Psychological Modelling and Action Recognition for Boxing Performance

Vishal Ravishankar
Computer Science Dept.
PES University
Bengaluru, India
email: v1sh4lrr@gmail.com

Itish Raj Shukla
Computer Science Dept
PES University
Bengaluru, India
email: itish2606@gmail.com

Nishat-E-Afroz M
Computer Science Dept
PES University
Bengaluru, India
email: nishatm1717@gmail.com

Anirudh Bhatta J A
Computer Science Dept.
PES University
Bengaluru, India
email: anirudh.bhatta@gmail.com

Surbhi Choudhary
Computer Science Dept.
PES University
Bengaluru, India
email: surbhi.choudhary@pes.edu

***Abstract*— Psychological dynamics in boxing remain largely unaddressed during real-time performance evaluation, despite their significant impact on decision-making and reaction time. Traditional analysis tools focus predominantly on physical performance, overlooking the internal cognitive states of athletes. This gap limits coaches' and psychologists' ability to intervene strategically during or after a match. The Boxing Psychological State Tracking System is a novel framework designed to analyze a boxer's psychological state throughout a fight by correlating physical performance indicators with inferred cognitive conditions. By leveraging action recognition and explainable Artificial Intelligence (AI), the system evaluates events such as knockdowns, strike patterns, and defensive behavior to infer mental states like fatigue, disengagement, and stress. Previous approaches to action recognition in sports have either ignored psychological interpretation or lacked transparency in decision-making, which this work addresses. Our solution involves a lightweight, vision-based system combining You Only Look Once, version 8 - nano variant (YOLOv8n) and SHapley Additive exPlanations (SHAP) to map strike behavior to inferred cognitive states. The system integrates these insights with a visual analytics interface, enabling coaches, sports psychologists, and athletes to understand and improve performance through a deeper awareness of psychological dynamics. Evaluation of the system demonstrates its potential to reveal meaningful psychological patterns using just two high-frequency strikes, providing a practical and explainable method for mental state assessment in combat sports.**

Keywords— Sports psychology; Action recognition; Boxing behavior; Convolution neural network.

I. INTRODUCTION

In competitive boxing, performance depends not only on physical ability but also on psychological resilience. Mental states like fatigue, disengagement, and stress can significantly affect a boxer's decision-making and reaction times, yet tracking these psychological states during a fight remains a difficult challenge. Current tools focus mainly on physical metrics, leaving a gap in real-time psychological analysis. This paper introduces the Boxing Psychological State Tracking System, a vision-based tool that infers a boxer's cognitive state by analyzing performance indicators such as strike patterns, knockdowns, and movement changes.

The primary difficulty in this domain lies in the absence of real-time tools that can interpret mental states from visible behavioral cues, complicating the ability to intervene during a match. By using computer vision models like YOLOv8 and Region-based Convolutional Neural Network (R-CNN) for action detection, and SHAP for interpretability, the system highlights behavioral cues linked to psychological states. Our research is driven by the challenge of mapping psychological states directly to physical indicators observed in real-time boxing matches. Our work investigates three main research questions: Can psychological conditions be inferred from observable fight behavior? Which physical cues best indicate mental fatigue or lapses? And how can these insights be effectively visualized to assist athletes and coaches? The system features a user-friendly interface that displays real-time plots, SHAP overlays, and psychological state indicators, making mental patterns easy to understand and act upon. The explicit purpose of this article is to present a framework that bridges the gap between action recognition and cognitive state inference, providing a novel approach to analyzing performance in combat sports. By bridging action recognition with cognitive modeling, our work contributes a novel and explainable approach to performance analysis in combat sports. One limitation of this approach is that it currently focuses on a limited set of strike types, and further work is needed to account for a broader range of actions and defensive maneuvers.

The paper is organized as follows: Section II reviews related work in the field of action recognition and psychological state tracking. Section III outlines the methodology behind the Boxing Psychological State Tracking System, including the dataset used and the models employed. Section V presents the results, followed by a discussion of the findings in Section VI. Section VII concludes the paper and Section VIII suggests directions for future work.

II. RELATED WORK

This section reviews prior research in the field. While no existing systems have specifically focused on tracking psychological states during boxing matches, our approach leverages advancements in action recognition. Models like YOLO, R-CNN, and other convolutional neural networks have been successfully used to detect and classify actions in video data. We build on these techniques, adapting them to analyze boxing-specific movements and perform psychological analysis in real time.

A. *Boxing Behaviour Recognition based on Artificial Intelligence CNN with sports psychology assistant*

The main intent of the research was to develop a mechanism to recognize boxing form with the help of AI-Convolutional Neural Network (CNN) and also to examine how the state of mind of athletes affects the accuracy and the effectiveness of behavior recognition. The research deployed a mixture of tools like psychological assessment survey and AI technologies to have an insight into fighters' psychological profiling and for the construction of boxing action classification and recognition algorithms. They developed the model using Bidirectional Encoder Representations from Transformers (BERT) fusion 3D-Residual Networks (ResNet) architecture, which provided the emotion conveying info along with action features. The model suggested in this research was very effective compared to the traditional models, the loss value, along with the accuracy and F1 values were improved, where the accuracy reached 96.86% [1].

