

PAPER • OPEN ACCESS

## Umpire Gesture Detection and Recognition using HOG and Non-Linear Support Vector Machine (NL-SVM) Classification of Deep Features in Cricket Videos

To cite this article: Suvarna Nandyal and Suvarna Laxmikant Kattimani 2021 *J. Phys.: Conf. Ser.* **2070** 012148

### You may also like

- [Hand motion tracking based on gesture understanding using leap gesture for virtual 3D batik gallery](#)  
A C Padmasari, M Hariadi and C Christyowidiasmoro
- [Extraction and physicochemical characterization of chitin and chitosan isolated from house cricket](#)  
E B Ibitoye, I H Lokman, M N M Hezmee et al.
- [Dynamic Gesture Recognition Based on iCPM and RNN](#)  
Chen Weihui, Fan Yichao and Zhang Ye

View the [article online](#) for updates and enhancements.



The Electrochemical Society  
Advancing solid state & electrochemical science & technology

243rd ECS Meeting with SOFC-XVIII

**More than 50 symposia are available!**

Present your research and accelerate science

Boston, MA • May 28 – June 2, 2023

[Learn more and submit!](#)

# Umpire Gesture Detection and Recognition using HOG and Non-Linear Support Vector Machine (NL-SVM) Classification of Deep Features in Cricket Videos

Suvarna Nandyal<sup>1</sup>, Suvarna Laxmikant Kattimani<sup>2,\*</sup>

<sup>1</sup>Department of Computer Science & Engineering

P. D. A. College of Engineering (Affiliated to Visvesvaraya Technological University, Belagavi -590018), Kalaburgi-585102, Karnataka, India.

<sup>2</sup>Department of Computer Science & Engineering, B.L.D.E.A's V.P. Dr. P. G. Halakatti College of Engineering and Technology (Affiliated to Visvesvaraya Technological University, Belagavi-590018), Vijayapur-586103, Karnataka, India.

*suvarna\_nandyal@yahoo.com, cse.suvarnalk@bldeacet.ac.in*

**Abstract.** Gesture Recognition pertains to recognizing meaningful expressions of motion by a human, involving the hands, arms, face, head, and/or body. It is of utmost importance in designing an intelligent and efficient human-computer interface. The applications of gesture recognition are manifold, ranging from sign language through medical rehabilitation, monitoring patients or elder people, surveillance systems, sports gesture analysis, human behaviour analysis etc., to virtual reality. In recent years, there has been increased interest in video summarization and automatic sports highlights generation in the game of Cricket. In Cricket, the Umpire has the authority to make important decisions about events on the field. The Umpire signals important events using unique hand on signals and gestures. The primary intention of our work is to design and develop a new robust method for Umpire Action and Non-Action Gesture Identification and Recognition based on the Umpire Segmentation and the proposed Histogram Oriented Gradient (HOG) feature Extraction oriented Non-Linear Support Vector Machine (NL-SVM) classification of Deep Features. Primarily the 80% of Umpire action and non-action images in a cricket match, about 1, 93, 000 frames, the Histogram of Oriented Gradient Deep Features are calculated and trained the system having six gestures of Umpire pose. The proposed HOG Feature Extraction oriented Non-Linear Support Vector Machine classification method achieves the maximal accuracy of 97.95%, the maximal sensitivity of 98.87%, the maximal specificity of 98.89% and maximal Precision of 97.02% which indicates its superiority.

**Key words:** Computer Vision, Machine Learning, Umpire Detection, Umpire Segmentation, Umpire Gesture Recognition, Histogram of Oriented Gradients (HOG), Non-linear Support Vector Machine Classifier.

## 1. Introduction

Content portrayal and naming of video successions have developed into critical examination regions with the objective of naturally depicting the subjective video. Separating significant rank semantics from video information is a troublesome issue. To start explanation, there should be information on the area of video to be prepared and a few restrictions forced on the kinds of scenes that can be examined. In sports video, for example, umpires in the sport can wear the sensors and have their



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

advancements recorded and analyzed all through the game [1]. Since sports recordings are unscripted in nature, it is a serious moving errand to create the features in sports recordings. An effective method to produce features in sports recordings is to recognize the occasions in that game. Since the occasions in various games are not comparable, we can't utilize a typical strategy for creating client favoured video cuts from sports recordings. Among the games, Cricket is a famous game having high viewership rating from one side of the planet to the other. The game term ranges from one day to five days. Removing the significant occasions from the cricket match-up assists with featuring the client's intrigued cuts with regards to a brief period. It is achievable to perceive various actions in the game from the novel signs showed by the Umpire [3]. Sports video has been broadly concentrated because of its high business potential. For a protracted sporting event, a couple of parts are alluring to the crowd. Age of features makes it conceivable to convey sports video over slender band organizations, like Web and remote organization. Since the significance of the total game drops before long the scores are known, the watchers are just intrigued with a compact adaptation of the game and it is important to foster a framework for programmed feature age [10].

