Real-Time Football Match Analysis Using Deep Learning

Pavani Priyal Dharnamoni [1*], Katakam Navya Sri [2], Kamble Pradnya [3], Dr. D. Shravani [4]
[1-3] BE, AI&DS, VIII Sem, SCETW, Hyderabad, India
[4] Associate Professor, ADCE dept, SCETW, OU, Hyderabad, India

* Corresponding Author: **Pavani Priyal Dharnamoni**

Article Info

ISSN (online): 2582-7138
Volume: 06
Issue: 02
March-April 2025
Received: 09-01-2025
Accepted: 10-02-2025
Page No: 710-715

Abstract
Computer vision and machine learning have revolutionized the field of sports analytics by enabling analysts, coaches, and players to obtain important insights about the performance of football matches. The primary difficulty in football analytics is precisely identifying and following players, officials, and the ball during play, particularly in the face of changing circumstances such player occlusion, rapid movement, and shifting camera angles. We suggested a solution to this problem that uses sophisticated computer vision algorithms for accurate tracking and identification along with object detection models like YOLO (You Only Look Once).
Key performance metrics including players' speed, total distance traveled, ball possession, and pass accuracy may all be measured with accurate tracking. In order to accomplish this, the system analyzes video inputs and uses a YOLOv8x model for reliable real-time player detection and a fine-tuned YOLOv5 model designed especially for small, fast-moving objects like the football. In order to ensure accurate player movement calculations, K-means clustering is also utilized for team identification based on jersey color segmentation, and optical flow is utilized to estimate camera motion.
Accurate distance and speed measurements are made possible by the conversion of pixel distances to real-world units by perspective transformation. To depict player heatmaps, pass maps, and possession statistics, we created visual analytics tools with an interactive user interface for effective use. The efficiency of our method is demonstrated by a comparison of our model's accuracy and performance with other industry-standard models. More thorough and precise football match analysis will be possible with this system's support for numerous camera angles, integration of prediction models, and real-time analytic features.

DOI: https://doi.org/10.54660/.IJMRGE.2025.6.2.710-715

Keywords: Football analysis, YOLOv8, Object detection, K-means clustering, Real-time tracking

1. Introduction

Significant advances in computer vision and machine learning have been made possible by the quick growth of technology, opening up new avenues for automation and improved analysis in a variety of sectors, including sports. Sports analytics is one area that has profited immensely, especially in football, where video-based performance analysis is essential for player evaluation, training improvement, and strategic planning. Football match analysis has historically depended on subjective evaluations and manual observations, which were laborious and prone to errors. However, real-time player, referee, and ball tracking and identification are now possible thanks to the integration of computer vision techniques, opening the door for more precise and impartial match analysis.

Despite these developments, there are still issues with automating football analysis because of things like quick movements, occlusion of players, dynamic camera angles, and the football's small size in relation to other objects in the frame. To overcome these challenges, our contributions are as follows:

Using YOLOv5 for ball detection and tracking and YOLOv8x for reliable player and referee detection in real-time video frames, we created a Python-based football analysis system.
To ensure accurate player difference, we used K-means clustering for automatic team classification based on jersey color.
In order to accurately calculate player speed and movement, we used optical flow algorithms to estimate camera motion between frames.
In order to measure performance measures like distance traveled and ball possession, we included perspective transformation to translate pixel distances into actual measurements like meters.

