On-Device Real-Time Analysis of Martial Arts Techniques Using Optimized Pose Estimation

***Dr.Bhagyashree Dhakulkar**
Assistant Professor
Department of Artificial Intelligence and Data Science
Ajeenkya DY Patil School of Engineering (SPPU)Pune, India
bhagyashree.dhakulkar@gmail.com

Alokraj Ahire
UG Student
Department of Artificial Intelligence and Data Science
Ajeenkya DY Patil School of Engineering (SPPU)Pune, India
alokrajahire@gmail.com

Piyush Gangurde
UG Student
Department of Artificial Intelligence and Data Science
Ajeenkya DY Patil School of Engineering (SPPU)Pune, India
akaom07@gmail.com

Abstract-This project introduces an AI assistant designed to enhance the training and performance of Mixed Martial Arts (MMA) athletes. By analyzing real-time and video-based body movements, the system helps athletes refine their techniques, improve accuracy, and gain a strategic advantage over opponents. The assistant analyzes posture, speed, and precise movements to provide actionable insights for strategy planning. Advanced analytics further optimize performance efficiency during matches. Traditionally, MMA athletes and coaches relied on manual methods to analyze fights, such as handwritten notes and limited footage, which often resulted in incomplete or delayed insights. This system addresses these limitations by automating the analysis process. It uses hybrid model which includes deep learning models to detect patterns, techniques, and weaknesses in both the athlete's and their opponent's performances and media pipe for pose estimation. Real-time feedback during training allows for immediate adjustments, while predictive analytics forecast fight outcomes based on historical data and performance. The application also focuses on injury prevention by identifying risks like poor technique or overtraining. Personalized feedback helps athletes refine their movements and adopt better strategies. The system is designed to be user-friendly, ensuring accessibility for athletes and coaches. By integrating video analysis with AI-driven insights, this tool bridges the gap between raw data and practical improvements, offering a modern solution to enhance MMA training and competitive performance.

Keywords- Performance enhancement, Real-time analysis, Movement tracking, Strategy development, Injury prevention.

I. INTRODUCTION

Recent years have witnessed significant growth in human activity recognition studies, with researchers exploring diverse sensing modalities and developing innovative computational methods for activity modeling and classification. This expanding body of work has yielded numerous technical approaches for accurately identifying and interpreting human movements[1]. Before the advent of modern technology, MMA athletes and their coaches relied on traditional methods to analyze performance, prepare for fights, and gain a competitive edge. Coaches and cornermen would take handwritten notes during fights to track opponent's techniques, patterns, and weaknesses. Access to fight footage was limited, making it difficult for athletes to study their opponent's techniques and strategies. There were no online forums, tutorials, or instructional videos to learn new techniques or gain insights from other experts. Coaches relied on their own experience, knowledge, and instincts to develop strategies and make decisions during fights. Athletes had to rely on their instincts, reflexes, and experience to make split-second decisions during fights. It is not that beneficial to the performance of the player though it is important but, due to technology and automation a lot of things get shifted from manually to technical. So does the athletes and their planning methods and strategies. By adapting to those new things, we can get benefited a lot in increasing efficiency and performance. The application will help in video-based analysis of the players move and strategies. AI driven assistant will predict a move from that video and real time updates and will provide better understanding. Using AI, Media pipe and the deep learning models we can achieve various profits. AI can analyze fight footage to identify patterns, techniques, and weaknesses in both the athlete's and their opponents' performances.AI-powered predictive analytics can forecast the likely outcome of a fight based on historical data, athlete performance, and other factors.AI-powered video analysis can provide detailed analysis of an athlete's technique, identifying areas for improvement and providing personalized recommendations. AI-powered video analysis can analyze an opponent's strengths, weaknesses, and tactics, providing valuable insights for strategy development also it can identify potential injury risks, such as overtraining or poor technique, enabling coaches and trainers to take proactive measures and can provide personalized feedback to athletes, helping them adjust their technique and improve performance. Ensuring that this application will be more efficient and user friendly.

