jerseys for clear differentiation in the footage (Team A: light blue jerseys, black shorts/stockings; Team B: red jerseys, white shorts/stockings. The field dimensions (105 m x 68 m) adhered to Fédération Internationale de Football Association (FIFA) guidelines. The venue is regularly used for matches and was in good condition at the time of the experiment. The natural grass playing surface is well maintained with clearly visible lines. Corner flags and nets for goalposts were installed. For injury prevention, a qualified football trainer led all participants in a proper 20-minute warm-up session before the match to elevate body temperature and engage core muscle groups. The match consisted of two 35-minute halves, officiated by qualified officials, and played according to standard rules. Participants were allowed to substitute themselves with other players if they wished. Weather conditions were favourable, dry with a light breeze, and at an ambient temperature of approximately 27 ^{o}C.

Aerial imagery data was captured using a (SZ DJI) drone equipped with a satellite positioning system capability (GPS/GLONASS) and advanced state-of-the-art safety sensors. The drone houses a 3-axis mechanical stabilization gimbal with a 12MP resolution camera capable of shooting a video up to 4K resolution video (4096 x 2160 pixels) at 24 frames per second. A dedicated vision system provided obstacle detection with a range of 0.7 m to 30 m. Three lithium polymer batteries juice up a maximum flight time of up to 60 minutes. Prior to the match, a pre-flight checklist ensured the proper function of the remote controller, batteries, camera, motors, and propellers. The flight test assessed for any irregular vibrations or movements. The pilot evaluated yaw, roll, and pitch movements at a height of 10 meters and a distance of 5 meters away from the home position for a successful flight test before actual recording. During the test match, a spotter assisted the pilot to ensure flight safety. The drone ascended vertically to the recording position and remained stationary throughout the match, offering a clear view of the entire playing field. The drone was positioned 60 meters high and 30 meters away from the long side of the pitch.

In this study, four players (n = 4), two from each team, were tracked simultaneously. Players were assigned unique identifiers (A0, A1, B0, B1) with a bounding box (a rectangular marker around each player) for tracking. The analysis began by defining the playing field dimensions and selecting the players. Once all players were identified, researchers initiated the program thus automatically prompted the algorithm to track their movements throughout the video. The system recorded the visual tracking performance by calculating the number of True Positives (TP), False Positives (FP), and False Negatives (FN). The X and Y coordinates of players in each frame were stored in a .CSV file format for data analysis.

F-measure score is a metric used to evaluate the performance of the tracking algorithm in terms of the accuracy of positive predictions (recall) and quantifying their occurrences (precision). The F-measure score is calculated from recall and precision value to assess the visual tracking performance [24]:

$$F - measure = 2\ x\ \left[\frac{Precision\ x\ Recall}{Precision + Recall}\right] \tag{1}$$

where,

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

and,

$$Percision = \frac{TP}{TP + FN} \tag{3}$$

True Positives (TP) represent the amount of tracking when the bounding box is successfully positioned on the Region of Interests (ROIs), false negatives (FN) represent the number of failed tracking, and false positives (FP) represent the sum of the misplaced bounding box.

Intra-operator reliability was investigated by repeating the data acquisition five times. We observed two dynamic measures that were derived from the X and Y coordinates; distance (m) and speed (km.h^{-1}). We calculated the absolute difference between the values obtained from each repetition and the mean value across the five measures [6]. Additionally, descriptive statistics (mean difference, coefficient of variation, and ICC) were computed to assess the consistency of the tracking process. The ICC for intra-operator differences was tested using a two-way random effects Analysis of Variance (ANOVA) model without replication. This model accounts for potential random effects arising from the measurement process (rater) rather than the players themselves. Measurement discrepancies were considered statistically significant if the 95% Confidence Interval (CI) did not include zero [25]-[27]. The formula to calculate the sample mean was,

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{4}$$

where, n = number of samples, x_i = each data point. Meanwhile, the standard deviation was determined as,

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}} \tag{5}$$

and 95% CI was calculated as,

$$CI = \bar{x} \pm z\frac{\sigma}{\sqrt{n}} \tag{6}$$

where, z = confidence level. The ICC was computed as,

$$ICC = \frac{var(\beta)}{var(\alpha) + var(\beta) + var(\varepsilon)} \tag{7}$$

where, $var(\beta)$ = differences in the rates, $var(\alpha)$ = differences in the rating scale and $var(\varepsilon)$ = differences in the rater. The systematic bias ± random error of player speed was assessed using Bland-Altman plot [28].

