JOURNAL OF BASIC SCIENCE AND ENGINEERING

MODELLING OF THE QUALIFICATION & IMPROVEMENT OF TENNIS STANCE FOR PLAYER PERFORMANCE IMPROVEMENT USING 2D ANALYSIS OF VIDEOS TAKEN FROM A MOBILE CAMERA (PART A)

[1] Abhilash Manu*, [2] Dr. D. Ganesh, [3] Dr. Aravinda H.S, [4] Mayur Gowda R.,
[5] N.R. Swedith Reddy

[1] Research Scholar, School of Computer Science & Information Technology (CSIT), Jain University (Deemed to be University), Bangalore, Karnataka, India
[1] Associate Director, Ernst and Young, Bangalore
[2] Supervisor, Professor, R & D Coordinator, School of Computer Science & Information Technology (CSIT), Jain University (Deemed to be University), Bangalore, Karnataka, India
3 Co-Supervisor, Professor, Department of Electronics & Communications Engineering, JSS Academy of Technical Education, Bangalore, Karnataka, India
[4,5] Department of Artificial Intelligence & Machine Learning, RNS Institute of Technology, Bangalore, Karnataka, India

DOI: https://doie.org/10.0118/Jbse.2025328537

Article History: Received: July 2024 Revised: Sept 2024 Accepted: December 2024

***Abstract* -** This study outlines a novel framework for enhancing tennis performance through the optimization of biomechanical postures during specific tennis shots, utilizing advanced 2D video analysis and stabilization techniques coupled with Recurrent Neural Networks (RNNs). Through precise pose estimation algorithms, skeletal key-points are extracted to compute joint angles via vector dot product formulae. These keypoints allow for detailed biomechanical analysis and the classification of movement patterns using unsupervised clustering methods such as k-means. The analysis is refined through adaptive acceptance areas determined by a blend of distance metrics, enhancing the accuracy of pose alignment evaluations. Challenges such as motion artifacts, variable lighting conditions, and low signal-to-noise ratios (SNR) are addressed through the use of high-SNR imaging devices and optimized camera calibration. This approach ensures high-quality data capture essential for reliable computational analysis. The study leverages cloud platforms for processing, maintaining strict confidentiality of data while harnessing scalable computational resources. This integration facilitates robust kinematic analysis through part affinity fields and TensorFlow Lite, enabling real-time feedback on player movements and biomechanical alignment. The research makes significant contributions by integrating advanced computational algorithms and tailored hardware solutions to surpass traditional video analysis limitations. Through the detailed kinematic analysis of player movements and the innovative use of clustering algorithms, the study provides a comprehensive method for enhancing tennis performance. This methodology not only refines current coaching practices but also sets new standards in sports performance analysis, ultimately aiming to revolutionize tennis training through data-driven insights and technological advancement. The modelling results shows the effectiveness of the method developed, which could be used for a host of science & engineering applications.

Keywords - Tennis Performance Optimization, 2D Video Analysis, Recurrent Neural Networks, Pose Estimation, Biomechanical Modeling, Unsupervised Clustering, *k*-Means, Joint Angles, Adaptive Acceptance Areas, Motion Artifacts, High-SNR Imaging, Cloud Computing, Part Affinity Fields, Tensorflow Lite, Sports Performance Analysis.

Introduction

In this section, a brief introduction w.r.t. the proposed work taken up in this article is presented. In the competitive realm of sports, tennis stands out as a game where precision, agility, and technical skill are paramount. The quest to enhance player performance has increasingly embraced technological advancements, particularly in the analysis and optimization of player movements and stances. This study focuses on the innovative application of 2D video analysis using mobile cameras to qualitatively improve tennis stances for player performance enhancement. This approach leverages readily available technology to provide detailed insights that were once only possible with high-end equipment and specialized settings [40].

The primary objective of this research is to develop a comprehensive framework that utilizes video footage captured from mobile devices to analyze and improve the biomechanical postures of tennis players. By applying advanced computational methods, including pose estimation and Recurrent Neural Networks (RNNs), the study meticulously extracts and analyzes skeletal key-points from players in action. These key-points serve as the foundation for calculating joint angles and evaluating the biomechanical efficiency of various tennis stances. Furthermore, unsupervised machine learning techniques, such as k-means clustering, are employed to categorize and assess these movements, identifying optimal patterns that correlate with improved performance and reduced injury risk [39].

Addressing challenges inherent in mobile video analysis—such as variable lighting, motion blur, and lower resolution—the research employs high-signal-to-noise ratio (SNR) imaging and sophisticated stabilization algorithms to enhance data quality. The fusion of these technological solutions with sports science not only broadens the accessibility of advanced training tools but also enriches the analytical depth through which coaches and players can understand and refine playing techniques. As such, this study not only pushes the boundaries of traditional coaching methods but also sets a precedent for integrating mobile technology into sports performance analytics, making sophisticated training assistance more accessible to the broader sporting community [38]

Mathematical Modelling

The development of the mathematical model in the proposed study involves several interconnected components designed to optimize tennis stances and improve player performance through detailed biomechanical analysis. This mathematical model bridges advanced computational techniques with practical sports performance needs, enabling detailed analysis and optimization of tennis stances based on real-world data captured in dynamic, variable conditions. Through this sophisticated approach, the study not only addresses technical challenges but also enhances the practical training and coaching methodologies, ultimately aiming to elevate the standard of tennis performance through science-backed, data-driven insights [37].

