IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

A Study Of Different Techniques Used In Sports Video Analysis

Manaal Waheed Khan,Maziya Sameen, Srijaanshu Chintala
Department Of Artificial Intelligence and Data Science and Computer Engineering
Stanley College Of Engineering And Technology For Women, Hyderabad, India.

ABSTRACT: Video analysis in sports has emerged as an essential resource for improving performance and tactics in different athletic domains. This study examines multiple algorithms and techniques applied in sports video analysis to identify patterns and strategies in video footage. We employ approaches such as tracking players and objects, offering crucial information to comprehend the dynamics of the game. The progress of computer assistance in sports is transformative. Our goal is to elucidate how these methods aid in player training, strategy formulation, and the evaluation of various techniques in regard to their advantages and disadvantages.

Keywords: Video Analysis, Real-time Tracking,Tactics Recognition, Highlight Extraction, Tracking Techniques, Game Performance.

1.Introduction

In recent years, the transformation of sports video analysis has been significant, driven by advancing technologies such as computer vision, machine learning, and artificial intelligence. These developments are transforming the ways athletes prepare, how teams formulate strategies, and how fans engage with the game. This review seeks to explore the different techniques employed in analyzing sports video, emphasizing aspects such as tactical play classification, player and ball tracking, highlight extraction, content insertion, and recognition of significant landmarks.The realm of sports is exceedingly intricate. Certain games, such as tennis and cricket, exhibit structured patterns, whereas others, including soccer and basketball, tend to be more fluid and unpredictable. This intricacy necessitates analytical methods that can match the rapid tempo of these sports. This review analyzes important studies in the area, focusing on how researchers categorize play sequences and detect emerging patterns that can enhance coaching techniques and training initiatives. For example, the US Tennis Association has recognized 58 successful patterns that can be essential for players aiming to enhance their performance.

Monitoring athletes and balls is another crucial aspect of sports analysis. Methods such as color clustering and trajectory-focused optimization have enhanced our ability to track movement on the field. The implementation of multi-camera setups has elevated this, enabling 3D visualizations that offer more profound understanding of gameplay mechanics.Nonetheless, the path of sports video analysis carries its own set of difficulties. Concerns such as tracking precision and the creation of computer-aided referee systems all need focus. This review will examine these challenges while also honoring the advancements that have been achieved.We will examine the current methodologies in use, identify gaps in our comprehension, and propose avenues for future research that may improve our analysis and appreciation of sports.

