International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 01 | January - 2023 Impact Factor: 7.185 ISSN: 2582-3930

Involuntary Table Tennis Recognition using Recurrent Neural Network on The Perception-Based Data

*Dr. Abhishek Sharma, **Utkarsh Nanda, ***Premansh Sharma, ****Barnabh Chandra Goswami
*Associate Professor, **Student, ***Student, **** Student
*Department of Electronics and Communication Engineering
*The LNM Institute of Information Technology, Jaipur, Rajasthan, India

Abstract ***- Sport performance analysis and sports practise are inextricably linked. It is critical to assist coaches in analysing and improving the performance of their athletes throughout training or game sessions. Because of technological advancements, notational analysis of video footage using multiple software packages is now possible. Unfortunately, the coach must manually recognise the acts before proceeding with additional analysis. The goal of this research is to create an automated system for recognising Table Tennis shots in widely available broadcasted videos using a pre-trained Recurrent Neural Network (RNN) approach. We provide a novel method for gathering video data from table tennis games and performing stroke detection and classification. Using the proposed setup, a diversified dataset encompassing video data of 4 basic strokes taken from 4 professional table tennis players, totalling 2000 films, was collected. With a validation accuracy of 97.02%, the temporal recurrent neural network model is created. Furthermore, the neural network generalises well over the data of a player who was not included in the training and validation datasets, categorising new strokes with an overall best accuracy of 91.12%. Several model architectures were trained for stroke recognition using machine learning and deep learning methodologies, and their results were compared and benchmarked. The model's inferences, such as performance monitoring and stroke comparison, have been examined. As a result, we are contributing to the creation of a computer vision-based sports analytics system for table tennis, namely a player's strokes, and is particularly insightful for performance enhancement.***

I. INTRODUCTION

Understanding human actions is an essential component of human social communication. We can say that we spend most of our waking hours monitoring and explaining the actions of others. Because the expansion of social economy promotes the advancement of computer technology, detecting technology must become more intelligent so that it can more precisely grasp human daily actions and replace humans to fulfil increasingly complex duties. It is critical to detect human activities properly, effectively, and correctly while designing automatic machines that interact with humans. A series of hardware and software facilities are used to load the human body recognition technology into the robot. Finally, intelligent machines will be able to interact and move autonomously in society, enabling the role of unmanned monitoring of complex tasks to be realised.

Deep learning is now the most powerful machine learning employed by many researchers for action recognition in computer vision. Action recognition has a wide range of applications, including video surveillance, sports performance analysis, rehabilitation, and virtual reality. Deep learning, a subtype of machine learning, is a new trend in action recognition due to its capacity to produce more precise results than the traditional machine learning method. The benefit of employing deep learning is that the network learns and pulls features automatically from raw photos rather than having to be manually extracted. Many earlier studies have been conducted in recognising action utilising deep learning methods such as RNN. The coach must manually evaluate the broadcast footage and key in the athlete's activity into the software packages one by one before further analysis in sports performance analysis utilising a vision-based approach. An automatic action recognition from broadcast footage is critical for providing a more efficient and accurate analysis to coaches. Smash is one of the actions that a player can perform during a match or game. It's a powerful overhead stroke with a downward trajectory. It is critical for coaches to analyse smash action performance to enhance smash technique. As a result, the goal of this research is to create an automated system for badminton smash and other badminton activities utilising a pre-trained Recurrent Neural Network (RNN) method using widely available broadcasted movies. In this project, RNN was created from aired video of a Table Tennis practice match to automatically recognise shots and other types of motions.

II. DATA SET COLLECTION

To obtain a large data set, a team of 10 individuals with varying levels of expertise in Table Tennis, ranging from beginner to professional, were assembled. Each player was instructed to execute all five shots 100 times, while video footage was captured at 240 frames per second for improved accuracy. As a result, the total number of frames utilized for training the Recurrent Neural Network Model amounted to 12,00,00 (10 players * 5 shots * 100 attempts * 240 fps), and the overall data set size exceeded 50 gigabytes.