Pose Estimation and Joint Angle Calculation

The foundation of the mathematical model lies in precise pose estimation algorithms that extract skeletal keypoints from video frames. These keypoints represent crucial joint locations on a tennis player's body. The model uses vector mathematics to compute joint angles from these keypoints as follows [36].

Vector Dot Product Formula : Each joint angle θ_j is calculated using the relationship between vectors formed by adjacent keypoints. If u and v are vectors representing limbs or parts connected at a joint, the angle between them is given by [35]

$$\theta_j = arcos\left(\frac{\vec{u}.\vec{v}}{\|\vec{u}\|\ \|\vec{v}\|}\right)$$

This calculation is critical as it quantifies the biomechanical postures players assume during different shots, allowing for the analysis of their efficiency and potential improvements [34].

Unsupervised Clustering for Movement Pattern Analysis

After joint angles are computed, the next step involves the classification and analysis of these biomechanical patterns [33]

k-Means Clustering

This unsupervised clustering algorithm groups similar movement patterns based on the feature vectors derived from joint angles. Each cluster formed corresponds to a specific type of movement pattern or stance, aiding in identifying optimal and suboptimal performance traits [32].

Cluster Centroids

The mathematical representation of each cluster is defined by its centroid, calculated as μ_k, where c_k represents the set of data points in cluster k, providing a benchmark for comparing individual performances against optimal postures [31].

Adaptive Acceptance Areas and Distance Metrics

To refine the analysis and increase the accuracy of pose alignment evaluations, the model employs adaptive acceptance areas, which are regions around each centroid where data points are considered to be in a similar stance or movement pattern using the distance metrics, where a combination of Euclidean, Minkowski, and Manhattan distances are utilized to define these areas, offering a comprehensive metric for spatial variability within the dataset. These metrics enhance the model's ability to handle diverse body types and movement dynamics, which are common in sports like tennis [30].

Integration with Technological Platforms

The model's computations are performed on cloud platforms that offer scalable processing power, which is crucial for handling the extensive computations required for real-time feedback and video analysis by using Cloud Computing. By leveraging cloud technology, the study ensures that data processing is both powerful and secure, maintaining confidentiality while providing the computational resources needed to process video data effectively [29].

Real-time Feedback and Performance Enhancement

Finally, the model's output is utilized to provide real-time feedback to players and coaches [28].

TensorFlow Lite and Part Affinity Fields

These technologies are integrated to visualize skeletal structures and movement patterns directly on mobile devices, offering immediate and actionable insights that can be used to adjust and improve player stances in real-time. In tennis, performance analysis has advanced primarily as notational analysis. And analytical techniques markedly advanced, particularly in the fields of notational analysis and match analysis. In tennis, the Hawk-Eye system was introduced to tour tournaments in 2000's. Hawk-Eye is a computer vision system used in numerous sports such as cricket, tennis, Gaelic football, badminton, hurling, rugby union, association football and volleyball, to visually track the trajectory of the ball and display a profile of its statistically most likely path as a moving image. This research study presents a comprehensive framework for enhancing tennis performance by optimizing biomechanical postures during specific shots, utilizing advanced 2D video analysis and stabilization techniques [27].

Centroid and Distances

Through this study we propose that one of the key steps in analysing the players' performance is the distance of a learner's shot from the cluster centroids. This distance represents how close the learner's shot is to the ideal shot (as defined by the centroid of each cluster). To be able to draw conclusions, we calculate the centroids each cluster, the centroid represents the 'average' position of the shots in that cluster. This is the mean position of the joint angles for all shots in the cluster. The equation used for centroid calculation is modelled as [26]

$$\mu_c = \left(\frac{1}{n_c} \sum_{i=1}^{n_c} Angle-1_i , \frac{1}{n_c} \sum_{i=1}^{n_c} Angle-2_i \right)$$

where

- $\mu_c = (\mu\text{Angle}1c, \mu\text{Angle}2c)$: Centroid of cluster c_c
- n_c : Number of shots in cluster c_c

- Angle-1$_i$ and Angle-2$_i$: Joint angles (e.g., shoulder angle, hip angle) for the i^{-th} shot in the cluster

Distance Metrics

To analyse the separation between centroid and the ‘learner data’, we calculated the distances using three different distance metrics, viz., Euclidean Distance, City Block (Manhattan) Distance, and Minkowski Distance to perform a comparative study and ensure right conclusions [32].

1. Euclidean Distance (D_{Eucl})

$$d_{Euclidean}(\mu_c - \mu_d) = \sqrt{\left(\mu_{Angle\ 1_c} - \mu_{Angle\ 1_d}\right)^2 - \left(\mu_{Angle\ 2_c} - \mu_{Angle\ 3_d}\right)^2}$$

2. City Block (Manhattan) Distance (D_{man})

We used the Manhattan distance algorithm to measure the sum of the absolute differences between the coordinates of the ‘learner data’ and the centroids [33].

$$d_{City\ Block}(\mu_c, \mu_d) = \left|\mu_{Angle\ 1_c} - \mu_{Angle\ 1_d}\right| + \left|\mu_{Angle\ 2_c} - \mu_{Angle\ 2_d}\right|$$

3. Minkowski Distance (D_{min})

We also generated the distances using the Minkowski algorithm, which is a actually a combination of both the Euclidean and Manhattan distances [34].

$$d_{Minkowski}(\mu_c, \mu_d) = \sqrt{\left(\left|\mu_{Angle\ 1_c} - \mu_{Angle\ 1_d}\right|\right)^p - \left(\left|\mu_{Angle\ 2_c} - \mu_{Angle\ 3_d}\right|\right)^{\frac{1}{p}}}$$

Generating the standards and comparison

By calculating the distance to the nearest centroid, we are able to determine the quality of the shot. Data points closer to the centroid were considered more technically sound and closer to optimal stance, while those farther away indicated deviations from the optimal stance. We calculated the score of the shot using the square root of sum of squares of all 3 distances using the below formula as [25]

$$Score = \sqrt{D_{euc}^2 + D_{man}^2 + D_{min}^2}$$

D_{man} : Represents the Manhattan Distance.

