Automatic event detection in association football using broadcast-derived tracking data

MILLS, Katie, WANG, Henry, CRANG, Zachary, BILLINGHAM, Johsan, DUTHIE, Grant, JOHNSON, Richard and ROBERTSON, Sam

Citation:

MILLS, Katie, WANG, Henry, CRANG, Zachary, BILLINGHAM, Johsan, DUTHIE, Grant, JOHNSON, Richard and ROBERTSON, Sam (2026). Automatic event detection in association football using broadcast-derived tracking data. Sports Engineering, 29 (1): 15. [Article]

Sports Engineering (2026) 29:15
https://doi.org/10.1007/s12283-026-00549-4

ORIGINAL PAPER

Automatic event detection in association football using broadcast-derived tracking data

Katie Mills[1,2] · Henry Wang[3] · Zachary Crang[4] · Johsan Billingham[1] · Grant Duthie[5,6] · Richard Johnson[4,7] · Sam Robertson[8,9]

Received: 22 September 2025 / Revised: 9 April 2026 / Accepted: 11 April 2026

Abstract

Player and ball tracking data derived from broadcast offers a cost-effective and scalable alternative to multi-camera optical tracking systems in football. However, the practical adoption of broadcast tracking systems depends critically on the accuracy of the data they produce. This study examined the accuracy of using broadcast-derived player and ball position data for automatic event detection for a 90 min match from the 2022 FIFA World Cup. The results were compared against events generated from a high-definition multi-camera optical tracking system (TRACAB Gen 5, ChyronHego, New York, USA) and manually tagged events from FIFA's Data Collection Unit. The results showed that broadcast-derived auto-events, particularly from Camera 1, have the potential to reach the accuracy of multi-camera optical tracking systems for certain events. The best-case performance for most events examined in this study either exceeded, matched, or fell within 0.05 F1-score of the multi-camera system performance. However, performance varied across the systems and was limited by instances of player visibility, ball tracking errors and subjectivity in event definitions, particularly around set pieces and shots. The findings of this study highlights both the capabilities and current limitations of broadcast tracking technologies, provides guidance for their appropriate use and informs future efforts to improve data accuracy in applied contexts.

1 Introduction

Advances in optical tracking systems have transformed performance analysis in football. Since 2017, optical tracking has been the fastest-growing category of Electronic Performance and Tracking Systems (EPTS) assessed by Fédération Internationale de Football Association (FIFA), increasing from fewer than 5 validated systems in 2018 to more than 40 by 2025 [1]. These systems are among the most accurate currently available, with reported positional errors around 0.15 m and speed errors around 0.15 $m \cdot s^{-1}$ [2]. However, these systems are often costly and require bespoke multi-camera set-ups, which limits their scalability across all levels of the game. An increasingly popular alternative is to extract the positional data of players and the ball from broadcast video. Since being formally recognized by FIFA as an EPTS category in 2021, the number of certified broadcast video systems has grown to 11 in 2025 [1]. Their appeal is largely due to their increased accessibility and minimal installation requirements, making them a popular choice for performance analysis, scouting, player monitoring, and fan engagement across a wide range of stadium environments. However, the usability of these systems

✉ Katie Mills
katie.mills@fifa.org

1 Fédération Internationale de Football Association, Zürich, Switzerland

2 Sports Engineering Research Group, Sheffield Hallam University, Sheffield, UK

3 MIT Sports Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

4 School of Exercise Science, Australian Catholic University, Brisbane, QLD, Australia

5 Global Sport and Movement Collaborative, University of Newcastle, Newcastle, NSW, Australia

6 School of Biomedical Sciences and Pharmacy, College of Health, Medicine, and Wellbeing, University of Newcastle, Newcastle, Australia

7 Sports Performance, Recovery, Injury and New Technologies (SPRINT) Research Centre, Australian Catholic University, Melbourne, VIC, Australia

8 Institute for Health and Sport (IHeS), Victoria University, Melbourne, VIC, Australia

9 School of Human Movement and Nutrition Sciences, The University of Queensland, Brisbane, QLD, Australia

Published online: 11 May 2026

depends (amongst other factors) critically on the accuracy of the tracking data they produce.

Previous work demonstrated that player position accuracy below 0.5 m can be achieved over a 90 min match from three commercially available broadcast systems [3]. While this validation addressed a key component of spatiotemporal analytics, realizing the potential of broadcast tracking data for auto-eventing, requires establishing ball tracking positional accuracy.

In addition to player measurement, ball tracking in football plays a crucial role in modern performance analysis, providing valuable technical and tactical insights into opposition team strategies, including patterns of possession, attacking build up and defensive transitions [4–6]. While previous research has demonstrated that detecting the ball in video footage can be achieved [7, 8], the ability to track its position on the field relative to a reference system (i.e., x, y, z co-ordinates) has not been investigated for a full match. This is primarily due to limited access to high-quality criterion ball positional data from multi-camera optical tracking systems for validation. Other studies have used high-speed 3D motion capture [9] and local positioning systems (LPS) [10], though such methods require specialized setups that are not implementable for full-field matches. Given the increasing reliance on broadcast footage for analysis, determining the accuracy of ball tracking from such footage would provide significant benefits for teams and analysts. For example, accurate ball tracking is a prerequisite for automatic event detection (e.g., kickoffs, line breaks, key passes, and shot trajectories), which in turn enables the identification of set pieces at specific locations (e.g., corner spot for corner kicks, center spot for kickoffs) and registration of passes, receptions, and interceptions by detecting changes in player possession from relative player–ball distances.