B. *Deep Learning for Micro-Expression Recognition: A Survey*

In this report, the mission is to conduct a complete survey of Deep Learning (DL) approaches toward Micro-expression Recognition (MER). The purpose of the article is to form a taxonomy for the field that illustrates all aspects of MER based on deep learning. It will be integrated with existing linguistic datasets and deep learning methodology, and also compare performances of the key DL methods. The paper makes a quick review of the related obstacles, e.g., difficulties in data gathering and annotation, data scarcity, and the dynamical MEs which are subtle, spontaneous and extremely fast. The bottom of the manuscript has a set of DL approaches which have been recommended to solve these issues and improve accuracy of MER. The paper proceeds further, indicating that deep learning has shown clearly the superb performances in MER too. The article notably mentions the still remaining issues against which of the MER mechanisms should be evaluated [2].

C. *A Video Based Human Detection and Activity Recognition –A Deep Learning Approach*

The goal is to achieve a match of the accuracy of existing systems used for Human Activity Recognition (HAR) systems that helped in the recognition of different human actions in different video clips. The proposed model is the CNN and it is a high-performance architecture, which is designed for pixels as inputs and aids in image recognition and other processing. The models are based on the CNN architecture and are trained on a set of human action's videos. Then, their performance is assessed using the standard set of videos. The suggested model surpasses the current status-quo shown by the above state-of-the-art on the HAR data set. The model shows a good performance and yields a high accuracy of 79.33% for the frame based and of 84.4% for the image-based measurement [3].

D. *Factors affecting concentration of attention in boxing athletes in combat situations*

This research focuses specifically on the determining the parts that play a key role in getting boxers to possess the right amount of concentration during the boxing matches. This research will talk about the self-assessment of athletes and coaches on attention skill to understand features of lack of focus and the role external factors (such as loud noises or crowd reaction) in recognition among competitors. As the analytical approach, there is the practice of interviewing boxers and the acquisition of their self-understanding data on how they perceive their concentration indicators. Training session beginnings and competitive team play allow for optimization processes which includes human oversight. This test is used to examine situations, which are responsible for disturbing players' attentional processes. On the other hand, coaches' assessments on the concentration of the boxers is a rated external metric that is used to assess attention. Data is gathered and analysed by using statistical methods including descriptive statistics analysis and non-parametric variance study methods to interpret it. The study revealed the role of certain factors that might have a negative effect on serious arousal of boxers, such as the perception of their opponent's superiority and the criticism by their trainer during the critical situations. These results, on the other hand, mean the high significance of the Self and external triggers bringing athletes to be either attentive or not [4].

E. *Relationship between selected psychological variables among trainees of combat sports*

The task of this study was to look into how the selected psychological parameters could influence the behaviour of the Trainees in combat sports- Aggression, Sports competition anxiety, and Sports Achievement motivation. Sports sciences laboratory conducted study with 10 male athletes in each of the three sports (Boxing, Wrestling, and Judo). The test subjects were exposed to Buss Perry Aggression Questionnaire (BPAQ), Sports Competition Anxiety Test (SCAT) and Sports Achievement Motivation Test (SAMT) which measure aggression, sports competition anxiety, and motivation towards sports achievement accordingly [5]. It was analysed using the tools of Descriptive Statistics and the procedure of Pearson Product-moment correlation coefficient. The outcome indicates that Sports Competition Anxiety is related with Sports Achievement Motivation ($r = -0.45$; p 0.05).

F. *A Unified Approach to Interpreting Model Predictions*

The SHAP framework of interpreting complex machine learning models addresses the critical challenge that several existing methods share in a unified theoretical foundation rooted in Shapley values from game theory. Model interpretability is critical in building trust, improving models, and understanding processes generally, especially in light of the increasing use of so-called "black-box" models such as ensembles and deep networks. SHAP defines a class of additive feature attribution methods that satisfy desirable properties such as local accuracy, missingness, and consistency that many prior approaches lack. By unifying six existing techniques, such as Local Interpretable Model-agnostic Explanations (LIME) and Deep Learning Important FeaTures (DeepLIFT), SHAP reveals their commonalities and resolves their limitations. New algorithms, such as Kernel SHAP and Deep SHAP, are proposed to estimate feature importances efficiently, which makes it suitable for both model-agnostic and model-specific contexts. SHAP outperforms existing methods in providing more accurate, computationally efficient, and human-intuitive explanations, as shown through experiments on decision trees, deep

 ISBN: 978-1-68558-268-5

networks, and user studies. Despite the computational costs for large datasets, SHAP's progress represents a significant step toward transparent and trustworthy AI, with ongoing work focused on further enhancing its scalability and explanatory power [6].

