COMPUTER VISION AND ARTIFICIAL INTELLIGENCE FOR AUTOMATED OFFSIDE DETECTION IN FOOTBALL: A SCOPING REVIEW

VISIÓN POR COMPUTADORA E INTELIGENCIA ARTIFICIAL PARA LA DETECCIÓN AUTOMATIZADA DEL FUERA DE JUEGO EN EL FÚTBOL: UNA REVISIÓN EXPLORATORIA

VISÃO COMPUTACIONAL E INTELIGÊNCIA ARTIFICIAL PARA DETECÇÃO AUTOMÁTICA DE IMPEDIMENTO NO FUTEBOL: UMA REVISÃO DE ESCOPO

10.56238/IIMultiCientifica-202

André Oliveira Silva[1], Reginaldo Alves de Sousa Junior[2], Leonardo Cardia da Cruz[3]

ABSTRACT

Offside decisions are among the most contentious moments in association football, with on-field error rates of 20–26%, and Video Assistant Referee (VAR) and Semi-Automated Offside Technology (SAOT) systems still rely on human interpretation. Recent advances in computer vision and artificial intelligence have prompted growing academic interest in automated offside detection, but no synthesis of the field's methods, datasets, metrics, and maturity is currently available. This scoping review mapped the volume and type of peer-reviewed literature on computer vision and artificial intelligence for offside detection in football published between 2020 and 2025, summarised the approaches, datasets, evaluation metrics, and reported performance of the included sources of evidence, and identified knowledge gaps and directions for future research. The review followed the PRISMA Extension for Scoping Reviews (PRISMA-ScR) and the Population, Concept, and Context (PCC) framework. Five electronic databases were searched in March 2026 for peer-reviewed articles and conference papers in English, Portuguese, or Spanish. Three reviewers screened independently in Rayyan; disagreements were resolved by discussion and by returning to the pre-defined criteria. Data were charted with a 17-field instrument. Six sources of evidence were included. Object-detection architectures of the YOLO family appeared in three sources, pose estimation in two, and a Transformer-based action-spotting model on SoccerNet-v2 in one. Only one source released a public offside-dedicated dataset, and none reported deployment in a professional stadium. The mapped literature shows an emerging field with bespoke datasets, heterogeneous evaluation protocols, and convergent gaps in automatic pass-event detection, robustness to occlusion, multi-camera integration, and SAOT-compatible pipelines.

[1] Undergraduate student in Data Science and Artificial Intelligence. Institution: Department of Computer Science, SENAI CIMATEC University. E-mail: andre.o.silva@aln.senaicimatec.edu.br
[2] Bachelor's degree in Electrical Engineering. Institution: Department of Computer Science, SENAI CIMATEC University. E-mail: reginaldo@aln.senaicimatec.edu.br
[3] PhD in Sciences and Informatics. Institution: Department of Computer Science, SENAI CIMATEC University. E-mail: leonardo.cardia@fieb.org.br

Keywords: Computer Vision. Offside Detection. Deep Learning. Association Football. Video Assistant Referee. Scoping Review.

RESUMEN

Las decisiones de fuera de juego se encuentran entre los momentos más polémicos del fútbol, con tasas de error en el campo del 20-26%, y los sistemas de videoarbitraje (VAR) y tecnología semiautomática de fuera de juego (SAOT) aún dependen de la interpretación humana. Los recientes avances en visión artificial e inteligencia artificial han impulsado un creciente interés académico en la detección automatizada de fuera de juego, pero actualmente no se dispone de una síntesis de los métodos, conjuntos de datos, métricas y madurez del campo. Esta revisión exploratoria mapeó el volumen y el tipo de literatura revisada por pares sobre visión artificial e inteligencia artificial para la detección de fuera de juego en el fútbol publicada entre 2020 y 2025, resumió los enfoques, conjuntos de datos, métricas de evaluación y rendimiento reportado de las fuentes de evidencia incluidas, e identificó lagunas de conocimiento y direcciones para futuras investigaciones. La revisión siguió la extensión PRISMA para revisiones exploratorias (PRISMA-ScR) y el marco de Población, Concepto y Contexto (PCC). En marzo de 2026 se realizaron búsquedas en cinco bases de datos electrónicas para encontrar artículos revisados por pares y ponencias de congresos en inglés, portugués o español. Tres revisores evaluaron de forma independiente en Rayyan; las discrepancias se resolvieron mediante discusión y volviendo a los criterios predefinidos. Los datos se analizaron con un instrumento de 17 campos. Se incluyeron seis fuentes de evidencia. En tres fuentes aparecieron arquitecturas de detección de objetos de la familia YOLO, en dos estimaciones de pose y en una un modelo de detección de acciones basado en Transformer en SoccerNet-v2. Solo una fuente publicó un conjunto de datos público dedicado al fuera de juego, y ninguna informó de su implementación en un estadio profesional. La literatura analizada muestra un campo emergente con conjuntos de datos específicos, protocolos de evaluación heterogéneos y brechas convergentes en la detección automática de pases, la robustez ante la oclusión, la integración multicámara y los flujos de trabajo compatibles con SAOT.

