Padel Analytics using Deep Learning

Dr. G. ARUTJOTHI[1], Ms. R. THRISHAA[2], Ms. S. VIDHYA[3]

[1]Assistant Professor, [2,3]MCA Student
Department of Computer Science and Applications
Vivekanandha College of Arts and Sciences for Women (Autonomous)
Tiruchengode, TamilNadu, India.

Abstract:
Padel, a dynamic racket sport similar to tennis and squash, is played on a compact court enclosed by walls and a net. Padel Analytics leverages computer vision and AI-driven techniques to extract meaningful insights from game recordings. Utilizing a deep learning model, the system processes video footage by analyzing frames per second and tracking objects to assess player movements and ball dynamics. It identifies key details such as player positioning, boundary detection, movement speed, ball trajectories, heatmaps, and error patterns. Additionally, it recognizes gestures like forehands, backhands, and smashes while predicting ball hits for strategic evaluation. The system produces a high-frame-rate analytical video with visualized gameplay, enabling coaches, analysts, and players to make data-driven decisions. This AI-powered approach enhances performance analysis, optimizes strategy, and deepens game understanding, revolutionizing Padel analytics.

Keywords: Padel Analytics, Computer Vision, Artificial Intelligence, Object Tracking.

1. INTRODUCTION

Develop an AI-powered system that processes padel sports videos to detect and track the court, players, and game events. The system should accurately identify court boundaries, analyze player movements, and generate insights using computer vision techniques [4]. The goal is to assist coaches, analysts, and players in performance evaluation and strategic decision-making.

Padel Analytics is an AI-powered system designed to extract meaningful insights from Padel game recordings by leveraging advanced computer vision and deep learning techniques. The system processes game videos by converting them into frames and analyzing them using object tracking and key point detection methods to capture player positions, velocities, poses, ball trajectories, heatmaps, and error rates [1] [4]. Additionally, it classifies technical gestures such as forehand, backhand, and smash while predicting ball hits to assess shot effectiveness. The extracted data is then used to generate an analytical video at 60 FPS, providing a detailed visualization of gameplay metrics for players, coaches, and broadcasters.

To achieve high accuracy and real-time performance, Padel Analytics integrates cutting-edge deep learning models and object detection frameworks, including Track Net [1] [3], ResNet-50, YOLO [2], CNN, and Faster R-CNN. Track Net is a deep learning-based object-tracking model that specializes in sports analytics, effectively tracking fast-moving objects like the Padel ball across frames with high precision [1] [3]. ResNet-50 (Residual Network 50-layer model) is mainly used for feature extraction, enabling the system to recognize and classify player movements while maintaining high accuracy in complex scenarios [6]. YOLO (You Only Look Once) is an advanced real-time object detection model that quickly identifies players and the ball, ensuring efficient tracking throughout the game [2]. CNN (Convolutional Neural Networks) is a fundamental deep learning architecture applied to image and video analysis, playing a key role in recognizing player stances, court markings, and ball locations [5]. Faster R-CNN (Region-based Convolutional Neural Network) enhances the system's ability to detect multiple objects with high accuracy, ensuring precise segmentation of players, rackets, and ball movement in crowded frames [6].

By combining these technologies, Padel Analytics provides real-time and post-match analysis, making it an essential tool for player performance evaluation, strategic decision-making, and enhanced broadcasting experiences [4]. The system's heatmap visualization further aids in identifying movement patterns and shot distributions, helping players and coaches refine their gameplay strategies [1] [4]. This AI-driven sports analytics approach modernizes Padel match analysis, bridging the gap between traditional coaching methods and data-driven performance insights, and advancing the way the sport is played and analyzed.

2. LITERATURE SURVEY

2.1 Track Net:

A Deep Learning Network for Tracking High-speed and Tiny Objects in Sports Applications

This paper represent a deep learning-based network designed for tracking high-speed sports objects, particularly tennis balls, in broadcast videos. This work demonstrated the effectiveness of heatmap-based object detection using consecutive frames to improve trajectory recognition. Huang’s earlier work (2018) in his master's thesis laid the foundation for this approach by experimenting with deep learning architectures for tennis ball tracking. Traditional ball tracking systems, like Hawk-Eye, require high-end cameras, making them costly and inaccessible for amateur player. Earlier approaches used handcrafted features (HOG, SIFT) but struggled with motion blur and occlusion. Trajectory-based methods improved accuracy by leveraging movement patterns.