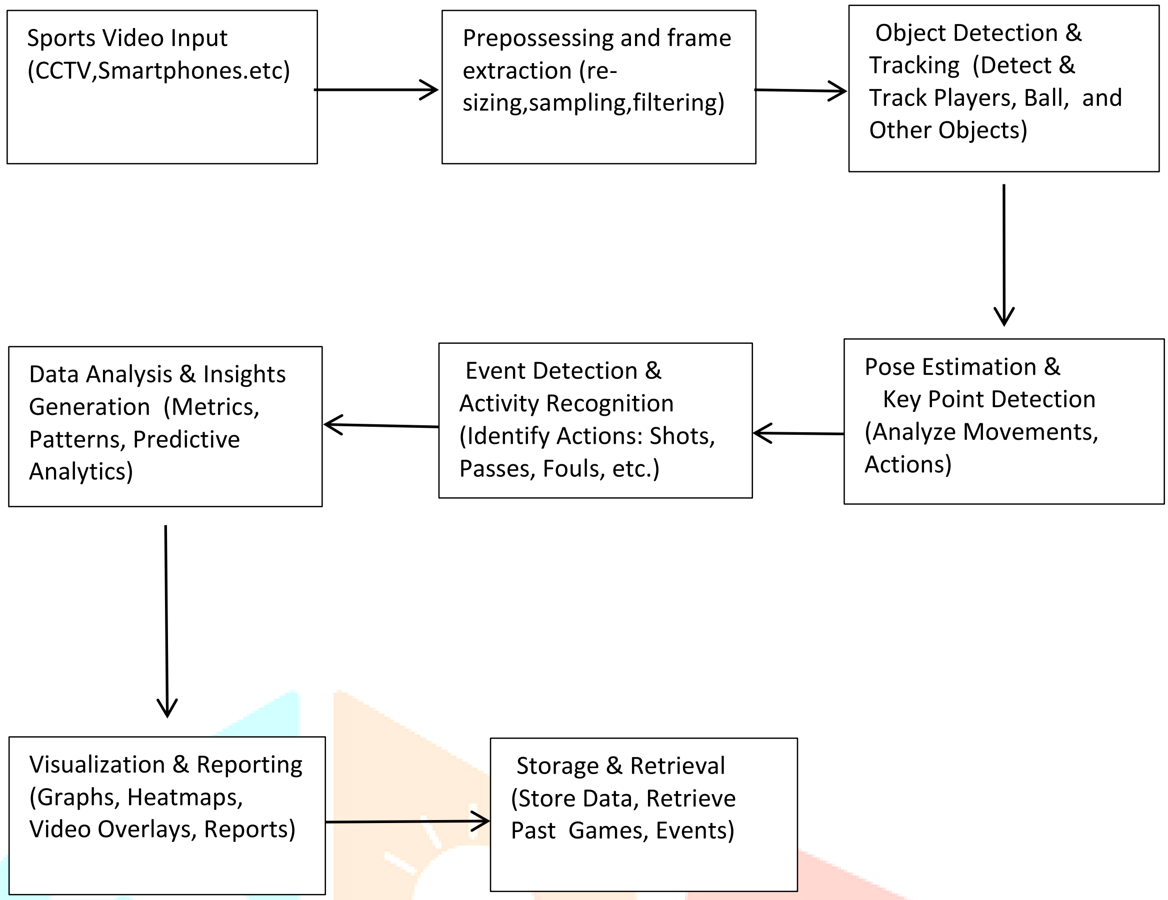

Fig : Process Diagram for Sports Video Analysis

This diagram illustrates a sports video analytics system that processes game footage to provide insights on player and game performance. The system extracts frames, detects and tracks objects (like players and the ball), and identifies key actions such as passes and goals. Advanced models analyze player movements and generate performance metrics, which are visualized in reports or overlays on video. The system combines computer vision and machine learning to deliver actionable insights and predictive analytics for players, coaches, and fans.

2.Applications in Sports Video Analysis

*1)Performance Evaluation:*Coaches and athletes utilize video analysis to examine game recordings, evaluate player performance, and pinpoint strengths and weaknesses. This input is essential for tactical planning and player growth.

2) Tactical Strategy Formulation: By examining gameplay trends and tactics, teams can enhance their strategies. Coaches can analyze rival tactics to create counter-strategies, boosting competitive edge.

3)Injury Prevention: Video analysis aids in recognizing movement patterns that could result in injuries. By tracking player biomechanics, coaches can develop specialized training programs to minimize injury risks.

4) Coaching and Skill Enhancement: Athletes can obtain tailored feedback through video analysis, enabling them to focus on particular skills and refine their techniques. Tools for motion analysis can assist in comprehending and rectifying technique.

5)Referee Assistance and Decision-Making: Tools such as Hawk-Eye and VAR (Video Assistant Referee) systems offer referees immediate video feedback, enhancing the precision and equity of officiating decisions.

6)Fan Engagement and Broadcasting: Highlight extraction algorithms generate captivating content for broadcasts, enabling viewers to witness significant moments as they happen. Improved viewing experiences, like virtual replays and analyses, maintain fan engagement.

7) Game Highlights and Summarization: Automated highlight creation assists in producing shortened versions of games for fans wishing to quickly update themselves, making sure crucial moments are efficiently included.

8) Player Scouting and Recruitment: Teams employ video analysis to assess prospective players. Reviewing game footage aids scouts in recognizing players' abilities, tactical understanding, and overall compatibility with the team.

9) Monitoring Player Health: Ongoing assessment of player movements can monitor physical state and fatigue, enabling improved management of player health and well-being.

10)Training with Virtual and Augmented Reality: Analysis of sports video can enhance VR and AR training settings, enabling athletes to train in simulated situations that mirror actual game circumstances.

3.Tasks and Techniques

3.1 Tasks:

1)Motion Change Region Detection in Sports Videos

Motion change region detection pinpoints regions of movement in sports videos by employing fuzzy clustering and interframe difference techniques to distinguish players from the backdrop. Other methods such as change detection in the wavelet domain and noise reduction improve accuracy, aiding in the tracking of dynamic movement in intricate, noisy settings.