III. METHOD

In this section, we describe the architecture and components of the proposed system, which integrates action classification, model explainability, and a user-facing visualization interface.

A. Action Classification Model Building

The first step involves selecting a model architecture capable of accurately analyzing fight sequences in images and video frames. Given the nature of the task—detecting and classifying rapid, overlapping movements—CNNs are a natural fit due to their strong performance in image-based tasks [3].

We initially considered R-CNN and Faster R-CNN [3], known for their ability to focus on specific regions of interest within an image using region proposals. However, these models are computationally expensive and not optimized for real-time analysis.

Instead, we selected YOLOv8n [7] due to its speed and accuracy in multi-object detection. YOLO is a one-stage detector that performs object localization and classification in a single forward pass, making it suitable for real-time video analysis. It is particularly effective at identifying multiple actions in complex scenes—such as various strikes and movements in a boxing match.

The model was fine-tuned on our custom dataset with annotated classes representing boxing actions such as jab, hook, uppercut, knockdown, and defensive movements. We adjusted hyperparameters such as network depth, learning rate, and number of epochs to optimize performance for our task.

B. Use of SHAP

To improve the transparency and trustworthiness of our model, we integrated SHAP [6] a game-theoretic approach to explain the output of machine learning models.

SHAP [6] assigns importance values (SHAP values) to different input features—in this case, regions of the image—indicating how much each part contributed to the model's final prediction. This allows researchers and behavioral analysts to understand not just what the model predicted, but why it made that decision.

The visual output from SHAP highlights the regions most influential in classifying specific actions, offering insight into both model performance and potential areas for improvement.

C. Visualization Interface

To make the results accessible and actionable, we developed a user-friendly visualization interface designed to support both analytical and interpretive tasks. The interface presents real-time plots of detected actions across a timeline [10], allowing users to track momentum shifts and behavioral patterns throughout a fight. It also includes overlay visualizations of SHAP explanations directly on the video frames, highlighting the specific regions that influenced the model's decisions. Additionally, the interface displays inferred psychological state indicators, derived from a combination of recognized actions and contextual cues. The system is also capable of identifying targeted attentional breaks—such as pauses in activity, delayed reactions, or shifts in defensive posture—by analyzing temporal patterns in the action data. Specifically, these moments are detected through noticeable fluctuations in the ratio of strikes over time [4], where a sudden drop or irregularity may indicate a lapse in focus or engagement. Such patterns are then mapped to behavioral scales representing cognitive disengagement, mental fatigue, or stress.

IV. BOXING DATASET

A. Data Collection

To design a boxing dataset that captured all the essential features, around 20 videos were downloaded from YouTube, with a broad cross section of weight categories, video quality, and sources to assure generalizability from the model. The videos were split into a series of frames and then reviewed frame by frame to find critical moments in the fighting processes. Many features related to the fights were captured and analyzed. LabelImg was the annotation tool that was used to draw bounding boxes around the region of interest and the annotations were stored in Extensible Markup Language (XML) format for better parsing and accessibility. All boxing videos, images, and annotations saved in Google Drive were further accessed from Google Colab to develop the model.

B. Dataset Preparation

LabelImg is the annotation tool that was used to label the frames individually. Repetitive frames were discarded to avoid overfitting. Different aspects of the fight were captured to help the model understand the nature of the fight better.

The labels are shown in Table I.

TABLE I. DATASET LABELS

ID	*Object/Action*	*Label*
0-1	Boxers	boxer1/boxer2
2-3	Successful Straight Punch	successfulstraightpunchboxer1/ successfulstraightpunchboxer2
4-5	Unsuccessful Straight Punch	unsuccessfulstraightpunchboxer1/ unsuccessfulstraightpunchboxer2
6-7	Successful Hook	successfulhookboxer1/ successfulhookboxer2
8-9	Unsuccessful Hook	unsuccessfulhookboxer1/ unsuccessfulhookboxer2
10-11	Cuts	boxer1cut/boxer2cut
12-13	Knockdowns	boxer1knockdown/ boxer2knockdown

These labels capture all the crucial aspects of the fight. The images were then flipped to increase the strength of the dataset and its ability to classify the frames. After this the images went through the standard process of resizing for uniformity and then normalizing. Normalizing would normalize the RGB values to a value between 0 and 1. This concludes the dataset preparation.

(a) successful straight punch (b) boxer knockdown

Fig. 1. Examples of different boxing actions.

As shown in Figure 1, key boxing actions such as successful straight punches and knockdowns are illustrated, which form the basis of our dataset labeling process.

V. RESULTS

In this section, we evaluate the performance of the model in action recognition.

