Once Upon a Goal: Towards Orientation-Based Shot Metrics in Football

Marc Gutiérrez-Pérez[a], Calvin Yeung[b], Keisuke Fujii[b,c], Antonio Agudo[a]

[a] *Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Carrer de Llorens i Artigas, 4-6, 08028, Barcelona, Spain*
[b] *Graduate School of Informatics, Nagoya University, Furo, Chikusa, 464-8601, Nagoya, Japan*
[c] *RIKEN Center for Advanced Intelligence Project, 1-1 Yamadaoka, Suita, 565-0871, Osaka, Japan*

Abstract

Sports analytics has been revolutionized by advanced tracking technologies, yet the integration of human pose estimation into performance metrics remains underexplored in football. Estimating the probability of scoring from a shot is a central task in football analytics and is commonly approached through expected goals (xG) models. Progress in this area, however, is often constrained by the limited availability of publicly accessible datasets that include fine-grained biomechanical information. In this work, we present xGHub, an open-source dataset of football shots enriched with player pose estimation, body orientation, and contextual features extracted from broadcast video. The dataset is generated using an automated pipeline for player detection and tracking, followed by an external verification process to ensure annotation reliability. As a use case, we analyze how pose- and orientation-related features can be incorporated into a standard xG modeling framework. Our results indicate that 3D orientation information is informative for specific subsets of shots, while its contribution is limited in others, reflecting the inherently non-linear nature of angular representations. This analysis serves to illustrate the potential and limitations of the released annotations. By making this dataset publicly available, we aim to support future research on the role of player biomechanics in shot analysis and related football analytics tasks.Sports analytics has been revolutionized by advanced tracking technologies, yet the integration of human pose estimation into performance metrics remains underexplored in football. Estimating the probability of scoring from a shot is a central task in football analytics and is commonly

approached through expected goals (xG) models. Progress in this area, however, is often constrained by the limited availability of publicly accessible datasets that include fine-grained biomechanical information. In this work, we present xGHub, an open-source dataset of football shots enriched with player pose estimation, body orientation, and contextual features extracted from broadcast video. The dataset is generated using an automated pipeline for player detection and tracking, followed by an external verification process to ensure annotation reliability. As a use case, we analyze how pose- and orientation-related features can be incorporated into a standard xG modeling framework. Our results indicate that 3D orientation information is informative for specific subsets of shots, while its contribution is limited in others, reflecting the inherently non-linear nature of angular representations. This analysis serves to illustrate the potential and limitations of the released annotations. By making this dataset publicly available, we aim to support future research on the role of player biomechanics in shot analysis and related football analytics tasks.

Keywords: Sports analytics, Expected goals, Human pose estimation, Football, Machine learning, Computer vision.

---

1. Introduction

Sports analytics has become an increasingly vital component in modern sports, transforming the way teams, coaches, and fans understand and optimize athletic performance. The proliferation of advanced tracking technologies [1, 2] such as player and ball tracking systems, has enabled the generation of rich, high-resolution data that provides unprecedented insights into the dynamics of sports competitions. This wealth of tracking data has revolutionized the way sports are analyzed, allowing more informed decision-making, enhanced player development, and the identification of strategic advantages [3, 4]. Leveraging these data-driven insights has become a key competitive edge, as teams strive to gain a deeper understanding of player movements, team tactics, and in-game patterns. The ability to accurately capture and analyze tracking data has become a crucial aspect of sports analytics, fueling innovations in areas such as player performance optimization [5], injury prevention [6, 7], and the development of advanced coaching strategies [4, 8].

While the proliferation of wearable tracking devices has been instrumental

in generating sports performance data, the use of computer vision techniques has emerged as a compelling alternative approach. By leveraging advanced computer vision algorithms, researchers and sports organizations can now extract tracking data directly from video footage, without the need for intrusive wearable sensors. This camera-based tracking approach offers several advantages, including the ability to capture data from multiple athletes simultaneously, the elimination of potential interference or disconnection issues associated with wearables, and the potential for retroactive analysis of historical game footage. Computer vision-based tracking leverages techniques such as object detection [9, 10, 11], object tracking [12, 13, 14], and camera calibration [15, 16] to accurately identify and monitor the movements of players, balls, and other key elements within a sports environment. This data-driven and non-invasive approach for tracking has become increasingly sophisticated, enabling the generation of rich, high-fidelity datasets that can provide deeper insights into athletic performance and team dynamics.

While advancements in sports analytics and computer vision have delivered unprecedented insights and improved experiences, the use of 3D player skeleton data obtained through human pose estimation remains in its early stages. Although current technology can accurately track player skeletons well enough to support semi-automated offside technology [17] in official broadcast feeds, most existing approaches in football (soccer) analytics continue to rely on 2D player positions or event-based data [5, 18, 19]. Furthermore, access to skeleton data remains highly restricted within the academic community due to the challenges and high costs associated with its acquisition.

Unlike traditional data sources, skeleton data provides a more comprehensive representation of player movements by capturing not only positional information but also body orientation and fine-grained actions. In recent years, several datasets that contain skeleton data have emerged in different sports contexts, including football [20, 21], basketball [22, 23], baseball [24], athletics [25, 26], golf [27], and combat sports [28], just to name a few. These resources have enabled significant advancements in various tasks, such as object tracking [29], action detection and segmentation [30, 31], and even automated competition performance scoring [32, 33]. Moreover, they have opened the door to the development of advanced pose-based performance metrics [34, 35, 36].

In football, as in other low-scoring sports, the scarcity of goals necessitates capturing more subtle features to accurately evaluate performance.

One prominent example is the expected goals (xG) metric [37, 38, 39], which assesses the quality of a scoring opportunity by estimating the probability of it resulting in a goal, based on historical data from similar shots. While in its most advanced implementations the goalkeeper and opponent players' positions are considered, we cannot overlook the fact that skeleton-based information such as the striker's orientation towards the goal or the goal's occlusion percentage is rarely considered [35]. Instead, two-dimensional proxies are commonly used by considering player locations into the planar field [40], offering only a simplified approximation as if we were looking at a football penny game. Therefore, such representations would assign the same goal probability to an overhead kick and a regular header, as well as to situations involving a standing goalkeeper and a goalkeeper already lying on the ground. Traditional xG models primarily rely on 2D positional data combined with annotated categorical features, such as the body part used or specific body postures. Their implementation typically ranges from conventional machine learning techniques, including logistic regression [41, 42], generalised linear models [43] and gradient-boosted trees [44, 40], to more recent deep learning-based approaches [45]. Recent methods [35] have incorporated skeleton data in xG estimation; however, neither the datasets nor the model architectures have been publicly released. To address this gap, we introduce xGHub, an open-source, comprehensive, and scalable collection of shot clips extracted from broadcast footage of the five major European football leagues. For each selected clip, we apply a combination of manual annotation and semi-automatic labeling to capture the 2D positions of all players visible in the camera view at the exact moment of the shot. Additionally, we annotate team information, the shooter's tracklet, the body part used, the shot type and the shot outcome. Furthermore, we estimate the shooter's orientation based on human pose estimation to provide a richer representation of the shooting context. xGHub is intended not only as a resource but also as a benchmark for advancing open-source research in goal probability estimation, while laying the foundation for practical applications in developing novel performance metrics for shooters, goalkeepers, and defenders alike. This makes it a valuable asset for both academic research and applied analytics in professional football. To model these newly generated features, we propose converting player skeleton data into orientation-based, single-valued numerical features. Specifically, we represent player orientation using angles derived from pose estimation, which are then fed into a gradient-boosted tree algorithm, such as XGBoost [46]. This approach ensures the interpretability of the solution,

requires fewer data compared to deep learning methods, and maintains the tabular data format commonly used in event-based football analytics. This work makes the following key contributions:

- xGHub: A comprehensive dataset of shot situations, purpose-built to advance research on shot-based performance metrics.
- A pipeline integrating open-source modules to automatically extract skeletal data from broadcast footage.
- A novel skeletal-based orientation feature compatible with traditional machine learning, with rigorous analysis of its impact on goal expectation modeling.

2. Related Work

We review prior research most relevant to our study, with a particular focus on football analytics, especially expected-goals (xG) modeling, and on studies that incorporate skeletal data in sports contexts.

2.1. Football Analytics and Expected Goals

Early approaches to football analytics date back to the 1960s, when Reep *et al.* [37] observed a stochastic relationship between the number of shots in a match and the resulting goals. Nearly three decades later, Pollard *et al.* [47] applied logistic regression to estimate the probability of a shot resulting in a goal, offering one of the first glimpses of what would later evolve into the xG metric. In the modern era, Green [42] proposed modeling shot probability to quantify which areas of the pitch are most likely to yield goals. Caley [48] introduced the shot matrix database, which statistically analyzed Premier League shots across eight zones to examine how shot location influences scoring probability. Extending beyond player location, StatsPerform [42] incorporated contextual categorical variables, such as play pattern and assisting pass type, into a logistic regression model, along with a competition-wise adjustment. StatsBomb [40] introduced the use of gradient boosting trees [49] and incorporated engineered features by replacing location-based discrete variables with continuous feature representations, modeling defenders and goalkeepers as 2D Gaussian distributions instead of fixed-radius circles. Building on player positional data obtained through

tracking methods. Anzer [50] further advanced xG modeling by incorporating player velocity as a feature, addressing the static limitations of earlier approaches. Mead *et al.* [51] applied a gradient boosting method that incorporated new features such as the body part used and the timing of the shot within the match, together with contextual variables including home/away status, game week, team goal difference, and both current and previous season rankings. More recently, Nipoti *et al.* [52] employed Bayesian models for uncertainty quantification and enabled in-season model training. Bandara *et al.* [53] employed a random forest model that incorporated features from preceding events, thereby capturing the temporal dimension of player actions. In terms of model explainability, several methods [45, 51, 54, 41] employ shapley additive explanations [55] to gain insights into the contribution of each feature to the model's decision-making process, while aggregated profiles have also been widely applied [44, 54]. Moreover, Van Haaren [56] utilized the explainable boosting machine algorithm [57], manually defining interactions between features. Delving into deep-learning techniques, some approaches [58, 45] have employed simple multi-layer perceptrons to estimate goal expectancy. More recently, Xu *et al.* [35] incorporated skeleton data into the task and implemented a graph neural network that models each shot as a spatiotemporal graph, effectively capturing the complex interactions between players body parts and the ball. However, neither the implementation code nor the dataset used have been publicly available. In contrast, our approach leverages an open-source dataset in which skeletal data, though limited to the shooter, is available and can be extended. Furthermore, we implement a minimal representation of shooter pose estimation that can be expressed in tabular form and is compatible with non–deep-learning methods, enabling lighter and more interpretable models.