Palabras clave: Visión por Computadora. Detección de Fuera de Juego. Aprendizaje Profundo. Fútbol. Árbitro Asistente de Vídeo. Revisión Exploratoria.

RESUMO

As decisões sobre impedimento estão entre os momentos mais controversos no futebol, com taxas de erro em campo de 20 a 26%, e os sistemas de Árbitro Assistente de Vídeo (VAR) e Tecnologia Semi-Automatizada de Impedimento (SAOT) ainda dependem da interpretação humana. Os recentes avanços em visão computacional e inteligência artificial têm impulsionado o crescente interesse acadêmico na detecção automatizada de impedimento, mas atualmente não existe uma síntese dos métodos, conjuntos de dados, métricas e nível de maturidade da área. Esta revisão de escopo mapeou o volume e o tipo de literatura revisada por pares sobre visão computacional e inteligência artificial para detecção de impedimento no futebol, publicada entre 2020 e 2025, resumiu as abordagens, conjuntos de dados, métricas de avaliação e desempenho relatado das fontes de evidência incluídas e identificou lacunas de conhecimento e direções para pesquisas futuras. A revisão seguiu a extensão PRISMA para Revisões de Escopo (PRISMA-ScR) e a estrutura de População,

Conceito e Contexto (PCC). Cinco bases de dados eletrônicas foram pesquisadas em março de 2026 em busca de artigos revisados por pares e trabalhos de conferências em inglês, português ou espanhol. Três revisores analisaram o material de forma independente no Rayyan; as divergências foram resolvidas por meio de discussão e retorno aos critérios predefinidos. Os dados foram mapeados com um instrumento de 17 campos. Seis fontes de evidência foram incluídas. Arquiteturas de detecção de objetos da família YOLO apareceram em três fontes, estimativa de pose em duas e um modelo de detecção de ações baseado em Transformer no SoccerNet-v2 em uma. Apenas uma fonte disponibilizou um conjunto de dados público dedicado ao impedimento, e nenhuma relatou implantação em um estádio profissional. A literatura mapeada mostra um campo emergente com conjuntos de dados personalizados, protocolos de avaliação heterogêneos e lacunas convergentes na detecção automática de eventos de passe, robustez à oclusão, integração de múltiplas câmeras e pipelines compatíveis com SAOT.

Palavras-chave: Visão Computacional. Detecção de Impedimento. Aprendizado Profundo. Futebol. Árbitro Assistente de Vídeo. Revisão de Escopo.

1 INTRODUCTION

The offside rule, codified as Law 11 of the Laws of the Game by the International Football Association Board (IFAB), is one of the most consequential and contentious decisions in association football. Adjudicating an offside event requires the assistant referee to determine, at the precise instant the ball is played, whether any part of an attacking player's body, except the hands and arms, is closer to the opponents' goal line than both the ball and the second-to-last defender. The simultaneous fixation on the line of the second-to-last defender and on the exact moment of the pass imposes a well-documented optical and perceptual burden, and on-field error rates between 20% and 26% have been reported for elite competitions [1, 2]. To mitigate these errors, the Video Assistant Referee (VAR) protocol has been progressively incorporated into international competitions, and the Semi-Automated Offside Technology (SAOT), introduced at the 2022 FIFA World Cup, combines multiple synchronised cameras with limb-tracking and a predictive ball-position model to support officials in real time [3, 4]. Despite their adoption, both VAR and SAOT still rely on human interpretation of multi-camera footage, reviews can extend over a minute, and the underlying algorithms have not been disclosed in the peer-reviewed literature.

The recent maturation of deep learning has accelerated the application of Computer Vision (CV) and Artificial Intelligence (AI) techniques to sports analytics, and a small but growing body of academic work has begun to address offside detection specifically. Architectures from the You Only Look Once (YOLO) family have been used for player and ball detection [2–4]; pose-estimation networks have been used to project the relevant body parts onto the field plane [1, 3]; multi-object trackers based on the Kalman filter have been used to maintain player identity across frames [2, 3]; and Transformer-based action-spotting models have been evaluated on the SoccerNet-v2 benchmark, in which offside-related actions are among the annotated classes [5]. Classical pipelines combining hand-crafted features with shallow classifiers have also been reported [6]. These contributions span distinct sub-problems, camera calibration, player and ball detection, pose extraction, pass-event identification, team classification, and final rule-based judgement and adopt heterogeneous datasets and evaluation protocols, which complicates direct comparison and obscures the overall maturity of the field.



Although several studies have explored computer vision and artificial intelligence techniques for supporting offside decisions in football, the evidence remains scattered across heterogeneous methodological approaches, datasets, and evaluation procedures. This heterogeneity hinders the comparison of results, obscures the maturity of existing solutions, and limits the identification of common challenges and research priorities. To address this gap, the present study adopts a scoping-review approach [7, 8] to systematically map the literature on automated offside detection, synthesise the technical approaches and evaluation practices reported to date, and identify directions for future research. The review was conducted and reported in accordance with the PRISMA-ScR guidelines [9] and the methodological recommendations of the Joanna Briggs Institute [10].

The objectives of this scoping review were therefore: (i) to map the volume and characteristics of the literature on computer vision and artificial intelligence for offside detection in football published between 2020 and 2025; (ii) to synthesise the technical approaches, datasets, evaluation metrics, and outcomes reported in the included studies; and (iii) to identify knowledge gaps and highlight directions for future research.