Results and Discussion

Visual performance

F-measure is a single metric value to evaluate the overall tracking performance by the harmonic mean between precision and recall. Table 1 summarizes the F-measure, precision, and recall values for the four players (A0, A1, B0, B1). Meanwhile, Figure 1 represents the visual representation of these values along with their mean across all five trials. In general, higher F- measure scores indicate better performance. The results showed player B0 achieved the highest score (mean F-measure: 0.89) followed by player B1 (mean F-measure: 0.84), player A0 (mean F-measure: 0.80), and player A1 (mean F-measure: 0.80).

Table 1: Visual tracking performances across all 5 trials

Sequence	Recall				Precision				F-Measure			
	A0	**A1**	**B0**	**B1**	**A0**	**A1**	**B0**	**B1**	**A0**	**A1**	**B0**	**B1**
Trial 1	0.71	0.68	0.87	0.77	0.93	0.92	0.96	0.95	0.81	0.78	0.91	0.85
Trial 2	0.69	0.69	0.84	0.71	0.92	0.91	0.96	0.95	0.79	0.78	0.90	0.81
Trial 3	0.69	0.68	0.84	0.80	0.92	0.91	0.96	0.95	0.79	0.78	0.90	0.87
Trial 4	0.71	0.76	0.84	0.78	0.93	0.94	0.96	0.94	0.81	0.84	0.90	0.85
Trial 5	0.73	0.74	0.79	0.71	0.92	0.94	0.95	0.95	0.82	0.83	0.86	0.81
Mean	0.71	0.71	0.84	0.75	0.92	0.92	0.96	0.95	0.80	0.80	0.89	0.84

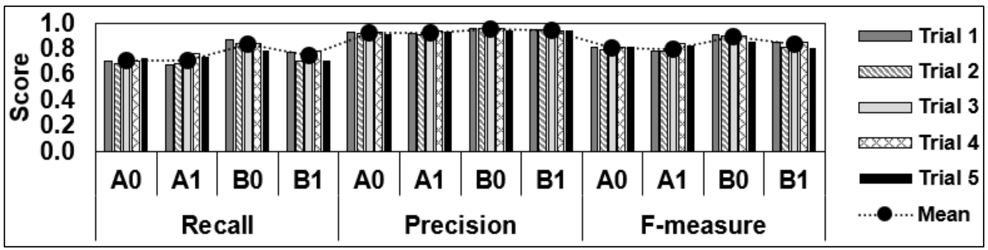

Figure 1: Recall, precision, and F-measure values across all 5 trials

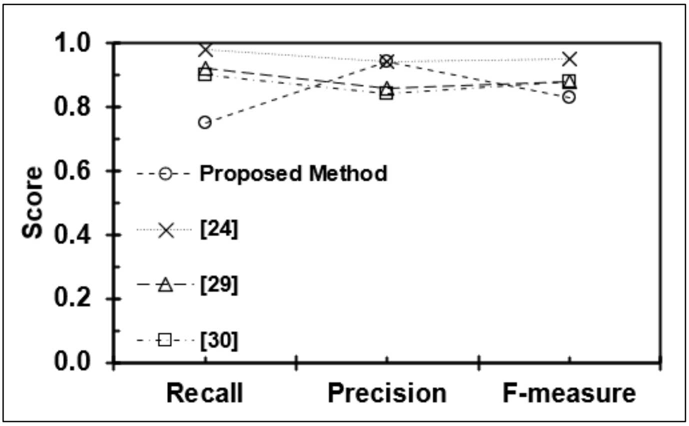

Figure 2: Recall, precision, and F-measure values between the proposed method and previous studies

This indicates the tracker was at best in following player B0. However, there is a slight difference in score between players B0 and B1, even though they are from the same team. This is because player B1 has a relatively lower accuracy of positive detections (recall: < 0.80). This could be due to the player positioning that might cause occasional occlusions. The tracker also performed better in following team B (F-measure: < 0.91) in comparison to team A (F-measure: < 0.84) showing the effect of jersey colour on the tracking efficiency. Team B who wore a red colour jersey had a distinct difference against the playing surface in the footage. Consequently, although not severe, the tracker suffers from a high number of failed detections (FN) as shown by the recall values (recall: < 0.84) and missed detecting players A0 (recall: < 0.73) and A1 (recall: < 0.76) more often. The proposed method is a precise tracker as shown by the low number of incorrect detections (FP) indicated by the high precision values (precision: > 0.92).