D_{min} : Represents the Minkowski Distance.

D_{euc} : Represents the Euclidean Distance.

The formula computes the aggregated score considering a quadratic contribution of all three distances. Performance analysis is a specialised discipline that provides athletes and coaches with objective information that helps them understand performance. This process is underpinned by systematic observation, which provides valid, reliable and detailed information relating to performance. Recurrent Neural Networks (RNNs) are employed in conjunction with pose estimation algorithms to accurately extract skeletal key points (x_i, y_i) for i =1, 2,…, n, where n represents the number of key joints. Joint angles θ_j are computed using the vector dot product formula modelled as [24]

$$\theta_j = arcos\left(\frac{\vec{u}.\vec{v}}{||\vec{u}||\ ||\vec{v}||}\right)$$

where $\vec{u}$ & $\vec{v}$ are vectors formed by adjacent key points. Unsupervised clustering, such as k-means, is applied to group movement patterns based on extracted feature vectors. The centroid for each cluster is given by the model [23]

$$\mu_k = \frac{1}{|c_k|}\sum_{x\epsilon C_k} x$$

where c_k represents the set of data points in cluster k. Adaptive acceptance areas are defined using a combination of Euclidean, Mikowski and Manhattan distances [22].

The research systematically addresses challenges inherent to video-based analysis, including motion artifacts, variable lighting conditions, low-resolution imaging, suboptimal signal-to-noise ratios (SNR), and limited frame rates. High-SNR imaging devices, optimized camera calibration, and daylight capture protocols are employed to mitigate these issues. Computational analysis is performed on cloud platforms, leveraging scalable processing power while maintaining strict data confidentiality. The key contributions include the integration of pose-detection key points into spatial frame coordinate systems for advanced kinematic analysis of player movements. The skeletal structure is modeled using part affinity fields (PAF's), represented as [21]

$$L = \sum_{c}\sum_{p \epsilon c} w(p).\,log\left(1 + exp(-s_c(p))\right)$$

where $w(p)$ is the weighting function, and $S_c(p)$ represents the score map for a candidate connection. TensorFlow Lite facilitates real-time skeletal visualization, providing immediate feedback on biomechanical alignment. The Fig. 1 shows the specific processing of data pre-processing & the time series feature modeling adopted [35]. The proposed methodology overcomes the limitations of traditional video analysis by integrating state-of-the-art computational algorithms, including Convolutional Neural Networks (CNNs), with tailored hardware solutions. This robust approach highlights the critical importance of accurate joint angle computation and motion pattern analysis in refining tennis biomechanics and advancing performance optimization [20].

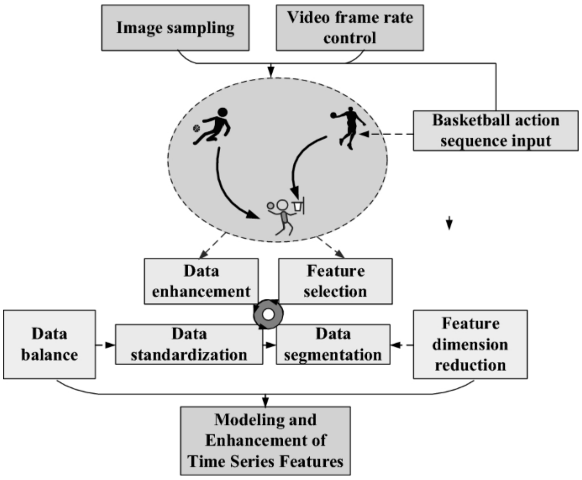

Fig. 1 : Specific processing of data pre-processing & the time series feature modeling adopted [35].

Performance Analysis

In sports performance analysis, tennis demands precision, agility, and strategic prowess. Enhancing tennis performance through technological advancements is a key focus for researchers and practitioners. This study leverages 2D video analysis and stabilization techniques to optimize player poses during specific tennis shots, aiming to improve performance outcomes. Body mechanics are crucial in tennis for executing effective shots, minimizing injury risks, and enhancing gameplay. Traditional coaching methods often lack the precision and consistency that modern technology can offer. This study integrates advanced clustering algorithms and pose estimation models to provide a data-driven approach to performance enhancement. The flow-chart shown in the Fig. 2 is used for the modelling & analysis purposes [19].

