National College of Ireland

CricNet: A Deep Learning Network to Enhance the Cricketing Analysis

MSc Research Project
MSc Data Analytics

Rohan Kanagal Sathyanarayana
Student ID: x19203829

School of Computing
National College of Ireland

Supervisor: Mayank Jain

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Rohan Kanagal Sathyanarayana		
Student ID:	X19203829		
Programme:	MSc Data Analytics	**Year:**	2023-2024
Module:	Research Project		
Supervisor:	Mayank Jain		
Submission Due Date:	14 – December - 2023		
Project Title:	CricNet: A Deep Learning Network to Enhance the Cricketing Analysis		
Word Count:	**8567**		
Page Count:	**37**		

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Rohan K S

Date: 14 – December - 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□✓

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

CricNet: A Deep Learning Network to enhance cricketing Analysis.

Rohan Kanagal Sathyanarayana

X19203829@student.ncirl.ie

Abstract

The present research examines the transformative effect using data and artificial intelligence perception upon strategy and training in cricket sport. CricNet, the idea and the primary objective that has been put forward, utilizes cutting-edge technologies involving convolutional neural network and transfer learning approaches to assess simultaneous streaming video as well as still images. VGG19, ResNet50, InceptionV3, MobileNet, EfficientNetB0, DenseNet121, Xception were adopted wherein InceptionV3 accomplished an amazing accuracy rate of **99.58%.** This deep learning-based outcome classifies cricket shots in real time, integrating gesture recognition employing Detectron mode to derive reliable angles of the body. All these perspectives are subsequently displayed to coaches and players to examine, enabling useful information towards enhancement of skills. CricNet not just exposes specific players' weakness but additionally allows coaches to take decisions based on data, which leads to improved training strategies. On top of that the technology facilitates in-game analysis permitting players as well as coaches to carry out strategic moves and modify their tactics all through the game.

Keywords: *Computer Visio, Deep Learning, Transfer Learning, Computer Vision, Gesture Recognition*

Chapter 1: Introduction

Millions of people across the globe enjoy playing cricket (Sampath et al., 2017). One use of computer vision is cricket shot identification with image gesture/posture detection, which can automatically identify a cricketer's pose in each image (Sharma and Aloimonos, 2016). Cricket instruction, injury prevention, and player performance analysis are just a few of the many potentials uses for this analysis (Lu, Haiping, et al.,2016 & Cao, Zhe, et al.,2017).

Figure 1. Illustration of a batsman playing a shot. The batsman's gesture is drawn as a reference for the correctness of the shot (Analytical Vidhya,2023).

The inherent position diversity of the human body is a significant obstacle to accurate stance identification in cricket shot images as can be seen in the above figure. When playing a shot, cricketers can strike with a variety of positions; these poses can change based on the shot type, the cricketer's own style, and the playing conditions. For instance, the position of the subject will vary depending on whether the shot is a drive, draw, or cut. Also, depending on the game scenario, the same cricketer could strike a different attitude for the same stroke (Sampath et al., 2017). Detecting the perspective of cricket shot image is not that easy, particularly if obstructions are present. The wicket, other players, or both might completely or partially block a cricketer's view of their own body during a match. As a result, a computer vision model might find it difficult to accurately identify the cricketer's posture (Sharma and Aloimonos, 2016). Regardless of all these barriers, cricket shot picture posture identification has come a long way in recent years. Deep Learning techniques and conventional computing techniques like computer vision, and hybrid methods are among those that have been suggested for this purpose (Moon, Gyeongsik, et al.,2020 & Yan.,2018). The challenging aspects of an effective cricketing shot can be demonstrated with the help of candle sticks as shown in Fig 1 that exhibits a batsman attempting the stroke. The batsman's gesture serves as an indicator for the shot's correctness, highlighting the challenges needed in successfully representing all the different postures seen during cricket shots.

Deep learning has been a game-changer for human pose estimate issues in computer vision in the last several years (Cao, Zhe, et al.,2017 & Moon, Gyeongsik, et al.,2020). 2D and 3D pose estimators that outperform classic computer vision approaches have been developed by utilizing deep convolutional neural networks and big, labelled datasets. There has been a dearth

of work on deep pose-based action identification for cricket batting poses, even though these poses display distinct kinematics and occlusion patterns (Yan.,2018 & Liu, Jun, et al.,2019).

Making a smooth transition, another intriguing avenue to explore is applying conventional computer vision techniques to the problem of cricket shot picture posture recognition. This discussion complements the exploration of shot classification, providing insights into the challenges associated with accurately identifying and validating the diverse postures exhibited during cricket shots. Given that deep learning models are considerably more accurate, standard computer vision techniques are more successful in handling obstacles. (Yan.,2018). Integrating deep learning with more standard computer vision techniques, hybrid algorithms are able to identify a cricket stroke from an image, based on an individual player's body position. Whereas compared to deep learning or traditional computer vision algorithms, hybrid techniques usually outperform and can yield better results (Moon, Gyeongsik, et al.,2020 & Yan.,2018).

This research project intention is to deal with the issues or challenges associated with cricket shot identification through the application of the deep convolutional neural network (CNN) approach (M.Jagadeesh , Sooraj Nikam.P , Sudharsan.V , Vishnu.S , Rithesh S and Sagar Y et al.,2023) . Alongside the primary objective is to develop an efficient and effective cricket shot classification framework which utilises visual representation that takes into consideration of different kinds of shots encountered in cricket. The proposed CNN approach aims to deliberatly integrate layers from both convolutional and pooling, having the objective of capturing crucial details vital for accurate cricket shot classification.

1.1 Motivation

Considering the cricket sports popularity which has approximately 2.5 billion worldwide spectators depends on manual shot monitoring. Like previously mentioned our research leverages deep convolutional neural network alongside conventional computer vision approaches (Moon et al., 2020; Liu et al., 2019). With the motivation of intelligently recognizing the different cricket shots through the images for improved analyzing instances, training purpose wherein individual can rectify the mistakes and work on to improve their skills. In addition, the idea which will contribute for the broader usage of Artificial Intelligence in the field of cricket analytics whilst extending the boundaries of automated analysis if the cricket shots motivated to consider this as the research assessment.

1.2 Problem Statement

Despite advances in sports technology, accurately detecting cricket shot types remains an open challenge. Existing computer vision techniques, like those based on motion vectors (Vats, Kuldeep, et al.,2019) or geometric features (Khan, Sirajum, et al.,2015), struggle to reliably classify the diverse array of strokes played by the batsmen. For instance, prior work by (Sharma and Aloimonos, 2016) achieved only 61% top-1 accuracy in classifying the six cricket shot types. The intricate structure and unpredictable nature of batsman striking the ball indicates the demand towards more robust solutions. Through rigorous benchmarking and model optimization, our goal is to surpass existing computer vision approaches. The outcomes could enable us to have more precise real-time shot analysis for the benefit of the players, coaches and multimedia applications. Overall, this research intends to enhance the adoption of modern

machine learning within sports analytics for a sport enjoyed by over 2.5 billion fans globally (Sampath et al., 2017).

1.3 Research Objectives

The goal of the research is to develop a novel automatic cricket shot classification technique by adopting algorithms based on deep learning which uses images of different cricket shots for the analysis. The positive and negative aspects associated with our suggested deep learning algorithm for identifying cricket shots using stationary images were carefully analyzed. The principal objective is to offer beneficial insights as an outline to future developments in picture-based cricket shot characterization. In summary, this research seeks to clear the path for establishing deep learning techniques that enhance automated cricket shot detection from RGB images, thereby contributing to enriched analytics, coaching, improving the skills of individual players, multimedia applications, and heightened fan engagement using the perfect gestures in the detection of the image analysis.

1.4 Research Questions

This research aims to develop a robust and accurate cricket shot detection system using deep learning. The specific research objectives are as follows:

1. How effectively can the proposed methodology, utilizing player pose images, classify diverse cricket shots? What role does the player's body posture play in accurately identifying different shot types?

2. To what extent do deep learning models, particularly Convolutional Neural Networks (CNNs), contribute to the accurate classification of cricket shots based on pose images? How do these models compare in terms of precision, recall, and overall performance?

3. How does incorporating pose information enhance the precision and robustness of cricket shot recognition compared to traditional image-based classification methods? Can distinctive features captured from player poses contribute significantly to the model's understanding of different shots?

4. In a practical cricketing scenario, how does the proposed system perform when applied to real-world match situations? What challenges and opportunities arise when deploying pose-based shot classification in professional cricketing environments?

1.5 Research Contribution

In this research we are proposing a novel network known as CricNet which will identify the cricketing shot a cricketer has played and based on that we will identify the gestures and postures of the cricketer. The angles between all the joints will then be projected for a better analysis. This whole system will then create an analysis of the shot being played and help the coaches and cricketing mentors to improve the shots played by the batsman.