An umpire is available, who has the power to settle on choices during game-play. Prior to the game starts, the umpire flips a coin and the groups need to call the throw. The group winning the throw chooses for bat or ball and the other group needs to concur with this choice [10]. Lamentably, Cricket is no special case and the pervasiveness of erroneous choices, disregarding the openness to different donning innovation, keeps on being perturbing to its wide fan base. Reasons range from the failure of the hardware (e.g. "blind spots" because of players or umpires hindering the perspective on cameras utilized in dynamic cycles) to human mistake by the on-field umpires. In global cricket, a Third Umpire talks with the on-field umpires utilizing remote innovation. All the more explicitly, the Third Umpire utilizes TV replays in circumstances, like questioned gets and limit encroachments, to fittingly prompt the on-field umpires. The Third Umpire is likewise called upon to settle on run-out choices [11]. Hand signal acknowledgment is one of the developing fields of examination today which gives a characteristic method of human-machine communication. Signals are a few types of activities that an individual communicates to communicate data to others without saying it. In our everyday life, we can notice not many hand signals oftentimes utilized for correspondence reason like approval, disapproval, triumph, headings and so forth Some normal models are in cricket where the umpire utilizes diverse hand motions to show various occasions that happened right then and there on the match, hand signals utilized by the traffic police, and so on [12].

Signals are known as expressive, significant body development which includes actual development of the fingers, hands, arms, head, face, or body with the purpose of 1) appointing significant data or 2) interfacing with the climate. As of late motions are generally utilized by people to associate with PCs and machines [9]. The human signal comprises of significant developments of hands, arms, face, head, or different appendages [2] [14]. This is a non-verbal method of correspondence and is appropriate for human-machine cooperation [2]. Sports authorities perform numerous signals which are demonstrative of what is happening in the game. Their signals can give something significant about a player, a group, or the whole game. On the off chance that the tokens of these authorities can be perceived, significant data can be determined. We allude to a signal as a purposeful activity whereby a piece of the body is moved in a predefined approach to show a particular occasion. Recognizing these occasions empowers programmed age of features and all the more significantly, rich, relevant naming of video. To tackle this issue we need to resolve the issues of sectioning ceaseless motion information and performing vigorous motion characterization [1]. Hierarchical Covered up Markov Model (HHMM) related to a filler model are utilized for dividing and arranging motions at varying levels. The HHMM permits us to think about signals as groupings of sub-motions, conceivably reusing the sub-motion parts for motions. The filler model permits us to possibly overlook obscure developments and consequently portion and arrange motions all the while from a constant stream [1]. Motion acknowledgment utilizing video grouping has been concentrated with significant premium because of its significant applications. Signal acknowledgment can be utilized for checking patients or senior individuals, observation frameworks, sports motion investigation, human conduct examination and so on. There are a few analysts chipping away at signal acknowledgment just as motion based human-PC cooperation (HCI) [2].

In this paper, an umpire action and non-action detection and classification is developed based on Histogram of Oriented Gradients (HOG) and Non-Linear SVM classifier. The general methodology of the proposed technique includes division, highlight extraction, and umpire activity and non-activity outlines arrangement. At first, the Umpire video frames are extracted from the Umpire Frames Segmented database and the 80% of Umpire Action and Non-action frames are selected manually and feature extraction is performed based on HOG and trained to Non-Linear SVM classifier. After that, the remaining 20% of the frames, feature extraction is performed using HOG and tested using pre-trained NL-SVM Classifier, which categories the Action and Non-action Umpire Frames.

The paper is coordinated as follow: In Area II, the Writing Study covering existing strategy will be examined. Area III presents the current difficulties. Segment IV proposes a cricket dataset portrayal. Segment V frameworks the general framework plan and approach for assessing the proposed dataset as a benchmark for umpire Signal identification and Acknowledgment for feature age. Section VI introduces the Umpire detection and Segmentation. The Section VII introduces the Training and Testing Phase of the classifier. Section VIII evaluates experiments and discussions. The Section IX discusses on results of experiments. Section X concludes the paper by providing guidelines for upcoming work.

