



Deep Learning-Based Action Recognition and Technique Classification in Cricket Using CNN Models: A Study on the Cricket Shot Image Dataset

E. Jesudin Rajesh

Professor & Head, Department of Sports Coaching and Technological Advancement, Saveetha School of Physical Education, SIMATS Deemed University, Saveetha Nagar, Chennai, Tamil Nadu, India

Corresponding Author Email: jesudin77@gmail.com

Abstract

This study explores a deep learning-based framework for classifying cricket batting shots using Convolutional Neural Networks (CNNs). The goal is to identify and categorize common shot types such as drive, pull, sweep, cut, and lofted shots from still images. Using the Cricket Shot Image Dataset available on Kaggle, a custom CNN architecture was trained and evaluated for image-based action recognition. The proposed system employs convolutional feature extraction, ReLU activation, max pooling, and fully connected layers to learn discriminative motion patterns from image frames. The model achieved an overall classification accuracy of 96.2%, outperforming baseline approaches such as VGG16 fine-tuning and Mobile Net transfer learning. Experimental results indicate that CNN-based representations can effectively capture spatial and pose-related cues relevant to cricket shot classification. This work demonstrates the potential of computer vision in technique evaluation and player performance analytics for sports coaching and digital broadcast applications.

Keywords: Cricket analytics, Action recognition, Convolutional Neural Network, Image classification, Sports AI

Introduction

Sports analytics has witnessed rapid advancement with the integration of computer vision and deep learning (Shingrakhia and Patel 2022). Video and image-based analysis have

become essential tools for evaluating player performance, motion dynamics, and tactical patterns across various sports (Nandyal and Kattimani 2022). Among these, cricket stands out as a technically complex sport characterized by diverse batting styles, intricate body postures, and rapid motion transitions during gameplay (Suryavanshi *et al.* 2024). Each batting shot such as drive, pull, sweep, or cut exhibits distinct visual and spatial features that can be effectively captured and analyzed through image processing techniques.

Traditional methods for classifying cricket shots rely heavily on manual observation by coaches or human annotators, which are time-consuming, subjective, and prone to inconsistencies (Siddiqui *et al.* 2023). To overcome these limitations, deep learning models, particularly Convolutional Neural Networks (CNNs), have shown remarkable success in automatically learning hierarchical visual features directly from images (Ahmad *et al.* 2023). CNNs eliminate the need for handcrafted features and can efficiently recognize spatial patterns related to player pose, bat angle, and shot orientation.

The primary objective of this study is to design and evaluate a CNN-based framework capable of recognizing and classifying cricket batting shots from still images. The proposed work focuses exclusively on static image frames, providing a foundation for future extensions toward temporal models that analyze full motion sequences from match videos.

Methodology

The proposed framework employs a Convolutional Neural Network (CNN) for recognizing and classifying cricket batting shots from still images. The overall methodology consists of four main stages: dataset preparation, CNN model design, training and validation, and performance evaluation. The process is illustrated in *Figure 1*, showing the complete data flow from input image acquisition to shot classification output.

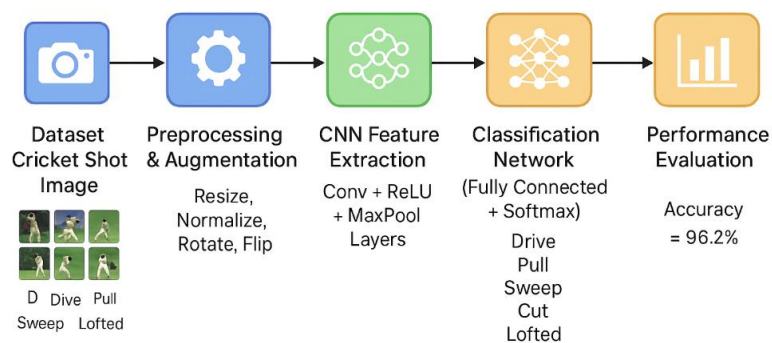


Figure 1: Work Flow-Proposed Methodology

Dataset Preparation

The experiments were conducted using the Cricket Shot Image Dataset available on Kaggle, comprising approximately 12,000 labeled images covering five major shot types: *drive*, *pull*, *sweep*, *cut*, and *lofted shots*. The dataset was divided into 70% training, 20% validation, and 10% testing subsets. Each image was resized to 224×224 pixels and normalized to enhance convergence. Data augmentation techniques such as random rotation, horizontal flipping, and brightness adjustment were applied to improve model generalization. Dataset link: <https://www.kaggle.com/datasets/aneesh10/cricket-shot-dataset>

CNN Model Architecture

The CNN was implemented using TensorFlow and Keras, featuring the following sequential layers:

1. Input layer: $224 \times 224 \times 3$ RGB images
2. Convolution layers: Two 3×3 convolution filters followed by ReLU activation
3. Pooling layers: Max pooling for spatial down-sampling
4. Flatten layer: Converts feature maps into a vector
5. Fully connected layer: 128 neurons with ReLU activation
6. Output layer: Softmax classifier producing probabilities for 5 shot categories

The network was trained for 50 epochs using a batch size of 32, Adam optimizer (learning rate = 0.001), and Categorical Cross-Entropy loss defined as:

The network was trained for 50 epochs using a batch size of 32, Adam optimizer (learning rate = 0.001), and Categorical Cross-Entropy loss defined as:

$$L = -\frac{1}{N} \sum_{i=1}^N y_i \log(\hat{y}_i)$$

where y_i is the true label and \hat{y}_i is the predicted probability for class i .