2 METHODS

2.1 PROTOCOL AND REGISTRATION

This scoping review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) guidelines [9] and the updated Joanna Briggs Institute (JBI) methodological guidance for scoping reviews [10]. The eligibility criteria were developed using the Population, Concept, and Context (PCC) framework. The review protocol was developed prior to study selection; however, it was not formally registered in a public repository. Consistent with the objectives of a scoping review, no formal critical appraisal of the included sources was undertaken, as the purpose was to map and characterise the available evidence rather than assess the effectiveness of specific interventions or methods.

2.2 ELIGIBILITY CRITERIA

Eligibility criteria were defined a priori following the PCC framework. The elements that guided the review question and eligibility criteria are presented in Table 1.

Table 1

Population, Concept, and Context (PCC) framework used to define the eligibility criteria.

PCC Element	Description
Population	Association football matches and related visual data, including players, officials, and match footage.
Concept	Automated and semi-automated offside detection, referee-assistance systems, and computational approaches supporting offside decision-making.
Context	Computer vision and artificial intelligence methods applied to football analytics and officiating.

Source: prepared by the authors.

Sources of evidence were considered eligible when they (i) reported the implementation or evaluation of a computational method directly related to the offside rule or to a closely associated VAR sub-task; (ii) were peer-reviewed journal articles or conference papers; (iii) were published between 1 January 2020 and 31 December 2025; and (iv) were written in English, Portuguese, or Spanish. Sources were excluded when they addressed sports other than association football, did not include computer-vision content, were purely conceptual or theoretical without an implementation or evaluation, were themselves reviews of any kind (systematic, scoping, narrative, or survey), did not address offside either directly or as part of a recognised football action-spotting benchmark, or for which the full text could not be obtained.

2.3 INFORMATION SOURCES

Five electronic databases were searched for peer-reviewed literature: Scopus, IEEE Xplore, ACM Digital Library, ScienceDirect, and SpringerLink. Searches were conducted on 9 March 2026 (IEEE Xplore) and 10 March 2026 (Scopus, ScienceDirect, SpringerLink, and ACM Digital Library). The review focused on peer-reviewed journal articles and conference papers; therefore, grey literature, theses, dissertations, and preprints were not considered. Table 2 summarises the databases searched and the number of records retrieved.

Table 2

Information sources and records retrieved.

Database	Search date	Records retrieved
IEEE Xplore	09 Mar 2026	16
Scopus	10 Mar 2026	31
ScienceDirect	10 Mar 2026	15
SpringerLink	10 Mar 2026	78
ACM Digital Library	10 Mar 2026	55
Total	–	**195**

Source: prepared by the authors.

2.4 SEARCH

The search strategy was organised around three concepts: (i) offside detection and VAR-related terminology, (ii) association football, and (iii) computer vision and artificial intelligence techniques. Terms within each concept were combined using the Boolean operator OR, and the three concepts were combined using AND.

Search strings were adapted to the syntax requirements of each database. Due to limitations on the number of Boolean operators supported by ScienceDirect, a simplified version of the search string was employed. An example of the complete search strategy and the corresponding simplified ScienceDirect query is presented in Table 3. The full search strings, applied filters, and retrieval counts are provided in Supplementary File S1.

Table 3

Concepts and keywords used in the search strategy.

Concept	Keywords
Offside and VAR	“offside detection”, offside, “offside line”, “Offside decision”, “video assistant referee”, VAR, “referee assistance system”
Football	“association football”, soccer, football
Computer Vision and AI	“computer vision”, “artificial intelligence”, “machine learning”, “deep learning”, “video analysis”, “image processing”, “player detection”, “player tracking”

Source: prepared by the authors.

2.5 SELECTION OF SOURCES OF EVIDENCE

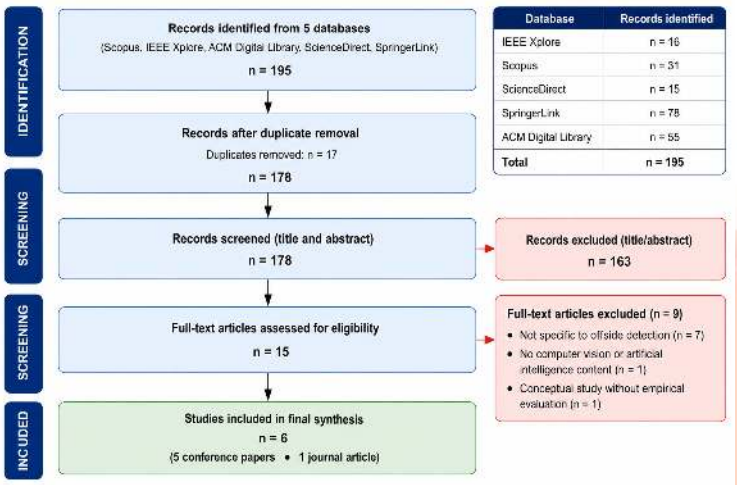
Records retrieved from the databases were imported into Rayyan for duplicate removal and blinded screening. Study selection was conducted in two stages. In the first stage, titles

and abstracts were screened against the predefined eligibility criteria. In the second stage, the full texts of potentially eligible studies were assessed using the same criteria.