1.6 Paper Outline

In the next section, we will discuss in detail the literature being done in the past and gather the information related to the different kinds of data being collected and the best of the models being used. This will follow up with the section on methodology (chapter 3) in which we will discuss the details of the different methods and techniques being employed. In the chapter on implementation, we will discuss the design of the proposed pipeline and the coding attributes. In the chapter on results and analysis we will discuss the different results and benchmark the model performances. We will draw conclusions and plan future work in the chapter on conclusion and future work.

Chapter 2: Literature Review

Early works (KristinaHost et al., 2022) in cricket shot recognition employed machine learning classifiers wherein professionals or researchers deliberately selected pictures or features via images or video frames they believed were significant for identifying the cricket shots. In (Hobday et al., 2022), the authors use bat centroid and hand cluster histograms along with audio cues to classify six shot types, achieving 61% accuracy. Similarly, (Iqbal et al., 2020) combines rotation, aspect ratio, motion vectors and other geometrical features to recognize three kinds of shots from broadcast match videos with 61% accuracy. While demonstrating initial progress, these techniques lack robustness due to not modeling batting stroke articulations.

More recent projects have started exploring deep learning methods for cricket shot detection tasks but without significant articulated pose components. For instance, in (Sharma and Aloimonos, 2016) the authors achieve 42% top-1 accuracy using a deep residual network trained on optical flow motion maps to classify four batting stroke types. While showing the applicability of deep models, their technique does not translate visual understanding into interpretable cricket batting pose terms.

2.1 Computer Vision Techniques and implementation in Identifying Cricket Shots

A rapidly developing area of computer vision with enormous potential for improving sports analytics and technology is automated recognition and analysis of cricket batting approaches (Sharma and Aloimonos, 2016). Classifying common batting stroke types and accurately capturing players' articulated stance and dynamic hand movements during shot execution are among the key focus areas (Saini et al., 2017).

Figure 2.1. Cricket shots classification

Effective vision-based modeling in this space needs to address unique complexities like the high dimensionality of potential batting poses, inter/intra-player variability across techniques and environments, potential for occlusions, as well as the distinct kinematics of specialized strokes played (Iqbal et al., 2020). For instance, distinguishing a cover drive from a straight drive requires detecting nuances in front-foot strides, bat swings relative to body etc. which classical techniques struggle with.

Recent breakthroughs in deep neural networks, especially convolutional architectures, show immense promise towards robust modeling of cricket batting gestures and semantics (Vats et al., 2019). By training deep CNNs on large, curated datasets of diverse batting gesture images, stronger feature representations can be learned for both classifying standard shots through output layers as well as precisely localizing anatomical key points on input frames (Cao et al., 2021).

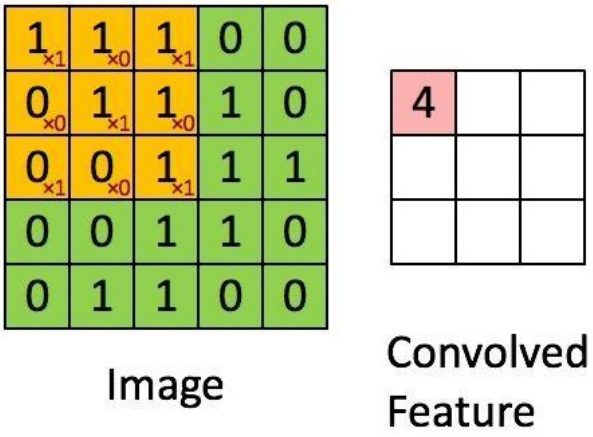


Figure 2.2. Convolution of a filter over 2D image (Subhash,2023)

Figure 2.2 convolution of a filter is a mathematical operation used in image processing and convolutional neural networks. It involves systematically applying a small matrix (filter) to an input image, capturing local features through element-wise multiplication and summation. This process aids in extracting patterns and hierarchies crucial for tasks like image recognition in deep learning. Integrating these dual objectives could enable shot-type-specific pose analysis to derive biomechanical insights e.g., examining elbow flexion differences between pull and flick strokes. Purpose-driven benchmarking of customized techniques is needed to transform computer vision's potential for automated cricket batting analytics into real-world impact.

2.2 Deep Learning Implementation for Cricket Shot Identification

Machine Learning and Deep Learning have been game-changers for Artificial Intelligence, making it possible for machines to learn and do things that were once unimaginable. These technologies aim to equip machines with intelligence akin to human cognitive capabilities (Goodfellow et al., 2016). In the realm of sports data analysis, the integration of neural networks and computer vision algorithms has become commonplace (James et al., 2013). A big question hangs over the sports world: how will deep learning truly reshape the way we crunch and analyze data? Deep learning is like a wild card in the game, its impact is a complete mystery.

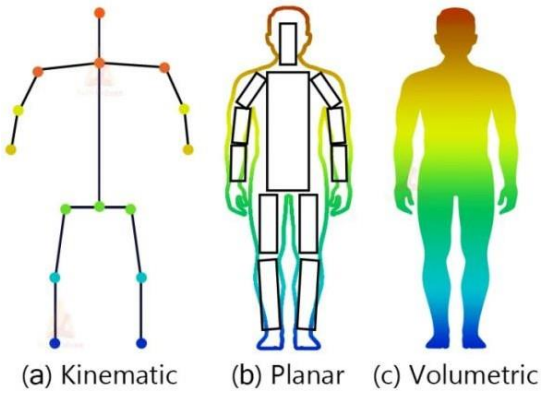


Figure 2.3. Types of poses in estimation model

Cricket, with its complex interplay of bowlers, batsmen, and fielders, features a wide range of shots employed by batsmen to accumulate runs. Identifying and categorizing these cricket shots is a challenging problem (Khan et al., 2015). To address this, researchers introduce a convolutional neural network (CNN) model known as "Stroke-Sense." This model's primary objective is to discern one of four distinct cricket shots from static images (Vats et al., 2019). The results achieved by researchers proposed model are indeed encouraging, boasting an accuracy rate of 83.2% (Vats et al., 2019). These results signify the effectiveness of their methodology and represent an improvement over prior attempts in this domain (Sharma and Aloimonos, 2016).

At the heart of this system is a Convolutional Neural Network (CNN) (Ye et al., 2021). This CNN boasts multiple layers, each with a specific role in the process. The outcomes achieved by this system are nothing short of remarkable. Fueled by the architectural prowess of the CNN, the system boasts an impressive overall accuracy rate of 91.5% (Vats et al., 2019). In addition to CNNs, modern object detection frameworks like Detectron2 (Wu et al., 2019) provide accurate human pose estimation capabilities that can further enhance shot analysis. By extracting skeletal key points (Figure 2.3) and limb rotations directly from video frames, detailed biomechanical insights can be derived interpretable by both players and coaches. Integrating such pose-driven analytics with learned shot classifiers can enrich the understanding of batting techniques and player performance.

2.3 Usage of Transfer Learning in cricket shot analysis.

Transfer learning has revolutionized the landscape of cricket shot analysis, ushering in a new era of understanding and performance optimization (Oquab et al., 2014). This innovative approach involves repurposing pre-trained neural network models such as VGG (Simonyan et al., 2015), ResNet (He et al., 2016), Inception (Szegedy et al., 2016), and MobileNet (Howard et al., 2017) that have undergone extensive training on diverse datasets. The adaptation of these models proves instrumental in effectively categorizing cricket shots, leveraging their inherent ability to identify fundamental attributes and repetitive patterns within a vast collection of images (Yosinski et al., 2014). The transfer learning process unfolds in several key stages (Tan et al., 2018). Initially, pre-trained models are employed as feature extractors, retaining their convolutional layers while discarding fully connected layers (Long et al., 2015). This feature extraction step is crucial in identifying intricate features within images, setting the stage for subsequent fine-tuning (Sun et al., 2016).

Fine-tuning represents the next pivotal phase. Preserved layers from the pre-trained models undergo adjustments using a specialized dataset comprising cricket shot photos (Kouw et al., 2018). This process allows the model to adapt its acquired characteristics to the distinct attributes of cricket shots. Through meticulous training, the model refines its understanding of cricket shots by adjusting its weights, ensuring a nuanced comprehension of the sport's intricacies. Knowledge dissemination plays a fundamental role in transfer learning. The revised model, drawing upon its proficiency in image recognition, builds upon an existing foundation rather than starting from scratch. This sharing of information ensures that the model benefits from the wealth of knowledge embedded in pre-trained models (Weiss et al., 2016). The training process involves exposing the adapted model to a manually annotated dataset containing diverse cricket shots. Back propagation and optimization iterations refine the

model's ability to accurately identify and classify these shots. The culmination of this training equips the model to tackle real-world scenarios, exhibiting a high level of accuracy in classifying cricket shots based on input images or video frames. The advantages of employing transfer learning in cricket shot analysis are noteworthy. By leveraging pre-training, the model can reduce its dependence on extensive datasets, as it harnesses background information acquired during initial training (Peng et al., 2017). This not only enhances the accuracy of shot categorization but also accelerates the training process, significantly reducing the time required for deep model training compared to starting from scratch.

