JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Football Match Analysis Using Yolo

Prof. S. M. Melasagare
Computer Engineering
SSPM College of Engineering (Mumbai University)
Kankavli, India

Jeet S. Shile
Computer Engineering
SSPM College of Engineering (Mumbai University)
Kankavli, India

Shashank J. Tipe
Computer Engineering
SSPM College of Engineering (Mumbai University)
Kankavli, India

Amrapali R. Ballal
Computer Engineering
SSPM College of Engineering (Mumbai University)
Kankavli, India

ABSTRACT

This project focuses on improving the analysis of football matches using advanced computer vision techniques. We are developing a model that will detect and track players, referees, and footballs in video footage with the **YOLO (You Only Look Once)** object detection model, which is well-known for its speed and accuracy. By leveraging YOLO features, multiple objects in complex scenes can be detected, which may form the basis for further analysis.

We will implement the **K-means clustering** algorithm for pixel segmentation. Using this method, we can segment players by their jersey color. The movement of objects within video frames will be tracked with the use of **optical flow** techniques to ensure accurate measurement of player movements. Our project will also determine the speed of each player and the distance covered throughout the match.

Index Terms: Yolo object detection , K-means pixel segmentation , Optical flow , Perspective Transformation , Ball possession, speed and distance covered.

I. INTRODUCTION

Artificial intelligence in today's world plays an important role in sports events. when it comes to analyzing athletic performance and pulling insights from the events as data and tech advance at speed, so do the teams and analysts using AI to understand players and sports on a deeper level [1].

Yolo (you look only once) is been used one of the latest object detection method. It used to identify

and track football players, refs, and even the ball during matches in real time. The Yolo is very fast and accurate at detecting objects which makes it great for keeping up with the action on the field. The goal was to take the standard YOLO model and retrain it with football-specific image datasets so it gets better at recognizing uniforms, players' faces, the refs and their signals, and of course the ball. That AI can pull from game footage for post-match analysis. More data equals more insights [2].

The Role of YOLO in Football Analysis:
The structure of YOLO allows to forecast several bounded boxes and class probabilities from one picture. The action happens very fast and a single frame often catches multiple points of interest from the dataset. YOLO is very fast enough to offer real-time analysis, giving teams the chance to make snap decisions in a game. Key moments - shots on goals, fouls, player positions, miss opportunities and errors in game play [3].

For YOLO to perform at its best, we should have or build a dataset featuring football situations from different leagues, weather types, and light conditions. This will help to train the model smart, enabling it to apply its knowledge across diverse settings effectively. The upgraded dataset will feature labelled video clips with players, referees, and the ball clearly marked. As training progresses, the model will be regularly scrutinized and adjusted for utmost detection precision. Our goal is to making a dependable tool for football analysis [4].

Team Categorization through Colour Analysis:
Important part of this work is to classify players based on their individual jersey colours. K-means clustering will handle this task. This autonomous learning method segments pixels and arranges colour data into obvious groups. Through this method, we can differentiate between opposing teams by studying the colour spread of jersey pixels in video frames [5].

The K-means process works by splitting the dataset into K clear clusters. Each cluster represents a certain colours range linked to a team's jersey colour. By allocating each player to a cluster based on their jersey colour, we can distinguish between opposing teams with ease. This method of distinguishing teams based on jersey colour is critical for determining essential game statistics like ball possession, allowing us to track which team is in control of the ball and for what duration during the match [6].

Addressing Camera Movement with Optical Flow:
Sports video analysis can be difficult due to camera movement, Such action can distort player tracking and object tracking. Football games are prime example where cameras shift and jerk to keep up with the play, These motions can break tracking sequences. Players can fall out of view, or their positions may not be read correctly. To combat this, we'll use methods known as 'optical flow' to gauge and correct any shifts in camera movement [7].

II. LITERATURE REVIEW

The study of literature review on football match analysis is to find research gaps and a unique examination of certain literature reviews.

