International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 10 | October - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

CRICKET POSE PREDICTION USING DEEP LEARNING

HARISH A.S, AAGASH SETHUPATHI N, VISHAL A.S

ABSTRACT

Cricket has a massive global following and is ranked as the second most popular sport globally, with an estimated 2.5 billion fans. Batting requires quick decisions based on ball speed, trajectory, fielder positions, etc. Recently, computer vision and machine learning techniques have gained attention as potential tools to predict cricket strokes played by batters. This study presents a cutting-edge approach to predicting batsman strokes using computer vision and machine learning. The study analyzes six strokes: pull, square-cut, cover drive, straight drive, flick, and sweep. The study uses the ResNet model to extract features from videos and to predict the strokes. The methodology involves data collection and pre-processing and using pre-trained model ResNet-34 CNN for cricket pose classification.The model captures video frames,preprocess them,predict the strokes.The study achieves an outstanding accuracy of 90.77% using the ResNet model, outperforming the other algorithms used in the study.. The study's results could help improve coaching techniques and enhance batsmen's performance in cricket, ultimately improving the game's overall quality.

KEYWORDS: *Cricket shot prediction, deep learning, Fastai CNN, Convolutional Neural Network, real-time video analysis, sports analytics.*

INTRODUCTION

Human pose estimation (HPE) is a rapidly developing field of research that employs computer vision techniques to estimate the positions of various human body components in images or video footage. Despite recent advancements in computer vision, accurately understanding human actions from visual data is still challenging. Human body movements are often driven by unique activities, making identifying and categorizing them accurately difficult. Understanding a person's body pose is crucial for identifying their actions, which is where HPE techniques come in handy. By recognizing and categorizing human body joints, such as the head, arms, and torso, HPE can capture coordinates for each joint that define a person's position.

In sports analytics, computer vision has become increasingly crucial for extracting valuable insights from various forms of visual data . Coaches and athletes can use computer vision techniques to track and analyze movement patterns during games or practice sessions, providing valuable performance feedback, identifying areas for improvement, and making strategic decisions [3]. Additionally, computer vision can be used for activity recognition, outcome prediction, and injury prevention. Using computer vision in sports can revolutionize how we analyze and train athletes, improving their performance and reducing the risk of injury.

One of the key challenges faced in this project is the accurate prediction of six distinct cricket strokes: the pull, square-cut, cover drive, straight drive, flick, and sweep. Each of these strokes requires a unique combination of body positioning, bat swing, and timing. Furthermore, the project addresses the real-time nature of the sport, where predictions need to be made within milliseconds to be useful. Achieving high accuracy in stroke prediction in such a dynamic environment is no small feat. To tackle these challenges, the project leverages state-of-the-art technology, particularly the ResNet model, a deep convolutional neural network. This model excels at extracting intricate features from video data, allowing for precise stroke classification. The project's methodology includes robust data collection, meticulous preprocessing of cricket pose images, and the utilization of a pre-trained ResNet-34 CNN. By capturing and analyzing video frames in real-time, the system can make rapid predictions, providing valuable insights into a batter's technique. Remarkably, the project attains an outstanding accuracy rate of 90.77% using the ResNet model, outperforming alternative algorithms. These results have the potential to revolutionize cricket coaching techniques, offering a new level of precision and feedback for players. Ultimately, this project not only enhances individual batting performance but also contributes to elevating the overall quality of the game itself.

- Cricket's Global Significance: Cricket is a widely beloved sport with an immense global following of approximately 2.5 billion fans, making it the second most popular sport worldwide.

- Complexity of Batting: Batters in cricket face the challenge of making quick and precise decisions based on various factors like ball speed, trajectory, and fielder positions, highlighting the complexity of the sport.

- Project Objective: This project seeks to leverage cutting-edge computer vision and machine learning techniques to predict specific cricket strokes made by batters. The study focuses on six key strokes: pull, square-cut, cover drive, straight drive, flick, and sweep.
- ResNet Model Utilization: The project employs the ResNet model, a powerful deep convolutional neural network, to extract intricate features from video data and classify cricket poses accurately. The ResNet-34 CNN is used as a pre-trained model to achieve high accuracy.
- Game-Changing Results: The project attains an impressive accuracy rate of 90.77% using the ResNet model, surpassing other algorithms examined in the study. These results have the potential to revolutionize cricket coaching techniques and enhance the performance of batters, ultimately elevating the quality of the game itself.

