

# Automated Commentary Based on Umpire Gesture Recognition

Abhishek Puppala, 11596928  
Jaswanth Kalyan Polavarapu, 11596929  
Kushal Krishna Vankalapati, 11587371  
Manisha Lokasani, 11609309

February 2024

## 1 Introduction and background

Nowadays, we can see that the sport is becoming more and more famous, which is being watched world wide by everyone. Now in the sector of sport, cricket is one of the most famous games played and well known till now. For every sport game, we follow each one of the rules that has been published. And our game cricket also consists of various kinds of rules for playing the game. This game consists of rules given by umpire decisions, based on certain actions, and the umpire displays the sign of being executed for the following action. These signs are so called gestures by the umpires. Here we are considering the eight kinds of signals or signs by umpire gestures, which are six, wide, four, no-ball, out, not-out, byes, leg-byes and no action class. For our project, we will be analyzing the umpire gesture in the image and classify it as one of the decisions for the action performed in the game, such as whether the player can be out or not-out and other gestures as mentioned above.

Currently we have collected some amount of umpire gesture image data needed, and now we are trying to further extend these image data of umpire gestures, by regenerating the images using the GAN algorithm techniques with some open source technologies such as Mid-journey, Jasper-AI and runway apis.

Our proposed project is image recognition or classification based, which is much more related to the data science field, more extended towards the deep learning field over machine learning, which helps me gain practical knowledge towards the real time scenario over the image classification techniques. From this project, by analyzing the deep learning techniques using the keras, py-Torch, spark modules, will be able to learn and implement the algorithm techniques for CNN related deep learning methodologies, how to analyze the image processing, with help of various layers, involving convolutional layer, pooling layer, hidden layers and output layer processing using the torch module. Also being able to learn to extract the features of an image and then map the label to a particular image with the help of a convolutional neural network model.

## 2 Statement

The main motive of our project is to analyze the cricket umpire images and label the umpire image based on their gestures and classify these images into one of the gestures or signals, such as out, not-out, etc with help of the deep learning convolutional neural network methodology by assigning the various layers such as pooling layer, convolutional layer etc. As per the project ideology, I will be comparing the model between classification and CNN models. This includes the processing by reading the images. Then for the images data, to standardize the file formats into the same format, followed by normalizing the pixel size to improve the model performance.

## 3 Related Work

The authors have predicted automated commentary based on umpire gestures using the hidden Markov model, which forecasts future states based on the observed series due to its flexibility. Additionally, in this research, K-means clustering is employed for dimensionality reduction, Random Forest is used to compare classifications, and KNN is applied to classify different classes to specific points. In this paper, distances are calculated in terms of Manhattan distance, Euclidean distance, and Minkowski [Medha Wyawahare et al., 2023]. [4].

G.S et al. developed an automated labeling approach for sports videos

based on umpire gesture recognition, utilizing a hierarchical hidden Markov model, but struggled to achieve a satisfactory model ratio. Meanwhile, Heikal, H et al. proposed a real-time 3D gesture recognition system using depth images and incorporating Naive Bayes and neural network classifiers. Bhansali, L and Narvekar, M employed the Gradient method for umpire gesture recognition, successfully identifying a group of six umpire gestures in cricket games; however, its performance in distinguishing gestures from continuous streaming was sub optimal. Various researchers explored methods such as multi-modal frameworks, dual-threshold techniques, and rule-based induction for summarizing sports videos and automating scorecards, each presenting unique strengths and limitations. [3].

A.Shahjal et.al has used the Haar cascade Algorithm to identify the human wrists from the video streams and then classifies the gesture using logistic regression based on the pixels intensity levels. And this is good for static gestures and found very poor for dynamic gestures and the system fails when it detects multiple subjects in the frames. D. T. John et.al has been proposed a pre-trained model from openCV called caffe which is used to extract the skeletal joints from the images. Hari et.al have proposed the intensity based projection profile of umpire gestures for detecting event like six, out and wide. M.H. Kolekar and K. palaniappam has been suggested a framework for event detection and classification from cricket videos by extracting audio and other low level features. West et .al have presented a technique by using accelerometers for video annotation. Hariskrishna et.al has proposed a scheme for summarizing the visual contents of the cricket videos . A.ravi et .al has a prior knowledge in the use of inception V3 network in event detection from the cricket videos for the purpose to match the highlights generation. [Vaishnavi K.Nair et.al ,2020] [2]