## 2. Literature Review

This segment uncovers the writing audit of a few techniques identified with umpire present identification, pitch outline order, batting shots acknowledgment, occasion, and action discovery are portrayed and broke down as follows: Chambers, G.S et al.[1] developed an approach for Automatic labelling of sports video using umpire gesture recognition using Hierarchical hidden Markov model. It considers gestures at different levels of abstraction and handles extraneous umpire movements. But however, it failed to use the filler model ratio to provide more insight. A Heickal, H et al.[2] presented an approach for real-time 3D gesture recognition using depth image adopting Naive Bayes and neural network classifier. It attained better accuracy as 98.11% using neural network and 88.84% using Naive Bayes method and failed to use the features, like finger positions. Hari, R. and Wilscy, M et al.[3] developed the K-means segmentation algorithm for occasion identification in cricket recordings utilizing power projection profile of Umpire signals. This algorithm effectively recognizes the occasions in the cricket video. This calculation yields an awesome outcome for put away cricket video, which can likewise be utilized progressively with devoted equipment support. The method does not consider excusals of a batsman by getting the ball by the defender of inverse side are not flagged as expected by the Umpire. Bhansali, L. and Narvekar, M [4] employed Gradient Method for Gesture recognition to make umpire decisions and It was capable of recognizing a group of six umpire gestures from the game of cricket and performs best once using a feature set. The main drawback of this method is, the performance of segmenting gestures from a stream of continuous gestures by selecting candidate gestures by the existence of movement was poor. Javed, A et al.[5] employed a Multimodal framework for audio-visual features for summarisation of cricket videos and does not consider longest match durations and broadcasting time concerns, selecting the test case was difficult. Javed, A et al. [6] developed a Dual-threshold-based method for automatic highlights generation from sports videos. This was more efficient as it does not rely on logo template recognition for replay detection. However, the performance of this method was poor. Premaratne, S.C. and Jayaratne, K.L [7] employed a Multi-model mining approach for identifying specific events uniquely. But, The detection rate of this model was poor. Javed, A et al.[8] developed Rule-based induction method for signifying the effectiveness in terms of video summarization. Due to the longest match durations and broadcasting time poses a challenging problem in this method. Md. Asif Shahjalal et al.[9] proposed an approach to Automate the Scorecard in Cricket with Computer Vision and Machine Learning using region of interest, a Haar-cascade-classifier and gesture is recognized using logistic regression, but dynamic gesture like four the efficiency was found to be poor. Kolekar, M.H. and Sengupta, S.[10] designed and developed a method for Semantic concept mining in cricket videos for automated highlight generation using a priori algorithm and SEQAD. But, more semantic concepts such as wide ball, no ball, type of wicket,

type of hit, etc will not be considered. S. Mitra and T. Acharya[11] proposed Gesture recognition pertains to recognize meaningful expressions of motion by a human, involving the hands, arms, face, head, and/or body. It is of utmost importance in designing an intelligent and efficient human-computer interface. Sigha, J. and Das K [12] proposed a system based on K-L Transform to recognize different hand gestures but it was not able to recognize gestures like STOP, FIVE, and THUMBS DOWN. Chambers G.S.et al. [13] developed a method where Actors who perform intentional gestures wear accelerometer sensors in the form of wrist bands. Gesture recognition is then performed in the sensor domain which avoids the problems associated with accurate image segmentation but officials require very lightweight and unobtrusive devices such as wrist bands. P. Premaratne et al. [14] proposed glove based systems to realize computer vision based hand gesture recognition without any sensors attached to the glove but different colour gloves are to be mined for different hand gestures. I. Gomez-Conde et al.[15] developed an approach to detect the persons in the scene and analyze different actions and gestures in real time using feature vector and histogram based algorithms. But a future report could join this data with present day consecutive Monte Carlo techniques to get more robust tracking.

### 3. Challenges

Gesture recognition is a difficult and challenging process using cricket videos. To recognize the gesture movement of umpire in the cricket videos poses several issues, hence recognizing the gestures of umpire is a tedious and complex process. Here are some of the challenges faced by the cricket videos and gesture recognition model.

In [13] the Hidden Markov Model methodology was presented for motion displaying with both confined signals and motions divided from a stream. It was capable of recognizing a set of 10 umpire gestures from the game of cricket and performs best when using a feature set. However, it selects only the candidate gestures.

In [6] double limit based strategy was displayed to recognize the frames from the information video. It does not rely on logo template recognition for replay detection, which makes it computationally efficient. It was not effective with more diverse dataset.

In [12] KL Transform was developed to recognize different hand gestures. It utilizes the point based characterization to distinguish which image the test picture has a place with. However, it only deals with the few gestures.

In [9] Haar course classifier was created to choose the locale of interest. It was end up being an exceptionally straightforward yet effective calculation for umpire's motion discovery. However, the efficiency found in the dynamic gesture was poor.

Markov model was introduced in [1] to solve the problem of robust gesture classification. It allows handling extraneous umpire movements. However, pause between regions of movement were not considered.

To build a learning system capable of learning a gesture classification problem from dynamic video data is significantly a complex issue in gesture recognition.

Given a test Cricket Video, our motivation is to Detect, Recognize and Classify as Umpire Action and Non-Action Gestures. The Non-Linear Support Vector Machine is trained on Histogram Oriented Gradient features calculated on 80% of manually selected frames to recognize six Gestures - "SIX", "NO-BALL", "WIDE", "OUT", "LEG-BYE", "FOUR"- used by the umpire in a Cricket game as Umpire Action Class and similarly 80% of frames to recognize Non-Action Class. The remaining 20% frames are tested and validated based on Knowledge Base created at training phase of the system for Event Detection.