In addition to player and ball tracking, computer-vision has been used for the detection of football events (e.g., passes, shots, set pieces), often called "action spotting" from match images or video. Existing methods frame event detection as a multiclass classification task, training deep learning architectures on extensive match datasets to categorize specific images or video clips into discrete events [11–14]. Traditional event collection methods relied on multiple human operators, representing obstacles for scalability and data democratization. Auto-eventing algorithms, however, encode football logic defined by the FIFA Football Language [15] and use tracking data to detect events without human intervention. Importantly, auto-event accuracy is closely tied to the quality of tracking data used by the algorithm, and lower-quality inputs will yield less accurate auto-events.

This study aimed to evaluate the accuracy of using broadcast-derived tracking data for automatic event detection. To address this, we first quantified the positional accuracy of ball tracking derived from three commercially available broadcast tracking systems and integrated this data with previously validated player tracking data (from the same providers and match) to generate events. Importantly, the objective was not to benchmark individual broadcast tracking providers, but rather to characterize the general performance and limitations of broadcast tracking technologies for automatic-event detection using different camera angles, in an applied, match-wide context.

2 Methods

The methods used in this study replicated those described in earlier work which evaluated the accuracy of player tracking using broadcast footage of three commercially available broadcast tracking providers [3]. In the present study, this approach was extended by extracting and assessing ball tracking data enabling the creation of a more complete dataset required for automated event detection. All procedures were approved by a Human Research Ethics Committee (ID: 2023-3480N).

Data were collected during a single group stage match of the 2022 Qatar FIFA World Cup. The match was filmed by the television broadcasters using multiple fixed video cameras positioned around the pitch at 50 frames per second (fps) and stored at 25 fps. The Tactical, Programme (PGM), and Camera 1 video feeds were obtained from the FIFA Data Platform at 1080p in MPEG-4 Part-14 (.mp4) file format. The Tactical feed is a wide-view angle of the pitch (Fig. 1A), captured from a fixed camera positioned on the top tier of the grandstand, at the halfway line. The purpose of this camera is to maintain 20 outfield players in shot at any one time, and as a result, the ball is made up of fewer pixels. The PGM feed (Fig. 1B) consists of multiple camera angles and is what is televised to the public (e.g., including score updates, graphics and other replays). Unlike other camera angles (Fig. 2A), the PGM feed (Fig. 2B) often offers a more focused and consistent view of the ball, as it tends to switch between different camera angles to zoom in on key events (e.g. free kicks), keeping the ball more prominently in the camera frame at a higher pixel resolution. The Camera 1 feed (Fig. 1C) captures a slightly tighter field of view compared to the Tactical feed. It makes up the majority of the PGM feed, except it does not have any graphical overlay (e.g., scoreboard), replays, or cut to different camera angles.

A combination of proprietary computer-vision and artificial intelligence (AI) techniques were used by the broadcast tracking providers to track the overall position (x, y, z coordinates) of the ball at 25 fps. Data was delivered in

Fig. 1 Broadcast footage from different feeds, captured at the same moment in time: **A** Tactical, **B** Programme (PGM), and **C** Camera 1 feed

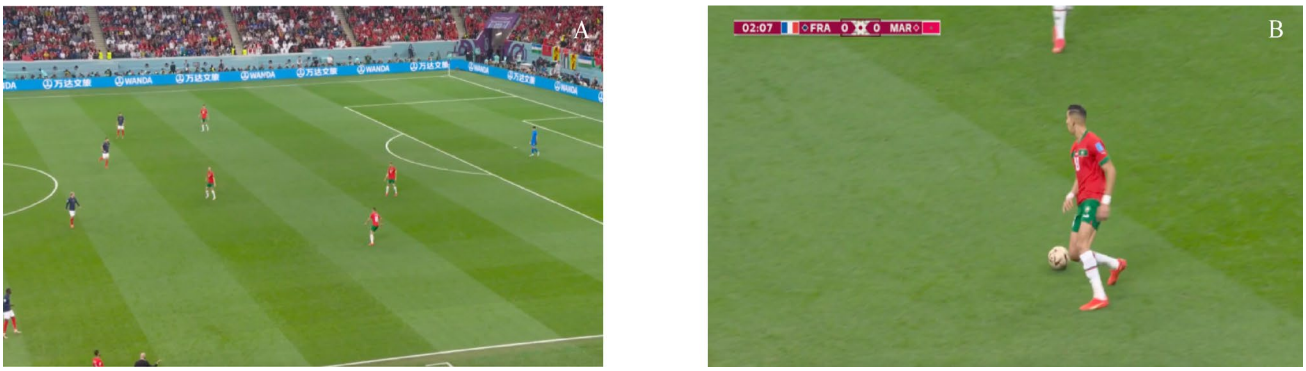

Fig. 2 Camera 1 **A** provides the main broadcast view, but certain passing sequences or events can be obscured. In such cases, the PGM feed **B** may switch to an alternative camera angle to improve ball visibility. Image captured at the same moment in time

comma-separated value (.csv) files in their raw sampling frequency, where each row provided an observation of ball position at 25 Hz.