2.2. Skeletal Data in Sports

Although the work of Xu *et al.* [35] was the first to incorporate skeleton data into the goal expectation task, earlier studies had already explored its use in other areas of football analytics. For example, Wear *et al.* [59] analyzed goalkeeper techniques in penalty and one-on-one situations using pose estimation, while Reddy *et al.* [60] more recently examined goalkeeper movement timing in penalty scenarios. In [61] was introduced a novel technique for estimating player orientation from monocular video recordings by mapping body pose parts. Building on that, they later proposed a method [34] that leveraged the orientation of offensive players together with the spatial

configuration of opponents to evaluate the feasibility of pass events. Yeung *et al.* [20] proposed the 3D shot posture dataset, consisting of annotated 2D pose sequences of shooting instances from professional football matches, and AutoSoccerPose, a semi-automated pipeline for extracting 3D pose sequences directly from professional football broadcast videos. However, their 2D-to-3D pose-lifting approach does not operate in true football-pitch world coordinates and is primarily designed to produce pose-based embeddings rather than metrically grounded pose representations. Schepers *et al.* [36] explored how player posture can improve our understanding of dribbling skills, capturing novel aspects like balance, orientation, and ball control. Recently, Fang *et al.* [62] present a novel method to predict football fouls with the combination of video, bounding box position, and pose information.

In this work, we collect broadcast videos of shot situations and develop a semi-automatic pipeline that integrates object detection, tracking, camera calibration, human pose estimation, and manual annotations. This pipeline produces player tracking data alongside the shooter's 3D pose, providing an open-source platform for studying pose estimation effect in shot-based football metrics, such as xG models.

3. xGHub Dataset

To establish a solid foundation for soccer shot analysis and to provide a benchmark for advancing research in this domain, we introduce *xGHub*, the first open-source dataset specifically dedicated to football shot evaluation. In the following sections, we describe the data collection process and present an overview of the dataset generation pipeline.

3.1. Dataset Collection

To construct the xGHub dataset, we collected untrimmed, full-match broadcast videos from the SoccerNet [63, 64, 65] and SoccerReplay-1988 [66] datasets, as well as from additional publicly available sources on the Internet. The collection comprises a total of 704 matches from the five major European football leagues[1], specifically from the 2015–16 and 2021–22 seasons. For each match, we obtained second-level timestamps with annotations for

[1]Premier League (England), La Liga (Spain), Bundesliga (Germany), Serie A (Italy), and Ligue 1 (France).

specific event types from publicly available datasets [64, 66]. In this work, we specifically focus on the shot-related timestamp annotations.

3.2. Semi-automated Data Curation

Given the nature of the dataset, we designed a modular pipeline that combines manual and automated tasks for dataset generation. The pipeline specifically includes: (i) temporal alignment, (ii) object detection, (iii) object tracking, (iv) camera calibration, (v) 2D pose estimation, and (vi) 3D pose estimation. The designed pipeline is illustrated in Fig. 1.

Temporal Alignment. First, we manually divide the full, untrimmed match videos into two halves, each beginning at kick-off. Note that videos obtained from SoccerNet [63] and SoccerReplay-1988 [66] are already pre-divided into halves. We then synchronize the second-level timestamps with the divided videos and manually refine the shot-event annotations to accurately identify the frame corresponding to the player-ball moment of contact, thereby obtaining frame-level annotations for each shot event. Finally, each shot is cropped into a one-second clip, with the player-ball moment of contact positioned at the middle frame. The duration of the clip was determined through extensive evaluation to minimize occlusions before and after the shot, thereby facilitating the processing of subsequent modules.

Object Detection. For the multi-object detection module, we employ RF-DETR [10], a state-of-the-art real-time object detection model. Specifically, we adopt the variant described in GSR-4 [67], in which the base RF-DETR [10] is fine-tuned on the SoccerNet-GSR dataset [68]. The resulting detector is capable of accurately detecting humans on the football field, without differentiating between specific match roles, such as players, goalkeepers, and referees.

Object Tracking. For the multi-object tracking module, we adopt a tracking-by-detection approach, aiming to identify football players and referees in video clips and generate corresponding tracklets. Specifically, we employ BoT-SORT [14], a robust state-of-the-art tracker that leverages both motion and appearance information. After applying the tracking module, we manually classify the generated tracklets into different categories. First, we select the shooter tracklet. The remaining tracklets are then assigned specific match roles, including a teammate flag indicating whether the tracklet belongs to the same team as the shooter, as well as goalkeeper and referee flags. Additionally, clips deemed unusable due to heavy occlusions on the

shooter tracklet, extreme motion blur, or low video quality, which could hinder accurate scene reconstruction, are removed from the dataset.

Camera Calibration. To accurately estimate the world-coordinate positions of detected players, we employ PnLCalib [16], a state-of-the-art optimization-based calibration pipeline that leverages a 3D football field model and a predefined set of keypoints. For a given camera, we denote the intrinsic matrix as $\mathbf{K} \in \mathbb{R}^{3\times3}$, the radial distortion coefficients as $\mathbf{k} \in \mathbb{R}^2$, and the extrinsic parameters as the rotation $\mathbf{R} \in SO(3)$ and translation $\mathbf{t} \in \mathbb{R}^3$. We summarize the complete set of parameters as $\mathbf{\Lambda} = (\mathbf{K}, \mathbf{R}, \mathbf{t}, \mathbf{k})$.

3D Pose Estimation To address the challenges posed by low-quality and blurred videos, we apply a preprocessing step to facilitate accurate pose estimation. Specifically, we employ NAFNet [69], an image restoration method, to enhance videos affected by noise or blur. For the 3D pose estimation task, we follow the approach in [70], where 2D body keypoints are first inferred using ViTPose [71], followed by HMR using a HMR2.0 architecture [70]. Additionally, PHALP [72] serves as a 3D tracking system, leveraging temporal information to predict the future state of each tracklet, including 3D appearance, location, and pose. This tracking module aids our pipeline in recovering 3D human meshes under challenging scenarios, such as occlusions or rapid body movements. By integrating the camera calibration module, the estimated 3D poses can be projected into world coordinates.

Additionally, we manually include extra annotations specifying the body part used (i.e., *right/left foot*, *head*, or *other*), the shooting technique employed (i.e., *regular*, *volley*, *backheel*, or *overhead kick*), and the play pattern (i.e., *regular*, *from free kick* and *from corner*). While the body-part annotation distinguishes between left- and right-foot usage, the dataset does not explicitly encode dominant versus non-dominant foot. Inferring foot dominance would require additional manual labeling or external metadata and is therefore beyond the scope of the current dataset, despite being a biomechanically relevant factor closely related to body orientation and a promising direction for future extensions of xGHub. Moreover, while the dataset includes non-direct set-piece situations, direct shots from penalties, free kicks, and corners are excluded, as these correspond to largely static play situations.

The xGHub dataset is released as a self-contained collection of annotation files and is available at `https://github.com/mguti97/xGHub` in an open-source manner. The corresponding video clips are available under a non-disclosure agreement that strictly prohibits commercial use and any form of unauthorized distribution.

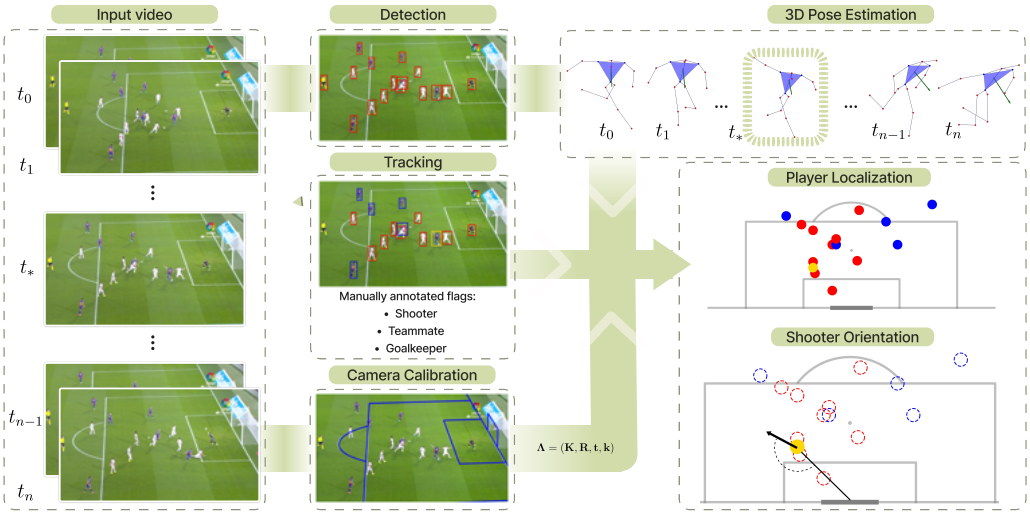

Figure 1: **xGHub generation pipeline.** From an input video with n frames, where t_* denotes the shot moment, RF-DETR [10] detects all humans on the field. BoT-SORT [14] generates tracklets, which are manually labeled to identify the shooter, teammates, and goalkeeper, while removing referee tracklets. Camera parameters $\mathbf{\Lambda}$ are obtained using PnLCalib [16]. The HMR2.0 & PHALP algorithm estimates 3D poses from the shooter tracklet. Combining camera calibration and tracking data enables 2D player localization and, together with 3D pose estimates, the extraction of 3D world-coordinate information such as shooter orientation. Blue and red dots represent teammate and opponent players, respectively, while the yellow dot denotes the shooter.