2)Sports Image Segmentation and Object Detection

Sports image segmentation utilizes SISA to recognize players and equipment, while Key Node Detection improves attention on smaller items. CRFs enhance edges, whereas residual blocks preserve accuracy, and motion control methods facilitate seamless motion transitions in video analysis.

3)Deep Learning for Tennis Stroke Classification

This method categorizes tennis strokes by employing CNNs for spatial characteristics and RNNs for recognizing temporal sequences. 3D CNNs enhance depth by capturing the dynamics of actions, while SSD identifies and softens faces in the video. Frame variations, optical motion, and feature selection techniques such as HOG and SIFT enhance accuracy in stroke classification by integrating spatial and temporal perspectives.

4)AlexNet and Transfer Learning for Sports Video Classification

This technique employs the AlexNet CNN model with transfer learning to categorize sports scenes, utilizing pre-trained features for better efficiency and precision. Augmented with three fully connected layers and dropout for regularization, the model prevents overfitting and generalizes effectively. Data augmentation broadens the feature space, while convolutional layers capture complex patterns, enhancing scene classification in sports clips.

5)Deep Learning Techniques for Soccer Video Analysis

This method employs CNNs for recognizing players and the ball, while utilizing RNNs to monitor player movements and forecast actions. Two-Stream Networks integrate spatial and temporal information for in-depth motion analysis, whereas YOLO and Faster R-CNN facilitate real-time detection of players and balls. Event detection models highlight crucial soccer game moments, streamlining analysis to provide insights into tactics and performance.

6)Sports Video Segmentation with Deep Learning Models

This approach employs CNNs to isolate athletes and equipment from intricate backgrounds, while RNNs are utilized to understand temporal relationships for monitoring player movements. 3D CNNs manage spatial and temporal characteristics, whereas SSD facilitates real-time face detection to ensure privacy. Frame variations and optical flow enhance segmentation by detecting slight shifts between frames.

7)Motion and Object Detection in Sports Using Conditional Random Fields

This method improves object boundaries by employing Conditional Random Fields (CRFs) following deep learning models to enhance the segmentation of players and equipment. A residual block design preserves semantic correctness, while quaternion interpolation guarantees smooth 3D motion tracking.Data preprocessing enhances training, improving accuracy in identifying small moving objects in sports videos.

8) Tennis Stroke Classification Using CNNs and RNNs

This method employs CNNs to capture spatial characteristics of tennis strokes, such as player positioning and racket movement, whereas RNNs manage temporal dynamics derived from consecutive frames. 3D CNNs integrate spatial and temporal information, improving classification precision for strokes such as forehand or serve. HOG and SIFT assist in feature selection, enhancing stroke recognition in video analysis.

9)Transfer Learning with AlexNet for Sports Video Scene Classification

This approach employs transfer learning utilizing a pre-trained AlexNet model to identify sports scenes by detecting spatial characteristics such as edges and textures. The model is improved with fully connected layers and dropout to mitigate overfitting. Data augmentation methods enhance generalization, enabling the model to effectively categorize dynamic sports scenes without the need for retraining entirely.

10)Soccer Video Analysis Using Deep Learning and Object Detection Models

This method employs CNNs to identify players and the ball, while RNNs monitor their movements as they progress. Two-Stream Networks merge spatial and temporal information to understand motion dynamics, while YOLO and Faster R-CNN facilitate real-time tracking. Data augmentation enhances the model's robustness, facilitating precise detection of players and balls in rapid game scenarios.

3.2 Techniques:

1)Detection of Motion Changes in Sports Videos

This research employs fuzzy clustering to distinguish moving athletes from the background and utilizes interframe difference to identify motion across frames. Change detection in the wavelet domain enhances precision by examining movement over various frequencies, while techniques for selecting features enhance detection, guaranteeing improved separation of foreground and background in noisy settings.

2) Techniques for Segmentation of Sports Images and Motion Control

This method employs the Sports Image Segmentation Algorithm (SISA) to accurately identify athletes and equipment in sports footage. The Key Node Detection Model emphasizes detecting smaller objects that conventional methods frequently overlook. Conditional Random Fields (CRFs) improve object boundaries by modifying pixel associations, increasing detection precision. Techniques for motion control, such as quaternion interpolation, facilitate seamless transitions within 3D settings, enhancing the overall analysis of video.