Three reviewers independently conducted the screening process at both stages. Disagreements were resolved through discussion and consensus, with the predefined eligibility criteria being revisited when necessary to ensure consistency in study selection. The overall study-selection process is summarised in Figure 1.

Figure 1

Flow of sources of evidence through the review, adapted from the PRISMA 2020 statement [11] to the scoping-review context.



Source: prepared by the authors, adapted from PRISMA 2020 [11].

2.6 DATA CHARTING

A structured data-charting form comprising 17 extraction fields was developed and piloted on a subset of the included studies prior to full data extraction, following the iterative approach recommended by the Joanna Briggs Institute (JBI) guidance [10]. The charting form captured information related to bibliographic characteristics, problem context, datasets, computer vision and artificial intelligence approaches, evaluation metrics, reported outcomes, degree of automation, solution maturity, study contributions, limitations, and research gaps. The domains and variables captured during the charting process are summarised in Table 4.

Table 4

Domains and variables included in the data-charting form.

Domain	Variables
Bibliographic information	Article title, authors, year of publication, country of the corresponding author, publication venue, and publication type.
Problem and context	Problem addressed, application context, and research objective.
Technical approach	Computer vision approach, geometric aspects, artificial intelligence technique, algorithm family, and specific models employed.
Dataset characteristics	Dataset name, source, sample size, and availability (public or private).
Evaluation and outcomes	Evaluation metrics, reported performance, degree of automation, and maturity of the proposed solution.
Evidence synthesis	Main contributions, reported limitations, and research gaps identified by the study authors.

Source: prepared by the authors.

Data extraction was performed independently and cross-checked by the reviewers to ensure consistency. The charted data were synthesised narratively and through descriptive frequency counts. Consistent with the objectives of a scoping review, no meta-analysis or statistical pooling of results was undertaken. The complete data-charting form and extracted data are provided as Supplementary File S2.

3 RESULTS

3.1 SELECTION OF SOURCES OF EVIDENCE

A total of 195 records were identified through database searching. After duplicate removal, 178 records were screened, of which 163 were excluded based on title and abstract assessment. Fifteen full-text articles were assessed for eligibility, and nine were excluded for not meeting the predefined criteria. Consequently, six studies were included in the final synthesis (Figure 1).

3.2 CHARACTERISTICS OF INCLUDED SOURCES OF EVIDENCE

The six included studies were published between 2020 and 2025, with publications distributed across conference proceedings and journal venues. Five studies were conference papers, whereas one was published in a peer-reviewed journal. The corresponding authors were affiliated with institutions located in Asia and Europe, namely India, Japan, Spain,



Denmark, China, and Indonesia. The principal characteristics of the included studies, including publication type, technical approach, datasets, and maturity level, are summarised in Table 5.

3.3 SYNTHESIS OF RESULTS BY THEMATIC CATEGORIES

The included sources of evidence were synthesised narratively across four thematic axes: the technical approach adopted, the datasets used for development and evaluation, the evaluation metrics reported, and the maturity of the proposed solution.

3.4 TECHNICAL APPROACHES

Object detection constituted the dominant technical paradigm among the included studies. The YOLO family was the most frequently adopted detection framework, appearing in three of the six sources, with implementations based on YOLOv4, YOLOv5, and YOLOv11 [2–4]. Beyond object detection, multi-object tracking based on the SORT algorithm and Kalman filtering was employed in two studies to preserve player identities across successive frames and support temporal consistency in decision making [2, 3].

A second recurring component was the geometric modelling of player positions relative to the pitch. Two studies incorporated pose-estimation techniques to identify relevant body parts for offside assessment, including the head, shoulders, hips, knees, and ankles [1, 3]. Similarly, two studies employed explicit camera-calibration procedures to project image coordinates onto the field plane, using either vanishing-point estimation derived from pitch markings or homography-based transformations [1, 2]. These geometric components were intended to improve the spatial fidelity of offside-line estimation and player localisation.

Not all included studies followed the conventional detection-and-geometry pipeline. ASTRA [5] addressed offside indirectly as one of the action classes included in the SoccerNet-v2 benchmark, adopting a Transformer-based action-spotting architecture rather than performing an explicit geometric offside decision. Likewise, Madake et al. [6] proposed a classical machine-learning approach based on Histogram of Oriented Gradients (HOG) features combined with Random Forest, Decision Tree, and K-Nearest Neighbours classifiers.

A notable finding is that only one study explicitly addressed pass-event detection, which is a fundamental requirement of Law 11 because the offside position must be evaluated at the instant the ball is played. Uchida et al. [2] identified this moment through the analysis of extrema in a smoothed ball-velocity signal, whereas the remaining studies either assumed the pass instant was known or did not describe how it was determined.