In the realm of cricket shot analysis, transfer learning emerges as a crucial tool. It allows for the adaptation of pre-trained models to address the distinctive complexities inherent in cricket shots. This approach not only enhances the accuracy and efficiency of shot classification but also serves as a valuable resource for coaches, players, and enthusiasts seeking deeper insights into the nuances and dynamics of the game. Ultimately, transfer learning stands as a cornerstone in advancing cricket shot analysis, bridging the gap between theoretical understanding and practical application.

Table 2.1: Overview of some related papers

Sl. No	Paper Name	Authors Name	Dataset Used	Models Used	Results Summary
1.	Cricket Shot Image Classification Using RandomForest	Mithelan Devanandan, Vithurson Rasaratnam, Manoj, Karthik Anbalagan, Narthanan Asokan, Rrubaa Panchendrarajan, Janani Tharmaseelan	Human body key points extracted with Media Pipe	Media Pipe Blaze Pose, Random Forest Classifier, Similarity estimation	• The new model with Media Pipe and Random Forest outperforms the previous CNN model by 5%, achieving an 87% F1-score. • Best for Straight Drive shots (95% F1), weakest for Scoop shots (83% F1). • Using 15 selected key points is better than arbitrary ones. • Simple cosine similarity used to match pro player images. Mobile app for shot analysis, classification, and tracking.
2.	Cricket shot classification using motion vector.	D Karmaker Chowdhury M S U Miah M A Imran M H Rahman	CricShot10	Motion Estimation, 3D MACH Filter, Angle-based classification	• The paper method analyzes multi-view cricket shot videos using motion estimation to extract optical flow and motion vectors. Ituses a 3D MACH filter for action recognition, classifying shots into 8 types based on motion patterns. • This approach improves accuracy over previous methods, which focused on body key points and couldn't capture the complexity of 3D cricket shot motions in videos.
3.	Shot-Net: A Convolutional Neural Network for Classifying Different Cricket Shots	Md. Ferdouse Ahmed Foysal, Mohammad Shakirul Islam, Asif Karim & Nafis Neehal	Shot-Net dataset	Shot-Net, Convolutional Neural Network	• The paper develops a deep CNN called Shot-Net that can accurately classify cricket shot images into 6 common shot types. • The strong results highlight the promise of using deep neural networks for advanced sports data analysis.
4.	Shot Classification of Field Sports Videos Using AlexNet Convolutional Neural Network	Rabia A. Minhas, Ali Javed, Aun Irtaza, Muhammad Tariq Mahmood and Young Bok Joo	The dataset includes 10 videos of 13 h from six major broadcasters, i.e., ESPN, Star Sports, Ten Sports, Sky Sports, Fox Sports, and Euro Sports	AlexNet Convolutional Neural Network	• The study used an eight-layered AlexNet CNN to classify shots in field sports videos. • It achieved an impressive accuracy of 94.07%, outperforming other methods like SVM and standard CNN. • This research demonstrates the effectiveness of their approach for shot classification in sports videos.

5.	Content-Based Cricket Video Shot Classification Using Bag-of-Visual-Features	M. Ravinder & T. Venugopal	http://cse.iitk.ac.in/~vision/dipen/	Bag-of-Visual-Features (BoVW), Support Vector Machine, K-Nearest Neighbors, Naive Bayes, Decision Tree	• The paper introduces a novel method for content-based classification of cricket video shots using the bag-of-visual-features approach. • This methodology involves collecting local features from training images, clustering them to create a visual vocabulary, and using this for training and classification. • The research focused on classifying different shot types in cricket videos, such as close-up view shots, distance view shots, and pitch view shots.
6.	Shot Classification and Semantic Query Processing on Broadcast Cricket Videos	DipenS.Rughwani	This dataset has a total of 256 abrupt and 84 gradual shot changes including 15 action replay scenes	SVM, BoVW, K-Means Clustering, SIFT, SURF, HOG, PCA, LDA, GMM, HMM, CRF, CNN, LSTM, SCD	• The results obtained in the paper suggest that the proposed approaches are effective in analyzing cricket videos and can provide valuable insights to users. • However, there is still scope for improvement in terms of accuracy and efficiency, and further research is needed to explore these areas.
7.	Classifying and Learning Cricket Shots Using Camera Motion	Mihai Lazarescu, Svetha Venkatesh & Geoff West	300 cricket shot videos. Split into two sets: training (240 videos) and testing (60 videos)	SVM, Random Forest, Gradient Boosting, Long Short-Term Memory (LSTM)	• This paper introduces a method for classifying cricket shots using camera motion parameters. • It avoids complex image segmentation and high-resolution images. • The approach also incorporates an incremental learning algorithm for efficient knowledge expansion. • The results demonstrate the system's strong performance in classifying four types of cricket shots.
8.	Activity Recognition for Quality Assessment of Batting Shots in Cricket using a Hierarchical Representation	Aftab Khan, James Nicholson, Thomas Plötz	cricket shot dataset of 20 cricket shot components collected from 6 professional cricket players using body sensor networks.	Hidden Markov Model (HMM) Support Vector Machine (SVM) Random Forest	• This study presents a cost-effective automated system for recognizing cricket shots using body-worn sensors. • It achieved an average F1-score of over 88% in identifying 20 unique shot components during a deployment study with 6 players. • The system offers a more affordable alternative to expensive vision-based systems for assessing batting quality.

Chapter 3: Methodology

3.1 Dataset

The Cricket Shot Dataset, housed in the specified directories, encapsulates a rich collection of 4,724 instances, categorically representing diverse cricketing shots. The dataset has been meticulously curated, with a well-organized structure that includes dedicated directories for specific cricket shot types. Within this comprehensive dataset (Figure 3.1), there is a focus on capturing various nuances of shot execution. Specifically, the dataset comprises images depicting cover drives, straight drives, and off drives, each exemplifying shots played on the front foot with distinctive angles. The dataset is substantial, containing 1224 images labeled as cover drives, 1120 images for leg glance-flick, 1260 images representing the pull shot, and another 1120 images illustrating the sweep shot. This meticulous organization and variety in shot types contribute to the richness and diversity of the dataset, providing a robust foundation for training and evaluating cricket shot recognition models.

(a)

(b)

(c)

(d)

Figure 3.1. Images capturing the player executing different shots: (a) drive shot, (b) pull shot, (c) leg glance shot, and (d) sweep shot.

This dataset serves as a valuable resource for cricket enthusiasts, analysts, and researchers seeking to delve into the intricacies of different shot types as seen from the above figure. With a well-organized structure and diverse shot representations, it enables a comprehensive exploration of cricketing techniques and playing styles. The dataset's richness lies not only in its volume but also in the detailed categorization, making it a versatile asset for cricket analytics, machine learning model training, and strategic insights into player performance. The dataset named 'Cricket_Data' can be accessed using the following link: Cricket_Data[1].

3.2 Convolutional Neural Networks (CNNs) and Transfer Learning Concepts

Behind stunning feats of image recognition and classification, like identifying objects in photos, lies a powerful tool called Convolutional Neural Networks. Inspired by the human visual cortex, CNNs excel at capturing spatial hierarchies and patterns in images, presenting a significant advancement over traditional machine learning algorithms (LeCun et al., 1989).

- **VGG19 (Visual Geometry Group 19):** VGG19 is a CNN architecture comprising 19 layers (Figure 3.2). It features convolutional blocks, each housing two or three convolutional layers followed by a max pooling layer, and three fully connected layers for image classification. VGG19's performance at ImageNet was nothing short of masterful, showcasing its ability to recognize images like no other (ILSVRC) (Russakovsky et al., 2015).

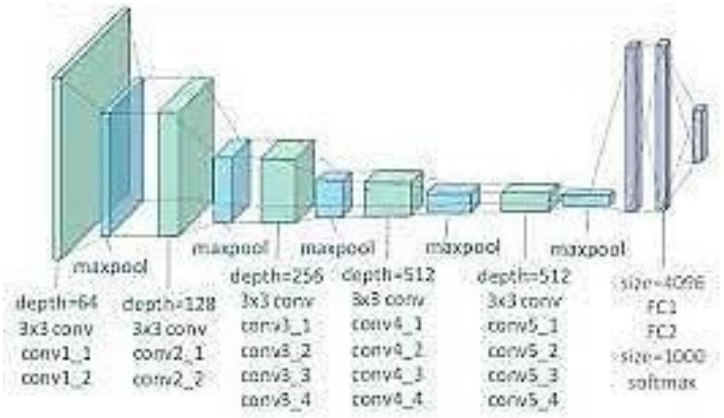


Figure 3.2. Architecture of VGG19
(Yufeng Zheng, Clifford K. Yang, Aleksey Merkulov, May 2018;
https://www.researchgate.net/figure/llustration-of-the-network-architecture-of-VGG-19-model-conv-means-convolution-FC-means_fig2_325137356)

- **ResNet50 (Residual Network):** ResNet50 introduces the concept of residual learning. This involves (Figure 3.3) shortcut connections that skip over certain layers, enabling the training of much deeper networks compared to their predecessors. Like a champion athlete

[1] https://www.kaggle.com/datasets/aneesh10/cricket-shot-dataset

on the image recognition podium, this architecture has crushed the competition in tasks like the prestigious ILSVRC (Deng et al., 2009).

