Enhancing Football Refereeing with AI: VARS and X-VARS for Assisted Decision-Making

The Video Assistant Referee (VAR) has revolutionized association football, enabling referees to review incidents on the pitch, make informed decisions, and improve fairness during the game. However, due to the lack of referees in many countries and the high cost of the VAR infrastructure, only professional leagues can currently benefit from it. To address these challenges, we introduced the Video Assistant Referee System (VARS), an automated solution for soccer decision-making using broadcast cameras. VARS was built upon the latest advancements in multi-view video analysis to provide real-time feedback to referees, helping them make informed decisions that can impact the outcome of a game.

While VARS effectively automates decision-making, it does not provide explanations for its decisions, making it difficult to interpret or trust its reasoning. To address this, we later introduced the eXplainable Video Assistant Referee System (X-VARS), a multi-modal large language model trained to analyze football videos from a referee's perspective. X-VARS can perform a wide range of tasks, including video description, question answering, action recognition, and conducting meaningful conversations based on football video content—all in accordance with the Laws of the Game for football referees.

To validate our VARS and X-VARS, we introduced two datasets. The first one, SoccerNet-MVFoul, is a video dataset of soccer fouls captured from multiple camera views, annotated with extensive foul descriptions by a professional soccer referee. Using this dataset, we benchmarked VARS to automatically determine whether an action constitutes a foul, assess the severity of a foul, and classify the type of foul (e.g., a tackle, holding, etc.). The second one, SoccerNet-XFoul, is a dataset of over 22,000 video-question-answer triplets annotated by more than 70 experienced football referees.

In this abstract, we present the results of a human study exploring the dynamics of when to provide VARS assistance to referee. We investigate how referees' behavior and performance vary across four conditions: (1) no VARS assistance, (2) always VARS assistance, (3) VARS assistance provided only when the model is confident, and (4) VARS assistance available upon referee request. Each condition involved 20 referees, who evaluated 25 multi-view videos of the same football action captured from three perspectives. For each video, referees were tasked with determining whether a foul occurred and, if so, assessing its severity.

Our results demonstrate that referees supported by VARS are significantly more accurate, quicker, and more confident in their decisions compared to referees making decisions independently. Additionally, referees have a statistically higher inter-rater agreement with VARS as support. These findings suggest that integrating AI assistance not only enhances individual referee performance but also promotes greater consistency in decision-making across referees.

In summary, we showed that VARS and X-VARS have the potential to significantly improve soccer refereeing by ensuring fairness and accuracy across all levels of play. VARS provides a reliable assistant in the decision-making process, while X-VARS demonstrates exceptional capabilities in explaining its decisions, paving the way for enhanced transparency and trust in football refereeing.