In general, the proposed method achieved a comparable visual performance (mean F-measure = 0.83) against the existing method in [24] (mean F-measure = 0.95) as in Figure 2. The author also produced the results for the method in [29] (mean F-measure = 0.88) and the Camshift [30] (mean F-measure = 0.88) algorithm. Worth mentioning that the proposed method measured 1365 frames of data across 5 trials at 27,300 visual data points, compared to the 200 frames of data across 7 trials by the existing method. The data source of the existing method was from the stadium-based camera [24]. Overall, the proposed method's performance was relatively in good agreement with the widely used stadium-based camera setup, while offering the advantage of analysing a larger data volume.

Intra-operator reliability performance

The proposed method measures distance, speed, and time for the purpose of testing reliability. Table 2 shows intra-operator reliability for four players across the 5 trials. Significant differences in distance covered were found between player A0 (mean ± SD: -2.66 ± 7.71 m; %MD: -1.75%), player A1 (mean ± SD: -10.90 ± 9.28 m; %MD: -6.06%), player B0 (mean ± SD: 11.92 ± 7.94 m; %MD: 5.33%), and player B1 (mean ± SD: 5.31 ± 16.41 m; %MD: 2.96%). In contrast, no significant difference in speed was reported between player A0 (mean ± SD: -0.18 ± 0.26 kmh^{-1}; %MD: -2.99%), player A1 (mean ± SD: -0.36 ± 0.37 kmh^{-1}; %MD: -5.06%), player B0 (mean ± SD: 0.32 ± 0.26 kmh^{-1}, %MD: 3.62%), and player B1 (mean ± SD: 0.12 ± 0.73 kmh^{-1}; %MD: 1.69%). Plots obtained using the Bland-Altman graph are presented in Figure 3. The figure represents the spread of data points for speed. Results show a better spread recorded by player A0 (95% CI = -0.69 to 0.33 kmh^{-1}) and player B0 (95% CI: -0.18 to 0.82 kmh^{-1}) compared to player A1(95% CI: -1.08 to 0.37 kmh^{-1}) and player B1 (95% CI: -1.31 to 1.56 kmh^{-1}). Even though, player B1 (bias: 0.12 kmh^{-1}) has an undesirable spread, the bias was near zero. Among all, player A0 (bias = -0.18 kmh^{-1}) recorded the smallest bias. Positive mean differences suggest that users became habituated with the tracker. Opposite results were achieved if negative values were observed.

As a comparison, the mean differences in distance covered (mean ± SD: -2.66 ± 7.71 m to 11.92 ± 7.94 m) by our proposed method were significantly smaller than the inter-unit reliability study of a 15 Hz GPS sport-vest of GPSports system [31]. The absolute mean differences (mean ± SD: 29.6 ± 4.2 to 31.2 ± 8.5 m) of the two units were achieved from a control experiment involving HSR (14.00 to 19.99 $km.h^{-1}$) and VHSR (> 20 $km.h^{-1}$) activities. Meanwhile, the mean difference in peak speed recorded by the proposed method (mean ± SD: 0.43 ± 0.3 $km.h^{-1}$) was slightly better than the peak speed among the LSR, HSR, and VHSR activities (mean ± SD: 0.55 ± 0.55 $km.h^{-1}$) [31] and when forty GPS units (GPSportsEVO: n = 13, mean ± SD: -0.86 ± 0.58 to 0.40 ± 0.29 $km.h^{-1}$; GPSportsHPU: n = 12, mean ± SD: -0.65 ± 0.58 to 0.54 ± 0.29 $km.h^{-1}$; and OptimeyeS5: n = 15, mean ± SD: -0.80 ± 0.58 to 0.43 ± 0.29 $km.h^{-1}$) were pushed together over 40 m tracks in a custom trolley on four different speed zones [32]. The intraclass correlations were acceptable, ranging from 0.54 to 0.66 with a value exceeding 0.7 being a good score. The % mean difference showed good to moderate results ranging from 1.75% to 6.06% for distance covered (mean difference: 0.93 m), 1.69% to 5.06% for speed (mean difference: -0.025 kmh^{-1}), 5.94% for overall distance and 1.28% for peak speed. These results align with the previous research [32] that found almost all distance, and speed, including acceleration/deceleration data, were deemed as good to moderate.