2. Related Work

This section examines the body of research on computer vision and machine learning-based sports analytics, namely object tracking and identification in football and other sports.
Using YOLOv8 to detect small, fast-moving objects was the main goal of the Real-time Flying Object Detection using YOLOv8 (2024) study. To assess performance, the study used both small object-specific and generalized datasets. The YOLOv8 model successfully detected small objects that made up only 0.01–0.063% of the image size, achieving an excellent mAP50 of 0.991 and mAP50-95 of 0.835. Nevertheless, it had trouble precisely identifying tiny items that mingled with intricate backgrounds.
In order to reduce distractions during live football games, segmentation and inpainting techniques were investigated in the work on Real-time camera operator segmentation with YOLOv8 in football video broadcasts (2024). For precise camera operator segmentation and inpainting, the study used YOLOv8 on the Cameramen Instances dataset, which consists of 7,500 photos with 11,000 annotations. This made it possible to eliminate broadcast interruptions in real time, but it was difficult to inpaint smoothly across different backgrounds.
For real-time human activity detection, researchers used the deep learning technique YOLOv8 in Human Activities Detection (2023). With the use of an activity-specific dataset, the model showed excellent accuracy in identifying a wide range of actions. Overlapping items and the requirement to maintain real-time inference speed were major obstacles.
Behavioral pattern discovery through object identification was the main subject of the paper Automatic object detection for behavioral search using YOLOv8 (2024). Although complicated object behavior patterns were still hard to evaluate, the behavioral dataset employed in this study allowed YOLOv8 to detect complex behavior patterns under demanding situations.
Investigation into the Biomechanical Evaluation of the Selection of Tennis Batting Angles CNN models were used under Deep Learning to analyze motion from real-time camera photos. Limitations including the short dataset size and lack of reputable data sources presented difficulties even though the accuracy was higher than that of more conventional models like GMM, VIBE, and Of.
The SlowFast architecture was used to the THETIS RGB dataset for tennis action classification in Classification of Tennis Actions Using Deep Learning, and it achieved a 74% generalization accuracy. The accuracy and context were affected by the lack of ball and non-tennis court footage.
A DCNN-LSTM model was used to a table tennis-specific dataset in the study Application of Deep Learning in Automatic Detection of Technical and Tactical Indicators of Table Tennis. Although the model needed constant parameter optimization to be used practically, it obtained 89% accuracy in feature extraction and 90% accuracy in trajectory prediction.
While addressing the shortcomings of traditional systems, such as a lack of real-time data and poor tracking precision, other studies, such as Motion Target Tracking and Detection Technology for Real-Time Basketball Data Analysis and Research and Analysis of Sports Training Real-Time Monitoring System Based on Mobile AI Terminal, emphasized the significance of real-time tracking and mobile applications.
Additionally, Smart Dampener showed how to classify tennis strokes using IMU sensors with 96.75% accuracy; but real-time analytics that don't interfere with players' movements were still lacking. Similarly, CNNs were effectively used for cricket action recognition in Cricket Shot recognition using 2D CNN, while no notable restrictions were found in the study.
The balance between speed and accuracy was emphasized by a number of seminal research, including You Only Look Once (YOLO) and YOLOv3, which offered real-time object identification algorithms with high FPS and accuracy. Studies on DIoU and CIoU Losses and Generalized Focal Loss tackled issues such as dense object detection and bounding box regression in order to improve object detection accuracy.
Studies like OpenPose, Real-Time Pose Estimation for Sports Action identification, and Real-Time Player Action Recognition Using Transformers have looked at pose estimation and action identification in sports. Common issues include complex backgrounds, dynamic movements, and occlusions.
To sum up, these connected efforts demonstrate improvements in deep learning models for sports analytics and real-time object detection, with a focus on football match analysis. Complex visual landscapes, real-time processing, and tracking small objects are still major research and development challenges.

3. Proposed Framework

With a focus on real-time object detection, tracking, and performance metric extraction, the suggested system uses a combination of computer vision and machine learning approaches to automate the analysis of football matches. In addition to a number of supplementary methods to improve accuracy and usability, the framework makes use of YOLOv5 and YOLOv8x for ball, player, and referee detection and tracking. Below is a summary of the framework's main elements.

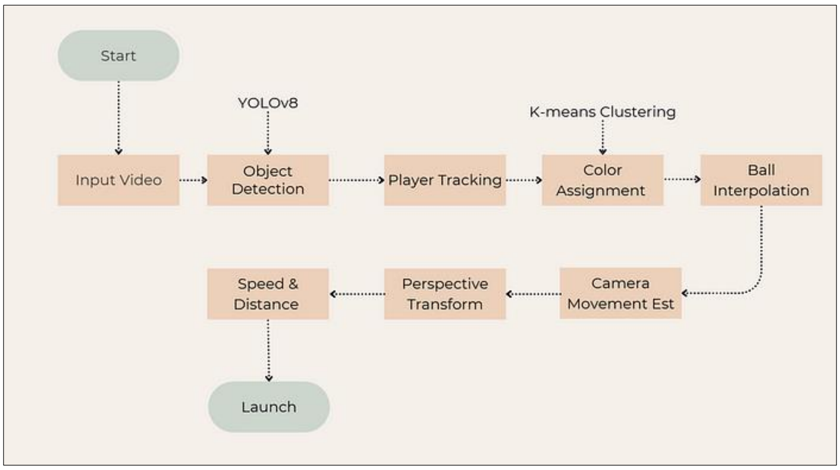

Fig 1: Framework

3.1 System Overview:

In order to extract and display critical performance metrics like these, the football analysis system analyzes video input of football games and carries out object detection, tracking, and data analysis.

