Football: Deep-learning based positional estimation of agents, automated real-time spatiotemporal tracking, & planar visualisation

- A work by Makesh Srinivasan

INSTITUTION'S INNOVATION COUNCIL (Ministry of Education Initiative) · V-NEST VIT STARTUP INCUBATION CENTRE WE NURTURE YOU TO FLY HIGH · VIT Vellore Institute of Technology (Deemed to be University under section 3 of UGC Act, 1956) CHENNAI

1 Introduction

With over a billion fans, football is experiencing a transformation through Artificial Intelligence (AI). Innovations like Video Assistant Referee (VAR) and Semi-Automated Offside Technology (SAOT) have enhanced decision-making and fairness, but AI's potential extends beyond these applications.

Utilising AI for visualisation, player tracking, and positional predictions can provide data-driven insights for improved tactics and strategies. However, research gaps exist, such as estimating player positions with moving cameras. This study aims to address this by developing an innovative method of analysing football matches using deep learning and computer vision techniques.

The results will contribute to football analytics, laying the groundwork for more accurate and efficient player tracking methods.

Embracing AI's potential will lead to a more engaging, fair, and data-driven future for football.

2 Objectives

AIM: The purpose of this research is to devise a novel way to track players and visualise them in a way that enhances the viewing experience as well as provide a means to collect data, analyse games and draw insights from the footages in real-time.

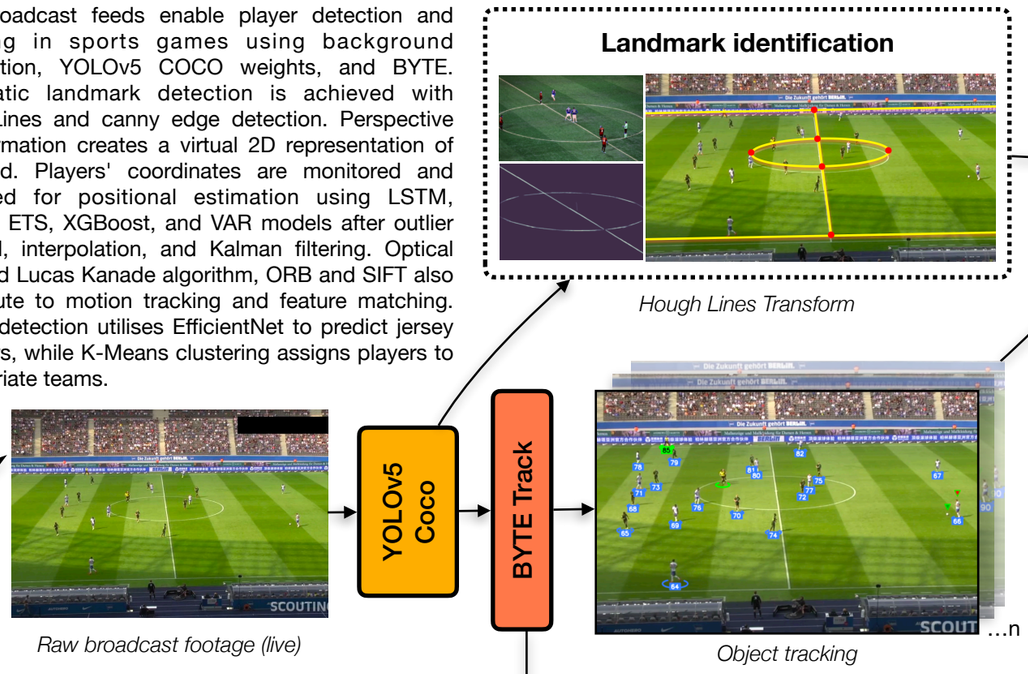
- **Detection of agents** } players, referees and ball are tracked
- **Tracking of agents** } in real-time with broadcast-feeds
- **Landmark detection** - corner, side-line, centre circle, etc.
- **Perspective transformation** - 2D plane with a top-down view
- **Statistical data collection** - record events, possessions, etc.
- **Future positional estimation** - Predict spatiotemporal location
- **Jersey number detection** - Identify player by number
- **Team assigner** - Detect the team of the players

Who benefits?: Football broadcasting companies, remote audiences and coaches & football clubs

Makesh Srinivasan (19BCE1717)
B.Tech Computer Science and Engineering
School of Computer Science and Engineering
Mentored by Dr. Premalatha M

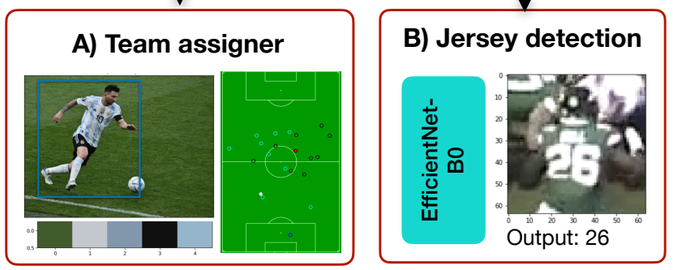
3 Methodology

Live broadcast feeds enable player detection and tracking in sports games using background subtraction, YOLOv5 COCO weights, and BYTE. Automatic landmark detection is achieved with HoughLines and canny edge detection. Perspective transformation creates a virtual 2D representation of the field. Players' coordinates are monitored and recorded for positional estimation using LSTM, ARIMA, ETS, XGBoost, and VAR models after outlier removal, interpolation, and Kalman filtering. Optical flow and Lucas Kanade algorithm, ORB and SIFT also contribute to motion tracking and feature matching. Jersey detection utilises EfficientNet to predict jersey numbers, while K-Means clustering assigns players to appropriate teams.



4 Additional experiments

Assigns players the teams based on dominant colour strategy (K-Means) and jersey detection uses the EfficientNet-B0 to classify jersey number



5 Results

LSTM's performances show that time series DL architecture with memory component makes better predictions than **XGB** or other than statistical forecasting models such as **VAR**, **ARIMA**, or **ETS***, allowing for off-camera positional estimations.

In conclusion, **BYTE** tracking and **YOLOv5** with custom coco weights achieve high performance, with over 90% confidence. **HoughLines** enables quick landmarking and **perspective transformation** for positional estimation. Team detection, jersey identification (using **EfficientNet**-B0, 0.85 accuracy), and **K-Means** algorithm for team assignment enhance visualisation. However, environmental factors like shadows and occlusions impact performance metrics.

Future steps include implementation of rolling predictions, experimental modules integration, and synthetic data generation for in-depth analytics. By transforming positional estimates and tracks, I can display them in 3D live footage. The goal is to develop an automated commentary system using action recognition, positional information, and natural language processing. Furthermore, we can train models to learn players' movements and strategies, which will benefit coaches, managers, and enhance AI engines in football video games like FIFA.

Performance metrics

		Confidence		
		Low	Avg	High
Object detection	COCO Std weights	0.4	0.6	0.7
	COCO custom weights	0.8	0.91	0.99
		Accuracy		
Jersey number detection	Sequential CNN	0.99		
	EfficientNet-B0	0.88		

* *Scan the QR code to learn more about my work*

Performance Measure