3) Classification of Sports Scenes Using AlexNet and Transfer Learning

This research applies the AlexNet CNN architecture utilizing transfer learning for classifying sports scenes. Utilizing a pre-trained AlexNet model, the system decreases training duration and computational expenses. Methods such as dropout and data augmentation aid in reducing overfitting and boosting the model's generalization capabilities, thereby enhancing classification accuracy.

4) Classification of Tennis Strokes Utilizing CNNs and RNNs

This research utilizes CNNs and RNNs to categorize tennis strokes in video recordings. CNNs detect spatial elements such as player positioning and racket movements, whereas RNNs evaluate temporal sequences to gain a clearer insight into stroke dynamics. Employing 3D CNNs further combines spatial and temporal dimensions, improving precision. Methods such as HOG, SIFT, and optical flow assist in detecting motion, enhancing accurate stroke classification.

5) Deep Learning Methods for Analyzing Soccer Videos

This research utilizes deep learning to examine soccer videos, using CNNs for detecting players and the ball, while RNNs monitor movements through time. Two-Stream Networks combine spatial and temporal information to capture motion dynamics, while object detection models such as YOLO and Faster R-CNN provide precise real-time tracking. Automated feature extraction and data augmentation enhance robustness, aiding in sophisticated event and strategy evaluation in soccer.

6) Segmentation of Athletes and Sports Equipment Employing CNNs and RNNs

This research utilizes CNNs to identify athletes and sporting equipment from video frames by extracting spatial characteristics in intricate sports environments. RNNs assess time-based relationships between frames, allowing for continuous monitoring of players. The integration of 3D CNNs includes both spatial and temporal information, improving segmentation precision, while SSD aids in real-time facial recognition to ensure player anonymity during live streaming.

7) Improved Object Detection and Tracking Utilizing CRFs and Residual Networks

This research enhances sports video segmentation by utilizing Conditional Random Fields (CRFs) to sharpen object edges, particularly for smaller elements such as balls and athlete limbs. Residual networks preserve semantic clarity, guaranteeing precise segmentation of different objects within each frame. Quaternion interpolation helps facilitate seamless 3D rotations, essential for fluid motion tracking, whereas data preprocessing enhances the dataset for precise segmentation and motion recognition.

8) Classification of Tennis Strokes Through CNNs and RNNs

This research employs CNNs to identify spatial characteristics of tennis shots, such as player positioning and racket movement, while RNNs examine the temporal dynamics throughout video frames. 3D CNNs integrate spatial and temporal information for improved accuracy in stroke classification. Techniques for feature extraction, such as HOG and SIFT, enhance stroke recognition by emphasizing important visual patterns.

9) Using AlexNet for Sports Video Scene Classification through Transfer Learning

This research utilizes transfer learning with the pre-trained AlexNet model to categorize scenes in sports videos. The model identifies spatial characteristics such as edges and textures, lowering computational complexity. Techniques for data augmentation, such as rotations and shifts, enhance generalization, whereas dropout layers avoid overfitting, rendering the model proficient at classifying dynamic sports scenes.

10) Soccer Video Analysis with Deep Learning

This paper utilizes CNNs and RNNs for detecting players and the ball in soccer videos, capturing both spatial and temporal dynamics. Two-Stream Networks process spatial and temporal data, while YOLO and Faster R-CNN enable real-time tracking. Data augmentation improves model resilience, allowing accurate tracking of fast-moving players and events.

Task	**Techniques Used**	**Dataset Used**	**Evaluation Metrics**	**Strengths**	**Limitations**
Object Detection	YOLO variants (e.g., YOLOv4-Tiny) and MobileNet-based architectures	Publicly available sports datasets	Precision, mAP, Recall, FPS	Fast detection, real-time performance	Lower accuracy on small/distant objects, high dependency on video quality
Classification	VGG-M , Cascade-CNN , SiamCNN	Annotated sports datasets	Intersection over Union (IoU), Precision, Recall, Confusion Matrix	Good for classification tasks, handles occlusion, multi-stage for accuracy	High computational cost, struggles in complex real-world scenarios
Pose Estimation	- CNN-based Pose Estimation - Vision Transformers	Custom pose datasets (e.g., cricket, left-arm spin bowling)	MPJPE (Mean Per Joint Position Error), PCK (Percentage of Correct Keypoints), Procrustes Alignment	Accurate posture analysis, robust pose tracking	Complex training and high computational cost, especially with Vision Transformers
Tracking	- Mean-Shift	Broadcast data	Accuracy of trajectory	High precision	Requires multi-