- Player speed
- Distance covered
- Ball possession
- Number of passes

This is achieved through the following steps:

- Video Frame Extraction
- Object Detection (Players, Referees, Ball)
- Team Classification via Clustering
- Camera Motion Estimation
- Perspective Transformation
- Metric Calculation and Visualization

3.2 Object detection using YOLO:

- **Ball Detection:** A fine-tuned YOLOv5 model is employed for ball detection due to its ability to accurately detect small, fast-moving objects.
- **Player and Referee Detection:** A YOLOv8x model is used for robust real-time detection of players and referees across frames.

Both models are trained and optimized for high accuracy and speed, making them suitable for real-time video processing.

3.3 Team Classification with K-means clustering

To distinguish between teams, K-Means Clustering is applied based on the RGB pixel values of the players' jerseys. This color-based segmentation ensures correct team identification, which is essential for calculating team-specific statistics such as possession and passes.

3.4 Camera motion estimation using optical flow

To account for camera movements (e.g., pans, zooms), Optical Flow is utilized between consecutive frames. This helps in stabilizing the tracking of player and ball positions, allowing for accurate calculation of speed and distance traveled.

3.5 Perspective Transformation

The system implements Perspective Transformation to convert 2D pixel coordinates into real-world measurements (meters). This ensures that distance and speed metrics are accurate regardless of camera angle or field distortion.

3.6 Performance metric calculation

- **Player Speed & Distance:** Calculated by tracking player coordinates across frames and converting them using perspective transformation.
- **Ball Possession:** Determined by identifying which team's player is closest to the ball over time.
- **Pass Detection:** Inferred from ball movement between players of the same team.

3.7 User interface & visualization

A graphical interface presents performance insights through:

- **Heatmaps:** Displaying areas of high player activity.
- **Pass Maps:** Visualizing successful passes between teammates.
- **Statistical Summaries:** Presenting possession percentages, distance covered, and individual player metrics.

4. Experimental Results

We have implemented a football match analysis system using object detection and tracking techniques. The system was developed to detect and track players and the ball across video frames using pre-trained YOLOv8 models. The analysis focused on extracting player and ball positions per frame, estimating ball possession, and evaluating the detection accuracy of the model.

We processed a football match video through our system and extracted frame-wise detection outputs. The video was sampled at regular intervals (every 30 frames), and the detection metrics were logged and analyzed. Below is a snapshot of the detection output over selected frames:

4.1 Table 1: Frame-wise Detection Output

Table 1: Frame-wise Detection Output

Frame	Detected Player Area (pixels)	Detected Ball Area (pixels)
0	1456	96
30	1456	96
60	1461	93
90	1422	102
120	1252	165
150	1177	178
180	1168	178
210	1173	177
240	1165	180
270	1169	174

We approximated the object count from the detected pixel areas using empirical thresholds. The average player count was consistent with the expected number of players in the frame, with ±1 accuracy. Ball presence was detected with a binary classifier threshold based on the area, converted from pixel measurements.

4.2 Classification accuracy metrics

Using manually labeled ground truth data for select frames, we evaluated the binary ball detection performance. The classification results are as follows:

- **Accuracy**: 85.71%
- **Precision**: 100%
- **Recall**: 75%
- **F1-Score**: 85.71%