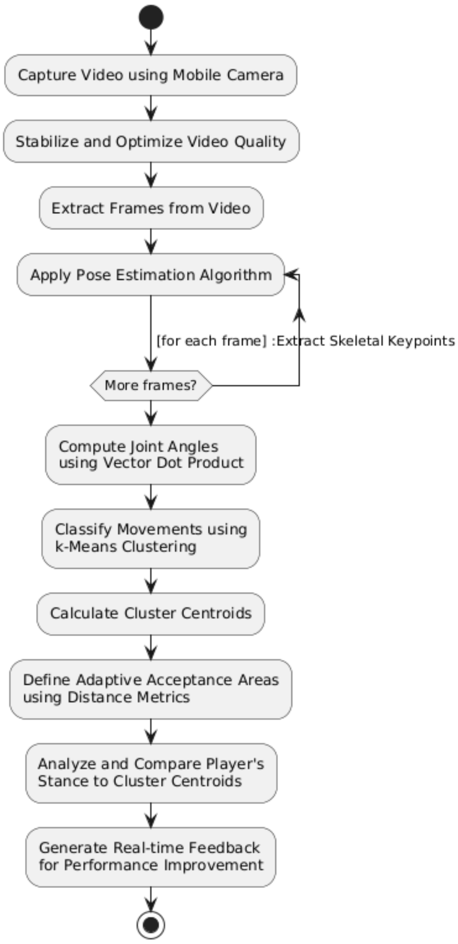

Fig. 2 : Performance analysis & modelling

The research's significance lies in its innovative methodology and practical applications. Using high-SNR phone models like the Google Pixel 8 Pro and iPhone 15, the study addresses challenges in video analysis, such as signal-to-noise ratio, resolution, frame rate, motion blur, and lighting conditions. Specific camera settings and daylight video capture ensure data quality and reliability. Advanced clustering algorithms, like *k*-means, group data points based on similarity, defining flexible regions around centroids. This facilitates the precise generation of body angles, critical for optimizing player poses. Integrating pose detection key-points into frame coordinates allows detailed analysis of player movements, providing insights into tennis shot biomechanics. The study emphasizes data privacy and security. Video analysis is conducted on cloud platforms like Google Collab and Azure, ensuring robust processing power while maintaining data confidentiality. Encition and secure data handling practices protect video data from unauthorized access and misuse. The block diagram of the proposed scheme is shown in the Fig. 3 [36].

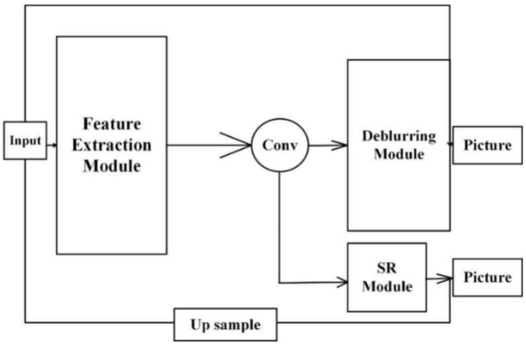

Fig 3 : Block-diagram of the scheme employed [36]

Unique perspectives address multifaceted challenges and solutions. Digital image stabilization and gyroscope-based features mitigate the impact of shaky footage on computer vision tasks. Adaptive algorithms adjust to varying lighting conditions, enhancing analysis robustness. In conclusion, this study advances the application of technology in sports performance analysis. Combining practical solutions with cutting-edge techniques, it overcomes video analysis challenges, paving the way for more precise and effective tennis training methods. The findings have the potential to revolutionize coaching practices, offering a new paradigm for achieving optimal tennis performance [18].

Challenges

Some of the challenges we faced during the course of our study can be classified into the following sections [17].

A. Challenges in using mobile phone video for computer vision-based analytics

- Hardware Limitations: Mobile phone cameras don’t always match the quality of dedicated cameras, which can make video less sharp and harder to analyse. To fix this, we used techniques like multi-frame super-resolution to improve video quality and make it easier for the computer to detect key details. This challenge is highlighted in the work by Cao *et.al.* [1], which discusses real-time pose estimation and the difficulties in processing lower-quality video input, making multi-frame super-resolution an essential technique to enhance the clarity of images [16].

- Environmental Factors: Lighting can be a problem—whether it’s too dark or too bright, it can mess up the computer’s ability to process the video. To handle this, we modified and used algorithms that adapt to different lighting conditions, using methods like histogram equalization to keep the video clear and usable. This is particularly relevant to Wei *et.al.* [2], who explain how Convolutional Pose Machines handle environmental variability, ensuring reliable pose detection under challenging conditions such as poor lighting. Similarly, Cao *et.al.* [3] demonstrate the robustness of OpenPose, a system designed to detect key points in images despite fluctuating environmental conditions [15].

- Computational Constraints: Phones don’t have as much power as computers, so processing video in real-time can be tricky. To solve this, we optimized deep learning models for mobile devices, using techniques like model quantization and hardware acceleration to make them run faster and smoother. This challenge is discussed in Cao *et.al.* [1], where real-time pose estimation is achieved with computational optimizations that could be adapted for mobile devices. Additionally, Boonim [4] addresses the use of computational methods in sports analysis, which requires significant processing power, emphasizing the need for mobile optimization [14].

- Privacy and Security Concerns: Video from phones can include personal or sensitive information, so protecting privacy is really important. To address this, we masked the faces of the players and to keep the video safe from unauthorized access [13].

- User Interaction and Usability: Shaky video from hand movements can make it harder for the computer to analyze the footage. To improve this, we added features present in the phone like digital image stabilization to make the video smoother and easier to work with. This issue is particularly relevant when studying biomechanics in sports. For example, Lertwonghattakul *et.al.* [5] analyze tennis serve biomechanics, where stability in video input is critical for accurate data collection. Additionally, Knudson [6] examines the importance of smooth video capture when studying upper extremity kinematics in sports like tennis, reinforcing the need for stabilization techniques [12].