Table 2: Intra-operator reliability of the visual tracking data across 5 trials

Sequence	Tracking	Difference				%MD	SD as CV%	ICC	95CI
	(mean ± SD)	**(mean ± SD)**	**Min**	**Max**	**95CI**				
Distance covered (m)									
Player A0	152.28 ± 5.50	-2.66 ± 7.71	-11.48	9.30	-17.77 ~ 12.45	-1.75	3.61	0.66	0.63~0.68
Player A1	179.81 ± 9.91	-10.90 ± 9.28	-18.76	0.60	-29.09 ~ 7.29	-6.06	5.51	0.58	0.56~0.60
Player B0	223.67 ± 9.97	11.92 ± 7.94	-0.46	27.65	-3.64 ~ 27.49	5.33	4.45	0.57	0.55~0.60
Player B1	179.17 ± 11.63	5.31 ± 16.41	-25.16	25.13	-26.84 ~ 37.47	2.96	6.49	0.54	0.52~0.57
Speed ($km.h^{-1}$)									
Player A0	6.02 ± 0.22	-0.11 ± 0.31	-0.45	0.37	-0.70 ~ 0.49	-2.99	3.65	0.66	0.64~0.68
Player A1	7.11 ± 0.39	-0.43 ± 0.37	-0.74	0.01	-1.15 ~ 0.29	-5.06	5.49	0.58	0.56~0.60
Player B0	8.85 ± 0.39	0.47 ± 0.31	-0.02	1.09	-0.14 ~ 1.08	3.62	4.41	0.57	0.55~0.60
Player B1	7.09 ± 0.46	0.21 ± 0.65	-1.00	0.99	-1.06 ~ 1.48	1.69	6.48	0.54	0.52~0.57
Overall distance (m)	183.74 ± 29.56	10.92 ± 8	0.03	27.65	5.51 ~ 16.33	5.94	16.1	0.15	0.0 to 0.7
Peak speed ($km.h^{-1}$)	33.69 ± 7.23	0.43 ± 0.31	0.001	1.1	0.22 ~ 0.65	1.28	21.5	0.15	0.0 to 0.7
Time spent (s)	239.72 ± 1.86	2.22 ± 1.5	0.20	5.2	1.15 ~ 3.29	0.93	0.78	-	-

*SD = standard deviation; Min = minimum; Max = maximum; CI = confidence interval; MD = mean difference; CV = coefficient of variation; ICC = intraclass correlation; m = meters; $km.h^{-1}$ = kilometres per hour; s = seconds.

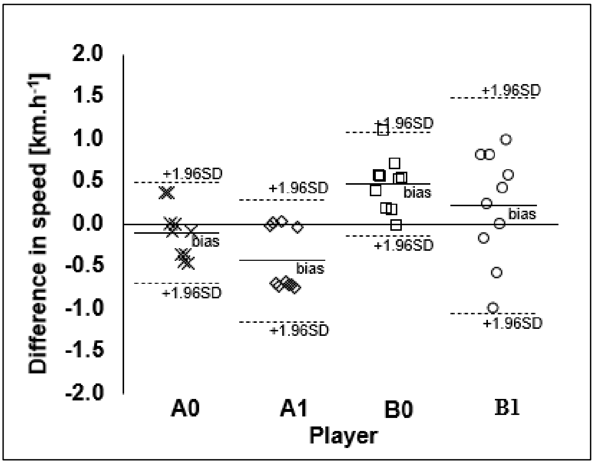

Figure 3: A Bland-Altman plot for the speed in a match measured across 5 trials. The bias (continuous line) and the 95%CI (dashed line) are also presented on the plot

For context, the measures for validity and reliability were rated as good (0% to < 5%) moderate (5% to 10%), or poor (> 10%) as defined by [33]. The rating was similar for both the % mean difference and coefficient of variation. Good to moderate CVs were observed for distance covered (CV: 3.61% to 6.49%) and for speed (CV: 3.65% to 6.48%). While slightly inferior, the dispersions of data around the mean were comparable to the two Optimeye S5 units [34] measuring total distance at the scapula and centre of mass (CV: 1.41% to 3.64%). However, these studies assumed that multiple GPS units were identical. No intra-reliability results of a single unit were reported. These studies measured the movement on predetermined circuits with controlled speed. The proposed method, however, measured the movement on an actual match. Each test subjects move randomly in speed (0.00 km.h^{-1} to 28.00 km.h^{-1}). Undesirable results should be expected. However, the proposed method recorded mostly comparable to slightly better results in both distance-based and speed-based reliability. In general, the measurement performance achieved is comparable even with the commercially available wearable sensors. Across all tests, almost similar time was spent (mean ± SD: 2.22 ± 1.50) during the analysis with overall time spent of 239.72 ± 1.86 s.