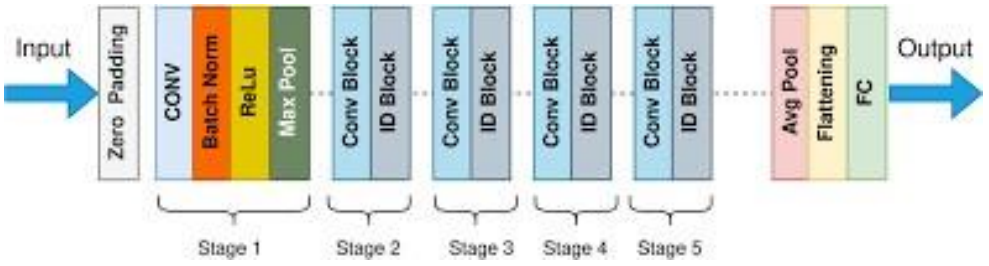


Figure 3.3. Architecture of ResNet.

- **MobileNet:** MobileNet is tailored for mobile and edge devices with a focus on computational efficiency (Figure 3.4). MobileNet gets the job done with fewer parts and less work. Its depth wise separable convolutions are like a smart way to recognize images,making it faster and easier than other networks (Howard et al., 2017).

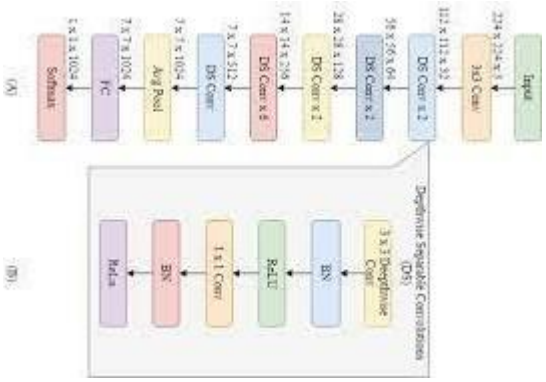


Figure 3.4. Architecture of MobileNet.

- **EfficientNetB0:** EfficientNetB0 is part of the EfficientNet family designed to strike a balance between accuracy and computational efficiency. It employs compound scaling, simultaneously adjusting the depth, width, and resolution of the network (Figure 3.5). This enables EfficientNetB0 to achieve top-tier performance in image classification tasks with reduced parameters and FLOPs compared to other CNNs (Tan et al., 2019).

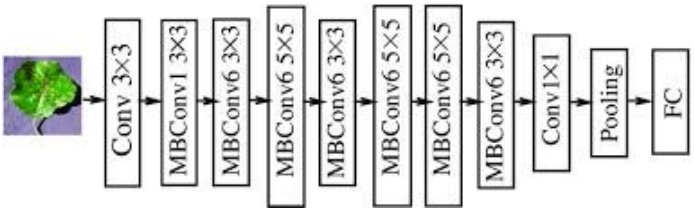


Figure 3.5. Architecture of EfficientNet.

- **DenseNet121 (Densely Connected Convolutional Network):** Proposed in 2017 by Gao Huang, Zhuoyu Heng, Gang Li, and Siwei Liu, DenseNet121 is built on dense connectivity (Figure 3.6). By connecting all the dots, this "full talk" network learns deeper details about images. This superpower lets it ace image classification tasks like a pro (Huang et al., 2017).

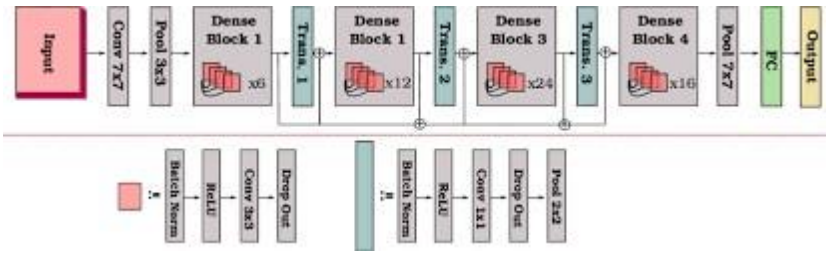


Figure 3.6. Architecture of DenseNet.

- **Xception (Extreme Inception):** Xception is an extreme version of the Inception module. It exclusively employs depth wise separable convolutions (Figure 3.7), enhancing efficiency compared to other CNNs using Inception modules. Xception achieves state-of-the-art performance in image classification tasks with fewer parameters and FLOPs (Chollet, 2017).

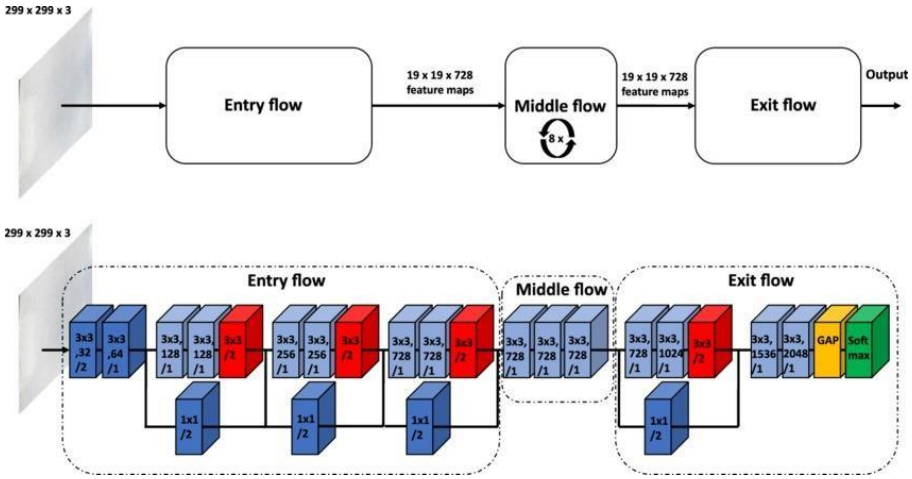


Figure 3.7. Architecture of Xception.

- **InceptionV3:** InceptionV3 is an advanced iteration of the Inception module. It incorporates multiple filter sizes in inception modules (Figure 3.8), allowing it to learn intricate features and achieve top-tier performance in various image recognition tasks, including ILSVRC (Szegedy et al., 2016).

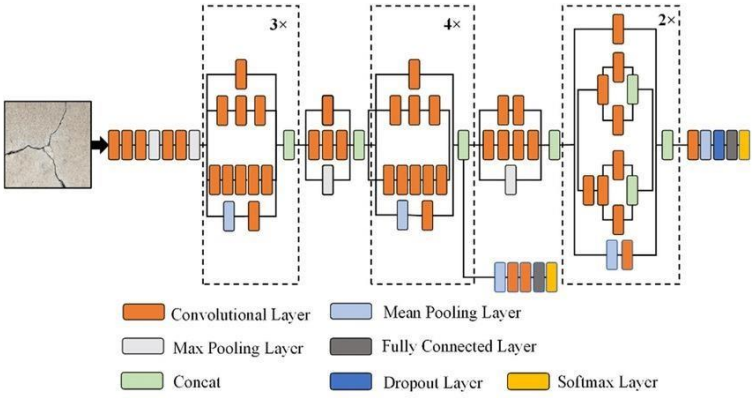


Figure 3.8. Architecture of Inception.

3.3 Model For Gesture Detection

Detectron2 is a state-of-the-art object detection library developed by Facebook AI Research (Wu et al., 2019). Built on the PyTorch framework, Detectron2 provides a flexible and efficient platform for developing, training, and deploying computer vision models. It features a modular and extensible design, enabling researchers and practitioners to easily experiment with various object detection architectures, including Faster R-CNN, RetinaNet, and Key point RCNN (Wu et al., 2019). Detectron2 also supports a wide range of tasks, such as instance segmentation and key point estimation. With a rich set of pre-trained models and easy-to-use APIs, Detectron2 has become a go-to choose for both academic research and industrial applications, empowering users to achieve high-quality results in object detection and related computer vision tasks (Wuet al., 2019).

3.4 Evaluation Metrics

In assessing the performance of our image classification models, we leverage a set of robust evaluation metrics that provide nuanced insights into their effectiveness. These metrics serve as crucial benchmarks for understanding the models' predictive capabilities and their ability to generalize to unseen data.