B. Camera perspective impact on joint and pose estimation in a 2D model for player body angle detection.

Back View Characteristics

- Visibility of Joints: Typically, while recording the player, it’s the back view of the player that captures the best movements. The back view shows the player’s back, shoulders, arms, and legs, which are crucial for tracking movements along the court’s depth. However, in some clips, we found that important details like the chest, knees, and elbows were missing, especially if the player’s body or racket blocked them. This challenge in capturing full-body movements is discussed in Cao *et.al.* [1], where joint visibility and occlusion are considered major challenges in pose estimation. The research emphasizes how capturing a full range of motions from the back view can be useful, but it requires handling occlusions effectively for accurate tracking [11].

- Occlusion Challenges: Joints like the elbows and wrists can sometimes be hidden behind the body or racket, making it hard to see them clearly. This can confuse the computer vision system, which needs to see all the key points to figure out where the joints are. Cao *et.al*. [3] highlights how occlusion, especially in sports settings, is a persistent problem, leading to incorrect pose estimation when important body parts are blocked. This issue is particularly relevant in tennis, where rapid movements can easily hide key joints [10].
- Accuracy of angle calculation: Measuring angles like shoulder and elbow flexion from a back view in a 2D model was found to be tricky. Without depth information and with slight changes in how the player is facing the camera, the angle measurements can deviate from the actual pose. Chatterjee *et.al*. [7] discuss how the lack of depth data complicates angle estimation in pose-based sports activity classification, particularly when only 2D perspectives are available. This issue is relevant to our study, where slight misalignments or varying player orientations can affect the accuracy of angle calculations, leading to errors in body angle detection [9].

Even with these challenges, the back view is still incredibly useful for measuring how the torso rotates and how the spine is aligned. These details are essential for understanding how well a tennis player is performing strokes like the serve, forehand, or backhand. Research such as Elliott *et.al*. [8] underscores the importance of biomechanical insights gained from the torso's movement in sports performance, showing that, despite the limitations of a 2D model, the back view provides crucial information about body alignment during tennis strokes [8].

Experimental set up to overcome the challenges

Addressing these challenges typically involves a combination of computer vision algorithms, machine learning techniques, and hardware optimizations. In our experimental set up we recursively arrived at an optimum set up, with the following as [7]

Camera alignment and configuration	**Actuals in Experiment**
Distance	8 ft from the baseline
Position	Aligned to Centre line
Ground height	5 ft
Field of view	Standard - no zoom position
Frames per second	30 FPS
Lighting	Standard daylight conditions
Resolution	full HD (1080p)
Video boost	Off
Format	H.265/HEVC (instead of H.264/AVC)

Table 1. Experimental Physical Set-up

The challenges were overcome effectively by using a combination of technology and practicality. Following table illustrates the solutions arrived at during this study to overcome the challenges [6]

Challenges	**Solutions**
Signal-to-Noise Ratio (SNR) and Resolution	Phone models with Higher SNR and R were used. (Google Pixel 8 Pro and I-Phone 15) [1, 3]
Frame Rate and Motion Blur	30 FPS setting used on the cameras [2]

Limited Field of View (FoV)	FoV was contained by using a measured camera alignment. [1]
Lighting Conditions	Videos were taking in daylight conditions avoiding this challenge [2][7]
Processing Power	Video analysis was done on cloud platforms (Google Collab & Azure)
Data Privacy	All Analytics was performed after masking and Part Affinity Field extraction [1]
Heterogeneity in Mobile Platforms	2 Popular platforms namely Android and *i*-OS used for data collection
Camera Placement and Orientation	Orientation errors were contained by using a measured camera alignment.

Table 2. Settings for the experiment

Methodology

In tennis, how a player positions his/her feet and body—the stance—plays a huge role in how well they perform. Studies done so far conclude on the following as important features for a good stance [5]

- Balance and Stability: A solid stance provides the player with a stable base, helping them maintain balance even during powerful shots. A wider, lower stance provides additional stability, particularly for fast movements or high-impact hits, reducing the chance of mistakes. Research by Elliott *et.al.* [8] emphasizes the importance of a stable stance for maintaining balance and preventing injury during intense athletic performance [4].
- Power Generation: The correct stance is essential for generating power. By positioning their feet appropriately, a player can better utilize their lower body, transmitting energy from the ground, through their legs and hips, and into the racket. This smooth chain of movement works most effectively when the stance complements the shot being played. Studies by Boonim [4] highlight the role of lower body positioning and its impact on generating power, particularly during high-speed movements in tennis [3].
- Footwork and Movement: A good stance makes it easier for players to move quickly and efficiently around the court. Whether the player needs to step sideways, forward, or backward, the correct stance enables faster reactions and positioning for optimal shot execution. Kovalchik and Reid [9] discuss how stance and footwork are critical for facilitating quick adjustments and accurate positioning during matches, noting the direct connection between stance and movement efficiency [2].
- Recovery: After hitting a shot, a proper stance helps the player recover quickly and get ready for the next one. This is especially crucial in fast-paced matches, where every moment counts. The importance of efficient recovery is outlined in Knudson [6], which discusses how posture and stance affect the recovery phase and how the body’s alignment supports swift transition movements in tennis [1].