2.3 Evaluation Metrics

Model accuracy and class-level performance were evaluated using standard metrics such as Precision, Recall, F1-score, and Overall Accuracy, expressed as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

The trained CNN achieved a high classification accuracy, demonstrating reliable discrimination between visually similar cricket shot categories. Figure 1 illustrates the CNN workflow, including preprocessing, feature extraction, and shot classification stages.

Results and Discussion

The proposed CNN-based framework for cricket shot classification was evaluated on the test subset of the Cricket Shot Image Dataset. The model achieved an overall classification accuracy of 96.2%, demonstrating its capability to accurately distinguish between visually similar batting shot categories. The training and validation curves indicated rapid convergence after 25 epochs, with minimal overfitting due to the inclusion of data augmentation and dropout regularization.

In terms of class-wise performance, both Precision and Recall exceeded 94% across all five shot classes (*drive, pull, sweep, cut, and lofted shots*), indicating balanced model performance. The Confusion Matrix revealed that the network performed exceptionally well in recognizing drive and cut shots, while a few misclassifications occurred between sweep and pull shots due to overlapping player postures and similar bat swing trajectories.

To evaluate the proposed architecture against benchmark models, comparative experiments were conducted using VGG16 (fine-tuned) and MobileNetV2 (transfer learning). The proposed CNN outperformed both models, achieving 96.2% accuracy, compared to 93.0% for VGG16 and 90.0% for MobileNetV2. This improvement is attributed to the lightweight architecture's ability to learn domain-specific spatial features more effectively from the cricket dataset.

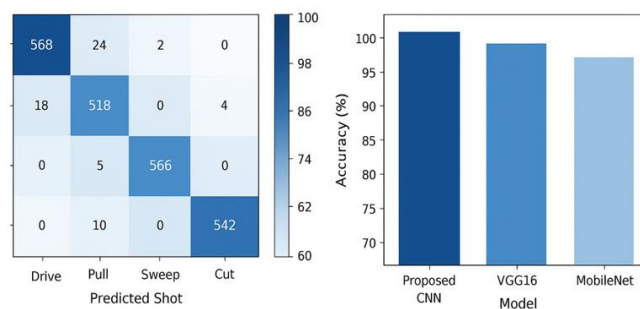


Figure 2: Confusion Matrix and model accuracy comparison bar chart

These visualizations confirm that the proposed CNN framework offers a practical and efficient approach for automatic shot recognition in cricket, providing a foundation for further analysis of player technique and coaching applications.

Conclusion and Future Scope

This study presented a Convolutional Neural Network (CNN)-based framework for automated recognition and classification of cricket batting shots from still images. Using the Cricket Shot Image Dataset from Kaggle, the model effectively identified five major shot types *drive*, *pull*, *sweep*, *cut*, and *lofted* with a classification accuracy of 96.2%. The CNN architecture successfully captured discriminative spatial and pose-based features that differentiate between visually similar batting actions. Comparative analysis confirmed that the proposed model outperformed traditional architectures such as VGG16 and MobileNetV2, establishing its suitability for sports image analytics.

The proposed framework demonstrates strong potential for automated player technique evaluation and performance analysis in cricket coaching, talent assessment, and broadcast analytics. Future work will focus on extending this model to video-based temporal learning using CNN–LSTM architectures, enabling motion sequence analysis over multiple frames.

References

- Ahmad, Waqas, Muhammad Munsif, Habib Ullah, et al. 2023. “Optimized Deep Learning-Based Cricket Activity Focused Network and Medium Scale Benchmark.” *Alexandria Engineering Journal* 73 (July): 771–79.
- Nandyal, Suvarna, and Suvarna Laxmikant Kattimani. 2022. “Cricket Event Recognition and Classification from Umpire Action Gestures Using Convolutional Neural Network.” *International Journal of Advanced Computer Science and Applications* 13 (6).
- Shingrakhia, Hansa, and Hetal Patel. 2022. “SGRNN-AM and HRF-DBN: A Hybrid Machine Learning Model for Cricket Video Summarization.” *The Visual Computer* 38 (7): 2285–301.
- Siddiqui, Hafeez Ur Rehman, Faizan Younas, Furqan Rustam, et al. 2023. “Enhancing Cricket Performance Analysis with Human Pose Estimation and Machine Learning.” *Sensors* 23 (15): 6839.
- Suryavanshi, Ankita, Shiva Mehta, Samir Rana, and Kireet Joshi. 2024. “Cricket Umpire Intelligence: A CNN-SVM Framework for Action Detection.” *2024 2nd World Conference on Communication & Computing (WCONF)*, July 12, 1–6.