A. Literature review

The study of literature on several research papers that have been done in a recent times on technologies such as yolo that is used in object detection which is used in large scale for sport events such as football, tennis, basketball, American football etc.

We have included a table that consist authors list, methods, result, conclusion and future scope.

Author and year	Methods	Results	Future scope
Cem Direkoglu, Melike Sah, Noel E. O'Connor, 2017 [1]	Focuses on detecting players in field sports using computer vision techniques.	Reports successful detection of players in various conditions, contributing to the understanding of player dynamics.	Proposes improvements in detection algorithms and their application across different sports.
Zhiyuan Yang, Yuanyuan Shen, Yanfei Shen, 2024 [2]	Develops a gesture recognition algorithm using YOLOv8s to identify referee gestures in football.	Achieves high recognition accuracy, enabling automated tracking of officiating decisions.	

Table continue here…

Author and year	Methods	Results	Future scope
Swetha Saseendran, Sathish Prasad Vetrivel Thanalakshmi, et al,2023 [3]	Analyzes player tracking data from football match feeds to derive performance insights.	Provides valuable metrics on player movements and positioning throughout matches.	Encourages the integration of AI for deeper insights and predictive analytics in player performance.
Mehran Rastegar Sani, 2023 [4]	Develops a tracking algorithm that addresses occlusion issues in soccer videos using advanced video processing techniques.	Achieves improved tracking accuracy under occlusion, facilitating better performance analysis.	Encourages further refinement of algorithms to handle complex scenarios in various sports settings.
Mohammad Mehedi Hassan, Stephen Karungaru, Kenji Terada 2024 [5]	Develops a robotic perception system to recognize player.	Demonstrates high accuracy in recognizing handball situations.	Suggests further research on expanding the system's application to various rule enforcement scenarios in sports.

Author and year	Methods	Results	Future scope
Wonjun Kim, Sung-Won Moon, Jiwon Lee, et al,2018 [6]	Introduces an adaptive multiscale sampling approach for tracking multiple players in soccer videos.	Achieves enhanced tracking accuracy and efficiency in complex video scenarios.	Encourages exploration of further adaptations for varying conditions in different sports.
Komorowski, Jacek et al., 2022 [7]	Graph-based multi-camera player tracking	Developed a comprehensive tracking system across multiple cameras	Future research may explore machine learning techniques for enhanced tracking performance.
Hongshan Liu, Colin Adreon, Noah Wagnon, et al, 2023 [8]	Utilizes a two-stage deep learning network for automated identification and indexing of players in videos.	Shows significant improvements in identification accuracy and indexing speed	Recommends further integration with live broadcasting technologies for real-time analysis.
Li, Guoqiang et al., 2020 [9]	New action recognition framework for video highlights summarization.	Improved highlight detection accuracy in sports videos	Future work could explore real-time applications and integration with broadcasting systems.
Keyan Zhao, 2024 [10]	Implements an upgradedYOLOv5 architecture for real-time detection of football and players using deep learning techniques.	Demonstrates significant improvements in detection accuracy and speed compared to previous models.	Explores applications in live game analytics and potential adaptations for different sports.

III. RESEARCH GAPS

1. Real-Time Multi-Object Detection and Tracking in Dynamic Environments: Existing systems often struggle with the real-time detection and tracking of multiple fast-moving objects, such as players and the ball, in football matches. High-speed movement, player occlusion, and changes in lighting and camera angles make it difficult for conventional systems to maintain consistent tracking. Although object detection models like **YOLO** have been widely adopted, their performance in the presence of these real-world challenges, particularly in sports analytics, is not thoroughly explored.

2. Accurate Player Team Classification: While **K-means clustering** allows for easy color-based classification, it can effectively distinguish between two teams based on their jerseys and kits.

3. Handling Camera Movement and Occlusion in Sports Analytics: One gap in research exists regarding how well systems can compensate for camera movement, such as panning or zooming, in football matches. **Optical flow techniques** are used to detect and account for camera movement, ensuring stable object tracking in dynamic environments.