Table 5

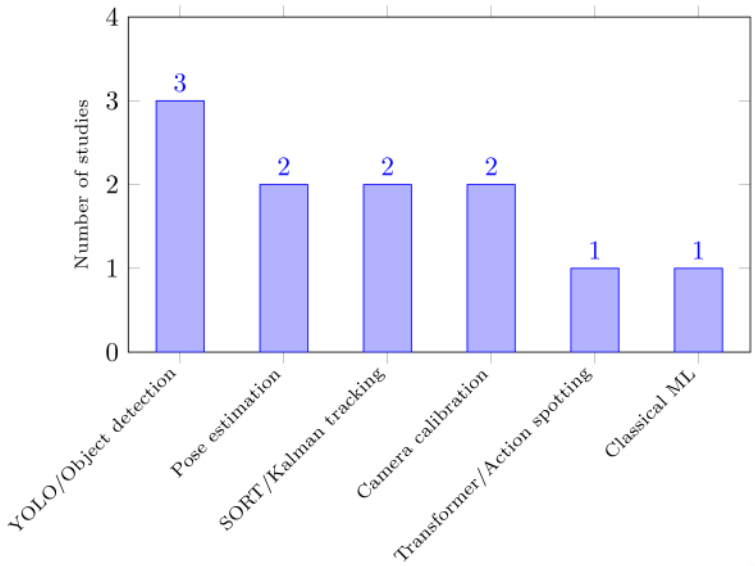
Principal characteristics of the included studies. PoC: proof of concept.

#	Year	First Author	Country	Venue	Technique Family	Dataset	Maturity
1	2020	Panse [1]	India	MMSports'20 (ACM)	Pose estimation and vanishing-point geometry	Own dataset; 500 broadcast images (public)	PoC; static-image evaluation
2	2021	Uchida [2]	Japan	MMSports'21 (ACM)	YOLOv5, SORT/Kalman filtering, and homography-based localisation	Own dataset; 3 full matches and 20 video clips (private)	Prototype evaluated on video sequences
3	2023	Xarles [5]	Spain / Denmark	MMSports'23 (ACM)	Transformer encoder–decoder architecture for action spotting	SoccerNet-v2; 550 matches (public benchmark)	Research benchmark
4	2023	Madake [6]	India	IEEE ICBDS 2023	HOG features with Random Forest, Decision Tree, and KNN classifiers	DFL Bundesliga Kaggle dataset; 1,500 augmented images	PoC based on a single match
5	2025	Zhang [3]	China	IJACSA	Hybrid CNN, RNN, YOLOv4, SORT, and OpenPose pipeline	Mixed real and synthetic datasets (size not reported)	Advanced prototype
6	2025	Indahsari [4]	Indonesia	IEEE ICIC 2025	YOLOv11 (Nano and Small variants)	Own dataset; 360 annotated images (private)	PoC

Source: prepared by the authors based on the included studies [1–6].

Figure 2

Frequency of technical approaches identified across the included studies. Categories are not mutually exclusive because some studies combined multiple techniques.



Source: prepared by the authors.

3.5 DATASETS

The datasets used by the included studies were predominantly bespoke and privately developed. Three studies relied on proprietary datasets that were not publicly released [2–4], limiting reproducibility and cross-study comparison. These datasets varied substantially in size and composition, ranging from 360 annotated images to collections of full-match recordings and mixed real–synthetic data.

Only one study [1] reported a publicly available dataset specifically designed for offside detection, comprising 500 broadcast images that generated 1,000 evaluation cases. This was the only dedicated public benchmark for offside analysis identified in the review. The remaining studies relied on existing public datasets originally developed for other football-analysis tasks, including SoccerNet-v2 [5], which contains 550 matches annotated with multiple action classes, and the DFL Bundesliga Data Shootout dataset, which was adapted for offside-related experimentation through data augmentation [6].

Overall, the identified datasets exhibited substantial heterogeneity in terms of size, accessibility, annotation protocols, and intended purpose, which complicates direct

comparison of reported performance and highlights the absence of a standardised benchmark for offside detection.

3.6 EVALUATION METRICS

The evaluation protocols adopted by the included studies exhibited substantial heterogeneity, reflecting the absence of a standardised benchmark and a common assessment framework for offside detection. Precision, recall, and F1-score were the most frequently reported metrics among studies evaluating the end-to-end offside-decision task [1–3, 6]. However, the reported values are not directly comparable because the underlying units of analysis differed considerably, ranging from individual images and video clips to action classes and intermediate pipeline components.

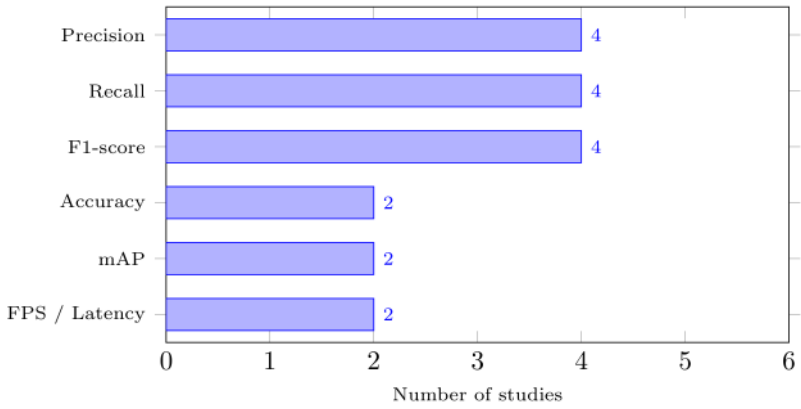
Among the studies performing explicit offside-decision evaluation, Panse and Mahabaleshwarkar [1] reported an F1-score of 0.85 under fully automatic goalkeeper detection, which increased to 0.98 when goalkeeper identification was performed manually, highlighting goalkeeper localisation as a critical source of error. Uchida et al. [2] reported an F1-score of 63.1% across twenty video clips, whereas Zhang et al. [3] reported an offside-detection accuracy of 99.85% together with an inference latency of approximately 150 ms per frame.