4. Expected Goal Modeling

In this section, we describe the machine-learning models, evaluation metrics, and feature engineering strategies employed for xG modeling.

4.1. Feature Engineering

Following [40], the features included in the models were designed according to the following principles. First, the number of non-continuous features is minimized, as discrete features can affect model behavior in borderline scenarios (e.g., the number of defenders within a specified area). Second, whenever feasible, monotonic constraints are imposed to ensure that the learned relationships adhere to an explicitly defined direction (e.g., the probability of scoring decreases with increasing distance to goal). While this approach does not generalize to all cases, particularly free-kicks and other set-piece situations, its implementation generally leads to improved results. The only

categorical features explicitly selected for this work are the body part used, the shot technique and the play pattern. Moreover, the *left/right foot* label is transformed into *inner/outer foot*, where the inner foot is defined as the leg closer to a line passing through the midpoint of the field along its length, connecting the centers of the two goals.

In addition, a set of more complex features was incorporated during the xG modeling stage. These features are discussed in detail in the subsequent subsections.

Distance and Angles. Consistent with most previous studies, we include as features the distance from the shot location to the midpoint of the goal, as well as the angle formed between the line connecting the shot position to the goal midpoint and the longitudinal axis of the field, defined by the line connecting the centers of both goals. Additionally, we incorporate the angle subtended by the two goalposts from the shooter's position as a strictly monotonic feature, complementing the angle-to-goal representation.

The following variables depend on player locations or are derived from them. We represent non-shooter players as 2D Gaussian distributions centered at their field coordinates, following [40].

Goal Face Occlusion. Based on this representation and following [40], we introduce a continuous metric termed *Goal Face Occlusion* (GFO), which takes values in $[0, 1]$ and quantifies the proportion of the goal face occluded from the perspective of the shot location. Given a specific player distribution on the field, we model an i-th player as a Gaussian distribution $\mathcal{N}$ centered at the player's position $\mathbf{d}_i \in \mathbb{R}^2$, with isotropic variance $\sigma^2 = 1$, corresponding to a standard deviation of 1 meter.

Let $\mathcal{G}$ denote the segment of the goal line delimited by the goalposts, and let $\mathbf{s}$ denote the shooter's location. For any point $\mathbf{g} \in \mathcal{G}$, we define the shooting direction as the ray $\ell(\mathbf{s}, \mathbf{g}) = \{\, \mathbf{s} + \tau(\mathbf{g} - \mathbf{s}) \mid 0 \leq \tau \leq 1 \,\}$ from shooting position $\mathbf{s}$ to goal line point $\mathbf{g}$. Along this ray, the occlusion contribution of the i-th player is given by the Gaussian field value as:

$$O_i(\mathbf{g}) \;=\; \max_{\tau \in [0,1]} \mathcal{N}\big(\mathbf{s} + \tau(\mathbf{g} - \mathbf{s}) \,\big|\, \mathbf{d}_i,\, \sigma^2 \mathbf{I}\big). \tag{1}$$

The total occlusion at $\mathbf{g}$ is then:

$$O(\mathbf{g}) \;=\; \max_i O_i(\mathbf{g}), \tag{2}$$

which corresponds to the maximum obstruction induced by defenders along the shooting direction. Finally, the GFO metric is defined as the normalized

integral of this occlusion function over the goal segment:

$$\text{GFO} \;=\; \frac{1}{|\mathcal{G}|}\int_{\mathcal{G}} O(\mathbf{g})\, d\mathbf{g}, \tag{3}$$

where $|\mathcal{G}|$ denotes the length of the goal segment. We compute GFO separately for the defenders and the goalkeeper, as well as jointly by considering all opposing players.

Shooter Pressure. Analogous to GFO, we define a circular region of radius $r = 1$ meter centered at the shooter's position. The *Shooter Pressure* (SP) metric is computed as the normalized mass of the player density distributions falling within this region. Formally, let the shooter be located at $\mathbf{s} \in \mathbb{R}^2$, and let n defenders have positions $\mathbf{d}_i \in \mathbb{R}^2$ for $i = \{1, \ldots, n\}$. Each defender is modeled by an isotropic Gaussian density $\mathcal{N}(\mathbf{d}_i, \sigma^2\mathbf{I})$ with $\sigma^2 = 1$, corresponding to a standard deviation of 1 meter.

Given the circle $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 : \| (x, y)^\top - \mathbf{s} \| \leq r\}$, the SP is defined as:

$$\text{SP} = \frac{1}{\pi r^2}\iint_{\mathcal{C}} \text{clip}\left(\sum_{i=1}^{n} \mathcal{N}\big((x, y)^\top \mid \mathbf{d}_i, \sigma^2\mathbf{I}\big),\, 0,\, 1\right) dx\, dy, \tag{4}$$

where $\text{clip}(z, 0, 1) = \min\{\max\{z, 0\}, 1\}$ denotes clamping the function values into the interval $[0, 1]$. This representation leads to a more natural definition of shooter pressure compared to previous approaches using closest defender distance [38]. Moreover, compared to [73], which proposes a directionally weighted decay function motivated by tactical considerations suitable for tracking data, our approach can be understood as a kernel density estimate evaluated locally around the shooter, suitable for static event data. Both GFO and SP metrics are illustrated in Fig. 2 to provide a clearer understanding.

Shooter Orientation. We propose two distinct formulations for *Shooter Orientation* (*SO*). The first approach, inspired by [61], leverages 2D pose estimation outputs from ViTPose [71]. Upper-torso keypoints (i.e., the shoulders) are selected for their relative stability during shooting movements, which involve rapid body rotations, and are projected into world coordinates ($Z = 0$) using the projection matrix obtained with PnLCalib [16]. The SO in this case, denoted by SO_{2D}, is defined as the relative angle between the normal vector to the line connecting the shoulders and the line joining the shooter position with the goal midpoint.

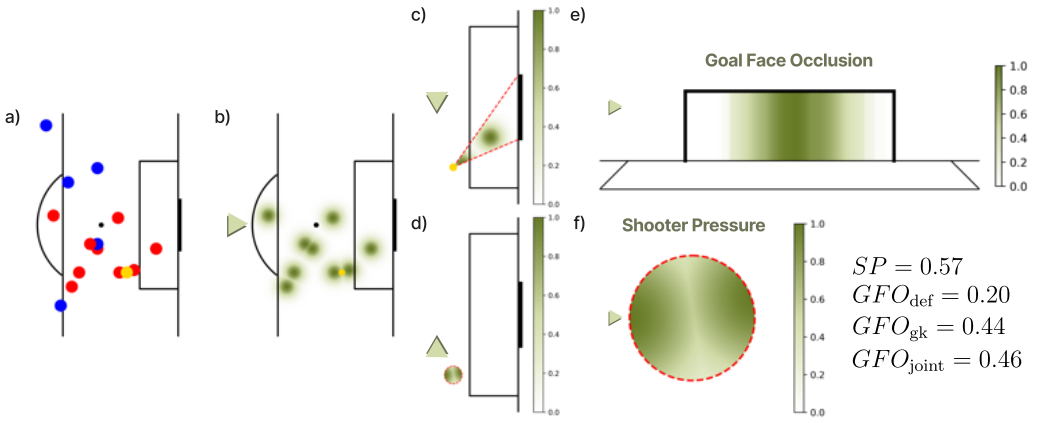

Figure 2: **Feature engineering process for continuous features. a)** Zenithal view of player positions at the shot frame. The shooter is depicted with a yellow circle, teammates in blue, and opponent players in red. **b)** Gaussian field modeling of opponent player distribution. **c)** Gaussian field masked with the shooter–goalposts cone. **d)** Gaussian field masked with a one-meter radius around the shooter's position. **e)** Projection of *Goal Face Occlusion* over the goal line. **f)** Qualitative view of *Shooter Pressure*.

As a consistency check, we also compute this angle using the hip and eye keypoint pairs. If the difference between the shoulder-based angle, θ_{sh}, and the hip-based angle, θ_h, exceeds 90°, we select the angle closer to that derived from the eye keypoints as the final orientation, with the latter serving only as an auxiliary discriminator.

The second formulation, denoted by SO_{3D}, builds on 3D pose estimation from [70]. In this case, a triangle is constructed from the world coordinates of the shoulders and the root hip joint, and the player's orientation is defined by the normal vector to this triangle. The orientation is parameterized as the relative rotation of this normal vector with respect to the shooter–goal direction, analogous to SO_{2D}. The ordering of the triangle vertices is defined such that the resulting normal vector points in the player's forward direction. The generation pipelines for both shooter-orientation formulations are illustrated in Fig. 3.