### 4. Cricket Dataset Description

For Umpire Gesture Detection and Recognition, we utilize One Day International World-Cup-2013 cricket video matches of range 2:12:38 seconds. Appropriately making a total of 20 accounts of at regular intervals and complete edges are 1, 93,000 open in the dataset. Test dataset SNWOLF traces are shown in figure 1. In our proposed approach, we haphazardly and manually selected 80% of Umpire Action and Non Action frames from total Dataset frames for Feature Extraction using

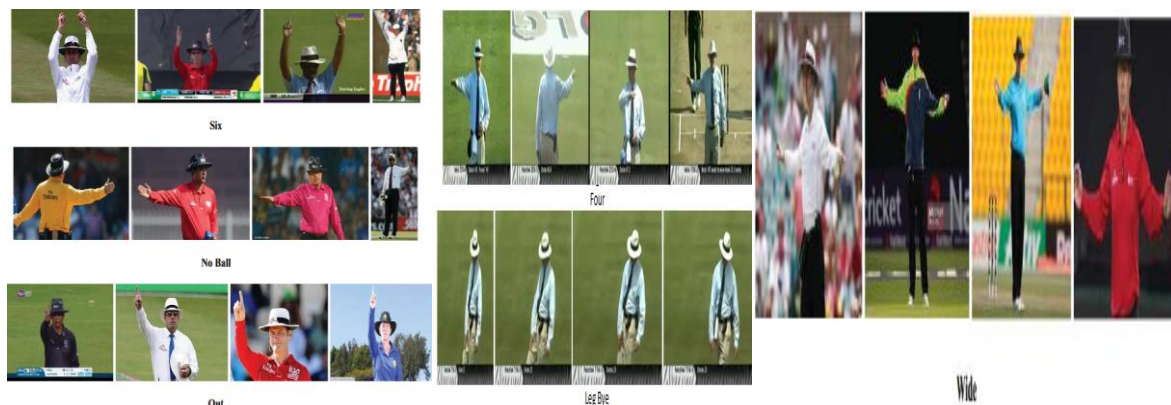
Histogram Oriented Gradient algorithm as shown in figure 2. & figure 3, respectively. Extracted HOG Features of Umpire Action and Non-Action frames are fed to train the Non-Linear Support Vector Machine for classification to create knowledge base. Remaining 20% frames are classified as Umpire Action and Non-action frames based on trained knowledge base of NL-SVM experiments.



**Figure 1.** Sample Cricket Video Dataset-SNWOLF-for classifying Umpire Action and Non-Action Frames

#### 4.1. Umpire Action Dataset

We have arbitrarily gathered 80% of pictures of umpires performing different activities relating to occasions, for example, "Six", "No Ball", "Wide", "Out", "Leg Bye", and "Four". These pictures have been gotten from One Day International World-Cup-2013 cricket video matches of length 2:12:38 seconds. The dataset includes six classes of information outlines. Every one of the six classes having a place with the six activities and one Non-Action class in which the umpire Action doesn't exist as displayed in figure. 3. The umpire Action class consists of 1040 images for all six classes. Figure.2. shows a portion of the pictures in the dataset for the six classes of occasions.

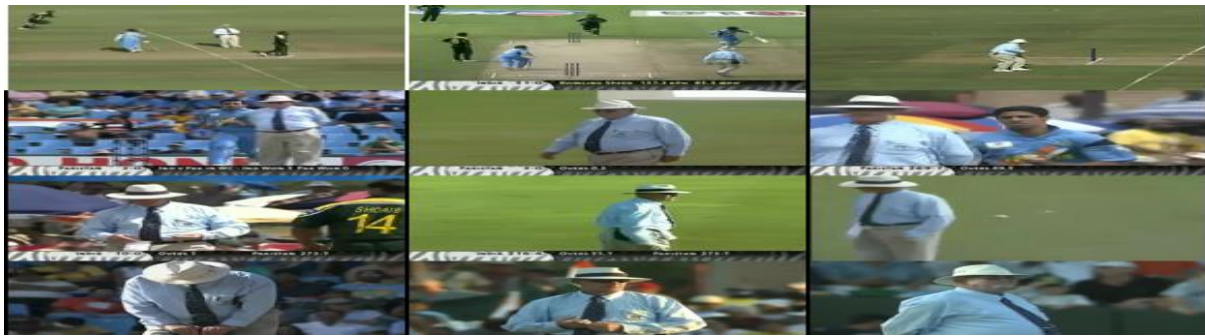


**Figure 2.** Illustrating Gesture images of Umpire Action Class such as "Six", "No Ball", "Wide", "Out", "Leg Bye", and "Four" in SNWOLF dataset.

#### 4.2. Umpire Non-Action Dataset

We have randomly collected 80% of images of umpires performing various umpire Non- actions not relating to occasions for example "Six", "No Ball", "Wide", "Out", "Leg Bye", and "Four". These photos have been gotten from ODI world-Cup-2013 cricket video matches of term 2:12:38 seconds. The dataset includes Non-Action class in which the umpire Action doesn't exist as displayed in figure 3.