2.1 Determining ball tracking accuracy

To establish the concurrent validity, providers were compared against a high-definition multi-camera optical tracking system, capturing data at 25 Hz (TRACAB Gen 5, ChyronHego, New York, USA). The system uses 12 fixed cameras elevated within the stadium infrastructure. This system has strong accuracy for measures of player position (RMSE = 0.08 m) compared to 3D motion capture [16] and for ball position (MAE = 0.07 m) compared to a Vicon camera system [17]. As per the author's previous work [3], tracking data from each provider, for each individual player, were temporally aligned by phase shifting the players' speeds in the TRACAB dataset to minimize the RMSE, with the median phase shift applied across all files. Spatial alignment was subsequently achieved by rotating the tracking data throughout 360 degrees in 0.1-degree increments and applying X–Y translations to minimize positional error. This alignment method was chosen to preserve the original scale of the data. The resulting transformation parameters were applied uniformly to both player and ball tracking data.

The statistical analyses were performed in RStudio (version 12.1; Posit, Boston, MA) using the R programming language (version 4.3.3, R Foundation for Statistical Computing, Vienna, Austria). To determine the agreement between the different providers and the multi-camera system, mean bias and the limits of agreement (LOA) were estimated from linear models built using the *stats::lm* function. Separate models were used for each feed (PGM, Camera 1, Tactical). Each model was run with an intercept only, with position error (i.e., difference to the multi-camera system) used as the outcome variable. From each model, mean bias ± 95% limits of agreement (LOA) were extracted [18, 19]. The root mean square error (RMSE) was separately calculated for position to quantify absolute error.

Linear models were also used to establish the influence of provider and camera feed (e.g., PGM vs. Tactical) on the accuracy of position. The RMSE for position was used as the outcome variable, with the interaction between provider and camera feed used as predictor variables. The main effects from each model were extracted using the *stats::anova* function. Where significant main effects were observed, post hoc tests were performed using the *emmeans*::*emmeans* function with a Tukey adjustment applied to account for multiple comparisons. Data are presented as mean ± SD; statistical significance was set at $\alpha = 0.05$.

2.2 Generating auto-events from broadcast tracking data

This study employed the Python implementation (version 3.9.12; Python Software Foundation, Wilmington, DE) of a deterministic, rules-based auto-eventing algorithm developed by Vida-Codina et al. [20]. A high-level description is provided below but it is recommended that the reader refer to the full paper for a comprehensive overview. The algorithm uses three inputs: x, y, z coordinate tracking data of the ball and x, y coordinates of players' center of mass, along with the frame-by-frame ball status ("live" or "dead"). It operates in two phases: asserting possession and detecting events. Possession occurs when the ball status is live and is determined based on the proximity of the ball to a player within a predefined distance *pz*, an adjustable hyperparameter representing the radius of the "possession zone". The algorithm utilizes event-specific triggers (player positions) and patterns (ball location) defined by additional hyperparameters based on football logic and the FIFA Football Language.

Set pieces are detected when the ball status switches from "dead" to "live" and assigned based on trigger and pattern criteria. For example, a corner kick is detected when the ball status switches from "dead" to "live" and a player and the ball share a possession zone, and both are positioned no further than 3 m away from a corner of the pitch (pattern). Similarly, a kickoff is detected when the ball is near the center spot (pattern), and all players are on their respective sides of the field (trigger). Notably, goals are detected by finding the kickoff that must succeed using the same logic. These auto-events are attributed to players via a set-piece-specific possession zone, defined by another set of hyperparameters.

Passes and crosses are tagged by detecting changes in ball possession and trajectory. For example, a pass from player A to player B occurs when the ball moves from player A's possession zone to player B's with a trajectory change, indicating a touch by B. Shots are defined as attacking players losing possession immediately followed by a goal, corner kick, goal kick, or goalkeeper's save. A save occurs when the goalkeeper gains possession within the penalty area immediately after a shot.

The auto-events evaluated for this study were set pieces (kickoffs, goal kicks, free kicks, corner kicks, and throw-ins), passes/crosses, shots, goals, and saves. Criterion event data were collected from the FIFA Data Collection Unit (DCU) dataset containing events manually tagged by 22 humans watching the match, each focused on one player. From the DCU data, there were a total of 4 kickoffs, 18 goal kicks, 28 free kicks, 5 corner kicks, and 31 throw-ins. In addition, there were 984 pass/cross events, 27 shots, 2 goals, and 3 saves. Tracking data were collected from a set of nine provider-feed configurations, which were systematically generated by pairing each of the three data providers with each of the three distinct data feeds (PGM, Camera 1, and Tactical). We generated auto-events for each configuration and employed default hyperparameters governing indicators of possession, and set pieces used by Vidal-Codina et al. [20]. As a benchmark, auto-events were generated using player and ball data from the multi-camera optical tracking system. Provider 1 supplied their own ball possession and status information, whereas Providers 2 and 3 did not and ball status was utilized from the multi-camera system. For validation, DCU events were matched with auto-events through iterative comparison, where matches were established when auto-events of identical type occurred within 3 s of the DCU event, consistent with the methods in Vidal-Codina et al. [20]. The F1 score, the harmonic mean of precision and recall, was computed for each event type to quantitatively assess detection accuracy. Subsequently, common sources of error were manually examined.

3 Results

3.1 Ball tracking

Table 1 presents the ball positional accuracy of each provider. There were significant main effects of both provider and camera feed on overall positional accuracy. The configuration yielding the best accuracy varied across providers, with the most accurate configuration from Provider 1 using the PGM feed (RMSE = 3.5 m; mean bias [LOA] = 1.5 m [− 4.7 to 7.7 m]). The best for Provider 2 was the Tactical feed (RMSE = 9.1 m; mean bias [LOA] = 5.0 m [− 9.9 to 19.9 m]), while for Provider 3 it was the PGM feed (RMSE = 5.5 m; mean bias [LOA] = 2.5 m [− 7.1 to 12.1 m]). The PGM feeds from Providers 1 and 3 demonstrated significantly better accuracy compared to their respective Camera 1 and Tactical feeds. Conversely for Provider 2, the Tactical feed had significantly better accuracy compared to the PGM and Camera 1 feeds. The spatial distribution of positional error across the pitch for the three providers is presented in Fig. 3. The Tactical feed exhibits more widespread errors, particularly in the attacking areas, which are located at the extremities of the camera frame.