Our boxing model provides a solution to understand the impact of mental health on a boxer's performance. By analyzing scientific data and body movements during a fight, the system is able to provide a better view of an athlete's condition during each round. Below, we discuss the main features, functions, and user feedback that demonstrate the effectiveness of this system. The system can also detect important physical signs that affect an athlete's mental and physical health, such as cuts. The recommendations suggest that this feature is particularly useful for performance. To help users understand the interaction between the mind and body, our system uses interactive visualization to present information. These features help coaches and athletes identify key points during competition where mental state changes significantly, allowing for a clear understanding of how this change impacts well-being. For example, users may notice an increase in their stress levels after a knockdown, or a boost in confidence after a successful strike such as a hook or a straight punch. By providing a clear annotation for each psychometric test, SHAP can help users identify which factors make up the majority of the sample output, thus providing confidence in the system's results. This feature will be of particular benefit to model developers who want to understand the logic behind each psychological test to make more informed decisions. The user-friendly photo upload function simplifies the analysis process, allowing instructors to seamlessly upload combat photos. Once downloaded, the system uses TensorFlow and OpenCV to identify and isolate frames which bring value or where changes are evident. This frame removal helps users focus on the most important issues, keeping analysis fast and relevant. User feedback suggests that these features increase usability, as they provide access to good, detailed information without long waiting times.

A. Action Classification Model Output

Figure 2 shows the model’s output on various images from different fights. We can see that the model is accurately able to classify all the different labels it has been trained on with high confidence.

Fig. 2. Image Action Classification Output.

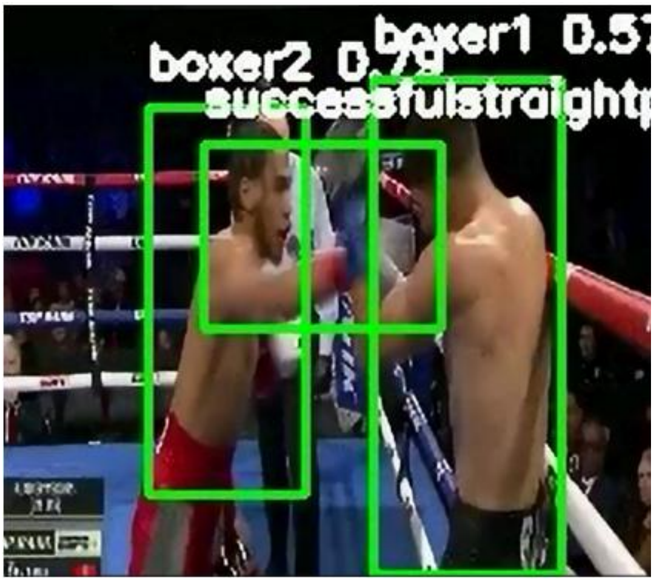

Fig. 3. Video Action Classification Output.

Figure 3 shows the model’s prediction on a fight video. Our model is able to perform near real time classification on any fight video with high accuracy. Figure 3 depicts boxer 1 and boxer 2 where boxer 2 throws a successful straight punch, our model was able to correctly classify this action.

B. *Action Classification Model Evaluation*

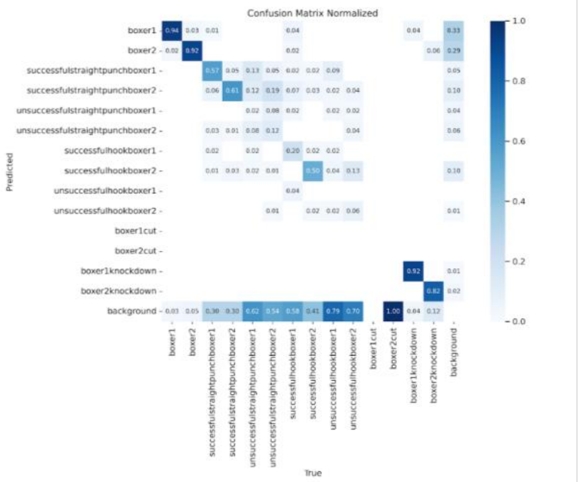

Fig. 4. Confusion Matrix.

Figure 4 shows the confusion matrix which can be used to evaluate the model's performance. Each cell in the matrix shows the proportion of samples classified into a particular category relative to the true category. Darker shades correspond to higher proportions which represents better classification performance of the model for that specific true-predicted pair. Classes like boxer1(true) vs. boxer1 (predicted) has a high proportion of correct classifications (~0.94), indicating strong performance of model for this class. Whereas, some confusion occurs between classifying the type of punch, straight or hook which indicates the need for a rich featured dataset. Specific events like boxer knockdown and cuts are classified well. This shows the overall performance of model on different classes.