Task	Techniques Used	Dataset Used	Evaluation Metrics	Strengths	Limitations
(Player & Object)	Tracking - Multi-camera Estimation (Hawk-Eye) - Color Clustering	(e.g., tennis, soccer)	estimation, IoU	for ball/player tracking, robust handling of occlusion	camera setups, struggles in complex environments with fast object movements
Tactics Analysis	- Pattern Recognition - Multi-modal Integration	USTA-based tennis datasets	Pattern Recognition Accuracy, False Positives/Negatives	Great for training, identifying recurring play patterns	Misses some patterns, high false positive rate can lower reliability
Highlight Extraction	- Audio-Visual Analysis - Event Detection	Broadcast sports data, YouTube clips (e.g., cricket highlights)	Precision, Recall, F1-score	Captures significant game moments using visual and audio cues	Dependent on audio quality, struggles in less noisy environments
Video Segmentation	- Fuzzy Clustering Algorithm - Time-domain and Frequency-domain methods	Real-world sports video footage	Spatial Accuracy, Noise Handling, Spatial Distortion	Effective in real-time segmentation, handles dynamic objects well	Sensitive to noise, complex motion handling issues
Computer-Assisted Refereeing	- Hawk-Eye Technology - Real-time Line Call Detection	Broadcast video from tennis, soccer	Decision-making accuracy, Response time	High accuracy in real-time decisions	Dependency on multi-camera setups
` **Pose Estimation &**			Pose Accuracy, Precision, Recall	Accurate detection of keypoints for sports movements	High computational cost, model training is complex
Tracking	- OpenPose - CNNs	Custom cricket datasets for left-arm spin bowling			
Action Recognition	- TensorFlow-based CNN Models	Custom sports datasets (e.g., handball actions)	Accuracy, Precision, Recall, F1-score	Good at recognizing sports-specific actions like handball jump shots	High complexity, dependent on robust dataset
3D Pose Estimation	- Two-Stage Deep Learning Models - 2D (e.g., Mask R-CNN) - 3D Models (e.g., EvoSkeleton,	Human3.6M, COCO 2017, custom datasets for handball	MPJPE, Procrustes alignment	Handles unseen environments well, high accuracy in 3D pose estimation	High computational cost, complex training

Task	Techniques Used	Dataset Used	Evaluation Metrics	Strengths	Limitations
	GnTCN)				
Scene Classification	- AlexNet CNN - Transfer Learning	Cricket videos (e.g., 6800 labeled frames from YouTube)	Accuracy, Precision, Recall, F1-score	High accuracy (99.26%) in scene classification	Struggles with complex scenes, requires large labeled datasets
Stroke Classification	- I3D - Twin Spatio-Temporal CNNs with Attention Blocks	TTStroke-21 dataset (table tennis)	Per-Class Accuracy, Confusion Matrix	Good at classifying different strokes, handles temporal data	Struggles with variations in strokes, complex setup
Sports Video Analysis	- Optical Flow - Multi-Camera Estimation (Hawk-Eye) - Edge Detection Techniques	Multiple sports datasets (e.g., tennis, soccer, baseball)	Trajectory Estimation Accuracy, FPS, Noise Handling, Edge Detection	Accurate object movement analysis, suitable for ball and player tracking	High dependency on camera quality, computational complexity for real-time applications

4.Advantages and Disadvantages

Advantages:

1) Automated Analysis and Insights

Computer vision provides automated tracking of players and the ball, estimating poses and recognizing actions, enabling coaches, analysts, and fans to swiftly and objectively obtain insights. For instance, action detection models in sports like basketball and soccer can record player actions, improving strategy development and minimizing manual evaluation.

2) *Capabilities for Real-Time Processing*

Systems that utilize deep learning and computer vision are progressively capable of processing data instantly, delivering up-to-date statistics, performance metrics, and game highlights during live contests. This is especially useful in sports broadcasts where there is a need for instant replay and analysis.