4. Perspective Transformation for Accurate Movement Measurement: Current systems typically provide pixel-based tracking data, which does not reflect real-world metrics such as distance covered or speed in meters. **Perspective transformation** is used to convert pixel movements into real-world measurements, allowing for the accurate measurement of the distance covered by players—an essential metric for performance evaluation.

---

5. Automated Ball Possession Calculation and Performance Metrics: Although tracking algorithms exist, few studies focus on the automated calculation of team-specific ball possession and performance metrics in football. Manually calculating these metrics is time-consuming and prone to error. Current automated systems often fail to incorporate real-time ball possession tracking with high accuracy. The challenge of detecting who has possession of the ball during crowded or fast-moving play has not been adequately addressed in research.

.

IV. PROBLEM STATEMENTS AND OBJECTIVES.

A. Problem statement.

- Players often overlap or block each other, yolo may struggle to detect and correctly identify objects.
- Complex background such as advertisement , crowds and field lines can sometimes confuse the model.
- players of both teams have to differentiate based on jersey colour's.
- Need for compensating for camera movement between frames, and ensuring consistent identity tracking.

B. Objectives.

1.Automated Detection and Tracking of Objects: Develop a system that uses a YOLO-based deep learning model to automatically detect and track players, referees, and the football in real-time video footage of football matches.

2.Team Classification via K-means Clustering: Implement K-means clustering to classify players into teams based on the dominant colors of their t-shirts, enabling team-specific analysis such as ball possession and movement.

3.Tracking Player Movement and Ball Possession: Track the movement of each player and the ball across multiple frames, ensuring consistent identity tracking and the ability to calculate the percentage of ball possession for each team.

4.Compensating for Camera Movement with Optical Flow:} Use optical flow techniques to estimate and compensate for camera movement between frames, ensuring accurate tracking of players and the ball even during camera panning or zooming.

5.Perspective Transformation for Accurate Distance and Speed Measurements:

Apply perspective transformation to convert pixel-based movements into real-world distances (in meters), allowing for precise measurement of distance covered by players and their speeds.

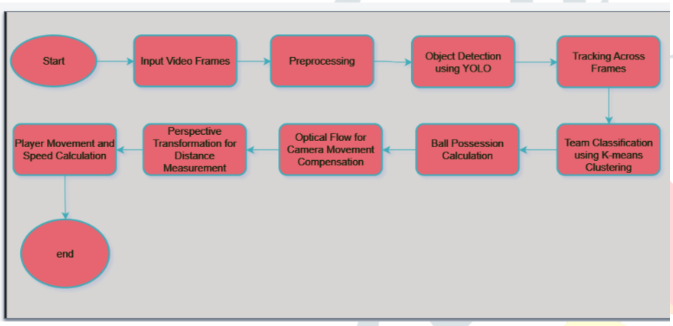
6.Real-Time Performance Analysis:

Provide real-time visualizations and statistical outputs for key performance indicators such as player speed, distance covered, and team ball possession, supporting post-match analysis and coaching decisions.

7.Evaluate System Performance:

Assess the accuracy of the detection and tracking system using evaluation metrics such as precision, recall, F1-score, Intersection over Union (IoU), and tracking accuracy, ensuring the system's robustness across different match scenarios and environmental conditions.

V. Proposed system



Block Diagram for Football Analysis System

A. Algorithm

1. Input Video Frames

- Read input video containing football match footage.
- Extract frames from the video at a consistent frame rate (e.g., 30 frames per second).

2. Preprocessing

- Resize each frame to the required resolution for the YOLO model (e.g., 416x416 pixels).
- Normalize pixel values for better performance of the YOLO object detection model.

3. Object Detection using YOLO

- For each frame, apply the YOLO object detection algorithm to detect:
  - Players
  - Referees
  - Football
- YOLO outputs bounding boxes, class labels, and confidence scores for each detected object.