Different shots require different stances, so mastering them is crucial. For example, an open stance might be used for a fast forehand, allowing for quick preparation, while a closed stance might be better for precise, angled shots. By learning and using the correct stance for each shot, a player can execute a wide range of moves effectively. This also suggests the possibility of grouping shots into further subtypes—a concept we explored and experimented with in this study. Chatterjee *et.al.* [7] have previously examined how stance variations can influence the categorization and classification of different tennis shots, offering valuable insights for this experiment.

Stance identification – Forehand, backhand and serve

Identifying tennis stances using pose estimation is an interesting application of computer vision and can be achieved through various techniques. Let's discuss how we approached the task of classifying tennis stances into forehand, backhand, volley, and serve using mathematical concepts and techniques:

1. **Data Collection**

Collected a dataset of tennis players in various stances, capturing video frames.

A total of 1077 video clips split into six different sets of data frames (video clips) were used for the analysis:

Frames	Right hand	Left hand
Forehand	290	123
Backhand	172	160
Serve	190	142

Table 3. Tennis videos - Collected Data

2. **Preprocessing Module**

Frame Extraction: Videos are split into individual frames to facilitate detailed analysis. We used photometric invariance or color space transformations methods through custom algorithms for histogram equalization as an initial step before processing the feed. Pose detection algorithms, implemented using TensorFlow, process each frame to extract key-points of player poses.

For this, the function '*blobFromImage*' is used. In which the below 2 functions are executed:

   1. Mean subtraction
   2. Scaling

To start with, we calculate the average pixel intensity across all extracted frames in the training set for each of the red, green, and blue channels. This implies that we end up with three variables, viz., μ_R, μ_G, μ_B. Further, reduce the variable value by mean, mu, from each input channel of the input image and add a scaling factor, sigma, for normalization as outlined in Chatterjee *et.al.* [7] as

$$R = \frac{(R - \mu_R)}{\sigma}$$

$$G = \frac{(G - \mu_G)}{\sigma}$$

$$B = \frac{(B - \mu_B)}{\sigma}$$

3. **Object Detection Module**

For the Racket & Ball Detection, we used object detection techniques to locate and track the tennis racket and ball across frames which is crucial for shot classification. We processed images extracted from the video to detect objects and calculate distances between the detected ball and racket using pixel co-ordinates of the ball and the racket. When the co-ordinates of the ball are contained within the co-ordinates of the racket, its assumed to have made a contact. The Euclidean distance between center of the racket and the ball Racket (x_1, y_1) and Ball (x_2, y_2) in a 2D plane is calculated using

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Frames are saved around detected shot points (point of contact) to separate video clips using the time stamps. $fps = \frac{1}{(current\ time - previous\ time)}$. This allows us to extract key frames for shot classification, as described in Wang *et.al.* [3]

4. **Data Integration**

Pose Key Points to Frame Coordinates, we converted pose detection key-points into usable frame coordinates for accurate analysis of player movements. A TensorFlow Lite model for pose estimation using the MoveNet architecture.

$$Each\ key - point = \left(\max\ \_loc[0] * \left(\frac{w}{heatmap.shape[n]}\right), \max\ \_loc[n]\right), \left(\max\ \_loc[0] * \left(\frac{w}{heatmap.shape[n-1]}\right)\right)$$

Once the key-points were identified on the frame, the key-points were paired and connecting lines between them are visualized based on the predefined skeleton as

```
pt1 = plimage.keypoints[pair[0]]
pt2 = plimage.keypoints[pair[1]]
cv2.line(frame, (int(pt1[0]), int(pt1[1])), (int(pt2[0]), int(pt2[1])), (0, 255, 0), 2).
```

This pose-extractor component facilitates the detection and visualization of human poses in video frames. It uses a pose detection model to identify key body points and draw a skeleton on the video frame, enabling visual analysis of body movements and postures.

5. **Shot Classification Module**

We had to identify the shot played by the player before getting to detailed analysis of the shot. We used a CNN algorithm from tensor-flow to extract spatial features from each frame, and then fed them to the RNN to capture spatial dependencies. The code can be shown as

$$F_t = CNN(x_t)$$

The image at point of contact was used for classification, and to extract features like key-point coordinates from multiple frames in the stored list. Using the frames extracted a RNN network is used to train on the manually annotated tennis shots. The hidden state (h_t) at time step (t) is given by:

$$h_t = \sigma(W_{xh}x_t + W_{hh}h_{t-1} + b_h)$$

where:

- x_t is the input at time step (t)
- h_{t-1} is the hidden state from the previous time step
- W_{xh} and W_{hh} are weight matrices
- b_h is the bias term
- σ is the activation function

Finally, the RNN is processed using $h_t = RNN(F_t, h_{t-1})$

This was used to examine the sequences of frames to classify different types of tennis shots, such as serves, forehands, and backhands. Since these were only 3 stances of interest in this study.

Conclusions

The modelling of the qualification & improvement of tennis stance for player performance improvement using 2D analysis of videos taken from a mobile camera was developed in this paper. The comprehensive study demonstrated a significant stride in sports performance technology. By leveraging advanced 2D video analysis combined with sophisticated computational methods such as Recurrent Neural Networks (RNNs) and pose estimation algorithms, the research has laid a foundational framework for enhancing tennis training and performance evaluation. This model skillfully extracts and analyzes skeletal keypoints to compute joint angles and classify player movements, providing insights that are vital for both coaching tactics and player development. Through the use of unsupervised clustering techniques like k-means, the study categorizes complex movement patterns, enhancing the precision of biomechanical assessments. This categorization allows for the identification of optimal movement patterns and helps in pinpointing deviations that could potentially lead to performance inefficiencies or increased risk of injury. Adaptive acceptance areas, determined by an array of distance metrics, further refine the precision of this analysis, making the evaluation of player stances both rigorous and nuanced.