Figure 1: Player Performing w Shot.

Figure 2: Player Performing x Shot.

Figure 3: Player Performing y Shot.

Figure 4: Player Performing z Shot.

III. WORKING OF CODE

The project below is based on video processing and deep learning, and it employs neural networks such as CNN (Convolutional Neural Networks) and the LSTM (Long Short-Term Memory) model. The reason for selecting these two models is that CNN is in charge of image processing and uses the convolution approach to find features in a model, whereas LSTM is in charge of feeding the output of one neural network to the other in order to assess the output based on past results. The following project is written in Python and employs tools such as tensor flow, scikit-learn, and OpenCV to build this neural network model.

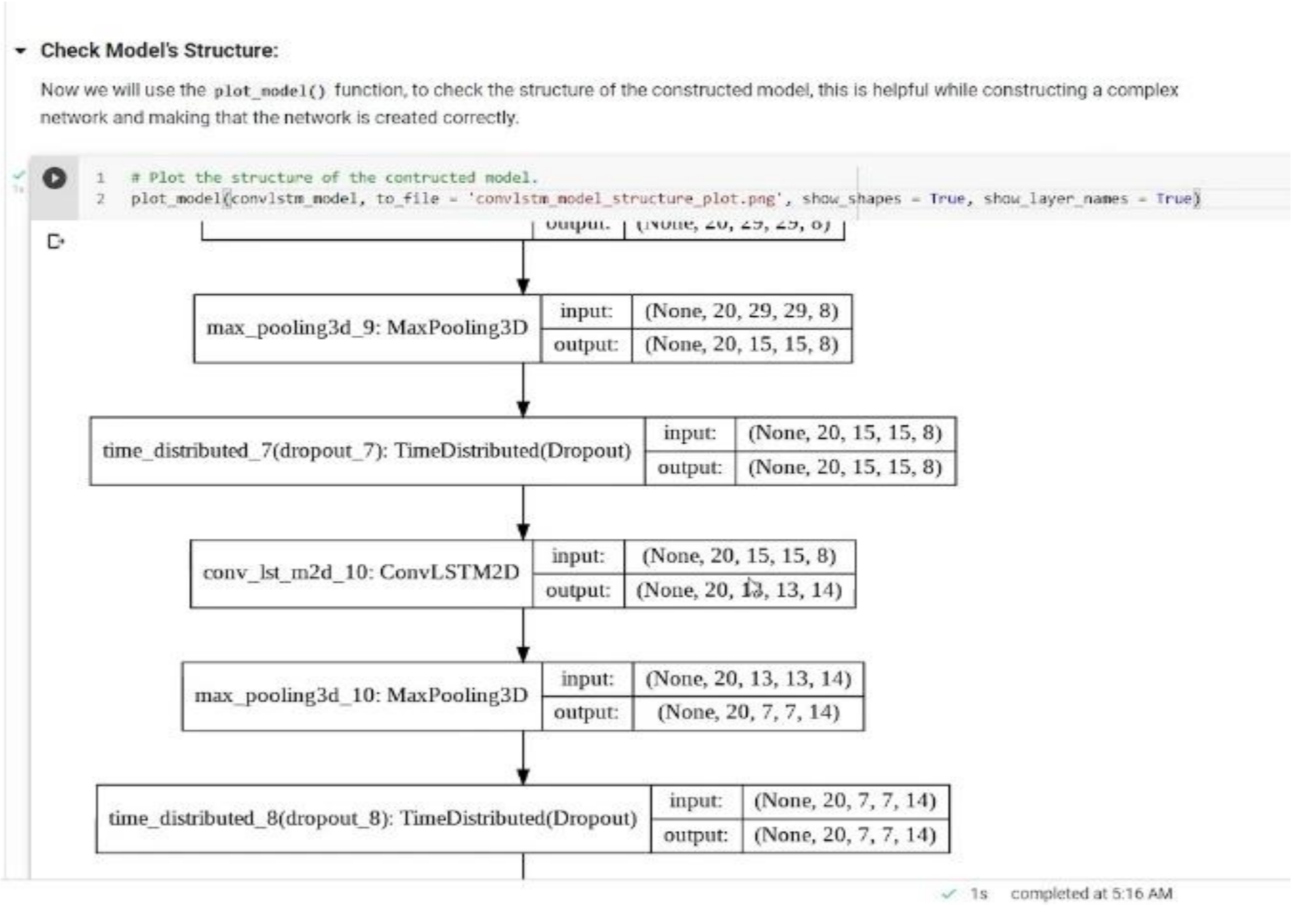

Figure 5: Model's Structure

The model is built on four CNN LSTM layers, with pre-processed data sent as input in the form of a NumPy array to the CNN model, which filters out the feature through its layer and then feeds the output to another LSTM layer for analysing the next image and processing the output based on the previous output. This method is repeated for each frame in a video, and the model is ready once all of the video frames have been fed. We kept an epoch of 50 and a batch size of four. So, in each epoch, four films will be fed into the model to prevent overfitting of the Neural network model.

IV. CONCLUSION

We investigated a strategy for categorising forehand table tennis strokes using multimodal data and neural networks in this paper. We intended to use the technology to deliver feedback to staff. We organised an expert interview with a table tennis coach to investigate the coach's impression of a computer-based training system that delivers feedback when mistakes are recognised during training. The instructor emphasised her scepticism regarding these technologies, particularly AI, because they are new and not extensively used. The coach, on the other hand, agrees that the method might help beginners exercise their essential techniques and operate as a support to further improve the learning outcome.