3) Improved Training and Injury Prevention

Pose estimation and motion analysis enable comprehensive tracking of athletes' biomechanics, assisting in injury prevention and enhancing performance by fine-tuning training programs. In activities such as gymnastics or martial arts, these tools can assist in identifying faulty posture, directing remedial measures.

4) Audience Involvement and Customized Material

Through video summarization and highlight creation, computer vision is able to generate tailored content, boosting fan involvement. Highlight reels, for example, simplify the process of experiencing crucial moments from games, which is very attractive to casual viewers.

5) Decision Making Based on Data

Computer vision allows for the quantitative assessment of individual and team performance, assisting coaches and players in making well-informed choices. Sophisticated metrics obtained from video data, such as player speed and response time, enable a more scientific method for coaching and strategy formation.

Disadvantages:

1) Elevated Computational Expense

Numerous computer vision models, especially those utilizing deep learning, demand substantial computational resources for both training and inference, leading to high expenses. This is especially difficult in real-time analysis where lag needs to be reduced for a smooth experience.

2) Ethics and Data Privacy

Concerns regarding player privacy and data ownership are crucial, particularly with the collection and broadcasting of live data. There are moral considerations about who possesses and can utilize this data, especially concerning minor athletes or amateur competitions.

3) Reliance on Extensive Datasets

Developing models for particular sports frequently demands large quantities of labeled data. Gathering and labeling these datasets is labor-intensive and costly, particularly for uncommon sports where limited resources exist.

4) Errors in Complicated or Dense Settings

In team sports such as football or hockey, occlusion (when one player blocks another) or intricate scenes can result in errors in tracking players or detecting events. This restricts the efficacy of these models in situations with crowded player arrangements or swift motion.

5) Issues of Generalization

Models developed for particular sports might not perform effectively in different ones. For example, a model designed for soccer might not perform well in tennis or volleyball because of variations in player dynamics, object traits, and court design. This limits the adaptability of computer vision systems for various sports uses.

5. Conclusions

Computer vision in sports is a fast-evolving domain, propelled by deep learning advancements such as CNNs and RNNs, which have greatly improved action detection, monitoring, and scene understanding in video evaluation. From a technical perspective, these developments provide unmatched accuracy, creating new opportunities for performance evaluation and immediate feedback. Coaches and sports analysts take advantage of the extensive data produced, enhancing strategic planning and player development by revealing insights that go beyond what standard video analysis can show. The applications cover a wide range of sports, including soccer and gymnastics, customizing insights to address unique sport requirements. Meanwhile, fans enjoy more captivating broadcasts, enriched by instant highlights, replays, and virtual reality features. These interactive elements, enabled by computer vision, enhance the spectating experience, making it more dynamic and customized to individual preferences. Nonetheless, as computer vision gains traction in athletics, ethical and privacy issues arise—especially regarding ongoing surveillance and data permission, impacting youth and amateur sports levels. In summary, computer vision in sports offers a data-driven, engaging experience advantageous for players, coaches, analysts, and fans. Utilizing methods such as object detection, action recognition, and pose estimation, it has transformed real-time sports analysis and interaction. Regardless of issues concerning computational expenses and privacy, the trajectory indicates a more accurate and engaging future for sports, influenced by data and technology.

References

1. Hamburger, C. (1995). Quasimonotonicity, regularity and duality for nonlinear systems of partial differential equations. *Annali di Matematica Pura ed Applicata*, 169, 321–354.
2. Sajti, C. L., Georgio, S., Khodorkovsky, V., & Marine, W. (2007). New nanohybrid materials for biophotonics. *Applied Physics A*.
3. Geddes, K. O., Czapor, S. R., & Labahn, G. (1992). *Algorithms for Computer Algebra*. Kluwer, Boston.
4. Broy, M. (2002). Software engineering — from auxiliary to key technologies. In Broy, M., & Denert, E. (Eds.), *Software Pioneers* (pp. 10–13). Springer, Heidelberg.
5. Cartwright, J. (2007). Big stars have weather too. IOP Publishing PhysicsWeb. http://physicsweb.org/articles/news/11/6/16/1. Accessed 26 June 2007.
6. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* (pp. 1097–1105).
7. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
8. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580–587).
9. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(6), 1137–1149.
10. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779–788).