Combat skill acquisition occurs through:

a) Kata: Solo pattern practice (cognitive encoding)

b) Kihon Kumite: Partnered pressure training (contextual adaptation)

This progression bridges theoretical technique and practical application.[1]

This paper [2]presents a machine learning framework for predicting UFC bout outcomes using the organization's comprehensive historical database. By strategically excluding direct performance indicators, we develop classifiers achieving 80.3-92.1% accuracy across validation sets. The model's architecture permits expansion for future matchup forecasting in MMA's dynamic competitive landscape. Artificial intelligence has emerged as a transformative interdisciplinary field, combining computer vision, machine learning, and data science to revolutionize martial arts training and analysis. This paper systematically reviews AI applications in martial arts, examining: style recognition techniques, training task automation, multimodal data acquisition methods, and (4) algorithmic innovations. Through comprehensive analysis of current research, we present a unified framework for intelligent martial arts systems, outlining the complete technological pipeline from data collection to performance evaluation. Our synthesis reveals significant progress in movement quantification, real-time feedback systems, and personalized training protocols enabled by AI advancements[3].

Human Activity Recognition (HAR) systems have become increasingly vital in modern applications due to their capacity to extract meaningful behavioral patterns from raw sensor inputs. By automatically identifying and classifying human movements, these systems enable deeper understanding of physical activities, facilitating advancements in fields ranging from healthcare monitoring to smart assistive technologies[4].

II.LITERATURE SURVEY

Sports entertainment now drives massive commercial value, pushing researchers to explore AI and machine learning for analyzing sports data. Over the past ten years, studies have increasingly focused on breaking down sports media content. Today's sports analytics deals with huge, varied datasets that many can access. The biggest challenge? Quickly pinpointing the most useful insights in this flood of information[5].

According to this paper [6]Blaze Pose is a streamlined AI model designed for fast human pose detection on smartphones. It identifies 33 body joints in real-time (30+ FPS on mid-range devices) using a smart blend of heatmaps and coordinate prediction.

Video sharing platforms now handle enormous volumes - over 300 new hours of content every minute. While this creates valuable data opportunities, the scale makes human-led analysis impractical, requiring automated solutions[7].

The field of 6D object pose estimation has gained significant traction, particularly for applications in autonomous navigation and robotic manipulation systems. While deep learning methods have become the predominant approach, there remains a critical need for comprehensive evaluation of contemporary neural architectures and their relative performance advantages across different use cases[8].

Spotting human actions has become incredibly important these days, with uses ranging from keeping public spaces safe to creating smarter video games. New tech advances now let us track movements from any angle, even in tricky conditions. Our system can pick up on these action patterns by analyzing how people move through space and time[9].

Our [10] approach combines three smart techniques to predict MMA outcomes:

1. We first identify key fighting styles by grouping similar fighters using their technical moves (K-means clustering)
2. We then test multiple AI models - including Random Forests, Neural Networks, and XGBoost - to see which predicts best
3. Finally, we combine all models' votes for smarter predictions

When testing with real UFC data, our combined system reached 65.5% accuracy - better than any single model. Detailed tests also proved that analysing fighting styles truly helps predictions. This gives both accurate results and clear insights into what makes fighters win[10] .

Statistical analysis of 4,129 MMA decisions (2003-2023) reveals 97.53% outcome concordance between standard round-aggregated scoring and judge-consensus methods, suggesting current MMA judging practices yield consistent results regardless of aggregation approach when applying the 10-Point Must System[11].

Metadata—the "data about data"—powers nearly every digital tool we use daily. Whether you're streaming music on Spotify, sharing Instagram photos, watching YouTube videos, managing finances in Quicken, or texting friends, metadata works behind the scenes. It includes details like creation dates, titles, tags, and descriptions that help systems organize content and users find what they need. This invisible layer is what makes searching, sorting, and sharing possible across all our apps and devices[12].

Single-camera football tracking systems struggle with partial coverage and player re-identification errors. Our multi-camera solution synchronizes feeds from stadium-mounted cameras, using cross-view correlation to maintain uninterrupted tracking and accurate player identification throughout matches[13].

We employ Long Short-Term Memory (LSTM) networks - a specialized recurrent neural architecture optimized for sequential data - to process mobile sensor time-series signals through a hybrid model where a dual-layer LSTM first captures long-term temporal dependencies in the raw sensor data, followed by convolutional blocks that extract localized spatial features from the LSTM outputs, enabling the combined modeling of both time-evolving patterns and their spatial relationships within the sensor readings[4].