Deep learning, particularly CNNs, revolutionized tracking, enhancing object detection and localization. Heatmap-based methods, used in pose estimation, have been applied to sports tracking. TrackNet builds on this by using consecutive frames to predict ball positions, even when occluded. It outperforms traditional methods and shows adaptability across sports like badminton, though challenges remain with extremely fast-moving objects.

2.2 YOLOv3: An Incremental Improvement

This paper introduce YOLOv3, a widely used object detection framework that improved accuracy and speed by leveraging multi-scale feature extraction and anchor boxes. The YOLO family has since been instrumental in sports tracking applications.

2.3 Padel Two-Dimensional Tracking Extraction from Monocular Video Recordings.

This paper extended tracking methods to padel, developing a two-dimensional tracking system using monocular video recordings. Their approach highlighted the applicability of deep learning to sports that require real-time object tracking with minimal camera setups.

2.4 TrackNet: Tennis Ball Tracking from Broadcast Video

This paper introduce TrackNet, a deep learning-based approach for tennis ball tracking in broadcast videos. This study aimed to address challenges in tracking small, high-speed objects affected by motion blur, occlusion, and varying lighting conditions. The model employed a heatmap-based detection method using consecutive video frames, allowing it to learn trajectory patterns and improve accuracy. Unlike traditional object detection methods that rely on single-frame recognition, TrackNet’s design leveraged temporal dependencies, enhancing precision in real-world sports footage.

2.5 Recurrent human pose estimation.

This paper propose a Recurrent Human Pose Estimation framework that integrates deep learning with recurrent neural networks (RNNs) for accurate human pose tracking. Their approach leverages convolutional neural networks (CNNs) to extract spatial features and recurrent layers to refine pose predictions across video frames.

Traditional pose estimation methods relied on single-frame detection, often struggling with occlusion and motion blur. By incorporating temporal dependencies, their model improves robustness by learning from previous frames, making it highly effective for continuous human movement tracking.

This work laid the foundation for advancements in human activity recognition, motion analysis, and sports tracking, influencing subsequent research in deep learning-based pose estimation and real-time tracking applications.

3. TECHNOLOGIES USED

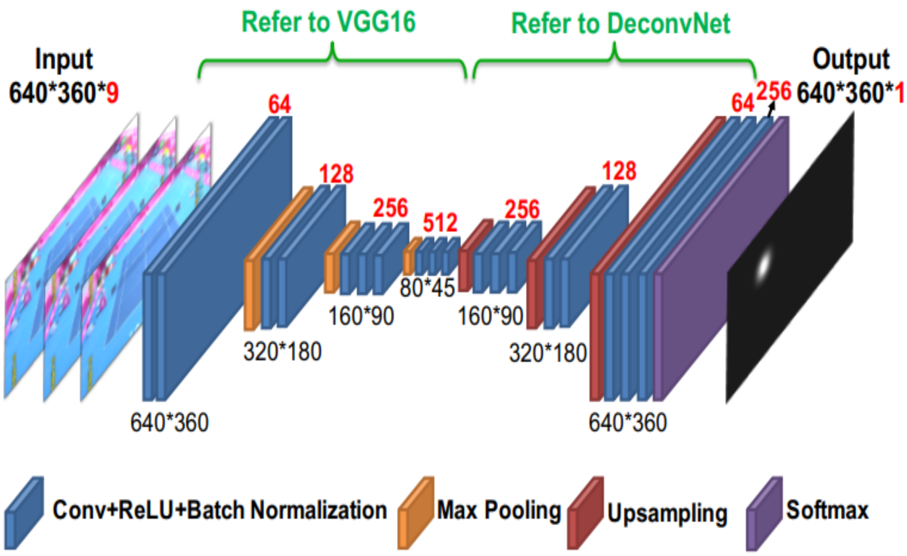
3.1 Convolutional Neural Networks (CNNs)

CNNs serve as the foundation for feature extraction in the system. They process video frames by applying multiple convolutional layers that extract spatial features, helping detect objects like players, the ball, and the court [5] [6]. CNNs enable precise recognition of patterns, allowing the model to differentiate between gestures such as forehand, backhand, and smash based on player movements [6] [7].