Basically, the group of specific umpire gestures is to identify the events like boundaries, wickets and runs which were captured by the sensor cameras. The computer system interprets the gestures, updating the scoreboard in real time with out any manual input. The streamlines of the scoring process reduces the errors and speeds up the updates .the technology improves the efficiency of the cricket matches, providing the accurate information to players, officials and viewers. It also simplifies the worlds for the scorers, allowing them to concentrate the other important aspects of the game. This approach represents the straightforward and an effective way for updating the cricket scoring. [Dhanusha T.John et.al, 2019]. [1]

## 4 Objectives of the study

In this research, we determine automated commentary in the game of cricket using eight various umpire gestures such as six, wide, four, no-ball, out, not-out, byes, leg-byes, and no-action class. To achieve this, we employ convolutional neural networks (CNNs) and compare results with several other classification models.

## 5 Research design and Methodology

Here we use a data set of size of 20000 images in which from 6000 images which will be collected from internet(like google images) and other sources, by using generative adversarial network (GAN) we will generate those 20000 and by using CNN where by using various deep learning algorithms, we first do feature extraction from the image and then classifying it. At first we will grayscale the images and we will add convolution layer where it applies multiple filters with each filter generating feature map which are combined to allow it to learn more patterns, and followed by pooling layer, fully connected layer and output layer before compiling the model, and we label those images according to image patterns from which we will get the evaluation metrics where we focus on accuracy by comparing it with previous papers. Audio commentary is provided through predefined templates.

## 6 Datasets to be used

Here the dataset is of images in which, we are taking datasets from google images and from previous research methods and by using GAN we will generate 20000 images. As CNN maximizes the advantage of a large dataset to understand the patterns in the images which helps for understanding and evaluating the model more, the data set is sufficient to support the research.

## 7 Contributions

1. **Kushal Krishna Vankalapati (Team Lead)** : Handling the team by assigning the required tasks and helping over every task by providing the contribution.

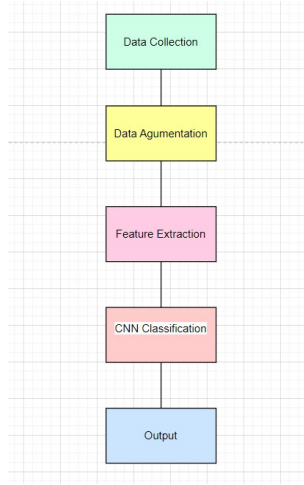


Figure 1: Umpire gesture processing

2. **Abhishek Puppala** : Helps over building the CNN model for image classification for the gestures.
3. **Jaswanth Kalyan Polavarapu** : Helps over the data pre processing of the image data.
4. **Manisha Lokasani** : Helps in the model hypothesis classification and comparing over the test accuracy results and providing the optimized model ideas.

## 8 Conclusion

In this paper, we utilize convolutional neural networks (CNNs) to identify automated commentary umpire gestures, comparing the outcomes with other classification models. Additionally, our research aims to surpass the accuracy achieved in prior studies by training our model on a dataset comprising 20,000 images.

## References

- [1] Dhanusha T John, Kavya S Kumar, Vaishak T Nair, P Visakh, and BR Poorna. Cricket scoreboard automation using umpire gestures. *Inter-*

*national Journal of Research in Engineering, Science and Management*, 2(7), 2019.

- [2] Vaishnavi K Nair, Raakhi Rachel Jose, Parvathy B Anil, Minnu Tom, and Lekshmy PL. Automation of cricket scoreboard by recognizing umpire gestures. *International Journal of Innovative Science and Modern Engineering*, 6:1–7, 2020.
- [3] Suvarna Nandyal and Suvarna Laxmikant Kattimani. Umpire gesture detection and recognition using hog and non-linear support vector machine (nl-svm) classification of deep features in cricket videos. In *Journal of Physics: Conference Series*, volume 2070, page 012148. IOP Publishing, 2021.
- [4] Medha Wyawahare, Amol Dhanawade, Shreyas Dharyekar, Asavari Dhole, and Mugdha Dhopade. Automating scorecard and commentary based on umpire gesture recognition. In *International Conference on Advancements in Smart Computing and Information Security*, pages 158–168. Springer, 2022.