Finally, all samples are normalized so that the attacking direction is consistently defined from left to right. Specifically, for shots taken toward the left goal, we apply a mirroring operation to all variables that require adjustment.

Note that temporality is not accounted for in the xG modeling; all of the metrics described above are computed based solely on the player-ball contact

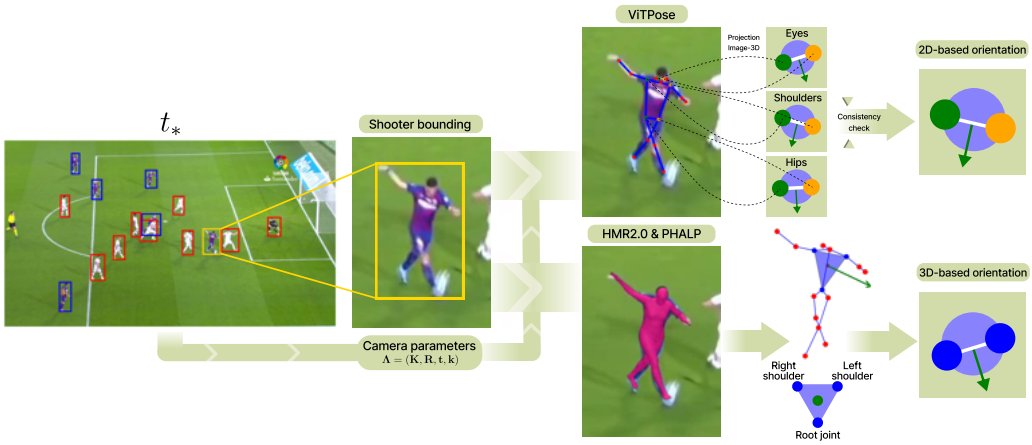

Figure 3: **Shooter-orientation generation.** The top and bottom pipelines correspond to the 2D- and 3D-based orientation generation processes, respectively. The green arrow denotes the orientation angle with respect to the pitch coordinate system.

at the shot frame.

4.2. Models

Across all evaluated models, we adopt a standardized data partitioning strategy, allocating 70% of the samples for training and 30% for testing. Given the inherent class imbalance, approximately one goal for every ten shots, the split is performed in a stratified manner to ensure that the relative proportion of goals is preserved in both subsets, thereby maintaining consistency for expected goals modeling.

We discuss two widely used models: logistic regression and XGBoost, the latter chosen to capture potential non-linear relationships in the data.

Logistic regression. Logistic regression is a statistical learning method that models a binary outcome variable as a function of input features using the logistic function [74]. Given an input feature vector $\mathbf{x} = [x_1, x_2, \ldots, x_M]^\top$ and a binary target variable $y \in \{0, 1\}$, logistic regression models the conditional expectation of y through the following function:

$$g(\mathbb{E}[y \mid \mathbf{x}]) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_M x_M = \beta_0 + \sum_{m=1}^{M} \beta_m x_m, \qquad (5)$$

where $g(\cdot)$ denotes the logit link function and the coefficients $\{\beta_m\}_{m=0}^{M}$ are estimated from the training data.

XGBoost. XGBoost is a gradient boosting framework that builds an ensemble of decision trees in a sequential manner, where each tree is trained to correct the residual errors of the previous trees [46]. The model optimizes a differentiable loss function, incorporating regularization to prevent overfitting. Formally, given K regression trees $\{f_k(\mathbf{x})\}_{k=1}^{K}$, the predicted output for an input $\mathbf{x}$ is given by:

$$\hat{y} = \sum_{k=1}^{K} f_k(\mathbf{x}), \quad f_k \in \mathcal{F}, \tag{6}$$

where $\mathcal{F}$ denotes the space of regression trees and each $f_k(\cdot)$ is learned sequentially to minimize the overall objective function.

4.3. Evaluation Metrics

Following previous work [56, 35], we employ the following metrics to evaluate the performance of the deployed models.

AUROC. The Area Under the Receiver Operating Characteristic Curve (AUROC) is a threshold-independent metric that quantifies a classifier's ability to discriminate between positive and negative classes [75]. It is computed as the area under the receiver operating characteristic curve, which plots the true positive rate against the false positive rate at various decision thresholds.

Brier Score. The Brier Score (BS) is a scoring rule that measures the accuracy of probabilistic predictions for binary outcomes [76]. It is defined as the mean squared difference between the predicted probabilities $\hat{y}_j$ and the true binary labels $y_j \in \{0, 1\}$:

$$\mathrm{BS} = \frac{1}{J} \sum_{j=1}^{J} (\hat{y}_j - y_j)^2, \tag{7}$$

where J is the total number of samples.

Logarithmic Loss. Logarithmic loss (LL) quantifies the deviation between predicted probabilities and actual binary outcomes, penalizing predictions that assign high probability to incorrect labels [74]. For a set of predictions $\hat{y}_j$ and true labels $y_j \in \{0, 1\}$, it is defined as:

$$\mathrm{LL} = -\frac{1}{J} \sum_{i=j}^{J} \left[y_j \log(\hat{y}_j) + (1 - y_j) \log(1 - \hat{y}_j) \right], \tag{8}$$

where J represents the total number of observations.

Expected Calibration Error. The Expected Calibration Error (ECE) evaluates how well predicted probabilities align with observed outcomes [77]. Predictions are grouped into L equally spaced bins according to their confidence scores, and the difference between the average predicted probability and the actual frequency of positive outcomes is computed for each bin. Formally, ECE is defined as:

$$\text{ECE} = \sum_{l=1}^{L} \frac{|B_l|}{J} \Big| \text{acc}(B_l) - \text{conf}(B_l) \Big|, \tag{9}$$

where B_l is the set of samples in bin l, $|B_l|$ is the number of samples in the bin, J is the total number of samples, $\text{acc}(B_l)$ denotes the fraction of positive labels in the bin, and $\text{conf}(B_l)$ is the mean predicted probability of the samples in the bin.

5. Experiments

In this section, we present our experimental evaluation, providing both quantitative and qualitative analysis. Additionally, we discuss the pipeline's limitations, potential applications and impact, and provide an assessment of the model's global explainability.

5.1. Experimental Details

Data. We use the collected dataset described in Section 3.1, comprising two full seasons from the five major European soccer leagues. The dataset is split into training (7394) and test (3169) samples, by using a 7:3 ratio.

For the subsequent experiments, we categorize the input variables used to train our models into five groups, α_{base}, α_{cat}, α_{ctx}, SO_{2D} and SO_{3D}. Detailed information regarding the feature types and their corresponding groups is provided in Table 1. A full description of these features can be found in Section 4.1.

Metrics. Following previous approaches [56, 35], we evaluate our goal expectation models using the metrics described in Section 4.3. Specifically, we report the AUROC, BS, LL, and ECE. For ECE, we adopt a binning strategy with 10 equally spaced bins.

Models. Logistic Regression and XGBoost are the models selected to evaluate goal expectation, as described in Section 4.2. We employ 10-fold

Features	Type	Feature group
shot distance	continuous	α_{base}
angle to goal	continuous	α_{base}
angle between posts	continuous	α_{base}
body part	categorical	α_{cat}
shot technique	categorical	α_{cat}
play pattern	categorical	α_{cat}
shooter pressure	continuous	α_{ctx}
goal face occlusion def	continuous	α_{ctx}
goal face occlusion gk	continuous	α_{ctx}
goal face occlusion joint	continuous	α_{ctx}
shooter orientation 2D	continuous	SO_{2D}
shooter orientation 3D	continuous	SO_{3D}

Table 1: Details of the features used to train our xG models.

cross-validation to assess the ability of the models to capture patterns in the data. In this procedure, the training set is partitioned into k mutually exclusive subsets (folds). The model is iteratively trained on $k-1$ folds and evaluated on the remaining fold, producing k performance scores that are subsequently averaged to obtain a robust overall estimate.

For Logistic Regression, all input variables are standardized to ensure comparable scales and improve model convergence. This step is not required for XGBoost, as tree-based algorithms are inherently scale-invariant.

For each algorithm applied to xG modeling, hyperparameters are optimized via grid search. This process involves specifying candidate values for each hyperparameter, exhaustively evaluating all possible combinations, and selecting the configuration that yields the best cross-validated performance. The BS is adopted as the optimization criterion, as it provides a principled balance between predictive accuracy and probabilistic calibration, while avoiding excessive penalization of rare events.

5.2. Results and Comparisons

To gain a comprehensive understanding of the effect of shooter orientation on goal probability estimation, we compare the performance of baseline models (Logistic Regression and XGBoost) on the test set. Each model incorporates features from the α_{base}, α_{cat}, and α_{ctx} groups, with orientation

information included separately through SO_{2D} and SO_{3D}, as shown in Table 2. For reference, we also include a constant-probability baseline that assigns a fixed goal probability equal to the overall shot conversion rate in the train set (11.52%). Uncertainty estimates are computed via bootstrapping, using 1,000 resamples of the test set to derive confidence intervals for each metric.

Incorporating the 3D-based shooter-orientation feature leads to consistent but modest improvements in the XGBoost model across all evaluated metrics, including AUROC, BS, LL, and ECE. While the direction of the effect is stable, the absolute magnitude of the gains is small and, in several cases, comparable to the corresponding uncertainty estimates. This indicates that 3D orientation does not introduce a strong global signal affecting all shots uniformly.