Conclusion and Recommendation

In this paper, we used a new computer vision system to track the movement of football players from drone footage by assessing its capability in terms of F-measure and intra-operator reliability. Based on the findings, we conclude that the proposed method is a capable tool to track football players. We assume the tracker is reliable and will produce repeated data consistently well as intended without significant failings. Currently, a comparison of visual performance and intra-operator reliability to other drone research is not possible due to a lack of existing studies. Although this study offers insight into a new finding that involved aerial-based image tracking using footage from a drone that was hovering away from a standard-sized football field, there are still some related issues that require investigation in future research. In our opinion, future works should be directed toward adding other computer vision metrics such as the F2 measure which places more weight on recall, and the F0.5-score which weighs more on precision. This allows the interpretation of different biases that prioritize recall and precision one over the other leading to identification of the truest positives and minimizing false positives. Another issue that is worth investigation is to assess the measurement accuracy and reliability of the proposed method against commercially used tracking technologies (i.e., radar gun, timing gates, and global positioning system (GPS)) by measuring the fundamental X and Y coordinates, distance, speed, and acceleration. Another promising issue is to report the performance using a larger number of participants and a diverse sample of football players.

Contributions of Authors

Khairul Imran Sainan, first/corresponding author: Conceptualization, Project management, Methodology, Investigation, Formal analysis, Validation, Writing – Original draft, Review and Editing. Ahmad Khushairy Makhtar, co-author: Supervision, Writing – Review. Zulkifli Mohamed, co-author: Supervision, Project oversight and management, Investigation, Validation, Writing – Review.

Funding

This work received no specific grant from any funding agency.

Conflict of Interests

The authors declare that they have no conflicts of interest.

Acknowledgment

This project is acknowledged and supported by the Sports Engineering Artificial Intelligence Centre (SEA-IC) UiTM and the College of Engineering, UiTM.

References

[1] P. Rahimian and L. Toka, "Optical tracking in team sports A survey on player and ball tracking methods in soccer and other team sports," *Journal of Quantitative Analysis in Sports*, vol. 18, no. 1, pp. 35–57, 2022. https://doi.org/10.1515/jqas-2020-0088

[2] A. Maglo, A. Orcesi, J. Denize, and Q. C. Pham, "Individual locating of soccer players from a single moving view", *Sensors*, vol. 23, no. 18, pp. 1-28, 2023. https://doi.org/10.3390/s23187938

[3] D. Linke, D. Link, and M. Lames, "Football-specific validity of TRACAB's optical video tracking systems", *PLOS ONE*, vol. 15, no. 3, pp. 1-17, 2020. https://doi.org/10.1371/journal.pone.0230179

[4] R. J. Aughey, K. Ball, S. J. Robertson, G. M. Duthie, F. R. Serpiello, N. Evans, B. Spencer, S. Ellens, E. Cust, J. Haycraft, and Johsan Billingham, "Comparison of a computer vision system against three-dimensional motion capture for tracking football movements in a stadium environment", *Sports Engineering*, vol. 25, no. 1, pp. 1-7, 2022. https://doi.org/10.1007/s12283-021-00365-y

[5] M. Buchheit, A. Allen, T. K. Poon, M. Modonutti, W. Gregson, and V. Di Salvo, "Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement, and GPS technologies," *Journal of Sports Sciences*, vol. 32, no. 20, pp. 1844–1857, 2014. https://doi.org/10.1080/02640414.2014.942687

[6] J. Mara, S. Morgan, K. Pumpa, and K. Thompson, "The accuracy and reliability of a new optical player tracking system for measuring displacement of soccer players," *International Journal of Computer Science in Sport*, vol. 16, no. 3, pp. 175–184, 2017. https://doi.org/10.1515/ijcss-2017-0013