3.4.1 Accuracy Score

The Accuracy Score serves as a foundational metric, quantifying the overall correctness of our models. It represents the ratio of correctly predicted instances to the total instances in our dataset. A higher accuracy score indicates a more reliable predictive performance.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

3.4.2 Precision Score

The Precision Score is a key metric that evaluates the accuracy of positive predictions. It is the ratio of correctly predicted positive observations to the total instances predicted as positive. Precision is particularly valuable in scenarios where minimizing false positives is critical.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

3.4.3 Recall Score

The Recall Score, also known as sensitivity, gauges the models' ability to capture all relevant instances of a particular class. It measures the ratio of correctly predicted positive observations to all actual positives, providing insights into the comprehensiveness of our models.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

3.4.4 F1 Score

The F1 Score strikes a balance between precision and recall, offering a harmonic mean of these two metrics. This provides a consolidated view of the model's performance, especially when there is a need to optimize both precision and recall simultaneously.

$$\text{F1 Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

3.4.5 Confusion Matrix

A Confusion Matrix is an indispensable tool for a detailed understanding of the model's performance. It presents a tabular representation of predicted versus actual classifications, breaking down the instances into true positives, true negatives, false positives, and false negatives. This matrix provides a granular perspective on the model's strengths and weaknesses, aiding in fine-tuning and optimizing its performance.

$$\text{Confusion Matrix} = \begin{bmatrix} \text{True Negative (TN)} & \text{False Positive (FP)} \\ \text{False Negative (FN)} & \text{True Positive (TP)} \end{bmatrix}$$

Chapter 4: Implementation

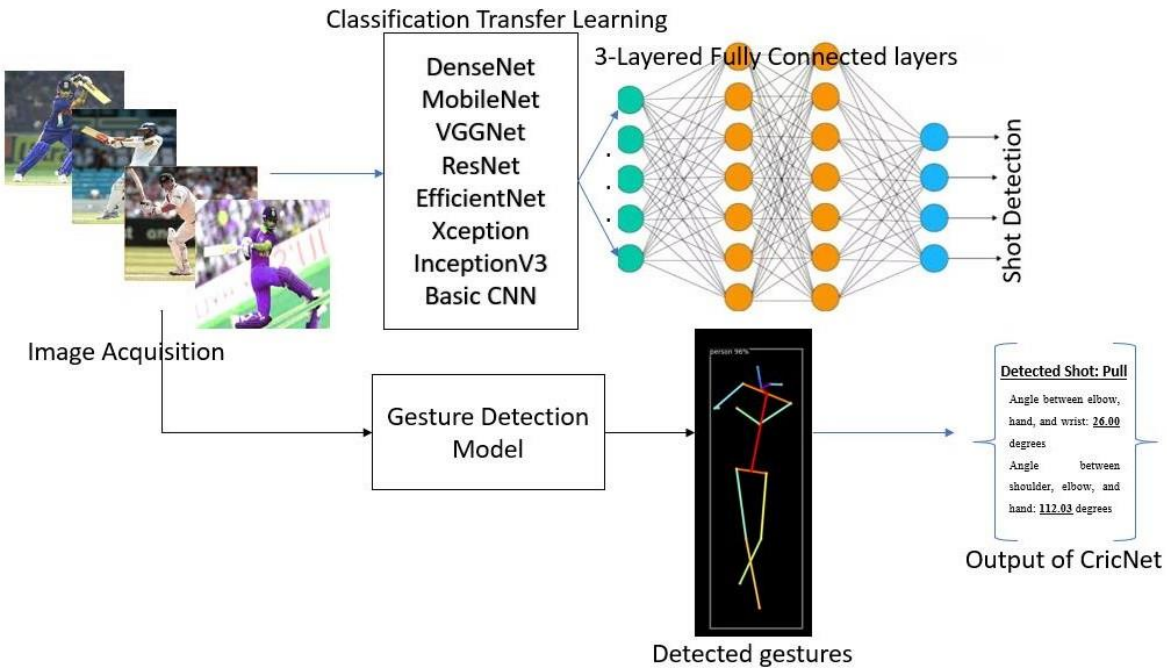


Figure 4.1. Implementation Framework

The above figure 4.1 is the implementation framework which follows the CRISP-DM methodology.

4.1 CRISP-DM

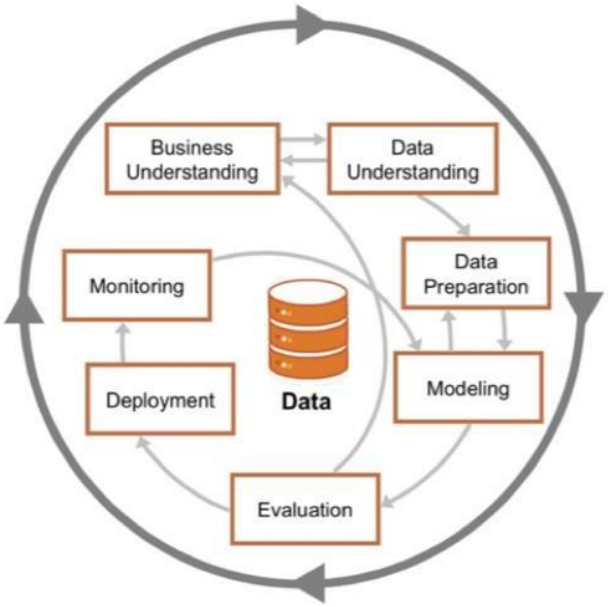


Figure 4.2. CRISP-DM architecture for the Cricket Analysis (Hotz, N,2023)

CRISP-DM, the Cross-Industry Standard Process for Data Mining (Figure 4.2), is a widely adopted framework comprising six phases. It starts with "Business Understanding," defining project goals. "Data Understanding" explores available data, while "Data Preparation" ensures its quality. "Modeling" involves algorithm selection, and the "Evaluation" phase assesses model performance. Finally, "Deployment" integrates models into the business, reflecting CRISP-DM's iterative and adaptive nature.

4.1.1 Business understanding

The Cricket Gesture Detection project addresses the demand for automated shot classification in cricket, leveraging advanced deep learning, primarily CNNs. By classifying cricket shots and detecting gestures, it aims to enhance viewer experiences, provide coaching insights, and advance sports analytics. With curated datasets, it combines technological innovation with real-time analysis potential. This data-driven solution not only revolutionizes player performance understanding but also creates business opportunities through collaboration with cricket boards and sports organizations. The project's comprehensive approach spans shot classification, gesture detection, and angle analysis, promising transformative impacts on the global cricketing community.

4.1.2 Data understanding

The project involves the analysis of a Cricket Shot Dataset comprising 4,724 instances, meticulously categorized into specific directories representing various cricketing shots such as drives, leg glances, pull shots, and sweeps. Each shot category offers diverse representations of player techniques, serving as a valuable resource for model training and cricket analytics. The dataset, organized with attention to detail, features images capturing the nuances of different shot types, providing a comprehensive foundation for the development of a robust deep learning model. The categorization and diversity within the dataset underscore its significance in advancing automated cricket shot analytics.

4.1.3 Data preparation

Data preparation is a crucial step in any machine learning project to ensure that the data is in a suitable format for model training. In the context of the provided code, the data preparation process involves loading images of cricket shots and setting up the dataset for training a convolutional neural network (CNN) model.

4.1.4 Modelling

In this we have used two kinds for framework. First to identify the correct cricketing shot and then to detect the gestures which will generate the insights of the shots as depicted in the previous figure. The code begins by loading gesture images from the specified path, where each subfolder represents a unique type of gesture. Using OpenCV, images are read and stored in the images list, while their corresponding labels are stored in the labels list. The code utilizes Detectron2, an object detection library, for gesture recognition. It sets up the environment, installs necessary packages, and configures a Key point RCNN model pretrained on the COCOdataset.

4.1.5 Evaluation

After training the models using various architectures, it is essential to evaluate their performance to choose the best model for the given task. The evaluation metrics employed here include accuracy, Precision, Recall, F1 Score. The evaluation results provide valuable insights into the performance of each model. The chosen model should exhibit high accuracy, precision, recall, and F1 Score, along with a well-balanced confusion matrix. Fine-tuning or further optimization can be considered based on these results to enhance model performance.

4.1.6 Deployment

This whole solution can be deployed over the cloud and can be leveraged as a mobile application to use in real time.

4.2 Algorithm Flow

Step 1: Image Acquisition – The images are taken from the standard datasets to train the model for the shot detection.

Step 2: The Shot detection model trains with the several algorithms present in the bag of the algorithms to generate the best model,

$$Model_{Selected} = \ argmax\{performance_k\} \ \forall \ k \ \in Models$$

Step 3: Generation of the gestures using the Decetron2

Step 4: Identifying the skeleton of the gestures detected.