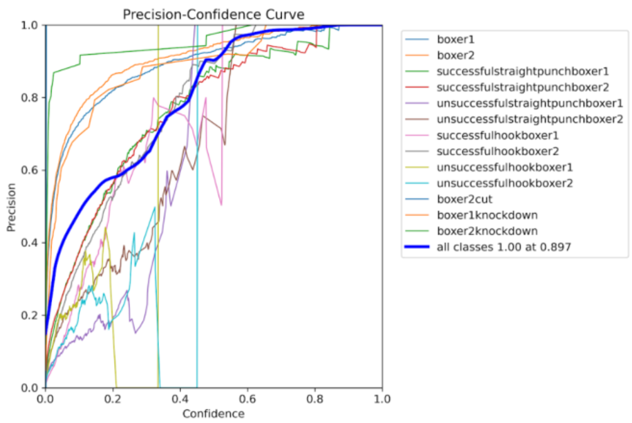

Fig. 5. Precision Curve.

Figure 5 shows the Precision-Confidence Curve which can be used to visualize the relationship between the model's confidence in its predicted and the precision achieved for various classes. Each curve shows how the precision changes as confidence threshold increases for a particular class. The blue line indicates that the model achieves perfect precision i.e., 1.0 for predictions when the confidence threshold is set to 0.897.

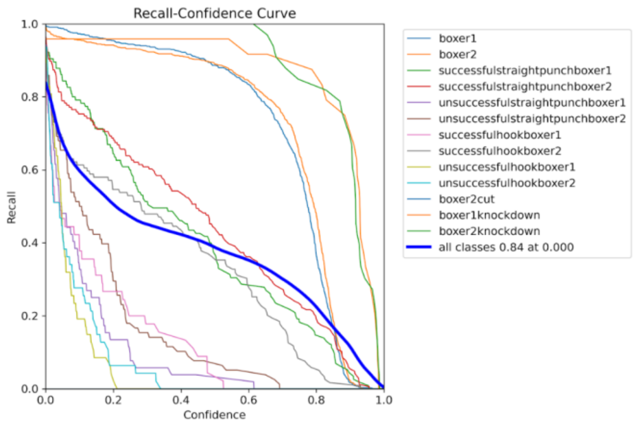

Fig. 6. Recall Curve.

Figure 6 shows the Recall curve for different classes. Recall measures the ability of the model to find all relevant instances in the dataset. It is calculated as the number of true positives divided by the sum of true positives and false negatives. A higher recall means that the model is capturing a larger portion of relevant cases. The curve starts at a higher recall levels when the confidence threshold is low which means that the model is more inclusive but less accurate. As the confidence threshold increases, the recall decreases.This indicates that the model is becoming more selective, thus more confident but potentially missing some true positive instances.

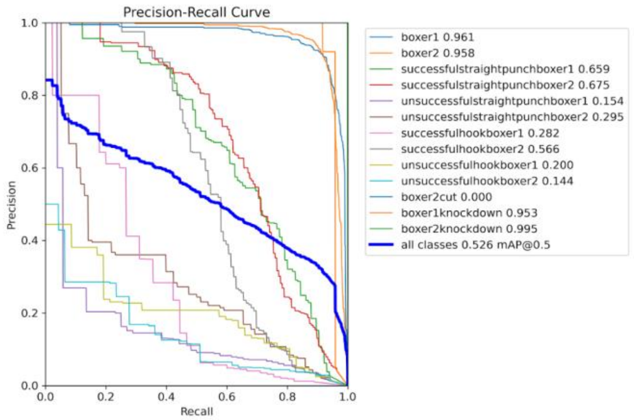

Fig. 7. Precision Recall Curve.

Figure 7 shows a Precision-Recall (PR) curve which is another way to evaluate the performance of the model. This helps in understanding the trade-off between precision and recall for each class at different thresholds of classification decision. At higher precision levels recall is generally low indicating the model being very selective, making fewer predictions but those predictions more likely to be correct. In the right side of the graph, recall is high and precision ends to be low which indicates that the model identifies most of the positive cases but also makes false positive errors. Understanding PR curve is crucial where the cost of false positives are different from tat of false negatives. When it comes to action classification, it is better to miss a positive case than incorrectly labelling a negative case as positive. This curve helps in finding the right balance based on specific needs.

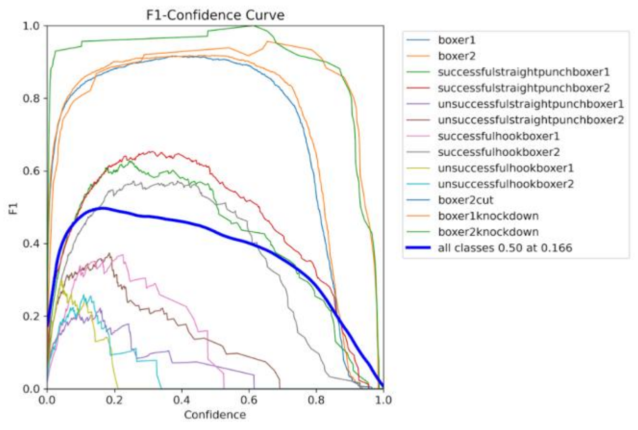

Fig. 8. F1 Curve.