The remaining studies employed task-specific evaluation protocols. ASTRA [5] reported Average-mAP values of 66.82 and 70.21 on the SoccerNet-v2 benchmark test and challenge splits, respectively, whereas Indahsari et al. [4] evaluated object-detection performance using mAP@0.5, mAP@0.5:0.95, precision, recall, and frames per second. In contrast, the 98.50% accuracy reported by Madake et al. [6] primarily reflects the team-classification stage of the pipeline rather than the end-to-end offside decision itself.

Figure 3 summarises the frequency of evaluation metrics reported across the included studies. Precision, recall, and F1-score were the most commonly adopted measures, although a wide range of task-specific metrics was also observed. Overall, the diversity of evaluation protocols, datasets, and task definitions prevented direct comparison of performance across studies and reinforces the need for a standardised benchmark and evaluation framework for offside detection.

Figure 3

Frequency of evaluation metrics reported across the included studies. Categories are not mutually exclusive because individual studies frequently reported multiple metrics.



Source: prepared by the authors.

3.7 MATURITY

The maturity of the proposed solutions ranged from proofs of concept evaluated on static images [1] or on a single match [6], through video prototypes evaluated on a small number of recorded matches [2, 4], to a more advanced laboratory prototype combining several deep architectures and reporting near-real-time inference on a high-end consumer GPU [3]. ASTRA [5] occupies a different position on this spectrum, as it constitutes a competitive submission to an established academic benchmark rather than a deployable referee-assistance system. No included source reported deployment, controlled stadium-environment testing, or independent reproduction of its results in a professional match setting.

4 DISCUSSION

4.1 PRINCIPAL FINDINGS

The present scoping review identified six peer-reviewed studies published between 2020 and 2025 that proposed computer-vision or artificial-intelligence methods related to offside detection in association football. The mapped literature suggests that the field remains emerging rather than consolidated. The identified studies cluster around two publication waves (2020–2021 and 2023–2025), with no eligible studies identified in 2022 or 2024. Moreover, the cumulative volume of six studies over a six-year period remains modest when



compared with more established areas of football analytics, such as event detection, player tracking, and pose-based performance analysis.

A second notable finding concerns the geographical distribution of research activity. Five of the six included studies originated from institutions in Asia (India, Japan, China, and Indonesia), whereas one study was jointly authored by researchers from Spain and Denmark. Interestingly, none of the included studies originated from institutions located in the professional leagues where Video Assistant Referee (VAR) and Semi-Automated Offside Technology (SAOT) are most extensively deployed. This observation suggests a possible disconnect between the academic geography of the problem and the industrial environments in which the technology is currently applied.

The reviewed studies converge on a broadly similar technical pipeline consisting of player and ball detection, player tracking, projection onto the pitch plane, and a rule-based decision stage. Within this pipeline, the YOLO family emerged as the dominant object-detection framework, while SORT combined with Kalman filtering was the most frequently adopted tracking approach. However, substantial differences were observed in how the offside rule itself was operationalised. Among the six studies, only Uchida et al. [2] explicitly addressed pass-event detection, and only Panse and Mahabaleshwarkar [1] explicitly modelled the playable-body-part requirement through pose estimation. The remaining studies relied on simplified positional heuristics, action-spotting formulations, or partially described hybrid architectures [3–6].

The reported performance metrics should therefore be interpreted cautiously. The included studies employed heterogeneous datasets, evaluation protocols, and units of analysis, preventing direct comparison of reported results. For example, the 98.50% accuracy reported by Madake et al. [6] refers primarily to a team-classification task rather than an end-to-end offside decision. Similarly, the 99.85% accuracy reported by Zhang et al. [3] was obtained on a dataset whose size was not disclosed and is accompanied by inconsistencies between the abstract and the main text regarding related performance values. Consequently, these figures should not be interpreted as directly comparable estimates of offside-detection performance.

In contrast, the F1-scores reported by Panse and Mahabaleshwarkar [1] and Uchida et al. [2], although numerically lower, were obtained under evaluation protocols that more



directly assessed the offside-decision process itself. Likewise, the Average-mAP reported by ASTRA [5] reflects performance on an action-spotting benchmark rather than on geometric offside detection. Taken together, these findings indicate that the current literature lacks a common benchmark, a standardised evaluation protocol, and a shared definition of the task itself. This methodological heterogeneity remains one of the principal barriers to establishing a reliable state of the art for automated offside detection.