Step 5: Output Generation

Chapter 5: Results And Analysis

5.1 Case 1: Cricket Shot Classification Analysis

The Cricket Shot Classification project automated shot categorization using diverse CNN architectures, contributing to sports analytics. With a global cricket fan base of 2.5 billion, the project addressed manual shot classification challenges, enriching viewer experiences and advancing sports analytics. The dataset, with 4,724 instances of categorized cricket shots, ensured robust model training. Data preparation involved systematic image loading and dataset setup for CNN training. Various CNN models, including DenseNet, EfficientNet, InceptionV3, ResNet, Xception, VGG19, and MobileNet, captured spatial patterns. Each model was configured, compiled, and summarized for comprehensive understanding. Evaluation metrics like accuracy, precision, recall, and F1 Score guided model selection. The chosen model

exhibited high accuracy and a balanced confusion matrix, providing valuable insights. In summary, the project showcased a holistic deep learning approach, revolutionizing cricket shot classification, and offering insights for further optimization. The performance comparison table for the shot's classification using different architectures are as below.

Table 5.1: Evaluation Matrix Table

Model	Accuracy	Precision	Recall	F1 Score
InceptionV3	99.58%	99.58%	99.58%	99.58%
MobileNet	99.37%	99.37%	99.37%	99.36%
EfficientNet	99.29%	99.30%	99.29%	99.30%
Xception	99.15%	99.16%	99.15%	99.15%
DenseNet	99.01%	99.03%	99.01%	99.02%
ResNet	98.80%	98.82%	98.80%	98.80%
VGGNet	97.95%	98.00%	97.95%	97.96%

5.1.1 Convolutional Neural Networks – Basic Architecture

Predicted images using the CNN model highlight its recognition capabilities.

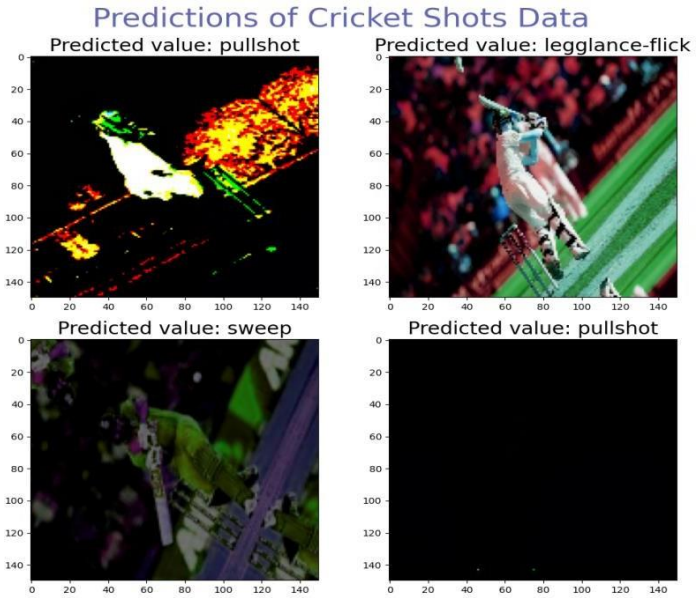


Figure 5.1. CNN predicted images.

Let's check how the transfer learning model with 3-layered neural network as the classifying part works.

5.1.2 DenseNet

Following are the DenseNet predictions, highlighting feature reuse and enhanced gradient flow.

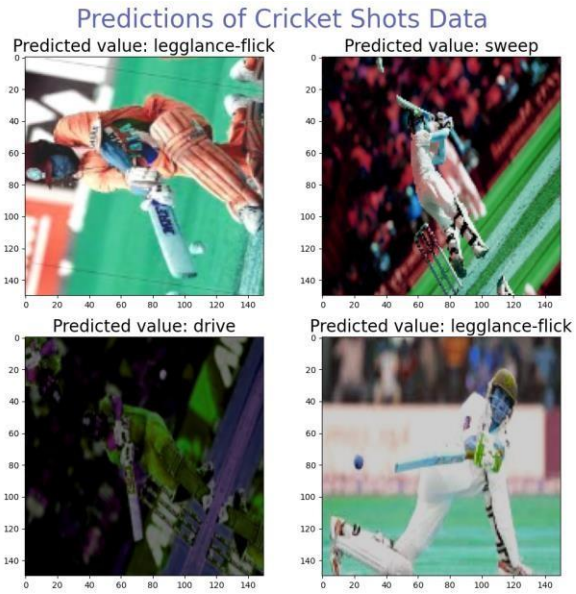


Figure 5.2. DenseNet predicted images.

The accuracy of DenseNet is found to be 99.01% with precision as 99.02%, recall as 99.01% and F1 score as 99.01%. When the loss curve and the accuracy curve is checked,

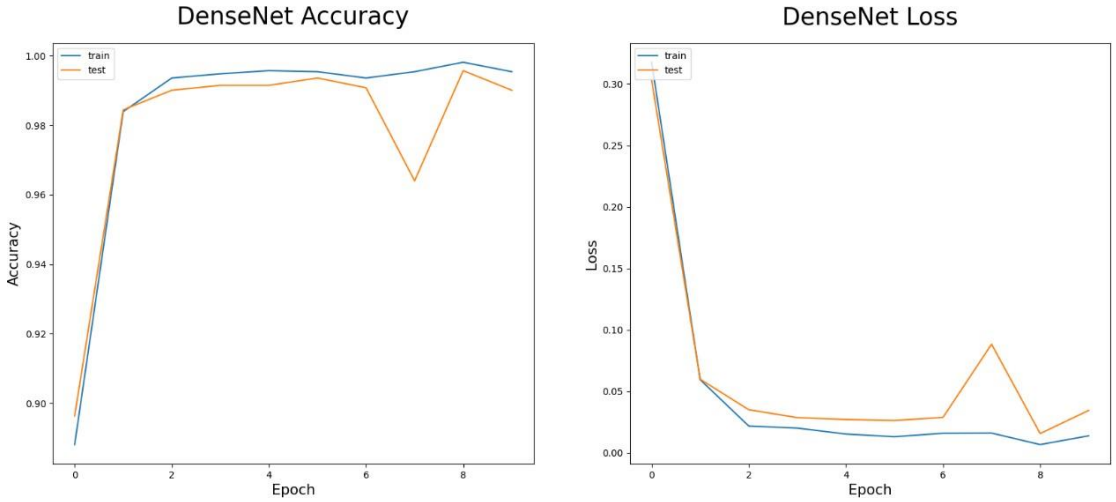


Figure 5.3. DenseNet accuracy and loss plot

Its densely connected convolutional network architecture excels in accurately classifying and detecting diverse gestures, showcasing strong feature reuse and gradient flow.

5.1.3 EfficientNet

EfficientNet's predicted images demonstrate its state-of-the-art performance with optimized model efficiency.

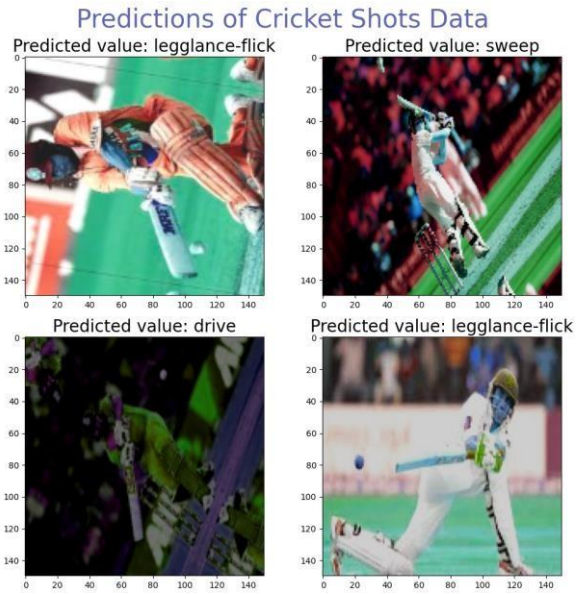


Figure 5.4. EfficientNet predicted images.

EfficientNet delivers commendable results, boasting 99.29% accuracy, precision, recall, and F1 Score. Its efficiency in scaling depth, width, and resolution makes it a compelling choice for gesture recognition tasks, achieving state-of-the-art performance with fewer parameters. The loss and the accuracy curve across the epochs are as below.