III.METHODOLOGY

A. Research Design-This study aims a qualitative approach to explore the development of an AI assistant for Mixed Martial Arts (MMA). The system aims to analyze and predict opponent's movements and provide actionable suggestions for counterattacks or improvements. The research utilizes a hybrid deep learning model, integrating pose estimation techniques and computer vision methods to detect and analyze fighters moves, weaknesses, and techniques. The model is made in such a way that it will analyze fight data and provide real-time feedback.

B. Data Collection-As there is lack of specific datasets for MMA fighters' poses and weaknesses, the system is trained on historical fight records and publicly available data from previous MMA matches such as video recordings and etc. Data Sources: The data is gathered from online fight records (e.g., fighter statistics, match outcomes, previous fight video footage). However, the quality of this data may be limited, and additional preprocessing will be required to enhance its relevance. Pose and Weakness Detection: For this project, we will focus on two MMA fighters. The AI assistant will analyze the movements of fighters and suggest counter-techniques, with a special focus on techniques like strikes, blocks, punches and submission attempts. Media Pipe is used for posing estimation, allowing the AI system to track key body joints and movements, extracting precise data on posture and body orientation during different fight phases.

C. Data Analysis Techniques-The core of the analysis relies on the combination of multiple technologies to extract meaningful results from the collected fight data:

D. Data Ingestion Layer - Detailed Overview

Input: Fight Clips with Metadata.
1.Fight Clips: The system allows fight clips to be uploaded in two ways, either manually by the users or retrieved from external sources such as online fight databases, sports archives.
2.Metadata: Each fight video comes with accompanying metadata, which includes essential information such as:

-Fighter Details

-Match Information

-Previous Stats

-This metadata is used not only for context but also for organizing the data and helping the AI correlate specific performance patterns across different matches or fighters.

3.Processing: OpenCV for Frame Extraction, Resizing, and Normalization

3.1 Frame Extraction:

a. OpenCV is used to break the fight video into single frames, usually at a specified frame rate (e.g., 30 fps). this is essential for detailed movement analysis.

-The system processes each frame to keep track on fighter's poses over time. By observing single frames, it becomes easier to perform pose estimation for each individual movement.

b. Resizing:

-Fight clips can be in different resolution. To ensure uniformity and speedy processing, OpenCV resizes the frames to a fixed dimension (e.g., 224x224 pixels).

c. Normalization:

-Frames are normalized to ensure uniform pixel intensity (no of pixels per unit area) across all images. Pixel values might be rescaled to a range between 0 and 1.

-Normalization helps in avoiding problems that could arise due to lighting conditions or other variations in the video. It also improves the accuracy and performance of models like CNN and LSTM by standardizing the inputs.

LSTM (Long Short-Term Memory) and CNN (Convolutional Neural Networks) are the main models used in the ML Model-Layer to analyze and predict temporary sequences and spatial movement patterns. The training phase includes learning from fight video datasets to understand the nuances of movement, timing, and techniques and accuracy.

A. LSTM (Temporal Analysis):

- LSTM is very useful for capturing the temporal dependencies in a fighter's movements. Fighting is sequential, where one action often leads to another, such as punches leading to blocks or punches leading to counters.

- The LSTM model is trained on sequences of data (successive actions after specified actions).

- The model learns to predict by analyzing the next action in a sequence. For instance, after a punch, the model predicts the next possible move can be a block, strike, or dodge

-Example: The model might predict that after a right-hand jab, the opponent will likely follow up with a cross.

B.CNN (Spatial Pattern Recognition):

- CNNs focus on spatial relationships between body parts at any moment, analyzing the patterns of the fighter's body in each frame.

- The CNN is trained on labeled video frames with actions like punches, kicks etc, to recognize specific spatial features, such as the position of hands, feet in different movements.

- The model learns to differentiate different techniques like punches, kicks, and defensive movements based on spatial qualities such as body posture and joint alignment.

-Example: The model might recognize that the fighter is punching the left jab based on the position of the shoulder, elbow, and hand.