3.2 Auto-eventing

Across both trials, systems that leveraged the Camera 1 feed consistently yielded the most accurate auto-events, followed by Tactical, with PGM weakest, which was largely due to broadcast replay artifacts. For several set piece and open play events, broadcast auto-events approached or

Table 1 Concurrent validity of ball software tracking systems to measure position (m) during football match-play in comparison to an optical tracking system

	Position error	RMSE (m)			Mean bias (m) ± 95% LOA		
		PGM	Camera 1	Tactical	PGM	Camera 1	Tactical
Provider 1	X	2.6	4.0	9.0	−0.3 ± 5.1	−0.3 ± 7.8	−0.7 ± 17.3
	Y	2.2	4.7	7.2	−0.03 ± 4.3	−0.7 ± 9.2	0.3 ± 14.0
	Z	0.9	0.9	1.4	−0.1 ± 1.7	−0.1 ± 1.8	0.1 ± 2.7
	Overall	3.5[ab]	6.3[b]	11.6	1.5 ± 6.2	2.5 ± 11.3	5.6 ± 19.6
Provider 2	X	8.9	9.0	6.9	−1.2 ± 17.4	−1.0 ± 17.5	−0.2 ± 13.5
	Y	6.7	6.4	5.7	−0.5 ± 13.1	−0.6 ± 12.6	−0.1 ± 11.2
	Z	1.7	1.7	1.7	−0.6 ± 3.1	−0.6 ± 3.1	−0.6 ± 3.1
	Overall	11.3[b]	11.2[b]	9.1	6.6 ± 18.0	6.9 ± 17.3	5.0 ± 14.9
Provider 3	X	4.4	4.3	12.2	−0.8 ± 8.2	−0.6 ± 8.4	0.1 ± 24.0
	Y	3.2	4.8	10.6	−0.5 ± 6.6	−0.1 ± 9.3	1.2 ± 20.8
	Z	0.8	1.0	1.5	−0.1 ± 2.1	0.003 ± 2.1	−0.1 ± 3.0
	Overall	5.5[ab]	6.6[b]	16.2	2.5 ± 9.6	2.9 ± 11.4	11.0 ± 23.6

RMSE; root mean square error, LOA; limits of agreement, PGM; PGM feed

a; significant difference ($p < 0.05$) to Camera 1 feed for the same provider

b; significant difference ($p < 0.05$) to Tactical camera feed for the same provider