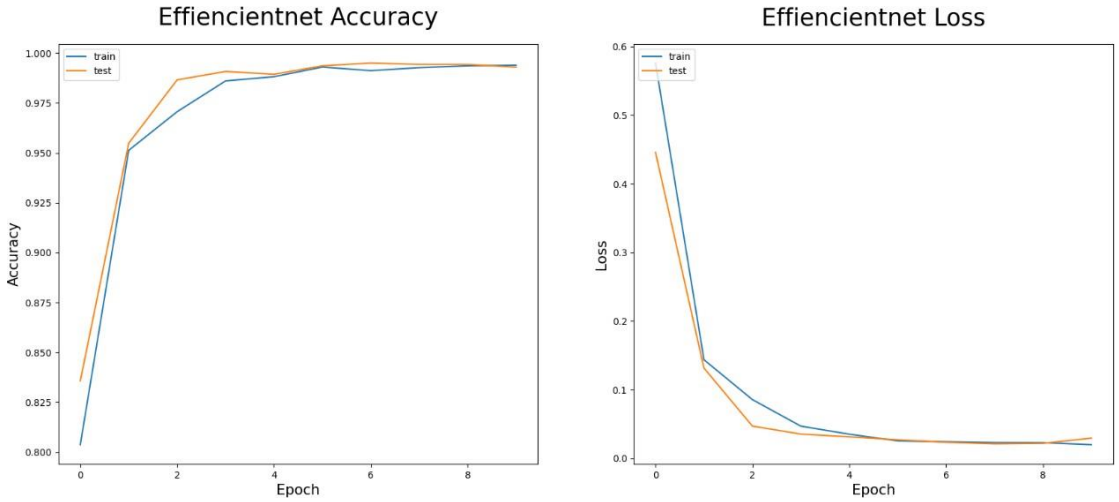


Figure 5.5. EfficientNet accuracy and loss plot

5.1.4 InceptionV3

InceptionV3's predictions exhibit its efficiency and high accuracy, thanks to inception modules with multiple filter sizes.

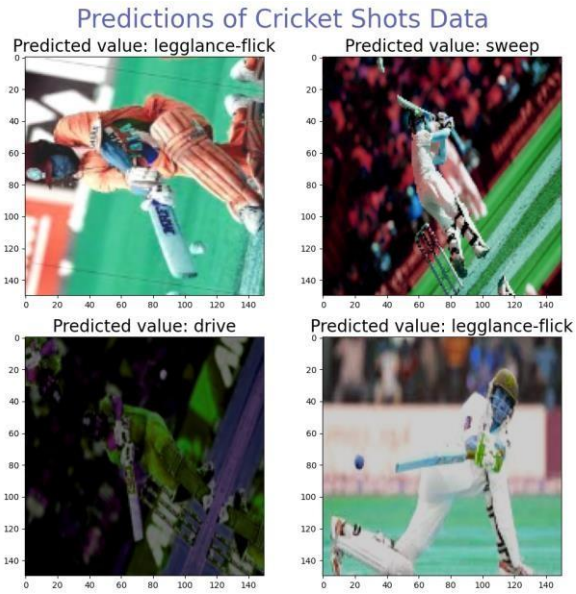


Figure 5.6. Inception predicted images.

InceptionV3 showcases exceptional performance with accuracy, precision, recall, and F1 Score all at an impressive 99.58%. This model's efficiency and precision in capturing intricate features within gesture images are evident, making it a robust choice for diverse gesture recognition tasks.

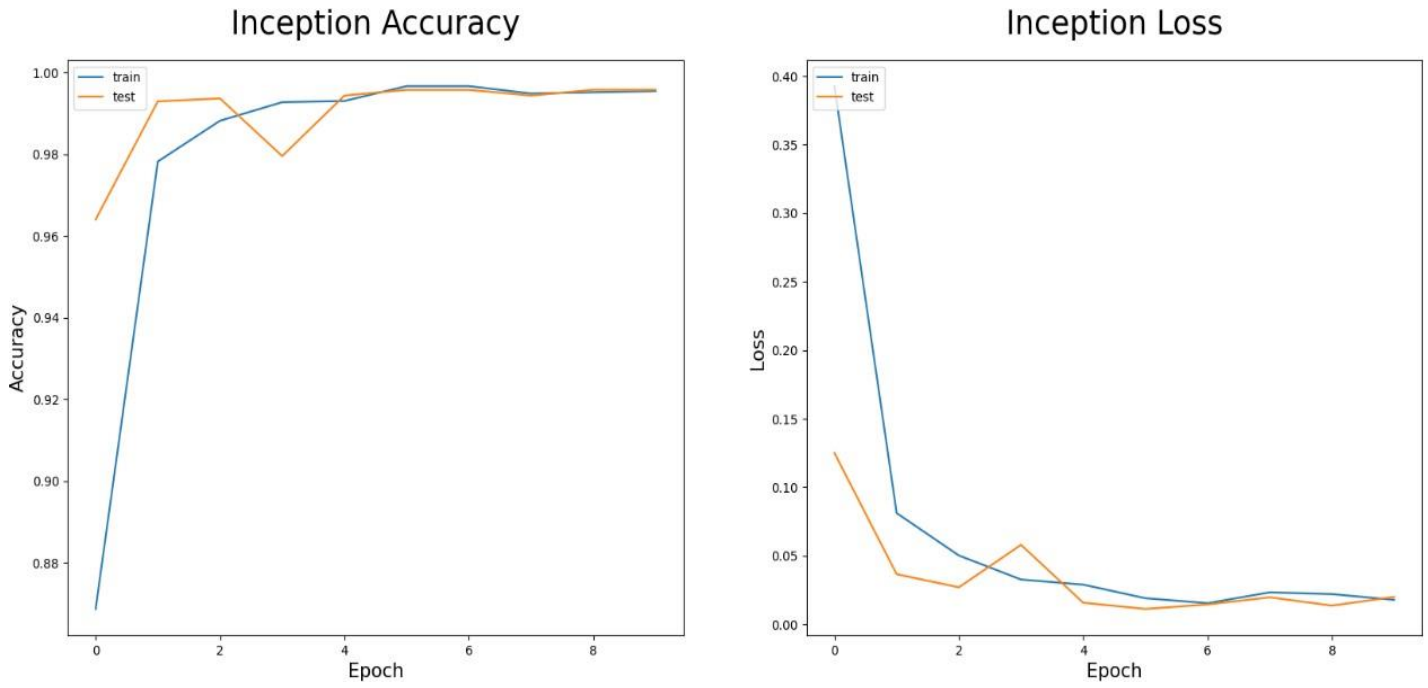


Figure 5.7. Inception accuracy and loss plot

5.1.5 ResNet

ResNet's predicted images illustrate its ability to ease training for very deep networks through residual connections.

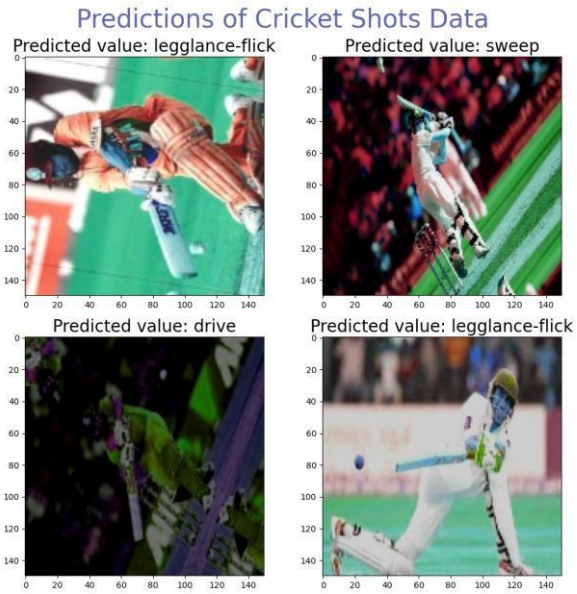


Figure 5.8. ResNet predicted images.

ResNet maintains a strong presence with an accuracy, precision, recall, and F1 Score of 98.20%. The incorporation of residual connections facilitates the training of deep networks, resulting in robust performance across various metrics.

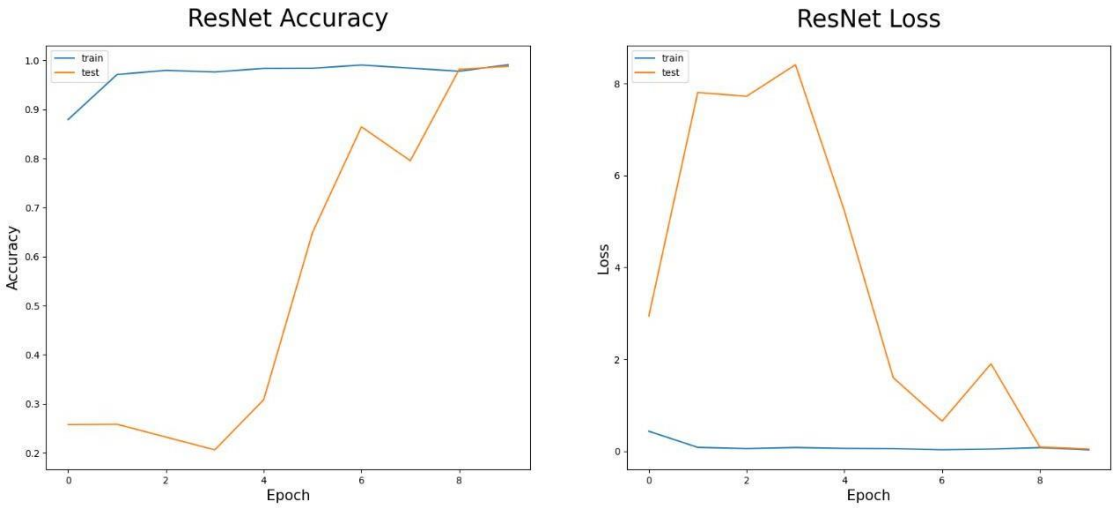


Figure 5.9. ResNet accuracy and loss plot

5.1.6 Xception

Experience Xception's predictions, an extension of Inception with depthwise separable convolutions for improved performance.