4. Tracking Across Frames

- Track each detected object (players, referees, ball) across consecutive frames using bounding box data from YOLO.
- Apply a tracking algorithm (e.g., SORT or DeepSORT) to maintain object identities as they move across frames.

5. Team Classification using K-means Clustering

- Extract t-shirt color pixels from the bounding boxes of the detected players.
- Apply K-means clustering to segment players into two teams based on the dominant t-shirt colors.
- Assign team labels to players based on the clustering result.

6. Ball Possession Calculation

- Calculate ball possession percentage for each team by tracking which team's players have the ball for a given duration.
- Use the tracking data and team classification results to identify the team controlling the ball.

7. Optical Flow for Camera Movement Compensation

- Apply optical flow (e.g., Lucas-Kanade method) to estimate camera motion between consecutive frames.
- Compensate for camera movement to maintain accurate tracking of player and ball positions.

8. Perspective Transformation for Distance Measurement

- Apply perspective transformation to the detected player positions to account for depth and field perspective.
- Convert player movements from pixel coordinates to real-world measurements (e.g., meters).

9. Player Movement and Speed Calculation

- Calculate the distance traveled by each player over time based on their transformed coordinates.
- Estimate player speed using the displacement over time.

10. Output Results

- Visualize tracked objects (players, referees, ball) by drawing bounding boxes and tracking lines on video frames.

- Display calculated ball possession percentage for each team.
- Output the total distance covered by each player and their speed over time.

VI. METHODOLOGY

1. Data Collection

• Gather video datasets of football matches for both training and testing purposes.
The dataset should include annotated data for players, referees, and balls.
• Manually label bounding boxes and class labels (players, referees, ball) in training videos using annotation tools like LabelImg.

2. YOLO Model Training

• Train the YOLO model (e.g., YOLOv5 or YOLOv8) on the annotated football match dataset.
• Use transfer learning to fine-tune a pre-trained YOLO model if training from scratch is infeasible.
• Perform data augmentation (rotation, scaling, etc.) to enhance model robustness.

3. Team Classification using K-means

• For each detected player, extract the pixel values corresponding to their t-shirt color.
• Apply K-means clustering to classify players into two teams based on the dominant colors in their t-shirts.
• Assign a team label to each player for further analysis (e.g., ball possession calculation).

4. Tracking using DeepSORT

• After detecting players and the ball, track these objects across multiple frames.
• Use a tracking algorithm like DeepSORT, which combines the YOLO bounding box detections with appearance features to track objects even under occlusion or rapid movement.
• Maintain unique IDs for players, referees, and the ball throughout the video.

5. Optical Flow for Camera Movement

• Apply the Lucas-Kanade or Farneback method to estimate the motion of pixels between consecutive frames.
• Use the calculated optical flow vectors to adjust the positions of tracked objects and compensate for camera panning or zooming.

6. Perspective Transformation

• Detect field markings or reference points in the video to create a homography matrix.
• Apply perspective transformation to project pixel coordinates of players and the ball onto a real-world coordinate system.
• Convert distances in pixels to meters, allowing for the calculation of real-world distances traveled by each player.

7. Evaluation and Metrics

• Evaluate the detection and tracking performance using precision, recall, F1-score, and IoU.
• Evaluate team classification accuracy by comparing the predicted team labels with ground truth.
• Validate movement and speed measurements using ground truth data for player positions and speeds.
• Compare ball possession percentages calculated by the system with manually annotated data to validate its accuracy.

VII. EXPERIMENTAL SETUP

A. Details about input systems

• Video Footage: High-definition (HD) or 4K video clips of football matches are the main input. These videos must contain footage where players, referees, and footballs are visible from different camera angles. To account for various scenarios (e.g., night matches, outdoor lighting), videos will be chosen under diverse lighting and weather conditions.

• Frame Resolution: Input videos should ideally be in 1080p or higher resolution to ensure precise detection of small objects (such as the football) and to accurately differentiate between players. Each frame of the video will be extracted and processed by the YOLO model for object detection.