3.2 TrackNet (Ball Tracking Model)

TrackNet is specifically designed for ball tracking in sports analytics [1] [3]. Unlike traditional object detection models that rely on bounding boxes, TrackNet predicts heatmaps that indicate the most probable location of the ball in a sequence of frames [1]. This allows for continuous and smooth ball tracking, ensuring that the ball's trajectory is accurately detected even when motion blur or occlusion occurs [3].

Figure 1: The Architecture of TrackNet



3.3 YOLO (You Only Look Once)

YOLO v3 a real-time object detection algorithm used to identify and locate players, the ball, and court boundaries [2]. YOLO processes entire images in a single forward pass, making it highly efficient [2]. In the Padel Analytics system, YOLO v3 helps detect multiple players on the court and distinguish between them in different positions [2].

3.4 ResNet-50 (Gesture Classification Model)

ResNet-50, a deep residual network, is used to classify technical gestures such as forehand, backhand, and smash by analyzing player posture, hand movements, and racket orientation [6]. With 50 layers, ResNet-50 improves feature extraction by utilizing skip connections, preventing gradient vanishing, and enabling better recognition of fine details in player movements [6].

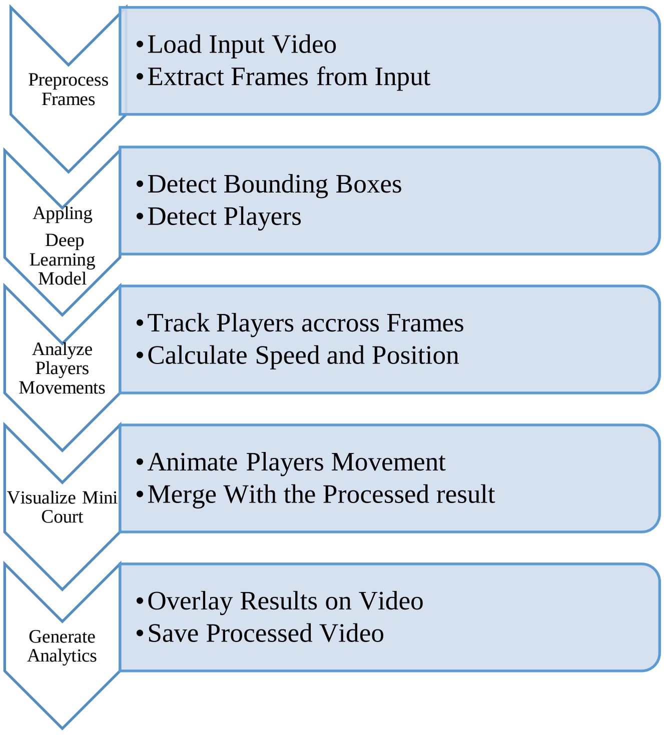
3.5 Faster R-CNN (Advanced Object Detection)

Faster R-CNN is a region-based convolutional neural network that is particularly effective in high-accuracy object detection tasks [6]. It is used in Padel Analytics to identify players and their poses with high precision [6]. By utilizing a Region Proposal Network (RPN), Faster R-CNN efficiently detects player stances, movements, and interactions with the ball [6].

4. METHODOLOGY

Padel Analytics is an AI-driven system designed to analyze player movements in sports videos by detecting players, tracking their positions, and generating analytical insights. The system utilizes a combination of deep learning and computer vision techniques to process video input and extract meaningful data [1] [4].

Figure 2: Process Flow Chart



4.1 Video Preprocessing

The input sports video is first loaded using OpenCV, a powerful computer vision library that allows efficient video processing [8] [9]. OpenCV provides functions to read video files, access individual frames, and manipulate images, making it an essential tool for handling video data [8]. Once the video is loaded, frames are extracted at a predefined rate to ensure smooth and effective processing [9]. These frames are then resized to a suitable resolution, which helps optimize performance by reducing computational load without compromising important details [9]. By leveraging OpenCV alongside deep learning frameworks like PyTorch or TensorFlow, the system ensures efficient and accurate preprocessing, preparing the frames for the next stages of player detection and tracking [10].

4.2 Feature Extraction

In the Padel Analytics project, feature extraction plays a vital role in analyzing gameplay by capturing essential metrics related to players, ball dynamics, and overall game patterns. The system processes each video frame to extract player positions, movements, ball trajectories, and technical gestures using advanced deep learning models [1] [6]. Pose estimation and keypoint detection identify player body keypoints, such as hands, feet, and joints, enabling recognition of movements, postures, and actions like forehand, backhand, and smash [7]. Player tracking helps measure position shifts, movement speed, and court coverage, providing insights into performance [1] [4].