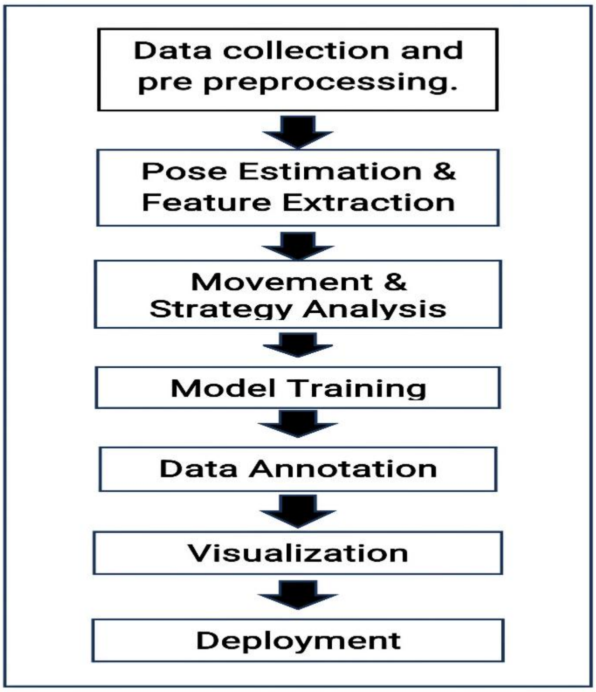
-Real-Time Action Detection: It can also identify defensive movements like blocking or evading an attack, by recognizing body postures like arms being raised to defend the head.

I. Pose Estimation-We use Media Pipe for accurate pose estimation and determination. Media Pipe helps track the key body joints of fighters (elbow, knee, finger joints, shoulder, waist, etc.), allowing the system to identify movements such as punches, kicks, dodges, and guard positions. This data is crucial for understanding the fight dynamics and providing tailored suggestions.

II. OpenCV- To enhance the visual recognition of fighter movements, OpenCV will be used for image and video processing. This includes frame-by-frame analysis of fight footage to track and compare movement patterns across different match scenarios.

III.Performance Metrics- The effectiveness of the AI assistant will be assessed through real-time analysis of the fighter's movements during simulated fights. The assistant's ability to detect weaknesses (e.g., exposed limbs, unguarded positions) and offer counter-strategies will be evaluated by tracking the accuracy and relevance of the suggestions provided.

D.System Workflow- Our system processes MMA videos frame-by-frame:



3.1 System Workflow

E.Algorithm

1.Start:Initialize the system resources such as libraries etc.

2.Data Collection:Gather fight videos from online sources such as YouTube, fight databases , etc. This data is then used for training the system.

3.Video Preprocessing:Extract frames from the videos at a consistent frame rate (e.g., 30 frames per sec).Resize the frames to a standard resolution (e.g., 224x224 pixels) for consistency.Normalize pixel values to a range [0, 1] to prepare the data for the machine learning models.

4.Pose Estimation:We Use MediaPipe for pose estimation, detecting key body joints (shoulders, elbows, knees, etc.) in each frame.Generate 2D coordinates of these landmarks(joints) for further analysis.

5.Feature Extraction:Calculate joint angles, speed, and distance between body parts (e.g., fist velocity during a punch).Track movement patterns across frames, identifying key features like combos or defensive movements.

6.Data Annotation & Model Training:Labeling the movements and techniques (e.g., jab, uppercut, hook, block, takedown) in the data we have extracted .Use Long Short-Term Memory (LSTM) to model temporary patterns in movement sequences .Use Convolutional Neural Networks (CNN) to analyze spatial patterns in individual frames (e.g., recognizing body posture or attacking type).

7.Movement & Strategy Analysis:Analyze movement patterns, detecting techniques and possible mistakes (e.g., dropped guard, poor posture).Identify opponent weaknesses by analyzing repetitive movements or openings in their defence.Provide feedback on improving technique and suggest counter-strategies based on detected weaknesses.

8.Visualization:Create heatmaps showing areas of vulnerability such as exposed head or torso in the opponent's stance.Display movement trajectories to visualize the path of punches, kicks, jabs, hook or other attacks.Overlay suggested corrections, such as raising the guard or adjusting positioning.

9.Deployment:Deploy the AI Assistant for real-time feedback during sparring sessions or post-fight analysis.Provide a user interface for coaches and fighters to interact with the system for detailed feedback.