[7] M. Taberner, J. O'Keefer, D. Flower, J. Phillips, G. Close, D. D. Cohen, C. Richter, and C. Carling, "Interchangeability of position tracking technologies; can we merge the data?", *Science and Medicine in Football*, vol. 4, no. 1, pp. 76–81, 2020. https://doi.org/10.1080/24733938.2019.1634279

[8] S. Ellens, D. Hodges, S. McCullagh, J. J. Malone, and M. C. Varley, "Interchangeability of player movement variables from different athlete tracking systems in professional soccer", *Science and Medicine in Football*, vol. 6, no. 1, pp. 1–6, 2022. https://doi.org/10.1080/24733938.2021.1879393

[9] M. Manafifard, H. Ebadi, and H. Abrishami Moghaddam, "A survey on player tracking in soccer videos," *Computer Vision and Image Understanding*, vol. 159, pp. 19–46, 2017. https://doi.org/10.1016/j.cviu.2017.02.002

[10] D. Linke, D. Link, and M. Lames, "Validation of electronic performance and tracking systems EPTS under field conditions", *PLOS ONE*, vol. 13, no. 7, pp. 1-19, 2018. https://doi.org/10.1371/journal.pone.0199519

[11] A. Redwood-Brown, W. Cranton, and C. Sunderland, "Validation of a real-time

video analysis system for soccer," *International Journal of Sports Medicine*, vol. 33, no. 8, pp. 635–640, 2012. https://doi.org/10.1055/s-0032-1306326

[12] M. Manafifard, "A review on camera calibration in soccer videos", *Multimedia Tools and Applications*, vol. 83, no. 6, pp. 18427–18458, 2024. https://doi.org/10.1007/s11042-023-16145-8

[13] H. Li, M. Yang, C. Yang, J. Kang, X. Suo, W. Meng, Z. Li, L. Mao, B. Sheng, and J. Qi, "Soccer match broadcast video analysis method based on detection and tracking", *Computer Animation and Virtual Worlds*, vol. 35, no. 3, p. e2259, 2024. https://doi.org/10.1002/cav.2259.

[14] M. R. Sani, "Multi object tracking in soccer video focusing on occlusion detection and resolving", *Multimedia Tools and Applications*, vol. 82, no. 23, pp. 35913–35947, 2023. https://doi.org/10.1007/s11042-023-14798-z

[15] X. Han, Y. Wang, S. Liu, and C. Jin, "Online multiplayer tracking by extracting temporal contexts with transformer", *Wireless Communications and Mobile Computing*, vol. 2022, no. 1, pp. 1-10, 2022. https://doi.org/10.1155/2022/6177973

[16] F. Zukieman, "*Kim Pan-gon Guna Drone Kenal Pasti Prestasi Setiap Pemain*," May 25, 2022. [Online]. Available: https://makanbola.com/kim-pan-gon-gunakan-drone-kenal-pasti-prestasi/ (Accessed Dec 26, 2023).

[17] M. S. Islam and M. Shahidul, "Introducing drone technology to soccer coaching," *International Journal of Sports Science and Physical Education*, vol. 5, no. 1, pp. 1–4, 2020. https://doi.org/10.11648/j.ijsspe.20200501.11

[18] F. T. Ferreira, J. S. Cardoso, and H. P. Oliveira, "Video analysis in indoor soccer using a quadcopter," *ICPRAM 2015 - 4th International Conference on Pattern Recognition Applications and Methods, Proceedings*, vol. 1, pp. 77–86, 2015. https://doi.org/10.5220/0005206000770086

[19] S. Karungaru, H. Tanioka, and K. Matsuura, "Soccer players real location determination using perspective transformation," *2022 Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems, SCIS and ISIS 2022*, pp. 1–4, 2022. https://doi.org/10.1109/scisisis55246.2022.10001939

[20] T. G. Russomanno, P. Blauberger, O. Kolbinger, H. Lam, M. Schmid, and M. Lames, "Drone-based position detection in sports—validation and applications", *Frontiers in Physiology*, vol. 13, pp. 1-11, 2022. https://doi.org/10.3389/fphys.2022.850512