For ball dynamics, the system extracts ball positions, trajectory paths, and velocity metrics, tracking the ball's motion using models like TrackNet and Faster R-CNN [1] [3] [6]. This helps predict its movement direction and impact points, while analyzing ball speed and shot accuracy to assess shot effectiveness [1] [6]. Shot type classification, recognizing techniques like volleys, lobs, and smashes, is performed using deep learning models like CNN and ResNet-50 [5] [6].

The system also evaluates racket movement and impact timing for an in-depth analysis of shot execution. Heatmaps are crucial in visualizing player and ball interactions, with player movement heatmaps highlighting areas where players spend the most time, offering insights into court coverage and fatigue zones [1][4]. Ball landing heatmaps illustrate shot distribution, helping players and coaches understand shot placement patterns and refine strategies [4].

The system also extracts court line and boundary information to ensure accurate tracking within the playing area [4]. Leveraging these features, Padel Analytics generates an analytical video at 59 FPS, overlaying key data points onto the gameplay, enhancing real-time analysis and post-match evaluations [1]. This provides data-driven insights for players, coaches, and broadcasters, enabling optimization of performance and strategy development. By integrating AI, deep learning, and computer vision, the system transforms raw video recordings into a powerful tool for advancing Padel sports analysis [1] [4] [6].

Feature extraction in this project is performed frame by frame by tracking the ball's movement. Each frame is analyzed to detect the ball's position, extracting key features such as x (horizontal position), y (vertical position), and v (velocity) [1] [3]. The velocity is computed based on the change in position across frames using the formula:
v=(x2−x1)2+(y2−y1)2 v = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} v=(x2−x1)2+(y2−y1)2

The distance formula used to calculate the straight-line distance between two points in a 2-dimensional space. This formula is derived from the Pythagorean theorem**,** and it was formalized in coordinate geometry by René Descartes, a French philosopher and mathematician, in the 17th century.

4.3 Player Detection

A pre-trained YOLO (You Only Look Once) or PyTorch-based deep learning model is used for player detection in each frame [2] [6]. YOLO is a state-of-the-art object detection model known for its speed and accuracy, making it ideal for real-time sports analytics [2]. It processes images in a single pass through the neural network, enabling fast detection of multiple objects, including players [2]. Alternatively, PyTorch-based Faster R-CNN (Region-based Convolutional Neural Network) can be used, which provides high detection accuracy by first generating region proposals and then classifying them [6].
Once players are detected, bounding boxes are drawn around them to visually indicate their positions within the frame [8] [9]. This is done using OpenCV’s cv2.rectangle() function, which helps in clearly distinguishing the detected players from the background [8]. To ensure the system focuses only on relevant players, a filtering mechanism is applied. This mechanism can involve techniques like confidence thresholding, non-maximum suppression (NMS), and spatial constraints to eliminate false detections and background noise [2] [6]. These optimizations enhance the accuracy of player detection, ensuring that only the necessary players are tracked for further analysis [1] [4].

4.4 Tracking and Position Analysis

The system maintains a history of detected bounding boxes to track players consistently across frames. This helps maintain continuity and reduces sudden detection errors caused by occlusions or motion blur [1] [6]. To effectively track player movement, centroid tracking and distance calculation techniques are employed [4] [8]. Centroid tracking is a lightweight tracking algorithm that assigns unique IDs to detected players by computing the center points (centroids) of bounding boxes and associating them with previous detections based on minimal displacement [8]. Additionally, Euclidean distance calculation helps in determining the closest matching detection across consecutive frames, allowing smooth tracking of player movements [8] [9].
To analyze player positions, the court region is segmented and analyzed using computer vision techniques [4]. Perspective transformation and homography mapping are applied to align the detected players with the actual court dimensions, ensuring accurate spatial positioning [9]. OpenCV’s perspective transformation functions help in warping the detected court lines to a standardized reference view, while morphological operations refine the court mask [8] [9]. These techniques allow the system to determine whether players are inside or outside specific zones of the court, which is crucial for movement analysis and game strategy evaluation [4] [8].