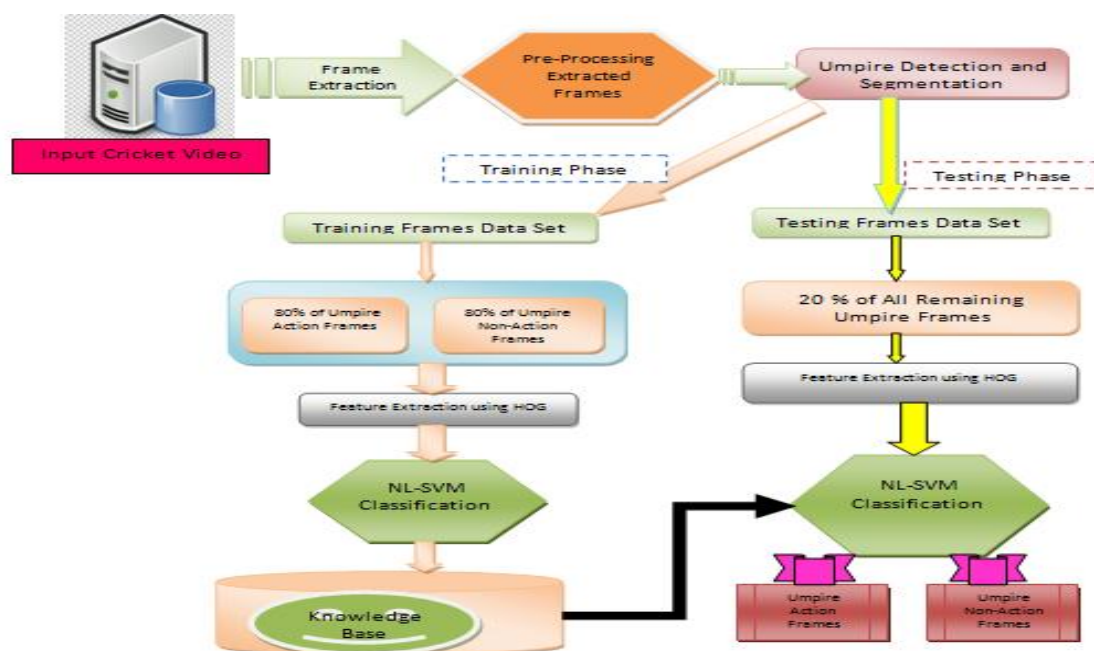




**Figure 3.** Illustrating Gesture images of Non-Action Umpire Class in SNWOLF Dataset

## 5. Proposed Methodology

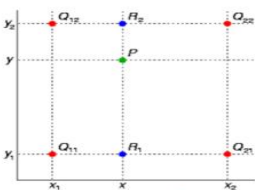
Our proposed strategy employs One Day International Cricket Video of length 2:12:38 seconds having 1,93,000 edges named SNWOLF Dataset. The 80% of frames are randomly selected for each Umpire Action and Non-Action classes from total SNWOLF dataset for NL-SVM training purpose and remaining 20% of frames are for Testing purpose. It is sufficient & thoroughly examined as information and yield will be Umpire Action and Non-Action class Classification out of Testing experiment using Knowledge Base generated out of Pre-trained network phase of Non-Linear Support Vector Machine. The figure 4. Exhibits the proposed methodology and suggested work.



**Figure 4.** A conceptual overview of Umpire Action and Non-Action Gesture Detection, Recognition and Classification using HOG and Pre-Trained Network as Non-Linear Support Vector Machine analysis for Event Detection.

### 5.1. Image Pre-processing

**5.1.1. Picture Resizing utilizing Interpolation strategy.** An insertion method that diminishes the visual twisting brought about by the partial zoom computation is the bilinear addition calculation, where the fragmentary piece of the pixel address is utilized to process a weighted normal of pixel splendor esteems over a little neighborhood of pixels in the source picture. Bilinear interjection produces pseudo goal that gives an all the more tastefully satisfying outcome by equation 1.

$$f(x, y) \approx \frac{f(Q_{11})}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y_2 - y) + \frac{f(Q_{21})}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y - y_1) + \frac{f(Q_{12})}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y_2 - y) + \frac{f(Q_{22})}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y - y_1).$$

(1)

**Figure.5.** Bilinear Interpolation for Resizing an Image

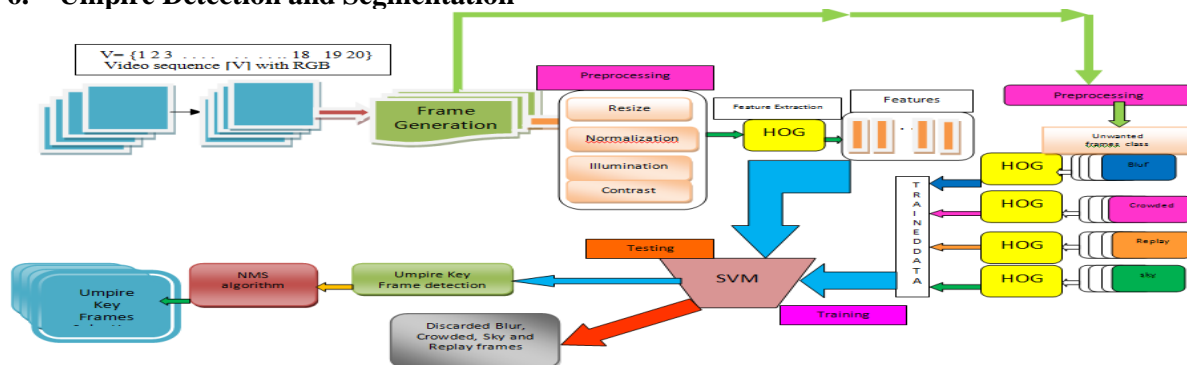
**5.1.2. Power Normalization (Histogram).** Standardization is a regular strategy that is used to overhaul fine detail inside an image. Each fragment in the joined histogram is dealt with as the proportion of all the picture power histogram respects up to and including that faint level, and from there on, it is scaled so the last worth is 1.0 using equation 2.