Figure 8 shows F1 curve for different classes, this curve combines precision and recall into a single metric. Classes like boxer 1 and boxer 2 maintain high F1 scores across confidence levels which indicates strong performance in these classifications. The blue line indicates that the average F1 score across all classes at a confidence threshold of 0.166 is 0.50.

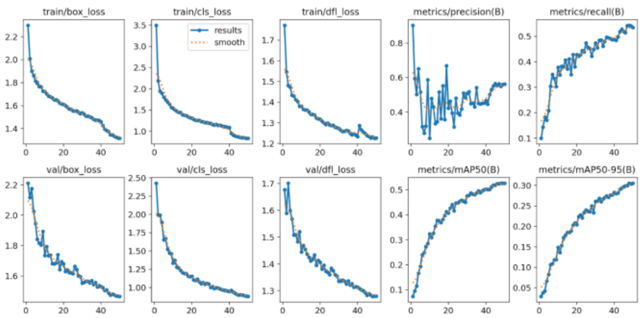

Fig. 9. YOLO Model Results.

Figure 9 provides a comprehensive overview of performance metrics over epochs during the training and validation phase of the model. The overall trend of decreasing loss and increasing precision and recall is a positive sign that the model is learning effectively from the training data and improving its prediction capabilities as training progresses.

C. *Psychological Analysis of Boxer's Performance*

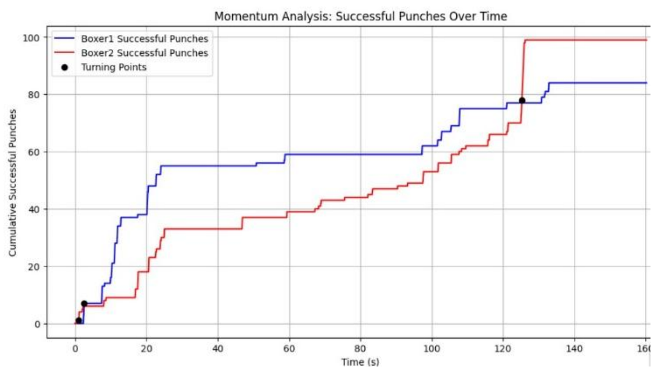

Fig. 10. Momentum Analysis: Turning Points.

Figure 10 shows the cumulative number of successful punches landed by two boxers in the match. The blue line tracks cumulative successful punches of boxer 1 and red line tracks cumulative successful punches of boxer 2, with the steepness reflecting the scoring pace. Steeper sections denote more rapid scoring of points. Black squares mark specific moments in the match labeled as "Turning Points." These indicate moments where the momentum of the match shifts significantly. This can be due to a critical strike, a tactical change, or other significant event that affects the dynamics of the contest. By comparing the trajectories of two lines we can infer that boxer 1 starts strong, gaining an early lead. However, boxer 2 increases his rate of successful punch as the match progresses eventually surpassing boxer 1. The black squares or turning points are crucial for coaches as they highlight moments when potential strategic changes might have influenced the match's outcome. Coaches can make use of this data to improve training regimens, focusing on shifting or maintaining momentum at critical stages of the match.

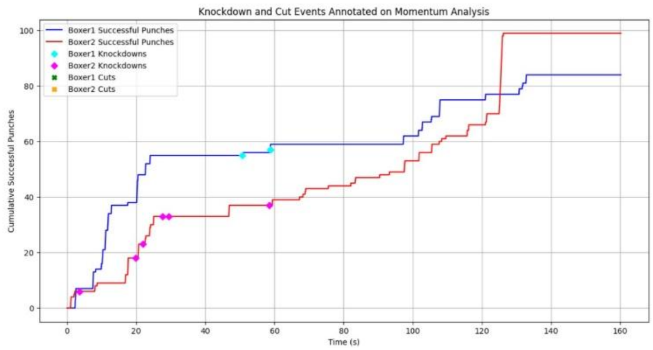

Fig. 11. Momentum Analysis: Knockdowns and Cuts Annotated.

Figure 11 expands on the earlier momentum analysis by including events like knockdown and cuts for understanding the impact of significant events on the momentum of a match. Initially, both boxer 1 and boxer 2 accumulate points at similar rates with boxer 1 taking a slight lead. The first significant event in the above graph is the knockdown of boxer 2. Despite this, boxer 2 begins to accumulate punches at higher rate. Shortly after boxer 1 experiences knockdowns around 50 to 60-second mark which marks a clear momentum shift in boxer 2's favor. At the 80-second mark another key turning point occurs where boxer 2 shows a surge in successful punches, significantly outpacing boxer 1. This can be correlated with strategic adjustments or a decline in boxer 1's defense. The match ends with boxer 2 having a clear lead in successful punches, possibly reflecting better stamina or strategy execution. This graph offers a detailed view of how events like knockdowns and cuts can impact the flow of match which can be used to train the boxers to enhance their performance in the future matches.