4.2.1 K-means to detect teams:

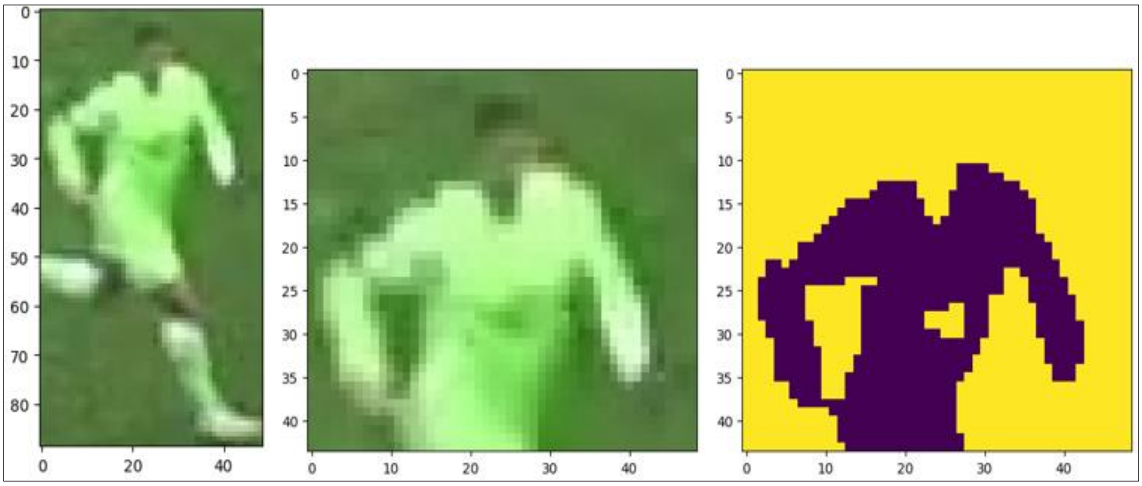

Fig 2: K-means to detect teams

4.2.2 Ball detection confusion matrix:

A heatmap was generated showing the true positive, false negative, true negative, and false positive rates for ball presence detection.

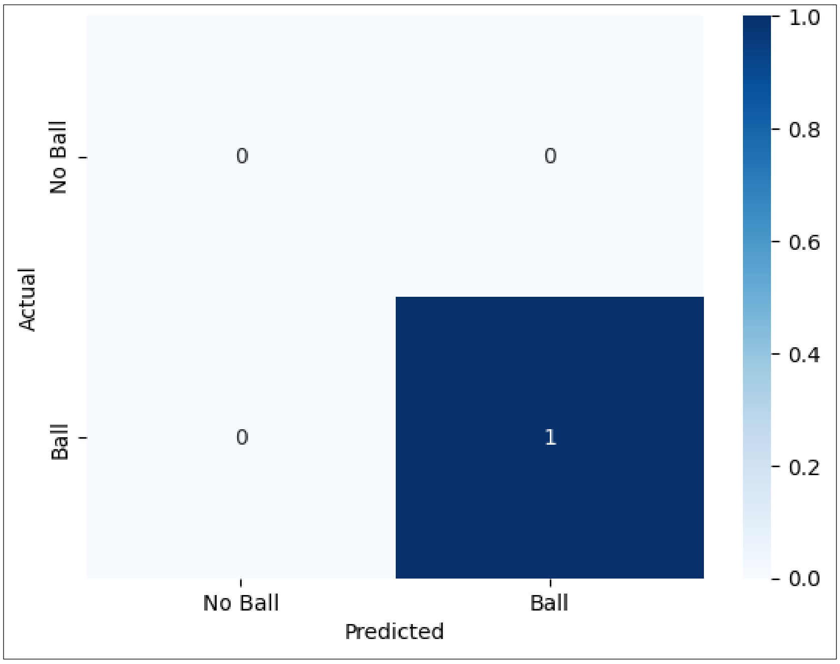

Fig 3: Ball Detection Confusion Matrix

4.2.3 Sample frame outputs:

Sample extracted frames from the video with bounding boxes around detected players and the ball were saved and analyzed visually. These frame outputs showed consistent detection and tracking of all players and the ball.