[21] T. G. Russomanno, P. Blauberger, M. Schmid, O. Kolbinger, and M. Lames, "Position detection in ultimate frisbee using drones," *Symposium der Sektion Sportinformatik und Sporttechnologie der Deutschen Vereinigung für Sportwissenschaft*, vol 13, pp. 1-11, 2022. http://doi.org/ 10.3389/fphys.2022.850512

[22] N. Ding, K. Takeda, W. Jin, Y. Bei, and K. Fujii, "Estimation of control area in badminton doubles with pose information from top and back view drone videos", *Multimedia Tools and Applications*, vol. 83, no. 8, pp. 24777–24793, 2024. https://doi.org/10.1007/s11042-023-16362-1

[23] A. Scott, I. Uchida, M. Onishi, Y. Kameda, K. Fukui, and K. Fujii, "SoccerTrack: A dataset and tracking algorithm for soccer with fish-eye and drone videos," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 3568–3578, 2022. https://doi.org/10.1109/cvprw56347.2022.00401

[24] H. Sabirin, H. Sankoh, and S. Naito, “Automatic soccer player tracking in single camera with robust occlusion handling using attribute matching”, *IEICE Transactions on Information and Systems*, vol. E98D, no. 8, pp. 1580–1588, 2015. https://doi.org/10.1587/transinf.2014EDP7313

[25] M. Lacome, A. Peeters, B. Mathieu, B. Marrier, C. Carling, and J. Piscione, “Can we use GPS for assessing sprinting performance in rugby sevens? A concurrent validity and between-device reliability study”, *Biology of Sport*, vol. 36, no. 1, pp. 25–29, 2019. https://doi.org/10.5114/biolsport.2018.78903

[26] M. Barr, T. Beaver, D. Turczyn, and S. Cornish, “Validity and reliability of 15 hz global positioning system units for assessing the activity profiles of university football players,” *The Journal of Strength & Conditioning Research*, vol. 33, no. 5, pp. 1-24, 2019. https://doi.org/10.1519/JSC.0000000000002076

[27] M. W. Hoppe, C. Baumgart, T. Polglaze, and J. Freiwald, “Validity and reliability of GPS and LPS for measuring distances covered and sprint mechanical properties in team sports”, *PLOS ONE*, vol. 13, no. 2, pp. 1-21, 2018. https://doi.org/10.1371/journal.pone.0192708

[28] G. Atkinson and A. M. Nevill, “Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine”, *Sports Medicine*, vol. 26, no. 4, pp. 217–238, 1998. https://doi.org/10.2165/00007256-199826040-00002

[29] H. Sankoh and S. Naito, “Free-viewpoint video rendering in large outdoor space such as soccer stadium based on object extraction and tracking technology,” *The Journal of The Institute of Image Information and Television Engineers*, vol. 68, no. 3, pp. J125–J134, 2014. https://doi.org/10.3169/itej.68.J125

[30] F. Meyer, “Color image segmentation,” *1992 International Conference on Image Processing and its Applications*, pp. 303–306, 1992.

[31] R. J. Johnston, M. L. Watsford, S. J. Kelly, M. J. Pine, and R. W. Spurrs, “Validity and interunit reliability of 10 hz and 15 hz gps units for assessing athlete movement demands,” *Journal of Strength and Conditioning Research*, vol. 28, no. 6, pp. 1649–1655, 2014. https://doi.org/10.1519/jsc.0000000000000323

[32] R. J. Brosnan, G. Watson, W. Stuart, C. Twentyman, C. M. Kitic, and M. Schmidt, “The validity, reliability, and agreement of global positioning system units - can we compare research and applied data?” *Journal of Strength and Conditioning Research*, vol. 36, no. 12, pp. 3330–3338, 2022. https://doi.org/10.1519/jsc.0000000000004139

[33] H. R. Thornton, A. R. Nelson, J. A. Delaney, F. R. Serpiello, and G. M. Duthie, “Interunit reliability and effect of data-processing methods of global positioning systems,” *International Journal Of Sports Physiology And Performance*, vol. 14, no. 4, pp. 432–438, 2019. https://doi.org/10.1123/ijspp.2018-0273

[34] Z. Akyildiz, F. M. Clemente, D. S. Xentu Rk, G. Rol, M. Yildiz, M. G. Nay, and Y. Ocak, “Investigation of the convergent validity and reliability of unit position differences of catapult s5 gps units in field conditions,” *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology.*, 2022. https://doi.org/10.1177/17543371221100592 10.1177/17543371221100592