RELATED WORK

- The paper “Enhancing Cricket Performance Analysis with Human Pose Estimation and Machine Learning” by Hafeez Ur Rehman Siddiqui, Faizan Younas, Furqan Rustam, Emmanuel Soriano Flores, Julién Brito Ballester, Isabel de la Torre Diez, Sandra Dudley, and Imran Ashraf was published in the journal Sensors in 2023. The paper discusses the use of human pose estimation and machine learning to enhance cricket performance analysis. Cricket has a massive global following and is ranked as the second most popular sport globally, with an estimated 2.5 billion fans. The paper presents a novel approach to analyzing cricket performance using advanced techniques such as human pose estimation and machine learning.

- The paper “Pose Recognition in Cricket using Keypoints” by Rahul Mili, N.R. Das, A. Tandon, S. Mokhtar, I.

- Mukherjee, and G. Paul was presented at the 2022 IEEE 9th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON). The paper discusses a technique for recognizing the gestures and poses of an umpire in cricket using keypoints generated through pose estimation. The researchers worked on a dataset termed SNOW for the detection of umpire poses in cricket. They tried to identify five signals: NO BALL, SIX, WIDE, OUT, and NO ACTION based on the umpire's pose from the frames of a cricket video. The experimental results showed that the accuracy of their proposed technique was 87%, and the evaluation metrics were quite promising compared to existing state-of-the-art works.

- The paper "Deterministic 3D human pose estimation using rigid structure" by Jack Valmadre and Simon Lucey was presented at the 11th European Conference on Computer Vision (ECCV) in 2010. The paper discusses a technique for estimating the 3D pose of a human using a rigid structure. The authors propose a deterministic approach to estimate the 3D pose of a human from a single image, which is based on the assumption that the human body can be modeled as a rigid structure. Their approach is able to estimate the 3D pose of a human even in challenging scenarios such as when the person is wearing loose clothing or when the image is taken from an unusual viewpoint.

- The paper "Human Pose Estimation and Object Interaction for Sports Behaviour" by Ayesha Arif, Yazeed Yasin Ghadi, Mohammed Alarfaj, Ahmad Jalal, Shaharyar Kamal, and Dong-Seong Kim was published in the journal Computers, Materials & Continua in 2022. The paper discusses the use of human pose estimation and object interaction for understanding sports behavior. The authors propose a novel approach to demonstrate that appearance features alone are not sufficient to predict human-object interaction (HOI). They detect human body parts using the Gaussian Matric Model (GMM) and detect objects using YOLO3. The interactions are linked with the human and object to predict actions. The experiments have been performed on two benchmark HOI datasets demonstrating the proposed approach.

- The dissertation "Cricket Batting Analysis Using Pose Estimations and Deep Neural Network" by Lakjeewa Wijebandara was published in 2020. The dissertation discusses the use of pose estimations and deep neural networks to analyze cricket batting. The author proposes a novel approach to extract features from the pose estimations of cricket players and use them to train a deep neural network for classification. The proposed approach is able to accurately classify different types of cricket shots and can be used to provide feedback to players on their batting technique.

3.PROPOSED METHODOLOGY

SAMPLE IMAGE OF DATASET

The study collected videos of eight different types of batsman strokes. The videos were preprocessed to remove any noise present to ensure accurate analysis.

The proposed algorithm for cricket stroke prediction combines computer vision and deep learning techniques to accurately classify and predict cricketing strokes from video data. The algorithm employs a pre-trained ResNet-34 convolutional neural network (CNN) as the feature extractor and a custom-designed stroke prediction model. The following paragraphs provide a detailed explanation of the algorithm.

The first step of the algorithm involves feature extraction from video frames. Each frame is passed through the ResNet-34 model, which has been fine-tuned on a curated dataset of cricket stroke videos. The output of the ResNet-34 model, which consists of high-level features, serves as the basis for stroke prediction. These features capture spatial information from individual frames, enabling the model to discern critical visual cues indicative of different cricketing strokes.