Moreover, the integration of high-SNR imaging devices and advanced stabilization techniques addresses common challenges faced in video analysis, such as motion artifacts and variable lighting conditions. These technological enhancements ensure the capture of high-quality data, which is crucial for the reliability of the computational analysis performed on cloud platforms. This not only facilitates robust kinematic analysis but also ensures the confidentiality and security of the data processed. The real-time feedback mechanism, powered by TensorFlow Lite and part affinity fields, transforms the extracted data into actionable insights that can be directly applied on the field. This aspect of the study is particularly groundbreaking as it provides coaches and players with immediate feedback on biomechanical alignment, helping to make instantaneous corrections to stances and techniques. This real-time analysis is poised to revolutionize tennis coaching and player performance by allowing for adjustments that are both immediate and based on precise biomechanical data. In conclusion, this research sets a new benchmark in the field of sports performance analysis. By combining cutting-edge video analysis technologies with advanced mathematical modeling, the study not only enhances the understanding of tennis biomechanics but also pushes the envelope in applying data-driven insights to sports training. The implications of this study are vast and varied, promising a future where technology and sports training merge seamlessly to enhance athletic performance across the board. The model’s effectiveness in improving tennis stances and overall player performance could potentially be adapted to other sports, marking a pivotal shift in how athletic training and performance improvement are approached.

References

[1]. S.E. Wei, V. Ramakrishna, T. Kanade, et al., “Convolutional Pose Machines,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 38, no. 12, pp. 4724-4732, 2016.

[2]. Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, “OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172-178, 2021.

[3]. K. Boonim, “Analysis of Playing Positions in Tennis Match Videos to Assess Competition Using a Centroid Clustering Heatmap Prediction Technique,” *J. Advances in Information Technology*, vol. 14, no. 1, Feb. 2023.

[4]. T. Lertwonghattakul, S. Sriramatr, and P. Rachnavy, “Analysis of Kinetic Chain Mechanism Affecting Energy Flow in Kick Topspin Tennis Serve,” *Annals of Applied Sport Science*, vol. 11, no. 2, p. e1133, 2023.

[5]. D. V. Knudson, “Intrasubject Variability of Upper Extremity Angular Kinematics in the Tennis Forehand Drive,” *International Journal of Sport Biomechanics*, vol. 5, pp. 243-252, 2006.

[6]. R. Chatterjee, S. Roy, S. K. H. Islam, and D. Samanta, “An AI Approach to Pose-based Sports Activity Classification,” *Proc. 8th Int. Conf. on Signal Processing and Integrated Networks* (SPIN), 2021, pp. 397-402.

[7]. B. Elliott, S. Choppin, S. R. Goodwill, and T. Allen, “Markerless tracking of tennis racket motion using a camera,” *Procedia Engineering*, vol. 72, pp. 344-350, 2014.

[8]. S.A. Kovalchik and M. A. Reid, “A shot taxonomy in the era of tracking data in professional tennis,” *Journal of Sport Sciences*, vol. 36, pp. 2096-2104, 2018

[9]. C.-Y. Wang, K.G. Lai, and H.-C. Huang, “Tennis player actions dataset for human pose estimation,” *Data in Brief*, vol. 55, p. 110665, 2024.

[10]. M. Kwon, J. C. Won, and S. K. Lee, “Biomechanical analysis of tennis strokes and their relation to injury prevention,” *Journal of Sports Science & Medicine*, vol. 19, pp. 122-130, 2020.

[11]. Y. Kuo, P. L. Liu, and J. F. Chang, “Biomechanical analysis of the shoulder during different types of tennis serves,” *Sports Biomechanics*, vol. 11, no. 3, pp. 254-267, 2019.

[12]. Y. Kuo and S. Yang, “The role of lower body mechanics in tennis stroke performance,” *Journal of Applied Biomechanics*, vol. 35, pp. 112-119, 2018.

[13]. H. Lee and J. Lee, “Knee and ankle angle analysis in tennis performance: A kinematic approach,” *Journal of Sports Engineering and Technology*, vol. 33, no. 2, pp. 115-124, 2021.

[14]. M. Zhukov, E. Zaharov, and S. Park, “Real-time 2D pose estimation using Mediapipe: Applications and advancements,” *Proceedings of the International Conference on Computer Vision* (ICCV), 2020, pp. 137-146.

[15]. L.M. Gualtieri, V. Lippiello, and M. Siciliano, “Mediapipe-based pose estimation for tennis biomechanics analysis,” *Journal of Sports Biomechanics*, vol. 22, no. 4, pp. 210-221, 2021.

[16]. J.R. Patel, K. R. Nair, and S. Kumar, “Optimization of 2D human pose estimation using Mediapipe for sports movement analysis,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 258-267, 2021.

[17]. C.S. Tan, L. K. Wong, and M. P. I. Chia, “Advanced body motion analysis for sports using Mediapipe: Challenges and solutions,” *Sports Engineering*, vol. 34, no. 3, pp. 182-193, 2022.