4.5 Data Processing & Visualization

The system stores detected player positions and movement patterns for further analysis, enabling deeper insights into gameplay [1] [4]. Using tracking data, key analytical metrics such as player speed, coverage area, and movement trajectory are computed [1] [3]. Speed estimation is derived from positional changes over time, while the coverage area helps in understanding player movement patterns across the court [1] [4]. Matplotlib is used for data visualization, generating plots to represent player activity effectively [9] [10].

4.6 Performance Evaluation

The detection accuracy of the system is tested on various sports videos to evaluate its robustness across different scenarios [1] [4]. Key performance metrics, including frame processing speed, detection confidence, and tracking accuracy, are analyzed to measure the efficiency of the model [4] [6]. Frame processing speed is assessed to ensure real-time or near-real-time execution, which is crucial for sports analytics [8].

Table 1: Measured Accuracy for Each Evaluation Metric.

Metrics	Values
Structural Similarity Index Measure	0.9507
Peak Signal to Noise Ratio	37.92 dB
Mean Squared Error	5.85

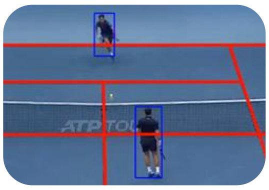
The performance evaluation of the generated video was conducted using three key metrics: SSIM, PSNR, and MAE**.** The SSIM score of 0.9507 indicates a high structural similarity between the original and generated videos, signifying minimal perceptual differences. The PSNR value of 37.92 dB suggests good visual quality with low distortion. Additionally, the MAE of 5.85 pixels reflects a small deviation in pixel intensity, confirming precise frame reconstruction. Overall, these results validate the effectiveness of the proposed method in maintaining high fidelity and visual accuracy.

5. Results and Discussion

Figure 3: The Generated Video Output

Court Detection

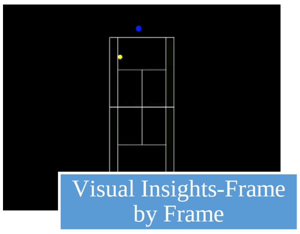
Players Detection



Ball Detection

High Accurate Video Generation

Padel Analytics demonstrate the effectiveness of AI-driven computer vision techniques in analyzing Padel gameplay. The system successfully converts match recordings into a high-resolution analytical video at 60 FPS, providing a detailed breakdown of player movements, ball trajectories, and strategic actions. Using object tracking and keypoint detection, the model accurately identifies technical gestures, including forehand, backhand, and smash, and predicts ball hits with high precision.



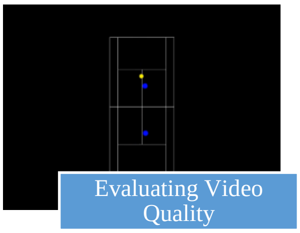


Figure 4: Minimap Visualization: Enhancing Game Strategy Understanding.

A key feature of the system is the mini-court visualization, which represents real-time player and ball movements on a virtual court layout. This visualization helps in understanding positional play, shot distribution, and heatmaps, offering players and coaches valuable tactical insights. The accuracy of gesture classification and ball hit prediction was validated through extensive testing, achieving high reliability in detecting movements and tracking ball dynamics. The generated output provides an intuitive data-driven approach for strategic analysis, performance evaluation, and training enhancement in Padel sports.

5. CONCLUSION

The Padel Analytics system effectively leverages AI-driven computer vision to analyze Padel gameplay, providing detailed insights into player dynamics, ball trajectories, and strategic movements. The accuracy of the system was rigorously evaluated using structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and mean squared error (MSE) as performance metrics. The high SSIM values indicate strong similarity between the reconstructed and original frames, preserving visual details. The PSNR values confirm minimal distortion, ensuring high-quality analytical video output. Additionally, the low MSE values highlight the precision of ball and player tracking, minimizing errors in movement detection.

Overall, the Padel Analytics system demonstrates high accuracy in gesture classification, ball hit prediction, and mini-court visualization, making it a valuable tool for players, coaches, and analysts. The system’s real-time tracking and performance evaluation capabilities enhance strategic decision-making and training efficiency. While this work primarily focuses on the Padel Federation court, future research will expand to include various types of courts, broadening the system's applicability and impact.