• Frame Rate: The frame rate should be consistent, typically 30 frames per second (fps) or higher. Higher frame rates provide better temporal resolution, which is essential for tracking fast-moving players and balls.

• Preprocessed Data: Before being fed into the model, the video frames will be preprocessed by normalizing the input data and resizing frames to fit YOLO's required input dimensions (e.g., 416x416 pixels). For team assignment, the color values of players' t-shirts will be extracted and clustered using K-means for segmentation.
All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

B. Performance Evaluation Parameters

• Tracking Accuracy: This metric evaluates how consistently the system can track a player's or ball's movement across multiple frames without losing the object. High tracking accuracy indicates that the system can maintain the identity of each object as it moves through the video, even during occlusions or rapid changes in position.

• Team Classification Accuracy: This evaluates the accuracy of the K-means clustering algorithm in correctly assigning players to their respective teams based on t-shirt color segmentation. The classification will be considered accurate if players are consistently grouped according to their team colors throughout the match.

• Precision: Precision focuses on the proportion of true positive predictions among all positive predictions, providing insight into the number of false positives. Precision is particularly important when false positives need to be minimized.

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} \quad (1)$$

• Recall (Sensitivity): Recall measures the proportion of true positive predictions among all actual positives. This metric is crucial when identifying every positive case is essential, even at the risk of more false positives.

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}} \quad (2)$$

• F1-Score: The F1-Score is the harmonic mean of Precision and Recall, providing a balanced evaluation of the model's performance. It is particularly useful when there is a trade-off between precision and recall.

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Or equivalently:

$$\text{F1} = 2 \times \frac{\text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \quad (4)$$

• Intersection over Union (IoU): IoU measures the overlap between the detected bounding box and the ground truth bounding box. This metric provides a clear indication of how well the detected object localization aligns with the actual object location. A higher IoU indicates better accuracy.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (5)$$

• Movement and Distance Measurement: Once perspective transformation is applied, the system's accuracy in calculating movement and distance will be validated using ground truth data (e.g., known distances or positions on the field). This validation will ensure that pixel-based measurements are correctly converted into real-world distances in meters.

• Optical Flow Stability: Optical flow will be evaluated based on its ability to accurately measure and compensate for camera motion. Misalignments or inaccuracies in motion estimation will be evaluated by comparing the system's output with known camera parameters and tracking error rates.

• Ball Possession Percentage: The percentage of ball possession for each team will be calculated using the tracking data, and its accuracy will be validated by comparing the results with manually annotated possession data from the video. This metric is crucial for assessing how much time each team controls the ball during a match.

VIII. SOFTWARE SETUP AND HARDWARE SETUP

A. Software Setup

- **YOLO (You Only Look Once):** The core object detection algorithm used for identifying and locating players, referees, and footballs. We will use YOLOv5 or YOLOv8, which are widely adopted versions with better performance than earlier versions.
- **OpenCV:** An open-source computer vision library that will be used for image and video processing tasks such as frame extraction, optical flow computation, and perspective transformation.
- **TensorFlow or PyTorch:** These deep learning frameworks will be used for training and fine-tuning the YOLO model on custom football datasets.
- **K-means Clustering (Scikit-learn):** Scikit-learn will be used to implement K-means clustering for color-based segmentation of players' t-shirts. The extracted pixel data from each player's t-shirt will be fed into the algorithm to classify players into teams.
- **Perspective Transformation (OpenCV):** This will be implemented using OpenCV to correct for distortion in video frames and convert pixel distances into real-world measurements.
- **Optical Flow (OpenCV):** The Lucas-Kanade or Farneback methods for optical flow will be used to detect camera movement and track player movements between frames.
- **CUDA and cuDNN:** GPU acceleration using CUDA (Compute Unified Device Architecture) and cuDNN (NVIDIA's deep neural network library) will be enabled to accelerate the deep learning computations in YOLO and speed up model inference.
- **Annotation Tools:** Tools such as LabelImg will be used for annotating the training datasets by manually labeling players, referees, and footballs in video frames.