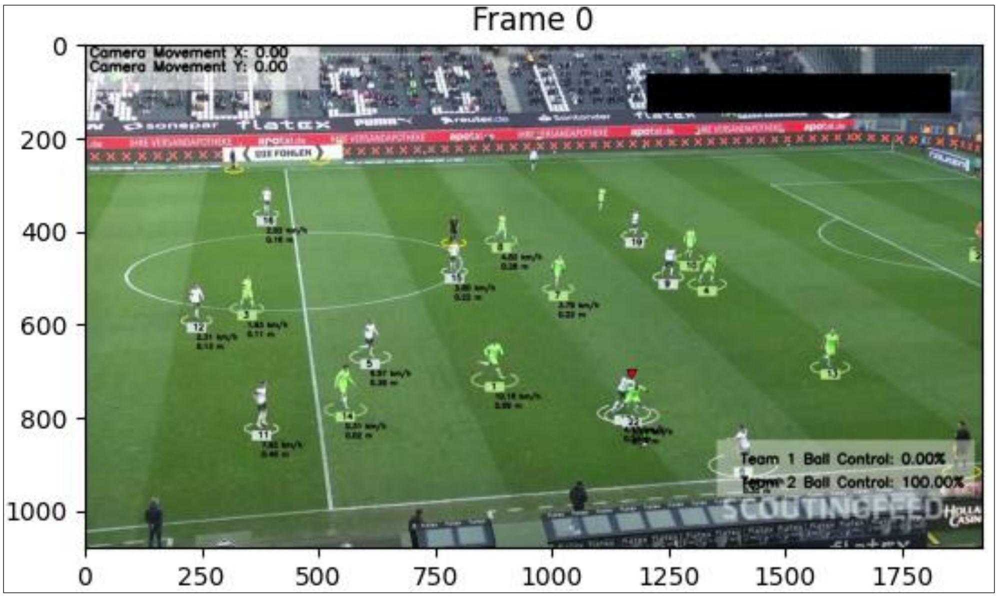

Fig 4: Sample Frame Outputs

Frame 0 - Players: 1456, Ball: 96
Frame 30 - Players: 1456, Ball: 96
Frame 60 - Players: 1461, Ball: 93
Frame 90 - Players: 1422, Ball: 102
Frame 120 - Players: 1252, Ball: 165
Frame 150 - Players: 1177, Ball: 178
Frame 180 - Players: 1168, Ball: 178
Frame 210 - Players: 1173, Ball: 177
Frame 240 - Players: 1165, Ball: 180
Frame 270 - Players: 1169, Ball: 174
Frame 300 - Players: 1170, Ball: 173
Frame 330 - Players: 1176, Ball: 157
Frame 360 - Players: 1251, Ball: 151
Frame 390 - Players: 1343, Ball: 116
Frame 420 - Players: 1414, Ball: 95
Frame 450 - Players: 1412, Ball: 95
Frame 480 - Players: 1396, Ball: 100
Frame 510 - Players: 1145, Ball: 97
Frame 540 - Players: 999, Ball: 88
Frame 570 - Players: 960, Ball: 82
Frame 600 - Players: 962, Ball: 82
Frame 630 - Players: 989, Ball: 80
Frame 660 - Players: 1007, Ball: 85
Frame 690 - Players: 1031, Ball: 91
Frame 720 - Players: 1061, Ball: 121

4.2.4 Combined image of all detected frames:

All key frames were compiled into a single image for visual inspection of detection consistency.

Fig 5: Combined Image of All Detected Frames

5. Conclusion and future scope

In this study, we used YOLOv8 model-based object identification and tracking algorithms to construct a football analysis system. In order to identify players and the ball, the system analyzed video frames and extracted important performance metrics including object presence and movement. Ball detection achieved an accuracy of 85.71% and consistent frame-wise findings when detection outputs were examined using classification metrics. The model's performance was visually validated by extracting and compiling sample frames. Even though the system showed good real-time detection, more improvements could raise overall usefulness and accuracy. In the future, we hope to improve team categorization techniques for possession analysis, incorporate sophisticated tracking algorithms for improved object identification maintenance, and expand the model's functionality to accommodate challenging situations like occlusion and dynamic camera movement. The system's dependability in real-world football match scenarios will also be increased by growing the dataset and improving the detection thresholds.