10.End:Conclude the analysis and save the results for future training or improvements.

IV.System Architecture Layer

1. Data Ingestion Layer:Input: Fight clips with metadata.Preprocessing: OpenCV for frame extraction, resizing, and normalization.

2. Pose Detection & Tracking:Pose Estimation: Media Pipe is used for 2D body landmark (joints) detection to track key body points.Tracking: Continuous tracking of body parts is done across frames.

3. Feature Extraction Layer:Features: Calculation of joint angles, speed, distance between body parts, and movement patterns.

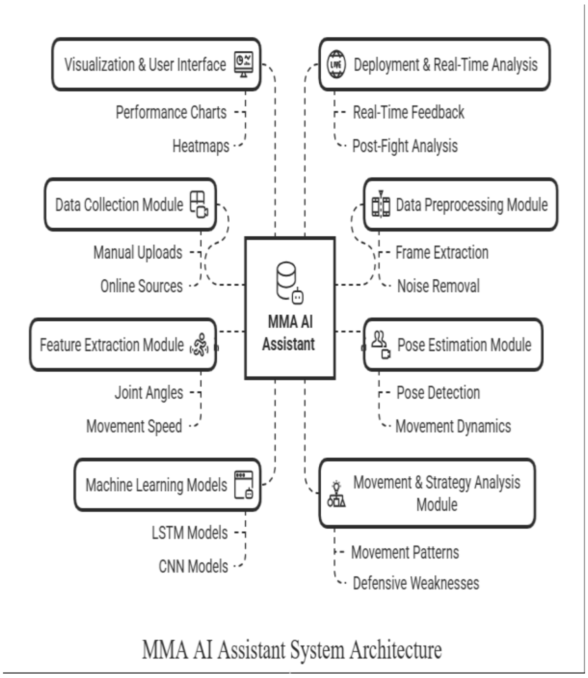
Action Recognition: Identifying offensive and defensive actions (punches, kicks, blocks) based on pose data.

4. ML Model Layer (LSTM & CNN):Training: LSTM for temporal analysis and CNN for spatial pattern recognition models trained on fight video datasets.Inference: The models analyze and predict temporal sequences and spatial movement patterns.

5. Strategy & Feedback Layer:Feedback Generation: Tactical insights on movement execution, timing, and strategy .Counter-Strategy: Suggest opponent weaknesses and corresponding strategies based on predictions.

6. Visualization & Interaction:Dashboard: Web interface (React/Angular) displaying interactive visualizations: performance charts, movement trajectories, and heatmaps.Visualization Libraries: D3.js/Plotly for data visualization.

7. Deployment & Real-Time Analysis:Cloud-based Deployment: Hosted on AWS or Google Cloud for scalability and accessibility.Real-Time Feedback: Real-time analysis of sparring videos or matches with immediate insights.



4.1 MMA AI Assistant System Architecture Diagram

V.CONCLUSION

The MMA AI Assistant presents a transformative approach to training and performance analysis in mixed martial arts by utilizing advanced AI technologies, including MediaPipe for pose estimation and LSTM and CNN for movement and technique analysis. This system offers detailed, real-time feedback, enabling fighters and coaches to assess technique execution, timing, and defence strategies, ultimately leading to enhanced performance and tactical decision-making.By using large datasets of fight videos, the system detects and analyzes various movements such as joint angles, speed, and attack/defense patterns, generating actionable insights. The ability to provide real-time analysis during training and post-fight analysis ensures that fighters can continuously refine their skills based on data-driven feedback.While the system helps in improving fighters strategies and overall training, challenges such as data quality and generalization across diverse fighting styles remain. However, with continued development and the integration of more diverse datasets, this AI-driven approach has the potential to revolutionize how coaches and fighters prepare for matches, offering personalized feedback and strategic recommendations.In conclusion, the MMA AI Assistant stands as a powerful tool in modern combat sports, with the potential to reshape how performance is analysed and optimized, making it an invaluable resource for both training and competition preparation.

VI . ACKNOWLEDGMENT

The authors would like to exprss their sincere gratitude to our mentor, Dr. Bhagyashree Dhakulkar, for her constant guidance, support, and invaluable insights throughout this project.Her insightful suggestions and constructive feedback significantly contributed to the successful completion of this research. Her expertise in artificial intelligence and continuous encouragement played a pivotal role in the development and success of this work. We are deeply thankful for her mentorship, which helped us navigate challenges and refine our approach. This project would not have been possible without her guidance.