Figure 12 represents the distribution of action between boxer 1 and boxer 2 across different rounds of a boxing match. Each pie chart depicts the percentage of total actions taken by boxer 1 and boxer 2. For trainers and coaches, these insights are valuable for assessing the resilience and recovery of a boxer after being knocked down. This can also help in strategizing training to enhance endurance and tactical responses to such events in the upcoming matches. Overall, these enhanced momentum analysis graphs serve as a tool for in-depth review and strategic planning in sports, particularly in boxing.

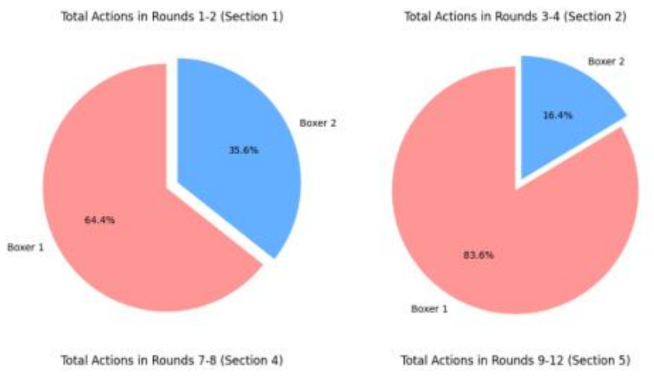

Fig. 12. Total Actions in Sections.

VI. DISCUSSION

As admitted, the current model is limited to recognizing only two primary strike types—straight punches and hooks. This design choice was intentional and guided by the objectives of the system. For our current goal of inferring psychological states, focusing on high-frequency, high-impact strikes like the straight and hook is sufficient. These are not only the most commonly thrown punches but also indicative of strategic intent, energy levels, and mental engagement, aligning with previous findings on how psychological factors shape athletic focus and execution in combat sports [4], [5], [13]. Thus, variations in their frequency, timing, and accuracy serve as strong proxies for cognitive states such as fatigue, disengagement, or stress.

Previous efforts in the domain of sports psychology have mostly focused on either physical performance metrics or indirect psychological assessments. Existing tools, such as traditional biometric sensors or motion tracking systems, fail to offer real-time, interpretable insights into the cognitive states of athletes. Moreover, while some systems incorporate physical actions like punch tracking, they often lack the ability to connect these actions to mental states, leaving a gap in understanding how psychological factors influence physical performance. Our system addresses this gap by leveraging real-time action recognition and explainable AI techniques [1], [6], [8] to correlate strike patterns with inferred psychological conditions, offering a unique, practical approach for performance optimization.

However, while our system offers significant advancements, it is not without limitations. The current approach is based on a narrow set of strike types, and there is a clear need to expand it to account for a broader range of actions, including defensive movements, counter-strikes, and combinations. Additionally, the system’s reliance on visual data means it could be enriched with additional inputs such as physiological signals (e.g., heart rate, galvanic skin response) for a more comprehensive understanding of a boxer’s mental state. These features are on our wish-list for future iterations of the system. Moreover, we plan to further refine our movement recognition model to improve accuracy and robustness in different lighting and environmental conditions. Another goal is to incorporate more advanced psychological tests that could enhance the system’s ability to identify specific mental conditions like anxiety or stress, which may not be fully captured by physical performance metrics alone.

The purpose of our contribution is to offer a practical, real-time tool for tracking mental states in combat sports, empowering athletes and coaches to make data-driven decisions. The system provides a unique combination of performance analysis and mental health monitoring, which can guide training regimens, optimize decision-making in competition, and contribute to athlete well-being by identifying early signs of mental fatigue or stress. By offering clear visualizations and intuitive feedback, the system bridges the gap between physical performance and psychological analysis, supporting athletes in maintaining peak performance while safeguarding their mental health.

Looking forward, we envision several areas of improvement. First, the addition of more complex strike patterns, defensive actions, and physiological signals will help enhance the system's accuracy and completeness. Second, improving the user interface through user feedback and further research will make the system even more intuitive and applicable across different sports. Finally, further usability surveys and structured testing with domain experts will ensure that the system continues to meet the practical needs of coaches, athletes, and sports psychologists. Ultimately, our goal is to create a tool that not only supports performance enhancement but also prioritizes the mental health and well-being of athletes, ensuring their longevity in the sport.