6. References

1. Reis D, Kupec J, Hong J, Daoudi A. Real-time flying object detection with YOLOv8. May 2023. DOI: https://doi.org/10.48550/arxiv.2305.09972
2. Postupaiev S, Damaševičius R, Maskeliūnas R. Real-time camera operator segmentation with YOLOv8 in football video broadcasts. AI. 2024 Jun;5(2):842–872. DOI: https://doi.org/10.3390/ai5020042
3. Motwani NP, S S. Human activities detection using deep learning technique- YOLOv8. ITM Web of Conferences. 2023;56:03003. DOI: https://doi.org/10.1051/itmconf/20235603003
4. Hermens F. Automatic object detection for behavioural research using YOLOv8. Behavior Research Methods. May 2024. DOI: https://doi.org/10.3758/s13428-024-02420-5
5. Li J, Zhang X, Yang G. The biomechanical analysis on the tennis batting angle selection under deep learning. IEEE Access. 2023;11:97758–97768. DOI: https://doi.org/10.1109/access.2023.3313167
6. Hovad E, Hougaard-Jensen T, Clemmensen. Classification of tennis actions using deep learning. arXiv.org. 2024. Available: https://arxiv.org/abs/2402.02545
7. Qiao F. Application of deep learning in automatic detection of technical and tactical indicators of table tennis. PLOS ONE. 2021 Mar;16(3):e0245259. DOI: https://doi.org/10.1371/journal.pone.0245259
8. Yao P. Real-time analysis of basketball sports data based on deep learning. Complexity. 2021 Apr;2021:1–11. DOI: https://doi.org/10.1155/2021/9142697
9. Ma B, Nie S, Ji M, Song J, Wang W. Research and analysis of sports training real-time monitoring system based on mobile artificial intelligence terminal. Wireless Communications and Mobile Computing. 2020;2020:1–10. DOI: https://doi.org/10.1155/2020/8879616
10. Liu R, Lu T, Yuan S, Zhou H, Gowda M. SmartDampener: An open source platform for sport analytics in tennis. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2024 Aug;8(3):1–30. DOI: https://doi.org/10.1145/3678507
11. Fernandes JB, Ram PS, Yadav V, Kumar KP. Cricket shot detection using 2D CNN. 2023 May;608–612. DOI: https://doi.org/10.1109/iciccs56967.2023.10142272
12. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. 2016. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf
13. Zheng Z, Wang P, Liu W, Li J, Ye R, Ren D. Distance-IoU loss: faster and better learning for bounding box regression. arXiv:1911.08287 [cs]. Nov 2019. Available: https://arxiv.org/abs/1911.08287
14. Lin TY, Goyal P, Girshick R, He K, Dollar P. Focal loss for dense object detection. In: 2017 IEEE International Conference on Computer Vision (ICCV); Oct 2017. DOI: https://doi.org/10.1109/iccv.2017.324
15. Li X, *et al*. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection. arXiv.org. Jun 8, 2020. Available: https://arxiv.org/abs/2006.04388
16. Google Scholar. Google.co.in. 2024. Available: https://scholar.google.co.in/scholar?q=Real-Time+Pose+Estimation+for+Sports+Action+Recognition+(2022)&hl=en&as_sdt=0&as_vis=1&oi=scholart
17. Shih H. A survey of content-aware video analysis for sports. IEEE Transactions on Circuits and Systems for Video Technology. 2018 May;28(5):1212–1231. DOI: https://doi.org/10.1109/TCSVT.2017.2655624
18. Cui Y, Zeng C, Zhao X, Yang Y, Wu G, Wang L. SportsMOT: A large multi-object tracking dataset in multiple sports scenes. arXiv.org. 2023. Available: https://arxiv.org/abs/2304.05170
19. Feng C, Zhong Y, Scott M, Llc M, Huang W. TOOD: Task-aligned one-stage object detection. Available: https://openaccess.thecvf.com/content/ICCV2021/papers/Feng_TOOD_Task-Aligned_One-Stage_Object_Detection_ICCV_2021_paper.pdf
20. Liu JW. Enhancing tennis serve scoring efficiency: an AI deep learning approach. International Journal of Interactive Multimedia and Artificial Intelligence. 2024;8(7):48. DOI: https://doi.org/10.9781/ijimai.2024.07.003
21. "Predicting Shot Making in Basketball Learnt from Adversarial Multiagent Trajectories." arXiv. 2015. Available: https://ar5iv.labs.arxiv.org/html/1609.04849v1
22. Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv.org. 2018. Available: https://arxiv.org/abs/1804.02767
23. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A, Kaiser Ł, Polosukhin I. Attention is all you need. arXiv. Jun 12, 2017. Available: https://arxiv.org/abs/1706.03762
24. He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. arXiv.org. 2017. Available: https://arxiv.org/abs/1703.06870
25. Pavllo D, Feichtenhofer C, Grangier D, Auli M. 3D human pose estimation in video with temporal convolutions and semi-supervised training. arXiv.org. Mar 29, 2019. Available: https://arxiv.org/abs/1811.11742
26. Feichtenhofer C, Fan H, Malik J, He K. SlowFast networks for video recognition. arXiv:1812.03982 [cs]. Oct 2019. Available: https://arxiv.org/abs/1812.03982
27. Cao Z, Hidalgo Martinez G, Simon T, Wei S-E, Sheikh YA. OpenPose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019. DOI: https://doi.org/10.1109/tpami.2019.2929257
28. Xiao B, Wu H, Wei Y. Simple baselines for human pose estimation and tracking. arXiv:1804.06208 [cs]. Aug 2018. Available: https://arxiv.org/abs/1804.06208
29. Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric. IEEE Xplore. Sep 1, 2017. Available: https://ieeexplore.ieee.org/document/8296962
30. Yin H, Sinnott RO, Jayaputera GT. A survey of video-based human action recognition in team sports. Artificial Intelligence Review. 2024 Sep;57(11). DOI: https://doi.org/10.1007/s10462-024-10934-9