$$I_N = (I - \text{Min}) \frac{\text{newMax} - \text{newMin}}{\text{Max} - \text{Min}} + \text{newMin} \quad (2)$$

**5.1.3. Difference Enhancement.** Picture contrast upgrade is a vital advance in picture handling [10]. In computerized picture preparing, the dark level picture contains a variety of advanced tally esteems from 0 to 255, is not the same as some other shading picture in light of the fact that each pixel requires less information, so contrast improvement kills the issue by developing the amazing degree of computerized upsides of the pixels.

$$O(x, y) = f(x, y) + (I \text{ mean} - G \text{ mean}) \quad (3)$$

## 6. Umpire Detection and Segmentation



**Figure. 6.** A calculated outline of our methodology. The suggested approach precisely sections the umpire key casings by utilizing disposing of obscure, swarmed, sky, and replay outlines in a video to keep away from non-useful edges. Separate examination to perform on the edge savvy quality withdrawal for four classes of casings utilizing HOG and pre-prepared organization Support Vector Machine used to check and group the leftover video as an input.

Specified an analysis Cricket Video, our motivation is to remove HOG components for the undesirable edges and because of this reality educate a direct Support Vector Machine model to look at to check the end outlines for grouping as displayed in figure 6. To enroll the umpire divided key frames required for umpire Action and Non-Action Gesture detection, recognition and classification for Event Detection in Highlights Summary Extraction for given Cricket Sport.

## 7. Training Phase-Classifer Design

For the purpose of Umpire Action and Non-Action Detection, Recognition and Classification on Cricket video SWOLF dataset, the classifier is designed to recognize a packaging containing an umpire Action versus a packaging that doesn't contain an umpire Action. Two sets of data were created. One set containing all umpire Action has a place with pictures from the SNWOLF dataset having a place with one class, and a subsequent set containing Umpire non-Action pictures having a



place with the other class. The Non-Linear Support Vector Machine Classifier is designed for the purpose of umpire Action and Non-Action Gesture recognition and classification for action detection. This classifier is prepared on 1040 umpire Action pictures having a place with six classes of occasions like Six, No Ball, Wide, Out, LegBye, Four, and 5399 Umpire Non-Action pictures, which are physically and haphazardly chose from the SNWOLF dataset.

The means engaged with planning the classifier are normal to the two Classes 1 and 2. The pre-handling is proceeded as the initial phase in preparing the pictures for the SNWOLF dataset. For the purpose of Umpire Action and Non-Action recognition and classification, the Umpire Detection and Segmentation will be done by eliminating unwanted images to keep more informative Umpire frames. Once Umpire candidate images are listed from SNWOLF dataset, the 80 % of Action and Non-Action Umpire images are manually and randomly selected to extract Histogram Oriented Gradient Features. The length of the feature vector is 34000. These features are forwarded to train the Non-linear Support Vector Machine to create Knowledge Base for the further classification to remaining 20 % frames in SNWOLF dataset and to give the best outcomes. Table 1. sums up the presentation of characterization through the disarray grid.

*Algorithm: Umpire Action/Non-Action Frames HOG features Extraction*

- Step1: Read Umpire Action /Non-Action images from Training Action /Non-Action Dataset folder
- Setp2: Set Training Ratio to 0.8
- Step3: Extract HOG features from true color or grayscale image I and returns the features in a 1-by-N vector for each image in Umpire Action/Non-Action Set
- Step4: Label “1” for each image in Action Dataset and Label “2” for image in Non-Action Dataset

*Algorithm: Extract HOG features for Testing Dataset at Testing Phase*

1. Read remaining 20% of frames from SNWOLF database
2. Pre-process the current frame
3. Detect the people in the frame
4. Detect the Umpire and Segment the Umpire frame
5. Extract the HOG features and assign class label for each frame in Testing Dataset
6. Forward HOG features and class labels of Testing Dataset to NL-SVM for Testing to classify into Umpire Action or Non-Action class by using Knowledge base created at Training Phase.
7. The test exactness is determined dependent on the excess 20% of the concealed information.
8. The Highest Result Accuracy is seen through confusion matrix

## 8. Experiments and Discussions.

NL-SVM Classifier is trained for umpire Action and Non-Action discovery. From the outcomes, it is obvious that Classifier has a decent presentation of precision achieved in testing phase as shown in figure 7. confusion matrix Action Frame Selection. The statics are generated while testing are Precision, Accuracy, Sensitivity/Recall and Specificity by Equations 4,5,6,7,8.