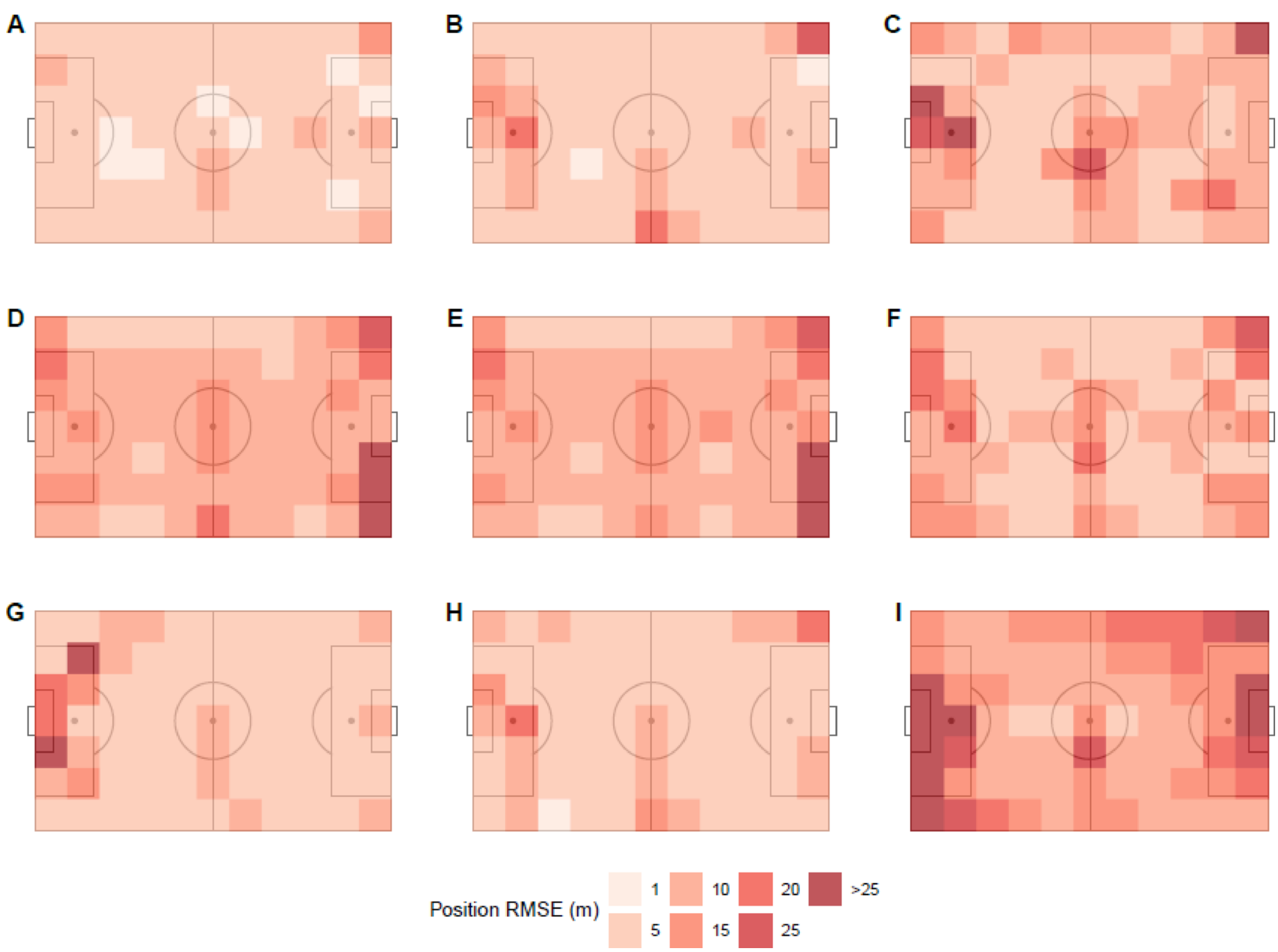

Fig. 3 Location of position error (m) for providers 1 (**A, B, C**), 2 (**D, E, F**) and 3 (**G, H, I**) for PGM (**A, D, G**), Camera 1 (**B, E, H**) and Tactical (**C, F, I**) feeds

Table 2 F1-Scores across all events, feeds, and providers

		Program			Camera 1			Tactical			Multi-camera
		Provider 1	Provider 2	Provider 3	Provider 1	Provider 2	Provider 3	Provider 1	Provider 2	Provider 3	TRACAB
Set Piece Events	Kickoff	0.4	0.67	0.67	0.86	0.86	1	0.67	1	1	1
	Goal Kick	0.42	0.62	0	0.97	0.76	0.88	0.8	0.94	0.29	0.94
	Free Kick	0.57	0.52	0.48	0.86	0.39	0.61	0.61	0.84	0.49	0.93
	Corner Kick	0	1	0.33	1	0.75	0.67	0.89	1	0.89	1
	Throw in	0.58	0.6	0.5	0.87	0.52	0.67	0.69	0.87	0.52	0.97
Passing Events	Pass/Cross	0.81	0.91	0.63	0.88	0.9	0.65	0.74	0.93	0.41	0.93
	Shots	0.07	0.24	0	0.11	0.12	0.07	0.06	0.4	0	0.48
Goals & Saves	Goals	0	0	0	1	0	0	0.67	0.67	0.67	1
	Saves	0	0	0	0	0	0	0	0	0	0.67

Table 3 Precision across all events, feeds, and providers

		Program			Camera 1			Tactical			Multi-camera
		Provider 1	Provider 2	Provider 3	Provider 1	Provider 2	Provider 3	Provider 1	Provider 2	Provider 3	TRACAB
Set Piece Events	Kickoff	1	1	1	1	1	1	1	1	1	1
	Goal Kick	0.83	0.82	0	1	1	1	1	1	1	1
	Free Kick	0.57	0.44	0.5	0.83	0.33	0.51	0.51	0.83	0.39	0.96
	Corner Kick	0	1	1	1	1	0.75	1	1	1	1
	Throw in	0.82	0.88	0.85	0.87	0.8	0.85	0.79	0.9	0.68	0.94
Passing Events	Pass/Cross	0.84	0.93	0.58	0.91	0.93	0.59	0.76	0.95	0.34	0.89
	Shots	0.5	0.57	0	0.25	0.29	0.5	0.14	0.62	0	0.67
Goals & Saves	Goals	0	0	0	1	0	0	1	1	1	1
	Saves	0	0	0	0	0	0	0	0	0	0.67

Table 4 Recall across all events, feeds, and providers

		Program			Camera 1			Tactical			Multi-camera
		Provider 1	Provider 2	Provider 3	Provider 1	Provider 2	Provider 3	Provider 1	Provider 2	Provider 3	TRACAB
Set Piece Events	Kickoff	0.25	0.5	0.5	0.75	0.75	1	0.5	1	1	1
	Goal Kick	0.28	0.5	0	0.94	0.61	0.78	0.67	0.89	0.17	0.89
	Free Kick	0.57	0.64	0.46	0.89	0.5	0.75	0.75	0.86	0.68	0.89
	Corner Kick	0	1	0.2	1	0.6	0.6	0.8	1	0.8	1
	Throw in	0.45	0.45	0.35	0.87	0.39	0.55	0.61	0.84	0.42	1
Passing Events	Pass/Cross	0.78	0.89	0.68	0.86	0.87	0.72	0.72	0.91	0.52	0.98
	Shots	0.04	0.15	0	0.07	0.07	0.04	0.04	0.3	0	0.37
Goals & Saves	Goals	0	0	0	1	0	0	0.5	0.5	0.5	1
	Saves	0	0	0	0	0	0	0	0	0	0.67

matched the multi-camera optical benchmark whereas shots and saves remained clear areas for improvement. For a more holistic approach, we emphasize the best F1-scores across feed–provider–trial configurations and reference the multi-camera optical tracking system benchmark to contextualize what is currently achievable. We report F1-scores for all events in Table 2, complemented by precision and recall metrics in Table 3 and 4, respectively.

