भारतीय विज्ञान संस्थान

SoccArt: Soccer Game Analysis

State of the Art Soccer Analysis Pipeline using Artificial Intelligence

Aditya Gupta*, Armaan Khetarpaul*, R K Shishir*, Sahil Chaudhary*, S. Sharath*, Umang Majumder*

Indian Institute of Science

Presented at 1st Conference on Applied AI and Scientific Machine Learning (CASML 2024)

Introduction

Automated soccer game analysis plays a crucial role in coaching, broadcasting, and fan engagement. However, existing solutions often lack integration. They rely on outdated techniques that fail to provide a complete, robust pipeline for tasks such as player tracking, team identification, and ball detection. To address these limitations, we present a state-of-the-art, end-to-end pipeline for soccer analysis. Our system combines YOLOv10 for detection, ByteTrack for tracking, and models like ResNet34 and ViTPose for jersey number recognition. Integrated with field-camera calibration and homography, it delivers accurate positional and identity data, enabling comprehensive game analysis in real-time.

Detection and Tracking

Video frames are processed using YOLOv10 to detect players, referees, and the ball. ByteTrack is employed for multi-object tracking, ensuring continuity of object identities across frames.

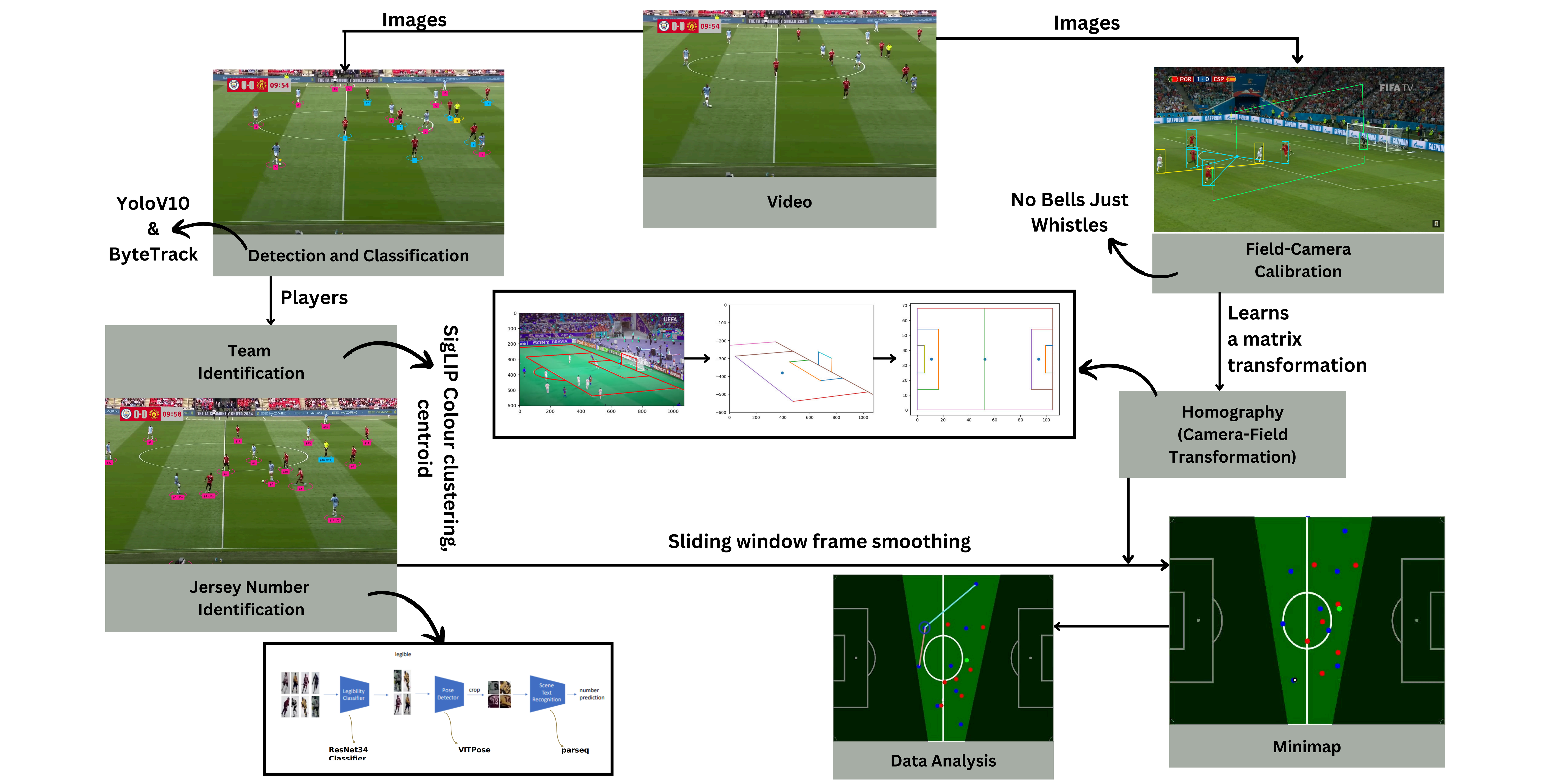
Team Identification

Detected players are assigned teams using SigLIP [1], which performs clustering of players based on their clothes. For goalkeepers, check their distance from the centroid of both teams.

Jersey Number Recognition

Player identification is achieved through jersey number recognition [2]. Models like ResNet34 and Vision Transformers for Pose Estimation (ViTPose) are used.

Pipeline



Calibration and Homography

Keypoints are detected on the field using models from No Bells Just Whistles [3] to map 3-D camera coordinates to a 2-D plane. A transformation matrix learned from this is used to perform homography on all detections. The resulting 2-D map of the field is used for analysis.

Interpolation and Smoothing

The ball's position is estimated through interpolation to account for detection inconsistencies. Additionally, a sliding window frame smoothing approach is applied to both ball and player trajectories to reduce noise and jitter in the final detection outputs.

Analysis

2-D map data is used to analyze player and ball movements throughout the game. Metrics like sprint patterns, ball possession and passing opportunities are calculated to provide tactical feedback. Team formations are visualised through voronoi diagrams.

Results

Player Detection AP Score (at IoU = 0.5)

Model	Average Precision (AP)
YOLOv5 (Roboflow)	0.810
YOLOv8 (PyResearch)	0.794
YOLOv10 (Ours)	**0.862**

Jersey No. Recognition Accuracy

Model	Accuracy
ZZPM	**92.85**
AIBrain Global Team	75.18
PARSeq & VitPose based (Ours)	79.31

Calibration Metric Score

Model	*acc@5*	*CR*	FS
SAIVA_Calibration	-	-	0.52
Sportlight	**0.766**	0.734	0.56
No-Bells-Just-Whistles	0.737	**0.775**	**0.57**

References

[1] X. Zhai, B. Mustafa, A. Kolesnikov, and L. Beyer, "Sigmoid loss for language image pre-training," September 2023.

[2] M. Koshkina and J. H. Elder, "A general framework for jersey number recognition in sports video," June 2024.

[3] M. Gutiérrez-Pérez and A. Agudo, "No bells just whistles: Sports field registration by leveraging geometric properties," June 2024.

Contact

adityapg@iisc.ac.in armaank@iisc.ac.in
shishirr@iisc.ac.in sahilc@iisc.ac.in
ssharath@iisc.ac.in umangm@iisc.ac.in