In contrast, the inclusion of the 2D-based orientation feature does not yield performance improvements for either model and, in some cases, slightly degrades results. This suggests that 2D pose–derived orientation, which relies on projected upper-body keypoints, is insufficiently stable or informative in broadcast footage to capture meaningful shot-related patterns, even when non-linear models are employed.

Logistic regression models do not benefit from the inclusion of either orientation formulation and occasionally exhibit slight performance degradation. This behavior is expected given the linear nature of logistic regression and the inherently non-linear geometric structure of angular orientation variables, which cannot be expressed as monotonic or additive effects without losing spatial meaning.

Taken together, these results suggest that the primary value of 3D-based shooter orientation does not lie in large average gains in predictive accuracy, but rather in its ability to capture meaningful differences in specific shot configurations that are otherwise treated as equivalent by pose-agnostic models. This interpretation is further supported by the explainability analysis in Section 5.4 and the qualitative and counterfactual case studies in Section 5.6, which demonstrate that orientation-aware models provide richer contextual reasoning and more nuanced assessments in particular edge cases.

5.3. Ablation Study on Feature Groups

To evaluate the contribution of each feature group to model performance, we train and assess an XGBoost model by progressively incorporating the

Model	SO_{2D}	SO_{3D}	AUROC ↑	BS ↓	LL ↓	ECE ↓
Baseline	-	-	0.5000	0.1019	0.3572	-
Logistic Regression	✗	✗	$0.7783_{\pm 0.0139}$	$0.0852_{\pm 0.0038}$	$0.2961_{\pm 0.0115}$	$0.0126_{\pm 0.0035}$
	✓	✗	$0.7782_{\pm 0.0139}$	$0.0852_{\pm 0.0038}$	$0.2961_{\pm 0.0115}$	$0.0126_{\pm 0.0034}$
	✗	✓	$0.7778_{\pm 0.0139}$	$0.0852_{\pm 0.0038}$	$0.2962_{\pm 0.0115}$	$0.0129_{\pm 0.0035}$
XGBoost	✗	✗	$0.7868_{\pm 0.0128}$	$0.0839_{\pm 0.0037}$	$0.2914_{\pm 0.0115}$	$0.0123_{\pm 0.0037}$
	✓	✗	$0.7857_{\pm 0.0129}$	$0.0840_{\pm 0.0037}$	$0.2917_{\pm 0.0114}$	$0.0123_{\pm 0.0037}$
	✗	✓	$\mathbf{0.7881}_{\pm 0.0126}$	$\mathbf{0.0836}_{\pm 0.0037}$	$\mathbf{0.2903}_{\pm 0.0113}$	$\mathbf{0.0120}_{\pm 0.0035}$

Table 2: **Performance comparison.** A constant goal-probability baseline is compared against Logistic Regression and XGBoost models. The SO_{2D} and SO_{3D} feature groups are evaluated separately to illustrate the performance impact of 2D-based and 3D-based orientation features on model capabilities.

feature groups, as summarized in Table 3. We start with a base model containing only the α_{base} features and sequentially add the remaining groups, concluding with the full model that includes the 3D-based orientation features.

The inclusion of categorical (α_{cat}) and contextual (α_{ctx}) features leads to clear and consistent improvements across all evaluation metrics, confirming their strong contribution to expected-goal modeling. Adding the 2D-based orientation feature, however, does not result in measurable performance gains and in some cases slightly degrades performance, suggesting that orientation estimates derived from 2D projected keypoints are insufficiently stable or informative in this setting.

Incorporating the 3D-based shooter-orientation feature yields further improvements across all metrics. While these gains are modest in absolute terms and often comparable to the associated uncertainty estimates, their direction is consistent and they produce the best overall performance among the evaluated configurations. This indicates that 3D orientation complements existing spatial and contextual features rather than acting as a dominant global predictor.

To address potential collinearity and its implications under monotonic constraints within the α_{base} feature group, Table 4 reports the performance of the XGBoost model when excluding the double-angle feature and when using either or both angle representations. Beyond the slight performance gain observed when both features are included, the two player-orientation features are both defined relative to the player's angle to the goal. Moreover,

α_{cat}	α_{ctx}	SO_{2D}	SO_{3D}	AUROC ↑	BS ↓	LL ↓	ECE ↓
✗	✗	✗	✗	$0.7548_{\pm 0.0140}$	$0.0896_{\pm 0.0041}$	$0.3097_{\pm 0.0121}$	$0.0127_{\pm 0.0035}$
✓	✗	✗	✗	$0.7694_{\pm 0.0135}$	$0.0864_{\pm 0.0039}$	$0.3004_{\pm 0.0119}$	$0.0123_{\pm 0.0035}$
✓	✓	✗	✗	$0.7868_{\pm 0.0128}$	$0.0840_{\pm 0.0037}$	$0.2914_{\pm 0.0115}$	$0.0123_{\pm 0.0037}$
✓	✓	✓	✗	$0.7857_{\pm 0.0129}$	$0.0840_{\pm 0.0037}$	$0.2917_{\pm 0.0114}$	$0.0123_{\pm 0.0037}$
✓	✓	✗	✓	$\mathbf{0.7881}_{\pm 0.0126}$	$\mathbf{0.0836}_{\pm 0.0037}$	$\mathbf{0.2903}_{\pm 0.0113}$	$\mathbf{0.0120}_{\pm 0.0035}$

Table 3: **Ablation study.** Performance analysis of the XGBoost model across different feature group configurations.

shot distance	angle to goal	angle between posts	AUROC ↑	BS ↓	LL ↓	ECE ↓
✓	✗	✗	$0.7736_{\pm 0.0125}$	$0.0853_{\pm 0.0037}$	$0.2965_{\pm 0.0111}$	$0.0133_{\pm 0.0038}$
✓	✓	✗	$0.7769_{\pm 0.0131}$	$0.0845_{\pm 0.0037}$	$0.2948_{\pm 0.0117}$	$0.0169_{\pm 0.0041}$
✓	✗	✓	$0.7879_{\pm 0.0126}$	$0.0839_{\pm 0.0037}$	$0.2905_{\pm 0.0113}$	$0.0121_{\pm 0.0036}$
✓	✓	✓	$0.7881_{\pm 0.0126}$	$0.0836_{\pm 0.0037}$	$0.2903_{\pm 0.0113}$	$0.0120_{\pm 0.0035}$

Table 4: **Ablation study.** Potential collinearity analysis of the XGBoost model across α_{base} feature group configurations.

tree-based models such as XGBoost are inherently robust to multicollinearity.

Calibration curves for a subset of the models are shown in Fig. 4, illustrating the effects of progressively enriching the feature set. The base model tends to overpredict low-probability shots, an issue that is mitigated by the addition of contextual features, resulting in a more balanced distribution of predicted probabilities and improved calibration in the upper probability range. Incorporating the 3D-based shooter-orientation feature leads to a further, albeit limited, refinement of calibration in this high-probability regime.

5.4. Global Explainability

Leveraging the use of tabular data and standard machine-learning methods (e.g., XGBoost), we employ SHapley Additive exPlanations (SHAP) [55], a game-theoretic framework for interpreting model outputs, to provide global model explanations and analyze the contribution of the proposed shooter-orientation modeling to goal-expectation predictions. Figure 5-left presents the global feature-importance plot, where the importance of each feature corresponds to the mean absolute SHAP value across all samples. The shooter-orientation variable appears among the top ten most influential features, ranking well above several widely used features in the literature.

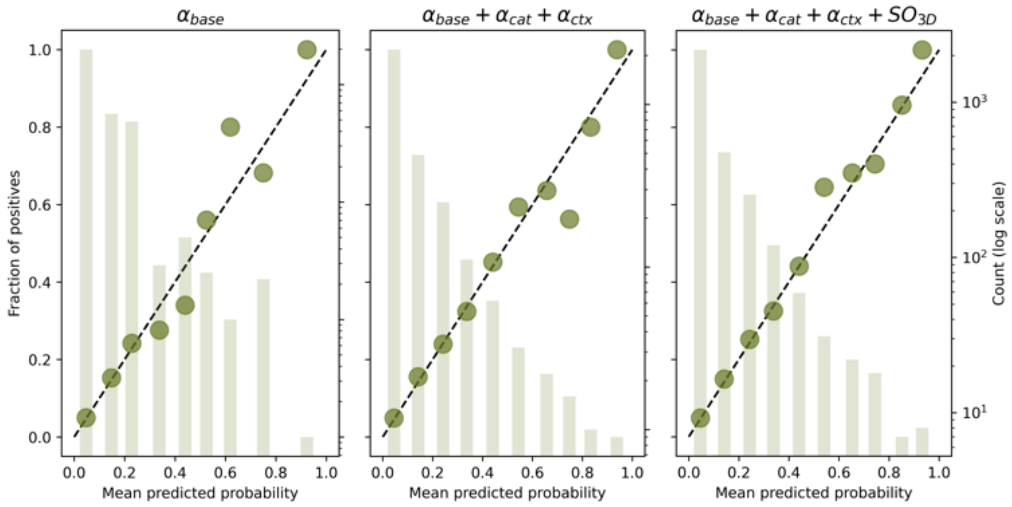

Figure 4: **Calibration curves** for the XGBoost models, shown from left to right as feature sets are progressively incorporated. The dashed black line indicates perfect calibration. The background histogram shows the number of observations per interval on a logarithmic scale.