The study collected videos of eight different types of batsman strokes. The videos were preprocessed to remove any noise present to ensure accurate analysis.

The proposed algorithm for cricket stroke prediction combines computer vision and deep learning techniques to accurately classify and predict cricketing strokes from video data. The algorithm employs a pre-trained ResNet-34 convolutional neural network (CNN) as the feature extractor and a custom-designed stroke prediction model. The following paragraphs provide a detailed explanation of the algorithm.

The first step of the algorithm involves feature extraction from video frames. Each frame is passed through the ResNet-34 model, which has been fine-tuned on a curated dataset of cricket stroke videos. The output of the ResNet-34 model, which consists of high-level features, serves as the basis for stroke prediction. These features capture spatial information from individual frames, enabling the model to discern critical visual cues indicative of different cricketing strokes.

Mathematically, the feature extraction process can be represented as follows:

$$\mathbf{F}\text{frame} = \text{ResNet34}\,(\,\mathbf{I}\text{frame}\,)$$

Where:

- **F**frame represents the feature vector extracted from a single video frame.
- ResNet34 denotes the fine-tuned ResNet-34 model.
- **I**frame is the input video frame.

Once the features are extracted from each frame, they are passed through a stroke prediction model. The stroke prediction model is designed as a multi-class classifier, where the classes correspond to different cricketing strokes (e.g., pull, square-cut, cover drive). The model employs a softmax activation function to compute the probabilities of each class for every frame. The cricketing stroke with the highest probability is predicted as the player's current stroke. Mathematically, the stroke prediction process can be represented as follows:

$$\text{Pstroke} = \text{Softmax}\,(\,W \cdot F\text{ frame} + b\,)$$

Where:

- **P**stroke is the predicted probability distribution over cricketing strokes.
- Softmax represents the softmax activation function.
- **W** and **b** are the model parameters learned during training. The proposed algorithm combines these steps to make real-time stroke predictions from cricket video data, providing valuable insights for players, coaches, and analysts.

Algorithm:

```
# Step 1: Data Preparation
Load and organize the cricket pose dataset
Split the dataset into training and validation sets
Preprocess the data (resize, data augmentation)

# Step 2: Model Building
Select deep learning architecture (Fastai with ResNet-34)
Create data loader for batch processing
Fine-tune the ResNet-34 model on the dataset

# Step 3: Real-time Pose Prediction
Initialize video camera
Loop:
   # Capture a video frame
   ret, frame = capture_frame_from_camera()

   if not ret:
      break

   # Preprocess the frame (resize)
   resized_frame = resize_frame(frame)

   # Convert the frame to Fastai Image format
   pil_frame = convert_to_pil_image(resized_frame)

   # Predict the cricket pose category using the model
   prediction = model.predict(pil_frame)

   # Calculate confidence score (optional)
   confidence_score = calculate_confidence(prediction)

   # Overlay predicted pose category (and confidence score) on the frame
   overlay_text = f"Pose: {prediction} (Confidence: {confidence_score:.2f})"
   overlay_text_on_frame(frame, overlay_text)

   # Display the frame with overlays
   display_frame(frame)

   # If 'q' key is pressed, exit loop
   if key_pressed('q'):
      break

# Step 4: Key Points Visualization (Optional)
If applicable:
   Extract keypoints or features from the frame
   Visualize keypoints on the frame

# Step 5: Termination
Release video capture
Close display windows
```

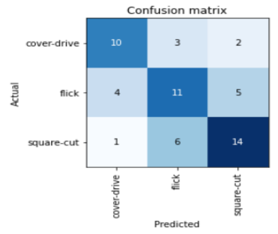
RESULTS AND DISCUSSION:

Results

Performance comparison with existing studies was also carried out in this study. For this purpose, models from existing studies for batsmen's stroke prediction are selected. The studies use images and video datasets to predict different types of strokes. These studies employ AlexNet, LSTM, and RF models for stroke prediction. Performance analysis given in table indicates that the RF used in the current study shows better performance with a higher number of strokes prediction and outperforms existing approaches.