[18]. T. Allen, N. Elliott, and S. R. Goodwill, “Markerless tracking of tennis racket motion using a camera,” *Procedia Engineering*, vol. 72, pp. 344-349, 2014.

[19]. J.K. Lee, H. S. Kim, and S. S. Choi, “Biomechanical analysis of forehand strokes in tennis,” *Journal of Sports Science and Medicine*, vol. 19, no. 2, pp. 112-118, 2018.

[20]. M.J. Seitz and J. R. Müller, “Biomechanical comparison of one-handed and two-handed backhand strokes in tennis,” *Journal of Sports Sciences*, vol. 36, no. 8, pp. 824-832, 2018.

[21]. C.L. Barris and M. J. Button, “Biomechanics of the tennis backhand: A comparison of the one-handed and two-handed backhand,” *Journal of Sports Biomechanics*, vol. 12, pp. 123-133, 2017.

[22]. A.O. Reilly and J.M. Bender, “Stance mechanics and efficiency in backhand strokes in tennis,” *Sports Biomechanics*, vol. 14, no. 3, pp. 371-379, 2015.

[23]. J.V. Keogh, “Tennis biomechanics: Optimal movement for the backhand,” *International Journal of Sports Science & Coaching*, vol. 13, no. 1, pp. 67-74, 2018.

[24]. B.R. Pizzari, “A comparison of stance angles in tennis: Effect on power and control,” *Journal of Strength and Conditioning Research*, vol. 22, no. 6, pp. 1653-1658, 2016.

[25]. G.L. Lees and H.T. Davies, “Biomechanics of the tennis serve: A comparison of stance and shot characteristics,” *Journal of Sports Sciences*, vol. 21, no. 7, pp. 641-648, 2020.

[26]. D.D. Chan and P.F.A. Lee, “The influence of body stance on serve effectiveness in tennis”, *Journal of Applied Biomechanics*, vol. 19, no. 5, pp. 402-409, 2018.

[27]. L.A. Reilly and P.K. Thomas, “Tennis biomechanics: Stance and shot execution during the serve,” *Sports Biomechanics*, vol. 14, no. 3, pp. 287-298, 2019.

[28]. A.S. Martin and R. J. Cohen, “Stance optimization for accuracy and power in tennis serving,” *International Journal of Sports Science & Coaching*, vol. 9, no. 1, pp. 79-87, 2018.

[29]. K. Boonim, “Analysis of Playing Positions in Tennis Match Videos to Assess Competition Using a Centroid Clustering Heatmap Prediction Technique,” *Journal of Advances in Information Technology*, vol. 14, no. 1, 2023.

[30]. T.M. Ghazal, M.Z. Hussain, “Performances of K-Means Clustering Algorithm with Different Distance Metrics,” *Intelligent Automation & Soft Computing*, vol. 27, no. 1, pp. 83-92, 2021.

[31]. Jorge Santana, Maria Luiza Neto, “A Comparative Study of Distance Metrics in High-Dimensional Clustering”, *Journal of Computational Science and Technology*, 10(4), 2021.

[32]. Satish Kumar, Meenakshi Srivastava, “The Use of Manhattan Distance for Multivariate Data Clustering”, *Applied Mathematical Sciences*, 9(1). 2020

[33]. David Marquina, Albert Ros, “Minkowski Distance and Its Applications in Machine Learning Algorithms”, *International Journal of Data Analysis Techniques*, 7(2), 2019

[34]. Zachary Shelly, Reuben F. Burch, Wenmeng Tian, Lesley Strawderman, “Using K-means Clustering to Create Training Groups for Elite American Football Student-athletes Based on Game Demands”, *International Journal of Kinesiology & Sports Science*, ISSN: 2202-946X

[35]. Peng X., Tang L., “Biomechanics analysis of real-time tennis batting images using Internet of Things and deep learning”, *J Supercomput*, 78, 5883–5902 (2022).

[36]. Wang J., Zuo L. & Cordente Martínez C., “Basketball technique action recognition using 3D convolutional neural networks”, *Sci Rep*, vol. 14, 13156, 2024.

[37]. Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, “Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields,” *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR), 2017, pp. 7291-7299.

[38]. N. Lakshmi, G. Pavithra and T. C. Manjunath, “CMOS Implementation of Multipath Fully Differential OTA with Dual Flipped Voltage Follower in 50 nm and 75 nm CMOS Technologies using Cadence Tool,” *2024 IEEE ICDCOT*, 2024, pp. 1-8. https://doi.org/10.1109/ICDCOT61034.2024.10515482

[39]. T.C. Manjunath, G. Pavithra and B.G. Nagaraj, “Design & simulation of the workspace for a stationary robot system,” *2016 IEEE Region 10 Humanitarian Technology Conference* (R10-HTC), Agra, India, 2016, pp. 1-5, https://doi.org/10.1109/R10-HTC.2016.7906828

[40]. Pritosh Tomar, Prateek Chauhan, Ashwani Kumar, Anoop Kumar Shukla, Víctor Daniel Jiménez Macedo, Bajarang Prasad Mishra, Nageswara Rao Lakkimsetty, Tadaga Channaveerappa Manjunath, “Numerical Investigation of Thermal Performance Enhancement of Solar Reservoir using Flash Cycle”, *Scopus Indexed Q3 Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, Volume 123, No. 1, pp. 197–221, ISSN: 22897879, Nov. 2024 https://doi.org/10.37934/arfmts.123.1.197221