Moreover, Fig. 5-right shows a scatter plot of SHAP values as a function of the shooter-orientation angle (in degrees). Since positive SHAP values indicate that a given orientation pushes the model toward a higher predicted goal probability, a clear and symmetric pattern emerges: positive SHAP contributions concentrate in the interval $SO_{3D} \in [-90°, 90°]$, as values outside this range correspond to shots taken with the goal effectively behind the shooter. Notably, orientations close to 0° tend to yield smaller SHAP values than orientations with larger absolute deviation, suggesting that a slight rotation relative to the direct goal-facing direction is associated with more favorable shooting outcomes.

Furthermore, in line with the outcomes discussed in the previous sections, we observe a stronger impact of shooter orientation in the SHAP contributions for shots with orientations lying outside the $[-90°, 90°]$ interval. In these cases, SHAP values tend to be strongly negative, pushing the model predictions toward lower scoring probabilities, and can reach magnitudes up to four times larger than those associated with orientations more directly facing the goal. This behavior indicates that shooter orientation acts primarily as a discriminative feature in extreme orientation scenarios, where pose-agnostic representations are more likely to assign overly optimistic goal

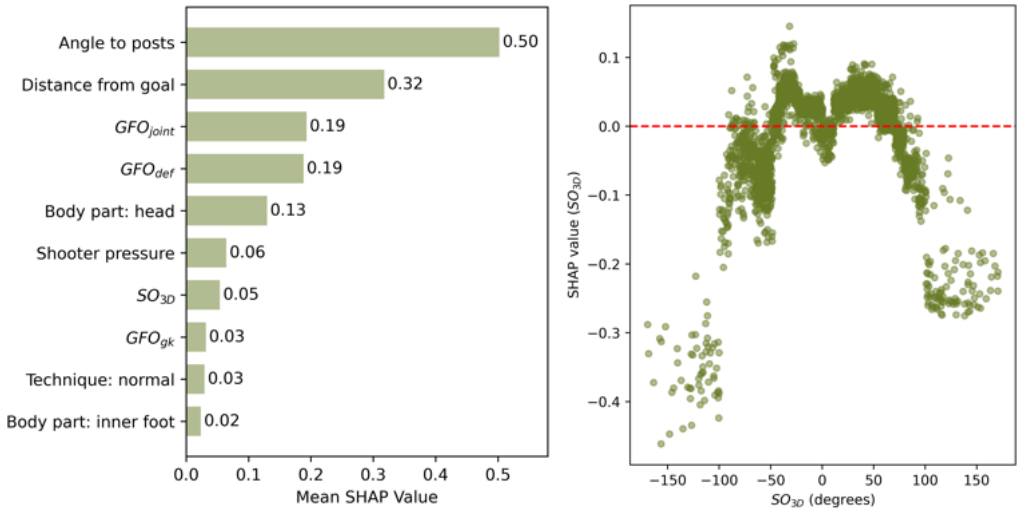

Figure 5: **SHAP global explainability. Left:** Global feature-importance plot, where the importance of each feature corresponds to the mean absolute SHAP value across all samples. **Right:** SHAP values as a function of the shooter-orientation angle (in degrees), with each point representing a test-set observation.

probabilities.

5.5. Limitations

The proposed pipeline relies on multiple modules for human pose generation, as illustrated in Fig. 1, which introduces several limitations. First, highly occluded scenes, such as corner kicks or free kicks, significantly reduce the reliability of the pipeline. While tracking accuracy is also affected, most modern methods handle moderate occlusions reasonably well, and offline techniques such as global tracklet association [78] can further mitigate this issue. However, pose estimation remains particularly sensitive to occlusions. Even when tracking is preserved, frames in which the shooter is partially or fully occluded tend to decrease the reliability of the estimated pose. The tracking-based design of PHALP [72] helps alleviate this limitation to some extent. To further reduce the impact of occlusions, we restrict each shot video to a total duration of one second. Second, pose reliability strongly depends on video quality. To ensure adequate keypoint accuracy, we restrict our analysis to videos with a minimum resolution of 800 pixels.

Finally, shooter bounding box selection is currently performed manually, limiting the pipeline's scalability to new data. However, xGHub annotations

can serve as training material for developing automatic shooter detection models.

5.6. Applications and Impact

The xG metric is widely adopted across both professional analytics and fan-oriented platforms. Football data providers and analysts routinely use xG as the foundation for a variety of derived performance indicators. However, we argue that neglecting orientation information introduces systematic biases, leading to potential misinterpretations. To illustrate that, we present practical examples showing how integrating orientation data refines xG-based assessments for both shooters and goalkeepers, resulting in more accurate and context-aware performance evaluations.

Shooter Evaluation. In the shot situation illustrated in Fig. 6, the baseline xG model assigns a probability of $xG = 0.030$. However, this estimation reflects an average over similar situations without considering shooter orientation. In contrast, the orientation-aware xG model, trained with 3D orientation data, predicts a slightly lower probability of $xG = 0.028$, which is consistent with the fact that the shooter is facing away from the goal. As a counterfactual experiment, we rotate the shooter's skeleton to face the goal, resulting in an updated prediction of $xG = 0.037$. This higher probability aligns with the improved shooting position and demonstrates the model's capacity to discriminate between different body orientations rather than averaging across geometrically dissimilar situations. This behavior highlights the added interpretability and contextual accuracy introduced by incorporating orientation information into xG estimation.

Goalkeeper Evaluation. Previous approaches [40] have employed xG to assess goalkeeper positioning in shot situations by placing the goalkeeper at different locations within the goal area and measuring the deviation from the optimal position. Using the same shot scenario illustrated in Fig. 6, we generate an *xG-suppression map* by computing the percentage change in xG relative to the original shot configuration as the goalkeeper's position is systematically varied across the goal area. We construct these maps for both the base and the orientation-aware xG models, denoted as ΔxG_{base} and $\Delta xG_{\text{orient}}$, respectively, along with their difference, as shown in Fig. 7. It is worth noting that lower values in the suppression map indicate higher goal-probability suppression. We can clearly observe the differences between the base and the orientation-aware models, with the latter exhibiting a visibly

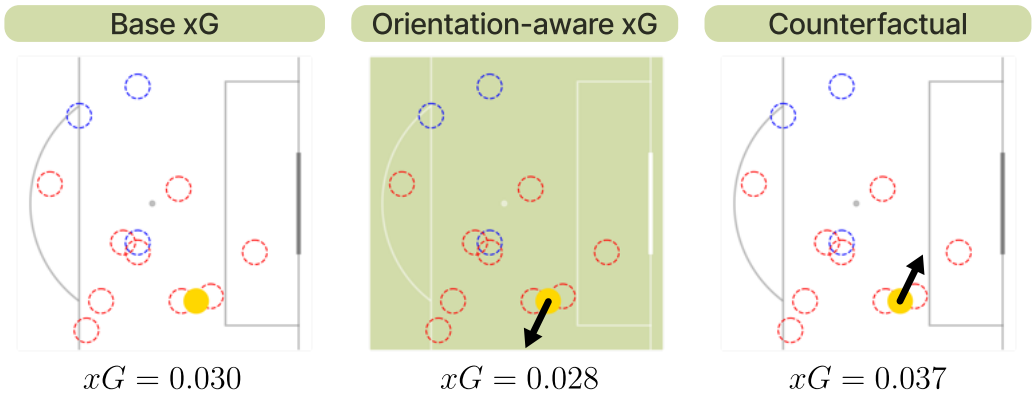

Figure 6: **Body orientation impact on shooter evaluation.** Comparison between the base and orientation-aware xG model inferences. Additionally, we introduce a counterfactual scenario by imposing a new shooter orientation to analyze model behavior. The black arrow indicates the shooter's body orientation. The solid yellow circle represents the shooter, while dashed circles denote other players, teammates in blue and opponents in red.

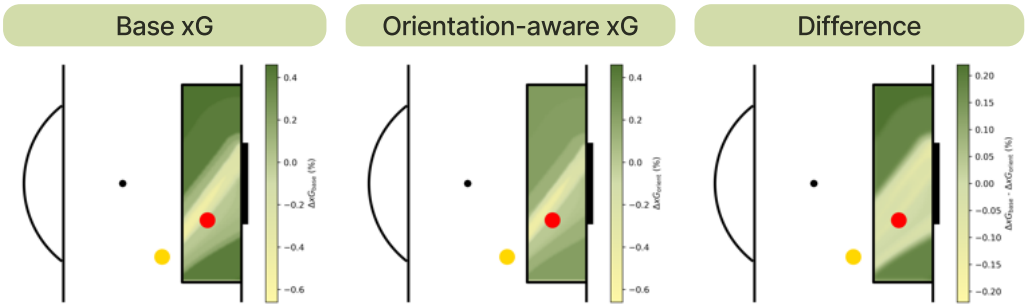

Figure 7: **Body-orientation impact on goalkeeper evaluation.** We construct *xG-suppression maps* using both the base model and the orientation-aware model, and additionally visualize their difference. The yellow circle indicates the shooter's location, while the red circle denotes the goalkeeper's position. Only the goal box is evaluated.

narrower xG-suppression band compared to the base one. This pattern indicates that the shooter's effective scoring possibilities are more constrained when body orientation is taken into account. The differences between suppression maps highlight the importance of incorporating orientation information for reliable goalkeeper performance assessment.