4.2 KNOWLEDGE GAPS

The most prominent gap identified in the reviewed literature is the absence of a canonical, large-scale, publicly available benchmark specifically designed for offside detection. The only dedicated public dataset identified was the 500-image benchmark released by Panse and Mahabaleshwarkar [1], whereas three of the remaining studies relied on private datasets that were not publicly released [2–4]. The other public datasets used in the literature, including SoccerNet-v2 and the DFL Bundesliga Data Shootout dataset, were originally developed for broader football-analysis tasks rather than for end-to-end offside assessment [5, 6]. Closely related to this limitation is the absence of a standardised evaluation framework. The included studies adopted different units of analysis, annotation procedures, and performance metrics, which complicates direct comparison of results and prevents meaningful quantitative synthesis.

A second group of gaps concerns the completeness with which the offside rule is modelled. Accurate offside assessment requires the simultaneous identification of the pass instant, the receiving player, and the projection of the playable body parts of the attacking player at that precise moment. However, only a small subset of the reviewed studies explicitly addressed these components. Goalkeeper identification, identified by Panse and Mahabaleshwarkar [1] as a major source of error, remains largely unresolved and is frequently handled through manual intervention or simplified heuristics. Likewise, multi-camera integration, which is a fundamental component of operational Semi-Automated Offside Technology (SAOT) systems, was not addressed by any of the included studies. Several studies also reported challenges related to occlusion, motion blur, illumination changes, and camera-angle variation [1–4], suggesting that robustness under realistic match conditions remains an open problem.



Finally, a substantial translational gap exists between academic prototypes and operational referee-assistance systems. None of the included studies reported deployment in professional competitions, controlled stadium-scale validation, independent replication of results, or long-term operational testing. Consequently, although the reviewed studies demonstrate the technical feasibility of automated offside analysis, the evidence base remains largely limited to proof-of-concept and laboratory-scale evaluations. Bridging this gap will require larger public datasets, standardised evaluation protocols, multi-camera validation, and closer collaboration between academic researchers, football governing bodies, and technology providers.

4.3 FUTURE RECOMMENDATIONS

The findings of this review suggest several priorities for future research on automated offside detection. The most immediate need is the development of a large-scale, publicly available benchmark specifically designed for the offside-decision task. Such a benchmark should include synchronized annotations of the pass instant, receiving player, playable body parts, and second-to-last defender position, enabling the establishment of a standardised end-to-end evaluation protocol. The adoption of common datasets and metrics would substantially improve the comparability and reproducibility of future studies.

A second priority concerns the development of more rule-faithful computational models. The reviewed literature indicates that temporal and geometric components of the offside rule are often addressed separately. Future systems could benefit from integrating temporal action-spotting approaches, such as those employed by ASTRA [5], with geometric body-part localisation and projection techniques similar to those proposed by Panse and Mahabaleshwarkar [1]. Such integration may enable more complete modelling of the rule and reduce reliance on simplified positional heuristics.

The review also highlights the need to narrow the gap between academic prototypes and operational referee-assistance technologies. None of the included studies incorporated multi-camera acquisition architectures comparable to those used in Semi-Automated Offside Technology (SAOT), nor did they report validation in professional-match environments. Future work should therefore explore multi-camera calibration, player tracking across viewpoints, and controlled field testing under realistic stadium conditions.



Finally, greater emphasis should be placed on research transparency and reproducibility. The public release of datasets, source code, trained models, and annotation protocols, together with consistent reporting of dataset characteristics, hardware configurations, and implementation details, would facilitate independent validation and accelerate progress towards robust and deployable offside-detection systems.

4.4 LIMITATIONS

This review has several limitations. First, the search was restricted to five major databases (Scopus, IEEE Xplore, ACM Digital Library, ScienceDirect, and SpringerLink) and did not include grey literature, preprint servers, or backward reference searching. Consequently, relevant studies published in technical reports, theses, or non-indexed venues may not have been captured. In addition, the search strategy was conducted in English and eligibility was restricted to studies published in English, Portuguese, or Spanish, which may have excluded relevant evidence reported in other languages.

Second, the review focused on studies published between 2020 and 2025 in order to capture recent developments in computer vision and artificial intelligence for offside detection. Although this time frame reflects the contemporary state of the field, earlier contributions based on classical computer-vision techniques may have been excluded. Furthermore, the review protocol was not registered in a public repository prior to study selection.

Finally, no formal critical appraisal of the included studies was conducted, consistent with the objectives of a scoping review and the recommendations of PRISMA-ScR. The purpose of this review was to map the available evidence rather than assess methodological quality or estimate intervention effects. The small number of included studies (n = 6) also limits the strength of generalisations that can be drawn from the synthesis. Consequently, the findings should be interpreted as an evidence map of an emerging research area rather than as a definitive assessment of the state of the art.

5 CONCLUSION

This scoping review mapped the peer-reviewed scientific literature published between 2020 and 2025 on the application of computer vision and artificial intelligence to offside detection in association football. Six studies met the predefined eligibility criteria, confirming that this remains a relatively small and emerging research field.

The reviewed studies converged on a broadly similar technical pipeline involving player and ball detection, player tracking, field projection, and rule-based decision making. The YOLO family and SORT-based tracking emerged as the most frequently adopted approaches, while pose estimation, camera calibration, and Transformer-based action spotting appeared as complementary techniques. Despite these common elements, substantial differences were observed in dataset characteristics, rule implementation, evaluation protocols, and reported performance metrics, limiting direct comparison across studies.