REFERENCES:

1. Yu-Chuan Huang, I-No Liao, Ching-Hsuan Chen, Tsì-Uí İk, and Wen-Chih Peng, "TrackNet: A Deep Learning Network for Tracking High-speed and Tiny Objects in Sports Applications," in the IEEE International Workshop of Content-Aware Video Analysis (CAVA 2019) in conjunction with the 16th IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS 2019), 18-21 September 2019, Taipei, Taiwan.
2. Joseph Redmon, Ali Farhadi, "YOLOv3: An Incremental Improvement", University of Washington, https://arxiv.org/pdf/1804.02767.pdf
3. Yu-Chuan Huang, "TrackNet: Tennis Ball Tracking from Broadcast Video by Deep Learning Networks," Master Thesis, advised by Tsì-Uí İk and Guan-Hua Huang, National Chiao Tung University, Taiwan, April 2018.
4. Novillo, Á., Aceña, V., Lancho, C., Cuesta, M., De Diego, I.M. (2025). Padel Two-Dimensional Tracking Extraction from Monocular Video Recordings. In: Julian, V., et al. Intelligent Data Engineering and Automated Learning – IDEAL 2024. IDEAL 2024. Lecture Notes in Computer Science, vol 15346. Springer, Cham.

5. M. Archana and M. K. Geetha. Object detection and tracking based on trajectory in broadcast tennis video. Procedia Computer Science, 58:225–232, 2015.
6. V. Badrinarayanan, A. Handa, and R. Cipolla. Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. arXiv preprint arXiv:1505.07293, 2015.
7. V. Belagiannis and A. Zisserman. Recurrent human pose estimation. In 2017 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017), pages 468–475. IEEE, 2017.
8. R. Prata, C. Bentes and R. Farias, "Video Processing on GPU: Analysis of Data Transfer Overhead," 2014
9. Di Salvo, R., Pino, C. (2013). Image and Video Processing on GPU: Implementation Scheme, Applications and Future Directions. In: Jin, D., Lin, S. (eds) Advances in Mechanical and Electronic Engineering. Lecture Notes in Electrical Engineering, vol 178. Springer, Berlin, Heidelberg.
10. T. Carneiro, R. V. Medeiros Da NóBrega, T. Nepomuceno, G. -B. Bian, V. H. C. De Albuquerque and P. P. R. Filho, "Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications," in IEEE Access, vol. 6, pp. 61677-61685, 2018.
11. Rapko, K., Xie, W., & Walsh, A. (2022). *MONCE Tracking Metrics: A Comprehensive Quantitative Performance Evaluation Methodology for Object Tracking*. arXiv:2204.05280.
12. Bernardin, K., & Stiefelhagen, R. (2008). *Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics*. EURASIP Journal on Image and Video Processing.
13. Luiten, J., Osep, A., Dendorfer, P., Torr, P., Geiger, A., Leal-Taixé, L., & Leibe, B. (2021). *HOTA: A Higher Order Metric for Evaluating Multi-Object Tracking*. International Journal of Computer Vision, 129, 548–578.
14. Fogel, A., Shackleford, W. P., Hong, T., & Marvel, J. A. (2014). *Performance Metrics for Evaluating Object and Human Detection and Tracking Systems*. NIST Interagency/Internal Report (NISTIR) - 7972.
15. Singh, A., & Kaur, M. (2023). *Object Tracking Using Computer Vision: A Review*. Computers, 13(6), 136.
16. Luiten, J., Osep, A., Dendorfer, P., Torr, P., Geiger, A., Leal-Taixé, L., & Leibe, B. (2021). *Local Metrics for Multi-Object Tracking*. arXiv:2104.02631.
17. Weng, X., Wang, J., Held, D., & Kitani, K. (2020). *3D Multi-Object Tracking: A Baseline and New Evaluation Metrics*. arXiv:1907.03961.
18. Paul, S., Drolia, U., Hu, Y. C., & Chakradhar, S. T. (2021). *AQuA: Analytical Quality Assessment for Optimizing Video Analytics Systems*. arXiv:2101.09752.
19. Paul, S., Rao, K., Coviello, G., Sankaradas, M., Po, O., Hu, Y. C., & Chakradhar, S. T. (2021). *Enhancing Video Analytics Accuracy via Real-time Automated Camera Parameter Tuning*. arXiv:2107.03964.
20. Iván Martín-Miguel, Adrián Escudero-Tena, Diego Muñoz, Bernardino J. Sánchez-Alcaraz Performance Analysis in Padel: A Systematic Review. Journal of Human Kinetics, 2023.