6. Conclusion

This work introduces xGHub, the first open-source dataset for football shot analysis that combines broadcast video clips with player tracking data, pose estimates, and detailed manual annotations. We present a modular pipeline integrating object detection, tracking, camera calibration, and human pose estimation, enabling the extraction of skeletal information directly from broadcast footage for shot-based performance modeling. Our experimental evaluation shows that incorporating 3D-based shooter orientation into gradient-boosted tree models yields consistent but modest improvements in expected-goal estimation. While these gains are limited in magnitude and often comparable to their associated uncertainty, they result in the best overall performance among the evaluated configurations. In contrast, 2D-based orientation features and linear models fail to capture the non-linear geometric structure of shooting situations and do not provide measurable benefits. Importantly, the main value of 3D-based shooter orientation lies not in large average improvements in global predictive accuracy, but in its ability to discriminate between geometrically distinct shot situations that are otherwise treated as equivalent by pose-agnostic models. This selective usefulness is reflected in the explainability analysis and qualitative case studies, where orientation-aware models enable more contextually grounded predictions and meaningful counterfactual reasoning. Several limitations remain, including sensitivity to video quality and occlusions and the current reliance on manual shooter identification. Addressing these constraints and extending pose-based modeling to additional match contexts constitute promising directions for future work. By releasing xGHub and our methodological framework, we aim to foster open research on skeleton-based football analytics and the development of biomechanically informed performance metrics.

Acknowledgment. This work has been supported by the project GRAVATAR PID2023-151184OB-I00 funded by MCIU/AEI/10.13039/501100011033 and by ERDF, UE and by the Government of Catalonia under Joan Oró FI 2024 grant.

References

[1] P. R. Kamble, A. G. Keskar, K. M. Bhurchandi, Ball tracking in sports: a survey, Artificial Intelligence Review 52 (3) (2019) 1655–1705.

[2] C. Yang, M. Yang, H. Li, L. Jiang, X. Suo, L. Mao, W. Meng, Z. Li, A survey on soccer player detection and tracking with videos, The Visual Computer 41 (2) (2025) 815–829.

[3] R. Beal, T. J. Norman, S. D. Ramchurn, Artificial intelligence for team sports: a survey, The Knowledge Engineering Review 34 (2019) e28.

[4] Z. Wang, P. Veličković, D. Hennes, N. Tomašev, L. Prince, M. Kaisers, Y. Bachrach, R. Elie, L. K. Wenliang, F. Piccinini, et al., TacticAI: an AI assistant for football tactics, Nature communications 15 (1) (2024) 1906.

[5] J. Fernández, L. Bornn, D. Cervone, A framework for the fine-grained evaluation of the instantaneous expected value of soccer possessions, Machine Learning 110 (6) (2021) 1389–1427.

[6] J. G. Claudino, D. d. O. Capanema, T. V. de Souza, J. C. Serrão, A. C. Machado Pereira, G. P. Nassis, Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: a systematic review, Sports medicine-open 5 (1) (2019) 28.

[7] N. Blanchard, K. Skinner, A. Kemp, W. Scheirer, P. Flynn, keep me in, coach!: A computer vision perspective on assessing acl injury risk in female athletes, in: IEEE Winter Conference on Applications of Computer Vision, 2019.

[8] K. Tuyls, S. Omidshafiei, P. Muller, Z. Wang, J. Connor, D. Hennes, I. Graham, W. Spearman, T. Waskett, D. Steel, et al., Game plan: What AI can do for football, and what football can do for AI, Journal of Artificial Intelligence Research 71 (2021) 41–88.

[9] Y. Tian, Q. Ye, D. Doermann, Yolov12: Attention-centric real-time object detectors, in: arXiv preprint arXiv:2502.12524, 2025.

[10] I. Robinson, P. Robicheaux, M. Popov, Rf-detr, https://github.com/roboflow/rf-detr, sOTA Real-Time Object Detection Model (2025).

[11] M. Gutiérrez-Pérez, A. Agudo, Soccernet-v3D: Leveraging sports broadcast replays for 3D scene understanding, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2025.

[12] N. Wojke, A. Bewley, D. Paulus, Simple online and realtime tracking with a deep association metric, in: IEEE International Conference on Image Processing, 2017.

[13] A. Pérez-Yus, A. Agudo, Matching and recovering 3D people from multiple views, in: IEEE Winter Conference on Applications of Computer Vision, 2022.

[14] N. Aharon, R. Orfaig, B.-Z. Bobrovsky, Bot-sort: Robust associations multi-pedestrian tracking, in: ArXiv, Vol. abs/2206.14651, 2022.

[15] M. Gutiérrez-Pérez, A. Agudo, No bells just whistles: Sports field registration by leveraging geometric properties, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2024.

[16] M. Gutiérrez-Pérez, A. Agudo, Pnlcalib: Sports field registration via points and lines optimization, Available at SSRN 4998149 (2025).

[17] FIFA, Semi-automated offside techonology, "https://inside.fifa.com/innovation/world-cup-2022/semi-automated-offside-technology" (2023).

[18] M. Bassek, R. Rein, H. Weber, D. Memmert, An integrated dataset of spatiotemporal and event data in elite soccer, Scientific Data 12 (1) (2025) 195.

[19] G. Capellera, A. Rubio, L. Ferraz, A. Agudo, Unified uncertainty-aware diffusion for multi-agent trajectory modeling, in: IEEE Conference on Computer Vision and Pattern Recognition, 2025.

[20] C. Yeung, K. Ide, K. Fujii, Autosoccerpose: Automated 3D posture analysis of soccer shot movements, in: IEEE Conference on Computer Vision and Pattern Recognition, 2024.

[21] T. Jiang, J. Billingham, S. Müksch, J. Zarate, N. Evans, M. R. Oswald, M. Polleyfeys, O. Hilliges, M. Kaufmann, J. Song, Worldpose: A world cup dataset for global 3D human pose estimation, in: European Conference on Computer Vision, 2024.

[22] S. A. Ghasemzadeh, G. Van Zandycke, M. Istasse, N. Sayez, A. Moshtaghpour, C. De Vleeschouwer, Deepsportlab: a unified framework for

ball detection, player instance segmentation and pose estimation in team sports scenes, in: arXiv preprint arXiv:2112.00627, 2021.

[23] K. Yamada, L. Yin, Q. Hu, N. Ding, S. Iwashita, J. Ichikawa, K. Kotani, C. Yeung, K. Fujii, TrackID3x3: A dataset and algorithm for multi-player tracking with identification and pose estimation in 3x3 basketball full-court videos, in: arXiv preprint arXiv:2503.18282, 2025.

[24] J. Bright, Y. Chen, J. Zelek, Mitigating motion blur for robust 3D baseball player pose modeling for pitch analysis, in: International Workshop on Multimedia Content Analysis in Sports, 2023.

[25] C. Yeung, T. Suzuki, R. Tanaka, Z. Yin, K. Fujii, Athletepose3D: A benchmark dataset for 3D human pose estimation and kinematic validation in athletic movements, in: IEEE Conference on Computer Vision and Pattern Recognition, 2025.

[26] T. Suzuki, R. Tanaka, C. Yeung, K. Fujii, Athleticspose: Authentic sports motion dataset on athletic field and evaluation of monocular 3D pose estimation ability, in: arXiv preprint arXiv:2507.12905, 2025.

[27] Z. Jiang, H. Ji, S. Menaker, J.-N. Hwang, Golfpose: Golf swing analyses with a monocular camera based human pose estimation, in: IEEE International Conference on Multimedia and Expo Workshops, 2022.

[28] H. F. Z. Khoiee, D. Labbe, T. Romeas, J. Faubert, S. Andrews, Multi-person physics-based pose estimation for combat sports, in: IEEE Conference on Computer Vision and Pattern Recognition, 2025.

[29] L. Bridgeman, M. Volino, J.-Y. Guillemaut, A. Hilton, Multi-person 3D pose estimation and tracking in sports, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019.

[30] R. Chatterjee, S. Roy, S. H. Islam, D. Samanta, An AI approach to pose-based sports activity classification, in: International Conference on Signal Processing and Integrated Networks, 2021.

[31] R. Tanaka, T. Suzuki, K. Fujii, 3D pose-based temporal action segmentation for figure skating: a fine-grained and jump procedure-aware annotation approach, in: ACM International Workshop on Multimedia Content Analysis in Sports, 2024.

[32] H. Li, Q. Lei, H. Zhang, J. Du, S. Gao, Skeleton-based deep pose feature learning for action quality assessment on figure skating videos, Journal of Visual Communication and Image Representation 89 (2022) 103625.

[33] V. Baghel, N. Rithihas, B. Srinivasan, R. S. Hegde, et al., Efficient boxing punch classification: fine-grained skeleton-based recognition made light, in: International Conference on Computer Vision and Computational Intelligence, 2024.

[34] A. Arbues-Sanguesa, A. Martín, J. Fernández, C. Ballester, G. Haro, Using player's body-orientation to model pass feasibility in soccer, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2020.

[35] Y. Xu, L. Bretzner, T. Wang, A. Maki, Skor-xG: Skeleton-oriented expected goal estimation in soccer, in: IEEE Conference on Computer Vision and Pattern Recognition, 2025.

[36] M. Schepers, P. Robberechts, J. Van Haaren, J. Davis, What makes a dribble successful? insights from 3D pose tracking data, in: arXiv preprint arXiv:2506.22503, 2025.

[37] C. Reep, B. Benjamin, Skill and chance in association football, Journal of the Royal Statistical Society. Series A (General) 131 (4) (1968) 581–585.

[38] P. Lucey, A. Bialkowski, M. Monfort, P. Carr, I. Matthews, quality vs. quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data, in: MIT Sloan Sports Analytics Conference, 2015.