**Figure 7.** (a).Confusion matrix for Umpire Action and Non-Action frame selection.(b)Images of Database of Umpire Action class -“Six”,”NoBall”,”Wide”,”Out”,”Leg-Bye”,”Four”. (c) Images of Database of Umpire Non-Action class.

$$\text{precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{Accuracy} = \frac{TP + TN}{P + N} \quad (6)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (7)$$

## 9. Results

At first, the HOG highlights were determined for 80% of each Umpire Action and Non-Action pictures. However, while calculating the HOG features for the Umpire Action and Non-Action images, which are contributions to the classifier, the cell size utilized is 16 x 16. This is done to lessen the quantity of estimations associated with HOG calculation, simultaneously keeping away from any corruption in the classifier execution.

The Radial Bias Kernel Function algorithm based NL-SVM classifier is implemented for the preparation and testing of the order stage. One Day International Cricket Match of duration 2:12:38 seconds having 1, 93,000 images Dataset named SNWOLF is utilized. It was picked in light of the fact that it gives a different subset of information base centered towards the various assignments engaged with a Umpire Gesture Recognition, namely People Detection, Umpire Detection, Umpire Segmentation and Classification for overall system evaluation. For preparing of classifier, it gives 1040 physically chose positive (or Umpire-Action) pictures and 5399 physically chose negative (or Umpire Non-Action) pictures. A classifier prepared with additional preparation information would be better proficient at precisely recognizing an Umpire Action and Non-Action images of remaining 20% data frames. Hence to bring in more Umpire Action and Non-Action present situations, the SNWOLF dataset utilizes Umpire competitor pictures of the Umpire Segmented dataset. The demonstrated work uses 80% of 1040 Action and 5399 Non-Action Umpire images for training of classifier to generate the knowledge base. For testing, the described work uses a total of 832 Umpire Action and 1079 Non-Action images by using knowledge base generated at training phase. Figure 9(b) and (c). Show two testing classes of images from both the Umpire Action and the Non-Action class of the SNWOLF Dataset. Table1. Sums up the significant boundaries relating to the preparation and testing of the Non-Linear SVM classifier and accomplishes a greatest edge of execution.

**Table 1.** Boundaries utilized in preparing and testing of Classifier

Factor	Rate
Fragment/Crop size	120 * 48
No. of exercise images	6439(1040 +ve & 5399 - ve)
No. of +ve test images	980
No. of -ve test images	5369
Cell-Size	16*16
Features vector length/cell	3400

Out of the 1040 positive test pictures, 980 edited Umpire Action pictures were delegated Umpire Action pictures by the executed component classifier pair, adding up to a genuine positive grouping exactness of 98.60%. With 1040 pictures, the classifier yielded a bogus positives grouping level of 2.97 % i.e., 60 Action pictures were misclassified as Non-Action, and table 2., gives a rundown of the outcomes acquired, while figure 8., gives a graphical portrayal of the equivalent.

**Table 2.** Classification Results

Classification	Total number of test images		Number of classified images		Percentage of Classification
	Positive	Negative	True	False	
True Positives (TP) (Umpire Actions classified as Umpire Actions)	980	60	980	NA	97.02%
False Negatives (FN) (Umpire Actions classified as Umpire Non-Actions)	30	NA	NA	60	1.10%
True Negatives (TN) (Umpire Non-Actions classified as Umpire Non-Actions)	NA	5369	5369	NA	98.89%
False Positives (FP) (Umpire Non-Actions classified as Umpire Actions)	30	NA	NA	30	2.97%

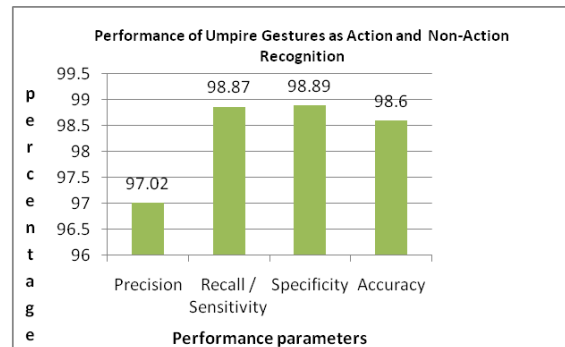
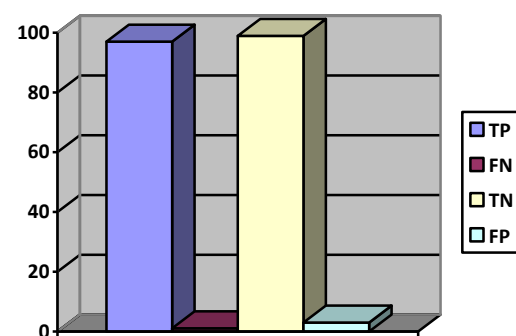
**Figure 8.** Graphical Representation of Classification Results

Table 3 assesses the classifier execution utilizing some ordinarily utilized measurements. From the chart in figure 9 and the boundary esteems acquired in table 3, it very well may be presumed that the executed HOG-NL-SVM include classifier pair shows promising outcomes, with a high evident positive characterization exactness and a negligible number of misclassification as 2.97%. The level of the Umpire Action which was wrongly arranged is addressed in figure 9 as "fake negatives".