Set pieces

Kickoffs Auto-eventing achieved perfect detection (F1 = 1.00) in multiple configurations, matching the multi-camera system. The best performance occurred for auto-events produced using the Camera 1 and Tactical feeds, where multiple providers achieved perfect detection. In contrast, kickoff detection produced using the PGM feeds were significantly worse as all providers detected only 1 or 2 kickoffs out of 4, a substantial drop compared to the other camera views, as reported in Table 2.

Goal kicks Systems using Camera 1 auto-events reached as high as F1 = 0.97, exceeding the multi-camera's 0.94. In the best case for the other feeds, Tactical auto-events achieved F1 = 0.94 and PGM auto-events reached F1 = 0.62. We find PGM, on average, was the worst feed for detecting goal kicks, with certain configurations that failed to detect any.

Free kicks Relative to the events generated from the multi-camera optical tracking system (F1 = 0.93), broadcast auto-eventing was competitive for detecting free kicks, as the best cases reached F1 > 0.85 on both Camera 1 and Tactical auto-events. In contrast, PGM auto-events performed worse, peaking at F1 = 0.57.

Corner kicks Multiple configurations across all PGM, Camera 1, and Tactical feeds achieved perfect detection (F1 = 1.00), matching the multi-camera system. Occasional complete PGM-based detection failures (no detections) highlight instability of that feed for corners.

Throw-ins Compared to the multi-camera's performance of F1 = 0.97, throw-ins were detected relatively effectively, with the highest F1-score reaching 0.87 from both Camera 1 and Tactical auto-events, whereas PGM auto-events peaked lower at F1 = 0.60.

Passes, crosses, and shots

Passes and crosses The detection of passes and crosses was widely competitive against the 0.93 F1-score of the multi-camera system, as all three video feed formats produced auto-events achieving F1 $\geq$ 0.90. Across all providers, average performance was best using the Camera 1 feed (mean F1 = 0.81), followed by the PGM feed (mean F1 = 0.78), and then the Tactical feed (mean F1 = 0.69). For comparison, the highest maximum F1 observed across all providers was 0.93 for Camera 1, 0.91 for PGM, and 0.90 for Tactical.

Shots Shots are already difficult to detect due to subjectivity about intent, resulting in many deflected shots being mistaken for passes. While the events generated from the multi-camera system yielded an F1-score of 0.48, the highest observed provider score was F1 = 0.24, coming from PGM auto-events. Improved detection of shots remains a known area for improvement in auto-eventing.

Goals & saves

Goals Multiple Camera 1 and Tactical feed configurations matched the perfect detection of goals from the multi-camera system. PGM, however, struggled to detect any goals due to TV replays during the subsequent kickoff, which are not present for Camera 1 and Tactical feeds. Additionally, across all video feed formats, some goals were misclassified as own goals, which stemmed from suboptimal player and ball tracking in the box where the last touch prior to the dead ball was assigned to the wrong team. Finally, across some video feeds, some goals are missed due to poor tracking of the ball near the center spot during the subsequent kickoff, resulting in the kickoff not being detected, and thus no goal being flagged.

Saves Saves were not detected in any instance by the broadcast-derived auto-events, although results were also modest for the multi-camera optical tracking system (F1 = 0.67), underscoring the need for improved save detection in existing auto-eventing algorithms.

4 Discussion

This study demonstrated that player and ball positional accuracy was strongly influenced by the type of broadcast feed, revealing trade-offs between tight, ball-centric views and wide, player-centric perspectives. Moreover, the findings demonstrated that broadcast-derived tracking data can yield auto-events with accuracy approaching that of established multi-camera systems for certain events only. These findings highlight the viability of emerging broadcast tracking technologies as scalable alternatives for selected analytical applications.

4.1 Ball tracking discussion

The findings indicate that current computer-vision and AI methods do not yet provide sufficient accuracy in determining the position of the ball from a single camera feed, with accuracy largely determined by camera feed characteristics and provider-specific model architectures. Algorithms leveraging the PGM feed generally yielded the best accuracy for ball position, compared to those using the Camera 1 and Tactical feeds. This is likely due to the PGM feeds' tight camera angles, that consistently keep the ball within frame and thereby maximize the pixel resolution of the object of interest. Whereas the wider Camera 1 and Tactical feeds reduced pixel resolution and impaired ball detection. This finding aligns with previous work that found the Tactical feed was generally superior when tracking the position of players [3], owning to the larger visual footprint of players relative to the ball. This underscores the need for careful consideration of feed selection by practitioners or other users of the data with their intended research objectives (ball-centric vs player-centric analysis) and must acknowledge the capabilities and limitations of each feed when interpreting findings or making performance decisions.

Although, some systems demonstrated relatively strong performance in tracking the ball using the Tactical feed, the overall position accuracy of the ball (RMSE = 3.5 to 16.2 m) remains far below the thresholds required for officiating (e.g., offside detection), or detailed tactical analysis (e.g., pass-completion analysis or shot locations). Alternative technologies, such as ball-integrated inertial measurement units (IMU) that leverage LPS or multi-optical camera systems, can provide much greater accuracy in comparison (RMSE = 0.05 to 0.35 m) [21, 22] but are resource heavy and require specialist technical knowledge, limiting their scalability across different levels of the game. Therefore, broadcast tracking systems may still be appropriate to support tactical analysis where precise spatial accuracy. For example, identifying zones of play, possession transitions, and ball-in-play time.