VII. CONCLUSION AND FUTURE WORK

In this work, we set out to bridge the gap between physical action recognition and cognitive state inference in competitive boxing. We developed the Boxing Psychological State Tracking System, a vision-based framework that uses deep learning models like YOLOv8 and R-CNN to detect key actions, and SHAP to provide interpretable insights into inferred psychological states such as fatigue, confidence, and stress. By analyzing behavioral cues like strike patterns, movement changes, and knockdowns, our system offers an accessible interface that visualizes mental dynamics in real time, supporting coaches and athletes in making more informed decisions during performance analysis. The system demonstrates high accuracy in classifying trained strike types and offers a novel, explainable perspective on mental health in combat sports.

Looking ahead, we aim to enhance the system’s capabilities and user experience in several key areas. Future work will consider a broader range of psychological factors beyond confidence and anxiety, such as opponent strength and team identity. For example, incorporating data on relative rankings or prior performance against specific opponents may provide deeper insights into mental state fluctuations. Additionally, we plan to expand the model’s action recognition capabilities to include more strike types such as jabs, enabling broader applicability. The system could also evolve to offer predictive insights based on historical data, guiding both training and tactical preparation. Furthermore, we envision extending the framework to other combat sports like fencing and mixed martial arts, where only the labeling methodology would need adjustment. Finally, to improve generalizability and robustness, we intend to expand the dataset beyond the current set of 20 fights, ensuring more diverse and representative coverage of combat scenarios.

REFERENCES

[1] Y. Kong and Z. Duan, “Boxing behaviour recognition based on artificial intelligence convolutional neural network with sports psychology assistant,” *Scientific Reports*, vol. 14, no. 1, pp. 7640, Apr. 2024.

[2] Y. Li, J. Wei, Y. Liu, J. Kauttonen, and G. Zhao, “Deep Learning for Micro-Expression Recognition: A Survey,” IEEE Trans. Affect. Comput., vol. 13, no. 4, pp. 2028–2046, Oct.-Dec. 2022.

[3] M. Dhar and B. Chakraborty, “A video based human detection and activity recognition–a deep learning approach,” Turk. J. Comput. Math. Educ., vol. 11, no. 1, pp. 551–559, 2022.

[4] M. C. Hernández, Y. S. Prieto, J. D. García, and M. S. Rodríguez, “Factors affecting concentration of attention in boxing athletes in combat situations,” Rev. PODIUM, vol. 15, no. 1, pp. 5–21, Apr. 2020.

[5] A. Sharma and P. Purashwani, “Relationship between selected psychological variables among trainees of combat sports,” J. Sports Sci. Nutr., vol. 2, no. 1, pp. 1–3, 2021.

[6] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Proc. 31st Int. Conf. Neural Inf. Process. Syst. (NIPS’17), Red Hook, NY, USA, 2017, pp. 4768–4777.

[7] P. Stefański, J. Kozak, and T. Jach, “The problem of detecting boxers in the boxing ring,” in *Recent Challenges in Intelligent Information and Database Systems*, E. Szczerbicki, K. Wojtkiewicz, S. V. Nguyen, M. Pietranik, and M. Krótkiewicz, Eds., *Communications in Computer and Information Science*, vol. 1716, Singapore: Springer, 2022, pp. 546–556.

[8] Y. Nohara, K. Matsumoto, H. Soejima, and N. Nakashima, “Explanation of machine learning models using shapley additive explanation and application for real data in hospital,” *Comput. Methods Programs Biomed.*, vol. 214, pp. 106584, 2022.

[9] V. Hudovernik and D. Skocaj, “Video-based detection of combat positions and automatic scoring in jiu-jitsu,” in Proc. 5th Int. ACM Workshop Multimedia Content Analysis Sports, Oct. 2022, pp. 55–63.

[10] J. Brindha and G. Nallavan, “A wearable biometric performance measurement system for boxing - A survey,” in 2022 IEEE World Conf. Appl. Intell. Comput. (AIC), Sonbhadra, India, 2022, pp. 536–540.

[11] R. Merlo et al., “Profiling the physical performance of young boxers with unsupervised machine learning: A cross-sectional study,” Sports (Basel), vol. 11, no. 7, p. 131, Jul. 2023.

[12] K. Nakamura, M. Uchida, and T. Sato, “Basic research on the primary prevention of boxing-related sports injuries with the development of a quantitative motion analysis software,” J. Phys. Ther. Sci., vol. 33, no. 6, pp. 495–498, Jun. 2021.

[13] D. Zhang, L. Bei, J. Wu, W. Li, and K. Zhang, “Effect of boxers’ social support on mental fatigue: Chain mediating effects of coach leadership behaviors and psychological resilience,” Work, 2023.

[14] R R. Kumar, “A comparative study of self-confidence among boxers and wrestlers of Hyderabad in India,” *Int. J. Phys. Educ. Sports Health*, vol. 2, no. 1, pp. 1–4, 2015.

 ISBN: 978-1-68558-268-5