[39] H. Eggels, R. van Elk, M. Pechenizkiy, Expected goals in soccer: Explaining match results using predictive analytics, in: The machine learning and data mining for sports analytics workshop, Vol. 16, Technische Universiteit Eindhoven, 2016.

[40] D. Vatvani, Upgrading expected goals, "https://www.hudl.com/blog/upgrading-expected-goals/" (2022).

[41] J. H. Hewitt, O. Karakuş, A machine learning approach for player and position adjusted expected goals in football (soccer), Franklin open 4 (2023) 100034.

[42] S. Green, Assessing the performance of premier league goalscorers, "https://www.statsperform.com/es/resource/assessing-the-performance-of-premier-league-goalscorers/" (2012).

[43] L. Schiavon, N. Sartori, et al., Bias reduced estimation of a fixed effects model for expected goals in association football, in: Smart Statistics for Smart Application, 2019.

[44] M. Cavus, P. Biecek, Explainable expected goal models for performance analysis in football analytics, in: International Conference on Data Science and Advanced Analytics, 2022.

[45] P. Madrero Pardo, Creating a model for expected goals in football using qualitative player information, Master's thesis, Universitat Politècnica de Catalunya (2020).

[46] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: ACM International Conference on Knowledge Discovery and Data Mining, 2016.

[47] R. Pollard, C. Reep, Measuring the effectiveness of playing strategies at soccer, Journal of the Royal Statistical Society Series D: The Statistician 46 (4) (1997) 541–550.

[48] M. Caley, Shot matrix i: Shot location and expected goals, "https://cartilagefreecaptain.sbnation.com/2013/11/13/5098186/shot-matrix-i-shot-location-and-expected-goals" (2013).

[49] J. H. Friedman, Greedy function approximation: a gradient boosting machine, Annals of statistics (2001) 1189–1232.

[50] G. Anzer, P. Bauer, A goal scoring probability model for shots based on synchronized positional and event data in football (soccer), Frontiers in sports and active living 3 (2021) 624475.

[51] J. Mead, A. O'Hare, P. McMenemy, Expected goals in football: Improving model performance and demonstrating value, Plos one 18 (4) (2023) e0282295.

[52] B. Nipoti, L. Schiavon, Expected goals under a bayesian viewpoint: uncertainty quantification and online learning, Journal of Quantitative Analysis in Sports 21 (1) (2025) 37–50.

[53] I. Bandara, S. Shelyag, S. Rajasegarar, D. Dwyer, E.-j. Kim, M. Angelova, Predicting goal probabilities with improved xG models using event sequences in association football, Plos one 19 (10) (2024) e0312278.

[54] M. Cavus, A. Stando, P. Biecek, Glocal explanations of expected goal models in soccer, in: arXiv preprint arXiv:2308.15559, 2023.

[55] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, 2017.

[56] J. Van Haaren, Why would i trust your numbers? on the explainability of expected values in soccer, in: arXiv preprint arXiv:2105.13778, 2021.

[57] H. Nori, S. Jenkins, P. Koch, R. Caruana, Interpretml: A unified framework for machine learning interpretability, in: arXiv preprint arXiv:1909.09223, 2019.

[58] H. Ruiz, P. Lisboa, P. Neilson, W. Gregson, Measuring scoring efficiency through goal expectancy estimation, in: European symposium on artificial neural networks, computational intelligence and machine learning, 2015.

[59] M. Wear, R. Beal, T. Matthews, T. Norman, S. Ramchurn, Learning from the pros: Extracting professional goalkeeper technique from broadcast footage, in: arXiv preprint arXiv:2202.12259, 2022.

[60] C. Reddy, W. Jeon, Identifying goalkeeper movement timing from single-camera broadcast footage through pose estimation: A pilot study, Applied Sciences 14 (13) (2024) 5961.

[61] A. Arbués-Sangüesa, A. Martín, J. Fernández, C. Rodríguez, G. Haro, C. Ballester, Always look on the bright side of the field: Merging pose and contextual data to estimate orientation of soccer players, in: IEEE International Conference on Image Processing, 2020.

[62] J. Fang, C. Yeung, K. Fujii, Foul prediction with estimated poses from soccer broadcast video, Sports Engineering 28 (2) (2025) 33.

[63] S. Giancola, M. Amine, T. Dghaily, B. Ghanem, Soccernet: A scalable dataset for action spotting in soccer videos, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018.

[64] A. Deliege, A. Cioppa, S. Giancola, M. J. Seikavandi, J. V. Dueholm, K. Nasrollahi, B. Ghanem, T. B. Moeslund, M. Van Droogenbroeck, Soccernet-v2: A dataset and benchmarks for holistic understanding of broadcast soccer videos, in: IEEE Conference on Computer Vision and Pattern Recognition, 2021.

[65] A. Cioppa, A. Deliege, S. Giancola, B. Ghanem, M. Van Droogenbroeck, Scaling up soccernet with multi-view spatial localization and re-identification, Scientific data 9 (1) (2022) 355.

[66] J. Rao, H. Wu, H. Jiang, Y. Zhang, Y. Wang, W. Xie, Towards universal soccer video understanding, in: IEEE Conference on Computer Vision and Pattern Recognition, 2025.

[67] S. Giancola, A. Cioppa, M. Gutiérrez-Pérez, J. Held, C. Hinojosa, V. Joos, A. Leduc, F. Magera, K. Sanchez, V. Somers, A. Xarles, A. Agudo, A. Alahi, O. Barnich, A. Clapés, C. D. Vleeschouwer, S. Escalera, B. Ghanem, T. B. Moeslund, M. V. Droogenbroeck, T. Abe, S. Alotaibi, F. Altawijri, S. Araujo, X. Bai, X. Bi, J. Cao, V. Chao, K. Czarnogórski, F. Deuser, M. Du, T. Feng, P. Frenzel, M. Fuchs, J. García, K. Habel, T. Hashiguchi, S. Hirose, X. Hu, Y. Hwang, R. Inoue, R. Itsuji, K. Iwai, H. Ji, Y. Ji, L. Jiao, Y. Kageyama, Y. Kamikawa, Y. Kanasugi, H. Kim, J. Kim, T. Kurihara, B. Li, L. Li, X. Li, Y. Lian, D. Liang, H. Lin, J. Lin, J. Liu, L. Liu, S. Liu, Z. Liu, Y. Lu, F. Méndez, H. Ma, W. Ma, J. Maksymiuk, H. Mantilla, I. Mathkour, D. Matthes, A. Motomochi, A. R. Muhammad, H. Nakayama, J. Oh, Y. M. Oo, M. Ortega, N. Oswald, R. Otsubo, F. Perez, M. Qi, C. Rey, A. Reyes-Angulo, O. Rose, H. Rueda-Chacón, H. Saito, J. Sarmiento, K. Sawafuji, A. Scott, X. Shen, P. Shrestha, J.-Y. Sim, L. Sun, Y. Sun, T. Suzuki, L. Tang, M. Tonouchi, I. Uchida, H. O. Velesaca, T. Wang, R. Watanabe, J. Wu, Y. Wu, S. Yamagishi, D. Yang, X. Yang, Y. Yang, H. Ye, X. Ye, C. Yeung, X. Yu, C. Zhang, D. Zhang, K. Zhang, Z. Zhao, X. Zhou, W. Zhu, J. Ziegler, Soccernet 2025 challenges results, in: arXiv preprint arXiv:2508.19182, 2025.

[68] V. Somers, V. Joos, A. Cioppa, S. Giancola, S. A. Ghasemzadeh, F. Magera, B. Standaert, A. M. Mansourian, X. Zhou, S. Kasaei, et al., Soccernet game state reconstruction: End-to-end athlete tracking and identification on a minimap, in: IEEE Conference on Computer Vision and Pattern Recognition, 2024.

[69] L. Chen, X. Chu, X. Zhang, J. Sun, Simple baselines for image restoration, in: European Conference on Computer Vision, 2022.

[70] S. Goel, G. Pavlakos, J. Rajasegaran, A. Kanazawa, J. Malik, Humans in 4D: Reconstructing and tracking humans with transformers, in: IEEE International Conference on Computer Vision, 2023.

[71] Y. Xu, J. Zhang, Q. Zhang, D. Tao, in: Vitpose: Simple vision transformer baselines for human pose estimation, 2022.

[72] J. Rajasegaran, G. Pavlakos, A. Kanazawa, J. Malik, Tracking people by predicting 3D appearance, location and pose, in: IEEE Conference on Computer Vision and Pattern Recognition, 2022.

[73] G. Andrienko, N. Andrienko, G. Budziak, J. Dykes, G. Fuchs, T. Von Landesberger, H. Weber, Visual analysis of pressure in football, Data Mining and Knowledge Discovery 31 (6) (2017) 1793–1839.

[74] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in python, Journal of Machine Learning Research 12 (2011) 2825–2830.

[75] T. Fawcett, An introduction to ROC analysis, Pattern Recognition Letters 27 (8) (2006) 861–874.

[76] W. B. Glenn, et al., Verification of forecasts expressed in terms of probability, Monthly weather review 78 (1) (1950) 1–3.

[77] M. P. Naeini, G. Cooper, M. Hauskrecht, Obtaining well calibrated probabilities using bayesian binning, in: Conference on Artificial Intelligence, 2015.

[78] J. Sun, H.-W. Huang, C.-Y. Yang, Z. Jiang, J.-N. Hwang, Gta: Global tracklet association for multi-object tracking in sports, in: Asian Conference on Computer Vision, 2024.