B. Hardware Setup

- **Hardware:** The computations were performed on a machine equipped with [specify hardware details, e.g., CPU, RAM, GPU if applicable], ensuring sufficient computational power for model training, validation, and evaluation. The hardware configuration supported efficient data processing and model optimization.

IX. CONCLUSION

In conclusion, this football match analysis project demonstrates how advanced computer vision techniques, such as YOLO for real-time tracking, can significantly enhance our understanding of the game. By accurately monitoring player movements and team strategies, the system provides valuable insights for coaches and analysts to make informed decisions. The growing impact of data in football opens the door to more strategic approaches to match analysis, fostering further development in the field.

X. ACKNOWLEDGEMENT

We sincerely acknowledge with a deep sense of gratitude to Project Guide Prof. S. M. Melasagare and Project Coordinator Dr. N. M. Shivsharan for their valuable guidance, genuine suggestions, and constant encouragement during the preparation of project synopsis work without which completion of this task would be a difficult task.

We are also thankful to all of our faculty members of Computer Engineering Department especially our head of department Prof. D. P. Mhapasekar and our respected principal Dr. D. S. Badkar who give us an idea of significant cooperation during the completion of this work.We are immensely grateful to all who were involved in this project work because, without their cooperation, inspiration, constant promotion, and useful suggestions, it would be impossible to complete this task and synopsis report within this allotted time.

XI. REFERRENCES

[1] Cem Direkoglu1 · Melike Sah2 · Noel E. O'Connor3. Player detection in field sports. Machine Vision and Applications, 2017. accessed: 5 December, 2017.

[2] Yuanyuan Shen* Zhiyuan Yang and Yanfei Shen. Football referee gesture recognition algorithm based on yolov8s. Frontiers in Computational Neuroscience, 2024. accessed: 19 February, 2024.

[3] Swetha PRABAKARAN Priyadharsini RAVISANKAR Swetha SASEENDRAN, Sathish Prasad Vetrivel THANALAKSHMI. Analysis of player tracking data extracted from foot-ball match feed. Romanian Journal of Information Technology and Automatic Control, 2023. accessed: 2 November, 2023.

[4] Mehran Rastegar Sani1. Multi object tracking in soccer video focusing on occlusion de-tection and resolving. Multimedia Tools and Applications, 2023. accessed: 5 February,2023.

[5] Stephen Karungaru Mohammad Mehedi Hassan and Kenji Terada. Robotics perception: Intention recognition to determine the handball occurrence during a football or soccer match. MDPI, 2024. accessed: 30 April, 2024.

[6] Wonjun Kim1 · Sung-Won Moon2 · Jiwon Lee2 · Do-Won Nam2 · Chanho Jung3. Multiple player tracking in soccer videos: an adaptive multiscale sampling approach. Machine Vision and Applications, 2018. accessed: 12 February, 2018.

[7] Jacek Komorowski and Grzegorz Kurzejamski. Graph-based multi-camera soccer player tracker. Institute of Electrical and Electronics Engineers (IEEE), 2022. accessed: 28 Sep, 2024.

[8] NoahWagnon2 Abdul Latif Bamba3 Xueshen Li1 Huapu Liu4 Steven MacCall4 YuGan1* Hongshan Liu1, ColinAdreon2. Automated player identifcation and indexing using two stage deep learning network. Nature Portfolio, 2023. accessed: 5 December, 2023.

[9] Guoqiang Li, Xin Li, and Cheng Yan. A new action recognition framework for video highlights summarization in sporting events. 2020. accessed: 28 Sep, 2024.

[10] Keyan Zhao1. Enhancing the performance and accuracy in real-time football and player detection using upgraded yolov5 architecture. International Journal of Computational Intelligence Systems, 2024. accessed: 10 June, 2024.