4.2 Auto-eventing discussion

Results indicated that systems that leveraged the PGM feed were generally outperformed by those leveraging the Camera 1 and Tactical feeds due to increased player and ball visibility. Camera 1 excelled at on-ball event detection with its focused and consistent view of the ball. While slightly less effective for some events, but better for throw-ins, Tactical views yielded more accurate player tracking [3], which suggests an optimal approach might be to combine Camera 1 and Tactical feeds. We note that for Providers 1 and 3, the Camera 1 F1-scores were greater than or equal to the corresponding PGM scores for all event types, as expected given that the PGM feed is derived from Camera 1 with additional production artifacts. Counterintuitively, this pattern does not hold for Provider 2, which may reflect internal processing steps that are not accessible to the authors.

The key sources of error in auto-eventing were rooted in suboptimal player and ball tracking, ball status errors, and fundamental difficulties in defining some events, such as shots. Firstly, tracking quality typically deteriorated during occlusions in congested areas like the penalty box, which affected goal attribution in the form of incorrect own goals. Insufficient ball tracking accuracy led to kickoffs mistaken for free kicks when the ball position deviated significantly from the center spot, and imprecision near the touchlines caused confusion between throw-ins and free kicks when the ball's tracked position was wrongly marked as in-or out of bounds. Detection of passing events when subjects are outside camera view and incorrectly tracked is limited, underscoring the role of player visibility, in line with previous research [23]. Additionally, the lack of ball height data impaired the detection of goalkeeper actions such as saves and claims. Alongside tracking challenges, some set-piece errors were attributed to incorrect ball status. Furthermore, when the ball status switched to live earlier than in reality, this resulted in the wrong set piece being detected due to an incorrect sequence of events prior. Finally, there are known limitations in the auto-eventing algorithm that hinder optimal performance, particularly for shots and saves. The reader should refer to Vidal-Codina et al. [20] for more details.

We also note that the performance of Provider 1 may be conservatively estimated relative to Providers 2 and 3, which benefit from "borrowing" multi-camera–derived ball live/dead status information. The fact that Provider 1 achieved these results using only its own broadcast-based possession signals provides a promising indication of the potential feasibility of a true single-source broadcast-derived event detection pipeline.

The results of this study demonstrate that broadcast tracking data has the potential to produce auto-events with comparable accuracy to those derived from established multi-camera systems, with best-case performance for most events exceeding, matching of falling within 0.05 F1-score of the optical tracking system benchmark. This is particularly noteworthy, given that this represents, to the authors knowledge, the first evaluation of broadcast-derived auto-eventing performance and constitutes an important step towards democratizing access to performance data beyond elite settings. The findings highlight an inherent trade-off between solutions that are optimal for player tracking to those that are optimal for ball tracking, suggesting that a hybrid configuration, which integrates either the Camera 1 or the PGM feed with a Tactical feed, may yield the most accurate outputs from existing broadcast infrastructure.

Nevertheless, several limitations must be acknowledged. The evaluation was restricted to a single match, and performance varied highly across the three providers examined, complicating direct like-for-like comparisons and indicating the critical influence of feed-dependent optimization within current industry practice. Comparisons are further challenged by the absence of universally standardized event definitions which introduces a degree of subjectivity into the benchmarking process. Fundamentally, while this study evaluated 9 key events, full match analysis typically requires more than 100 distinct events to capture the complete complexity of a match. Achieving reliable large-scale adoption will therefore require further exploration and validation across a broader range of event categories.

Several areas emerge as priorities for future research and industry development. The methodology presented in this research not only enables investigation of these directions but can also be extended to explore alternative alignment approaches between tracking data obtained from two different sources. First, improvements in extrapolation algorithms are essential, as tracking errors increase substantially when players or the ball are occluded. Second, moving beyond simple center-of-mass representations towards three-dimensional pose estimations from monocular video would provide richer information for auto-eventing. Third, the optimal camera configuration for broadcast tracking warrants investigation: are current broadcast camera positions sufficient, or could modified placements better balance accuracy and visibility? Finally, higher-resolution video, such as 4 K, has the potential to improve ball detection by increasing pixel density, that may yield meaningful gains in tracking accuracy from Tactical feeds.

Together, these directions underscore the importance of continued collaboration between researchers, technology providers and governing bodies, to enhance the accuracy, scalability and interpretability of broadcast tracking technologies. These collaborations will accelerate the impact of broadcast tracking systems on performance analysis, officiating and fan engagement across all levels of the game.

5 Conclusion

Tracking data derived from broadcast cameras, for *certain events*, data formats, and providers, has the potential to produce auto-events with comparable accuracy to those derived from multi-camera systems; however, performance was found to be highly sensitive to the type of camera feed and quality of the underlying tracking data.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12283-026-00549-4.

Acknowledgements We would like to thank the broadcast tracking providers participating in this research. We also express gratitude to Ramzi Bensaid and Dr. Ferran Vidal-Codina for their guidance on the auto-eventing task, as well as to Christina Chase and Professor Peko Hosoi for their mentorship and guidance.

Author contribution KM, JB, and SR guided the overall research design and defined the research objectives. KM provided the video data to broadcast tracking providers, coordinated collaboration between HW, ZC, GD, and RJ, and drafted the manuscript with input from all authors. HW conducted the auto-eventing algorithms and analysis using the player and ball data prepared and validated by GD, RJ, and ZC. All authors contributed to interpreting results and revised the final manuscript.

Funding This research was funded by FIFA. All data subjects were informed ahead of competition that “Optical player tracking data will be collected…, and used for… research, and development purposes”.