**Table 3.** Classification performance parameters

Classifier Performance Parameters	Formula	Value	Percentage
<b>Precision</b> (Positive predictions that are correct)	$TP / (TP + FP)$	0.9702	97.02
<b>Recall / Sensitivity</b> (Positive labeled segments that were predicted as positive)	$TP / (TP + FN)$	0.9887	98.87
<b>Specificity</b> (Negative labeled segments that were predicted as negative)	$TN / (TN + FP)$	0.9889	98.89
<b>Accuracy</b> (Predictions that are correct)	$(TP + TN) / (TP + TN + FP + FN)$	0.9860	98.60

**Figure 9.** Non-Linear SVM classifier Testing Dataset

## 10. Conclusion and Future Work

In this paper, execution of two key squares in the Umpire Gesture Detection, Recognition and Classification structure, specifically include extraction and grouping are introduced. Histogram Oriented Gradient highlights are carried out for include extraction, with a cell size of 16x16 (for computational accelerate) and a proficient standardization technique (for enlightenment invariance). A soft-margin Non-Linear SVM, based on the HOG algorithm is used for implementation of the classification module. The classifier uses a subset of the SNWOLF Cricket Sports Dataset, which is specifically aimed at the training and testing of the classification stage. The outcomes got show high Umpire Action and Non-Acton Umpire order exactness (TP) of 98.60 % and a general arrangement precision (TP+TN) of 97.95 %. Henceforth, the carried out include classifier gathering can go about as a quick and strong structure block for a total Umpire Gesture Recognition and Classification. Anyway the future work can likewise be planned to additionally ad lib the consequences of the introduced work, measures like expanding the measure of preparing information of the classifier and utilizing an appropriate confirmation methodology — to lessen the quantity of false positives incurred—can be embraced.

## References

- [1] Chambers G S, Venkatesh S and West G A 1969 *Automatic labeling of sports video using umpire gesture recognition* In Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR), Springer, Berlin, Heidelberg, pp. 859-867, August 2004
- [2] Heickal H, Zhang T and Hasanuzzaman M 2015 *Computer vision-based real-time 3D gesture recognition using depth image* International Journal of Image and Graphics, vol. **15** no. 01, p.1550004
- [3] Hari R and Wilsy M 2014 *Event detection in cricket videos using intensity projection profile of Umpire gestures* IEEE India Conference (INDICON) pp 1-6
- [4] Bhansali L and Narvekar M 2016 *Gesture recognition to make umpire decisions* International Journal of Computer Applications, vol. **148**, no. 14
- [5] Javed A, Irtaza A, Malik H, Mahmood M T and Adnan S 2019 *framework based on audio-visual features for summarization of cricket videos* IET Image Processing, vol. **13**, no. 4, pp.615-622
- [6] Javed A, Bajw K B, Malik H and Irtaza A 2016 *An efficient framework for automatic highlights generation from sports videos* IEEE Signal Processing Letters, vol. **23**, no. 7, pp.954-958
- [7] Premaratne S C and Jayaratne K L December 2017 *Structural approach for event resolution in cricket videos* In Proceedings of the ACM International Conference on Video and Image Processing, pp. 161-166
- [8] Javed A, Bajwa K B, Malik H, Irtaza A and Mahmood M T October 2016 *A hybrid approach for summarization of cricket videos* IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), pp. 1-4
- [9] Shahjalal M A, Ahmad Z, Rayan R and Alam L December 2017 *An approach to automate the scorecard in cricket with computer vision and machine learning* IEEE 3rd International Conference on Electrical Information and Communication Technology (EICT), pp. 1-6
- [10] Kolekar M H and Sengupta S 2010 *Semantic concept mining in cricket videos for automated highlight generation* Multimedia Tools and Applications, vol. **47**, no. 3, pp.545-579
- [11] Mitra S and Acharya T 2007 *Gesture recognition: A survey* IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. **37**, no. 3, pp. 311–324
- [12] Singha J and Das K 2013 *Hand gesture recognition based on Karhunen - Loeve transform* arXiv preprint arXiv: 1306.2599

- [13] Chambers G S, Venkatesh S, West G A and Bui H H 2004 *Segmentation of intentional human gestures for sports video annotation* IEEE International Multimedia Modeling Conference, Proceedings, pp. 124-129
- [14] Premaratne P 2014 *Historical development of hand gesture recognition* Springer, pp. 5–29
- [15] Gomez-Conde I, Olivieri D, Vila X A and Orozco-Ochoa S 2011 *Simple human gesture detection and recognition using a feature vector and a real-time histogram based algorithm* Journal of Signal and Information Processing, vol. **2**, no. 04, p. 279