VII .REFERENCES

[1] K. Xia, J. Huang, and H. Wang, "LSTM-CNN Architecture for Human Activity Recognition," *IEEE Access*, vol. 8, pp. 56855–56866, 2020, doi: 10.1109/ACCESS.2020.2982225.

[2] B. Wu and J. Zhou, "Video-Based Martial Arts Combat Action Recognition and Position Detection Using Deep Learning," *IEEE Access*, vol. 12, pp. 161357–161374, 2024, doi: 10.1109/ACCESS.2024.3487289.

[3] N. Ćenanović and J. Kevrić, "Mixed Martial Arts Bout Prediction Using Artificial Intelligence," in *Advanced Technologies, Systems, and Applications VII*, vol. 539, N. Ademović, E. Mujčić, M. Mulić, J. Kevrić, and Z. Akšamija, Eds., in Lecture Notes in Networks and Systems, vol. 539. , Cham: Springer International Publishing, 2023, pp. 452–468. doi: 10.1007/978-3-031-17697-5_36.

[4] Y. Pang, Y. Wang, Q. Wang, F. Li, C. Zhang, and C. Ding, "Applications of AI in martial arts: A survey," *Proc. Inst. Mech. Eng. Part P J. Sports Eng. Technol.*, p. 17543371241273827, Oct. 2024, doi: 10.1177/17543371241273827.

[5] H.-C. Shih, "A Survey of Content-Aware Video Analysis for Sports," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 5, pp. 1212–1231, May 2018, doi: 10.1109/TCSVT.2017.2655624.

[6] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann, "BlazePose: On-device Real-time Body Pose tracking," 2020, *arXiv*. doi: 10.48550/ARXIV.2006.10204.

[7] W. Li, "Deep Learning Based Sports Video Classification Research," *Appl. Math. Nonlinear Sci.*, vol. 9, no. 1, p. 20230029, Jan. 2024, doi: 10.2478/amns.2023.2.00029.

[8] Z. Fan, Y. Zhu, Y. He, Q. Sun, H. Liu, and J. He, "Deep Learning on Monocular Object Pose Detection and Tracking: A Comprehensive Overview," *ACM Comput. Surv.*, vol. 55, no. 4, pp. 1–40, Apr. 2023, doi: 10.1145/3524496.

[9] A. K. Singh, V. A. Kumbhare, and K. Arthi, "Real-Time Human Pose Detection and Recognition Using MediaPipe," in *Soft Computing and Signal Processing*, vol. 1413, V. S. Reddy, V. K. Prasad, J. Wang, and K. T. V. Reddy, Eds., in Advances in Intelligent Systems and Computing, vol. 1413. , Singapore: Springer Nature Singapore, 2022, pp. 145–154. doi: 10.1007/978-981-16-7088-6_12.

[10] J. Yin, "Data-Driven MMA Outcome Prediction Enhanced by Fighter Styles: A Machine Learning Approach," in *2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)*, Zhuhai, China: IEEE, Jun. 2024, pp. 346–351. doi: 10.1109/MLISE62164.2024.10674447.

[11] V. Berthet, “Improving MMA judging with consensus scoring: A Statistical analysis of MMA bouts from 2003 to 2023,” 2024, *arXiv*. doi: 10.48550/ARXIV.2401.03280.

[12] J. Riley, *Understanding metadata: what is metadata, and what is it for*. in NISO Primer series. Baltimore, MD: National Information Standards Organization, 2017.

[13] B. Zheng, “Soccer Player Video Target Tracking Based on Deep Learning,” *Mob. Inf. Syst.*, vol. 2022, pp. 1–6, Jul. 2022, doi: 10.1155/2022/8090871.

[14] Y. Pang, Q. Wang, C. Zhang, M. Wang, and Y. Wang, “Analysis of Computer Vision Applied in Martial Arts,” in *2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, Guangzhou, China: IEEE, Jan. 2022, pp. 191–196. doi: 10.1109/ICCECE54139.2022.9712803.