Data availability The data used for this study was collected by FIFA and collaborating broadcast tracking providers. Due to media and data rights, the datasets are not publicly available, but can be requested by contacting the authors of this article.

Declarations

Conflict of interests Two co-authors serve as Guest Editors for the Topical Collection on Football Research II in Sports Engineering and had no role in the blind peer review process of this paper.

References

1. FIFA (2025). FIFA Quality Programme Database. Version 5.3.1.3. Accessed 29.07.2025
2. FIFA (2024) Football Research & Standards. Learnings from a Decade of Assessing Electronic Performance & Tracking.
3. Crang ZL, Johnston RD, Mills KL, Billingham J, Robertson S, Cole MH, Weakley J, Hewitt A, Duthie GM (2025) Concurrent validity of computer-vision artificial intelligence player tracking software using broadcast footage arXiv https://doi.org/10.48550/arXiv.2508.19477
4. Wang S, Qin Y, Jia Y, Igor KE (2022) A systematic review about the performance indicators related to ball possession PLoS ONE 17(3):e0265540 https://doi.org/10.1371/journal.pone.0265540
5. Goes FR, Brink MS, Elferink-Gemser MT, Kempe M, Lemmink KAPM (2020) The tactics of successful attacks in professional association football: large-scale spatiotemporal analysis of dynamic subgroups using position tracking data J Sports Sci 39(5):523–532 https://doi.org/10.1080/02640414.2020.1834689
6. Forcher L, Altmann S, Forcher L, Jekauc D, Kempe M (2022) The use of player tracking data to analyze defensive play in professional soccer–a scoping review. Int J Sports Sci Coach 17(6):1567–1592
7. Kamble PR, Keskar AG, Bhurchandi KM (2019) A deep learning ball tracking system in soccer videos. Opto-Electron Rev 27:58–69
8. Kim J-Y, Kim T-Y. (2009) Soccer ball tracking using dynamic kalman filter with velocity control. In: 2009 Sixth International Conference on computer graphics, imaging and visualization IEEE: p. 367–74.
9. Ball K, Haycraft J, Bright L, Robertson S, Billingham J, Evans N (2022) Ball tracking in football. ISBS Proceedings Archive. 40:62
10. Blauberger P, Marzilger R, Lames M (2021) Validation of player and ball tracking with a local positioning system. Sensors Basel 21:1465
11. Darapaneni N, Kumar P, Malhotra N, Sundaramurthy V, Thakur A, Chauhan S, Thangeda KC, Paduri AR. (2022) “Detecting key Soccer match events to create highlights using Computer Vision,” arXiv preprint arXiv:2204.02573, [Online]. Available: https://arxiv.org/abs/2204.02573
12. Deliège A, Cioppa A, Giancola S, Seikavandi MJ, Dueholm JV, Nasrollahi K, Ghanem B, Moeslund TB, Van Droogenbroeck M. (2021) SoccerNet-v2: A dataset and benchmarks for holistic understanding of broadcast soccer videos. In: Proc. IEEE Conf. Computer. Vis. Pattern Recognition. (CVPR) Workshops.
13. Giancola S, Amine M, Dghaily T, Ghanem B (2018) Soccernet: A scalable dataset for action spotting in soccer videos. In: Proc. IEEE Conf. Computer Vis. Pattern Recognition (CVPR) Workshops, pp. 1711–1721.

14. Zhou X, Kang L, Cheng Z, He B, Xin J (2021) Feature combination meets attention: Baidu soccer embeddings and transformer based temporal detection," arXiv preprint arXiv:2106.14447, [Online]. Available: https://arxiv.org/abs/2106.14447
15. FIFA. (2022) The FIFA Football Language, FIFA Training Centre. [Online]. Available at: https://www.fifatrainingcentre.com/en/game/performance-analysis/football-language-analysis/the-fifa-football-language.php
16. Linke D, Link D, Lames M (2020) Football-specific validity of TRACAB's optical video tracking systems. PLoS ONE 15:e0230179
17. FIFA (2022). FIFA EPTS Test Report. TRACAB GEN 5 (Live Data). Chyronhego AB.
18. Parker RA, Weir CJ, Rubio N, Rabinovich R Pinnock H, Hanley J (2016) Application of mixed effects limits of agreement in the presence of multiple sources of variability: exemplar from the comparison of several devices to measure respiratory rate in COPD patients. PLoS ONE 11:e0168321
19. Parker RA, Scott C, Inácio V, Stevens NT (2020) Using multiple agreement methods for continuous repeated measures data: a tutorial for practitioners. BMC Med Res Methodol 20(1):14
20. Vidal-Codina F, Evans N, El Fakir B (2022) Automatic event detection in football using tracking data. Sports Eng 25:18 https://doi.org/10.1007/s12283-022-00381-6
21. FIFA (2025). FIFA Innovation. Resource Hub. Kinexon LPS. Available at: https://inside.fifa.com/innovation/resource-hub?id=776fb2814f8643de8dec1d6fda96b982
22. FIFA (2025). FIFA Innovation. Resource Hub. Hawk-Eye Innovations Limited. Player and Ball Tracking. Available at: https://inside.fifa.com/innovation/resource-hub?id=810f07b9d0694f0c9e43b653cbc51031
23. Bassek M, Theiner J, Ewerth R, Memmert D, Raabe D (2025) Broadcast analytics—an evaluation of video-based tracking systems with constrained player visibility. Sci Med Footb https://doi.org/10.1080/24733938.2025.2533808