In the findings of this study, the pre-trained CNN model is the best model to utilise among others to automatically recognise the smash action and send information to the coach for additional notational analysis. The study demonstrates that a deep learning approach can be used to construct an automated action recognition system. With a restricted dataset, a pre-trained model may be utilised to automatically distinguish the smash action from online broadcasted videos.

To summarise, automated action detection based on vision-based data employing deep learning could assist coaches, players, and the sports institution itself. Our work, however, plainly has significant limits. Regardless of this, we hope that our research can help to improve the deep learning technique in sports practise. The research to improve the findings is already underway.

V. REFERENCES

[1] Lodziak, T. How Long Does It Take to Get Really Good at Table Tennis? 2017. Available online: https://www.tabletenniscoach.me.uk/how-long-take-get-really-good-table-tennis/ (accessed on 29 April 2021).

[2] Larcombe, B. The Four Basic Table Tennis Strokes. 2012. Available online: https://www.experttabletennis.com/forehand-drive-in-table-tennis/ (accessed on 29 April 2021).

[3] Liu, R.; Wang, Z.; Shi, X.; Zhao, H.; Qiu, S.; Li, J.; Yang, N. Table Tennis Stroke Recognition Based on Body Sensor Network. In Internet and Distributed Computing Systems; Montella, R., Ciaramella, A., Fortino, G., Guerrieri, A., Liotta, A., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 1–10. [Google Scholar].

[4] Tabrizi, S.S.; Pashazadeh, S.; Javani, V. Comparative Study of Table Tennis Forehand Strokes Classification Using Deep Learning and SVM. IEEE Sens. J. 2020, 20, 13552–13561. [Google Scholar] [CrossRef].

[5] Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. Neural Comput. 1997, 9, 1735–1780. [Google Scholar] [CrossRef].

[6] https://www.sciencedirect.com/science/article/abs/pii/S0736585318307676.

[7] https://www.researchgate.net/publication/292304908_Highly_flexible_self-powered_sensors_based_on_printed_circuit_board_technology_for_human_motion_detection_and_gesture_recognition.

[8] https://rdcu.be/c3tYV.

[9] https://www.researchgate.net/publication/307616450_A_concise_tutorial_on_human_motion_tracking_and_recognition_with_Microsoft_Kinect.