Strokes	Strokes	Model	Accuracy
Glance, drive, block, and cut	4	AlexNet	74.33%
Cut, cover drive, straight drive, pull, leg glance, scoop	6	RF	87%
Backward and forward	2	LSTM	100%
Strokes and gameplay	2	AlexNet	96.66%
Straight drive, on drive, cover driver, cut, pull, sweep, flick, back foot punch	8	RF	99.7%

The performance of AlexNet is moderate with a 0.841 accuracy score. However, the best performance is obtained by the RF, which has a 0.997 accuracy score.



This is a confusion matrix used to evaluate the performance of a machine learning model. The matrix is a square grid with four rows and four columns. The rows and columns are labeled with the names of the classes that the model is trained to predict: “Cover-drive”, “flick”, “square-cut”.The cells of the matrix contain the number of predictions made by the model for each combination of actual and predicted classes. The diagonal cells, which represent correct predictions, are shaded in blue. The off-diagonal cells, which represent incorrect predictions, are shaded in white.

DISCUSSION:

The use of human pose estimation holds several strategic advantages in sports. It can be used by coaches to train players better and enhance their sports performance. Cricket, being the second most popular sport in the world, is liked and followed by billions of people around the globe. Consequently, coaches and players are continuously striving for excellence. The use of deep learning techniques to predict batsmen’s strokes can be very influential and useful in this regard. This study collects video data for different strokes and proposes a deep-learning approach for stroke prediction. Different important features are extracted from the preprocessed video data to train deep learning models. After training and testing the models, it was found that the ResNet model achieved the highest accuracy among all the other machine learning and deep learning models. The Resnet model has a higher accuracy , recall, F1 score, Cohen’s kappa, and geometric mean than other models. Additionally, the log loss was lower for the ResNet model. Overall, the results indicate that the ResNet model is the most suitable model for classifying the batsmen’s strokes in the cricket pose dataset. It can accurately identify the various types of strokes played by the batsmen with high accuracy and recall, making it a reliable tool for stroke analysis.

CONCLUSIONS:

Cricket stroke classification is proposed in this study using a deep learning approach to enhance the performance of batsmen. Several machine learning and deep learning algorithms are used to benchmark the newly collected video dataset. A novel video dataset is created that contains six types of strokes from cricket batsmen. The proposed approach is implemented along with other machine learning models to analyze its efficiency. Experimental results demonstrate that the proposed approach outperforms other models employed in this study. The ResNet model achieves a 0.9 accuracy, on an 80:20 data split. The proposed model achieves the highest accuracy and classifies the six cricket strokes on a video dataset. This study demonstrates the significant impact of emerging technologies like computer vision and machine learning on sports. As these technologies advance, we can expect even more sophisticated and accurate predictions of batsman strokes and other critical aspects of cricket. In the future, more strokes will be added to the video dataset. We also intend to incorporate more features into the training process, such as angel measurement, acceleration, etc., to further improve the accuracy

REFERENCES:

- Desmarais, Y.; Mottet, D.; Slangen, P.; Montesinos, P. A review of 3D human pose estimation algorithms for markerless motion capture. *Comput. Vis. Image Underst.* 2021, *212*, 103275.

- Rupapara, V.; Rustam, F.; Aljedaani, W.; Shahzad, H.F.; Lee, E.; Ashraf, I. Blood cancer prediction using leukemia microarray gene data and hybrid logistic vector trees model. *Sci. Rep.* 2022, *12*, 1000.

- Sid diqui, Hafeez Ur Rehman, et al. "Enhancing Cricket Performance Analysis with Human Pose Estimation and Machine Learning." Sensors 23.15 (2023)

- Mili, R., Das, N.R., Tandon, A., Mokhtar, S., Mukherjee, I. and Paul, G., 2022, December. Pose Recognition in Cricket using Keypoints. In 2022 IEEE 9th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) (pp. 1-5). IEEE.

- Arif, Ayesha, et al. "Human pose estimation and object interaction for sports behaviour." (2022).

- Devanandan, Mithelan, et al. "Cricket Shot Image Classification Using Random Forest." 2021 3rd International Conference on Advancements in Computing (ICAC). IEEE, 2021.