Three major gaps were consistently identified across the literature: the absence of a large public benchmark specifically designed for offside detection, the lack of a standardised end-to-end evaluation framework, and the limited validation of proposed methods under realistic operational conditions. These gaps currently represent the main barriers to establishing a reliable state of the art in automated offside detection.

Future progress will likely depend on the development of shared benchmark datasets, standardised evaluation protocols, greater transparency and reproducibility, and closer collaboration between academic researchers, football governing bodies, and technology providers. Such efforts may help bridge the gap between proof-of-concept research and operational referee-assistance systems deployed in professional football.

- **Supplementary information.** The following supplementary materials accompany this article:
- **Supplementary File S1**: Full electronic search strategies executed in each of the five databases.
- **Supplementary File S2**: Complete data-charting table for the included studies (17 variables per study).
- **Supplementary File S3**: PRISMA-ScR checklist completed for this study, with page references.

ACKNOWLEDGEMENTS

DECLARATIONS

- **Funding**: No funding was received for conducting this study.
- **Conflict of interest**: The authors declare no competing interests.
- **Ethics approval**: Not applicable. This study did not involve human participants, animals or human samples.
- **Consent to participate**: Not applicable.
- **Consent for publication**: All authors consent to publication.
- **Data availability**: All data extracted and analysed during this study are presented within this article and its supplementary materials.
- **Materials availability**: Not applicable.
- **Code availability**: Not applicable.
- **Author contribution**: All authors contributed to the conception and design of the study, literature search, screening process, data extraction, manuscript preparation, and critical revision of the manuscript. All authors read and approved the final manuscript.

REFERENCES

[1] Panse, N., Mahabaleshwarkar, A.: A dataset & methodology for computer vision based offside detection in soccer. In: Proceedings of the 3rd International Workshop on Multimedia Content Analysis in Sports (MMSports'20), pp. 19–26. Association for Computing Machinery, Seattle, WA, USA (2020). https://doi.org/10.1145/3422844.3423055

[2] Uchida, I., Scott, A., Shishido, H., Kameda, Y.: Automated offside detection by spatio-temporal analysis of football videos. In: Proceedings of the 4th International Workshop on Multimedia Content Analysis in Sports (MMSports'21), pp. 17–24. Association for Computing Machinery, Virtual Event, China (2021). https://doi.org/10.1145/3475722.3482796

[3] Zhang, Q., Yu, L., Yan, W.: AI-driven image recognition system for automated offside and foul detection in football matches using computer vision. International Journal of Advanced Computer Science and Applications **16**(1), 1191–1198 (2025) https://doi.org/10.14569/IJACSA.2025.01601114

[4] Indahsari, A.N., Setiyani, L., Wati, D.F., Solikin: Implementation of YOLOv11 for automatic offside detection in football. In: 2025 Tenth International Conference on Informatics and Computing (ICIC), pp. 1–6. IEEE, Karawang, Indonesia (2025). https://doi.org/10.1109/ICIC68054.2025.11309588

[5] Xarles, A., Escalera, S., Moeslund, T.B., Clapés, A.: ASTRA: An action spotting TRAnsformer for soccer videos. In: Proceedings of the 6th International Workshop on Multimedia Content Analysis in Sports (MMSports'23), pp. 93–102. Association for Computing Machinery, Ottawa, ON, Canada (2023). https://doi.org/10.1145/3606038.3616153

[6] Madake, J., Thokal, D., Ullah, M.A., Bhatlawande, S.: Offside detection for better decision-making and gameplay in football. In: 2023 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS), pp. 1–6. IEEE, Naya Raipur, India (2023). https://doi.org/10.1109/ICBDS58040.2023.10346449

[7] Arksey, H., O'Malley, L.: Scoping studies: towards a methodological framework. International Journal of Social Research Methodology **8**(1), 19–32 (2005) https://doi.org/10.1080/1364557032000119616

[8] Levac, D., Colquhoun, H., O'Brien, K.K.: Scoping studies: advancing the methodology. Implementation Science **5**, 69 (2010) https://doi.org/10.1186/1748-5908-5-69

[9] Tricco, A.C., Lillie, E., Zarin, W., O'Brien, K.K., Colquhoun, H., Levac, D., Moher, D., Peters, M.D.J., Horsley, T., Weeks, L., Hempel, S., Akl, E.A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M.G., Garritty, C., Lewin, S., Godfrey, C.M., Macdonald, M.T., Langlois, E.V., Soares-Weiser, K., Moriarty, J., Clifford, T., Tunçalp, Ö., Straus, S.E.: PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. Annals of Internal Medicine **169**(7), 467–473 (2018) https://doi.org/10.7326/M18-0850

[10] Peters, M.D.J., Marnie, C., Tricco, A.C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C.M., Khalil, H.: Updated methodological guidance for the conduct of scoping reviews. JBI Evidence Synthesis **18**(10), 2119–2126 (2020) https://doi.org/10.11124/JBIES-20-00167

[11] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ **372**, n71 (2021) https://doi.org/10.1136/bmj.n71