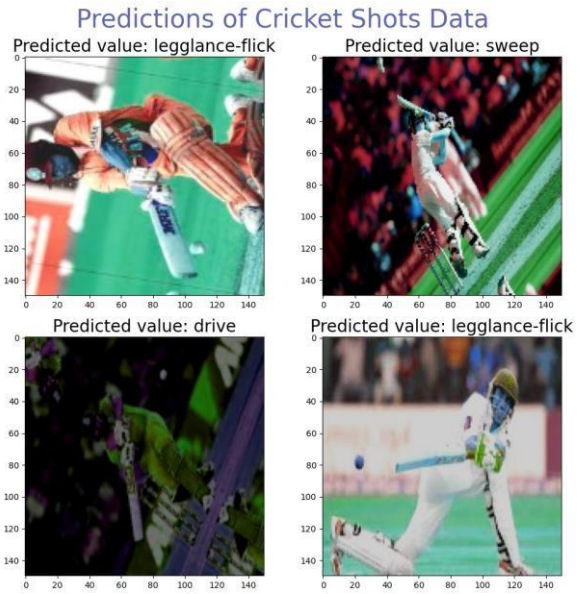


Figure 5.10. Xception predicted images.

Xception demonstrates robust performance with an accuracy of 98.77% and consistent precision, recall, and F1 Score of 98.77%. Its reliance on depthwise separable convolutions contributes to its adaptability and efficiency in handling complex gesture recognition tasks.

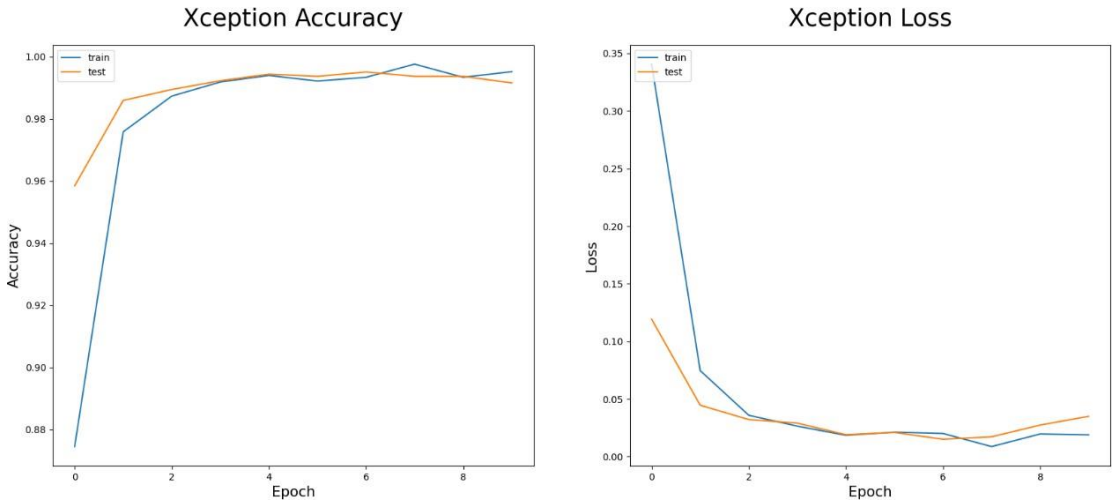


Figure 5.11. Xception accuracy and loss plot

5.1.7 VGG19

VGG19's predicted images showcase its simple and uniform architecture, featuring small 3x3 convolutional filters and max-pooling layers.

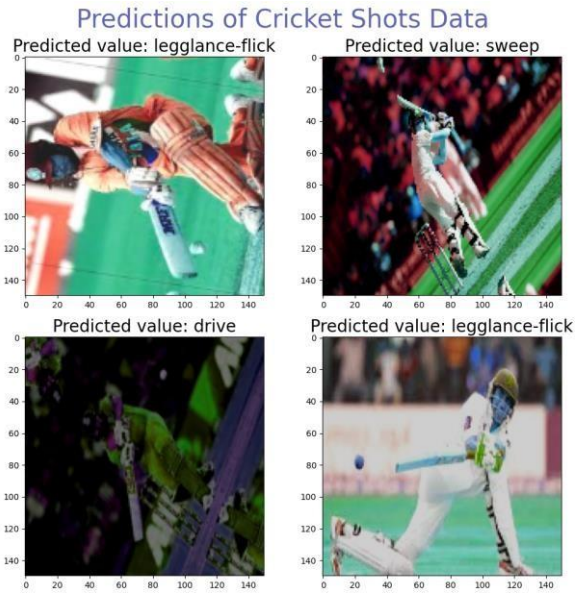


Figure 5.12. VGG19 predicted images.

VGGNet achieves solid performance, with an accuracy, precision, recall, and F1 Score of 97.96%. It's simple and uniform architecture with small convolutional filters contributes to its effectiveness in recognizing various gestures.

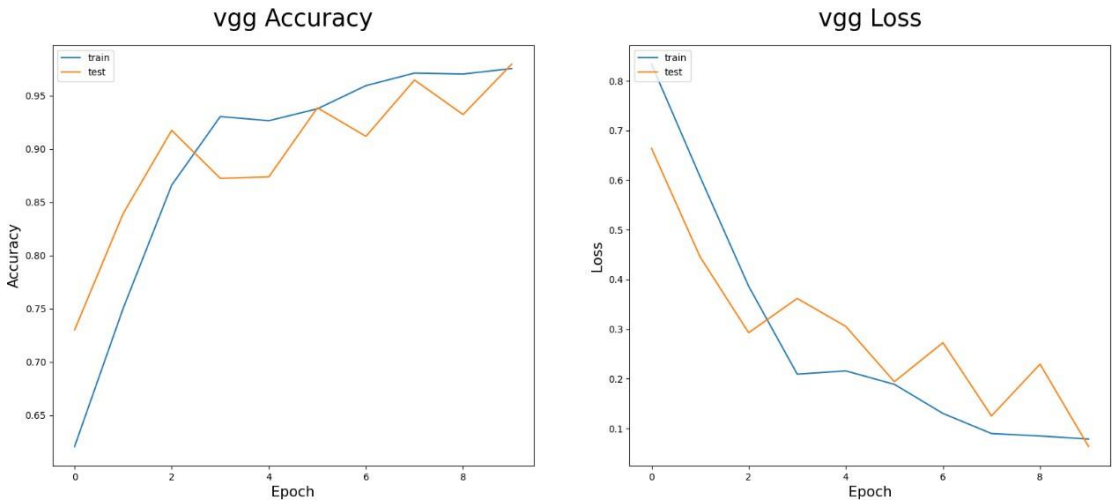


Figure 5.13. VGG19 accuracy and loss plot

5.1.8 MobileNet

MobileNet's lightweight architecture is evident in its predicted images, designed for mobile and edge devices with depthwise separable convolutions.

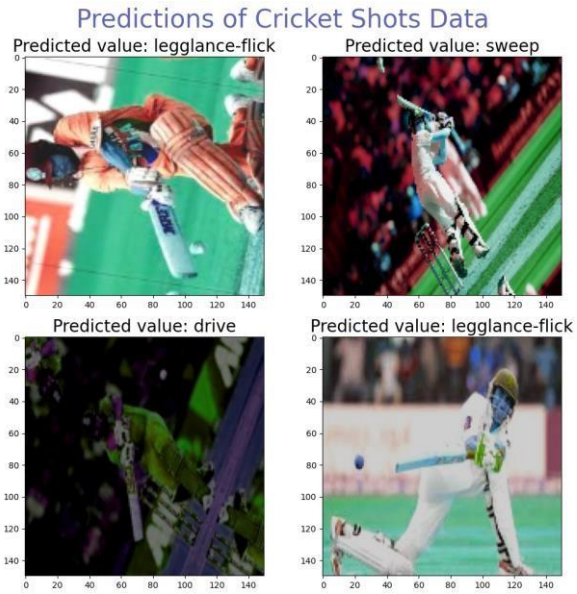


Figure 5.14. MobileNet predicted images.

MobileNet achieves strong results across the board, with an accuracy, precision, recall, and F1 Score of 99.37%. Designed for mobile and edge devices, MobileNet's lightweight architecture with depthwise separable convolutions proves effective in various applications.

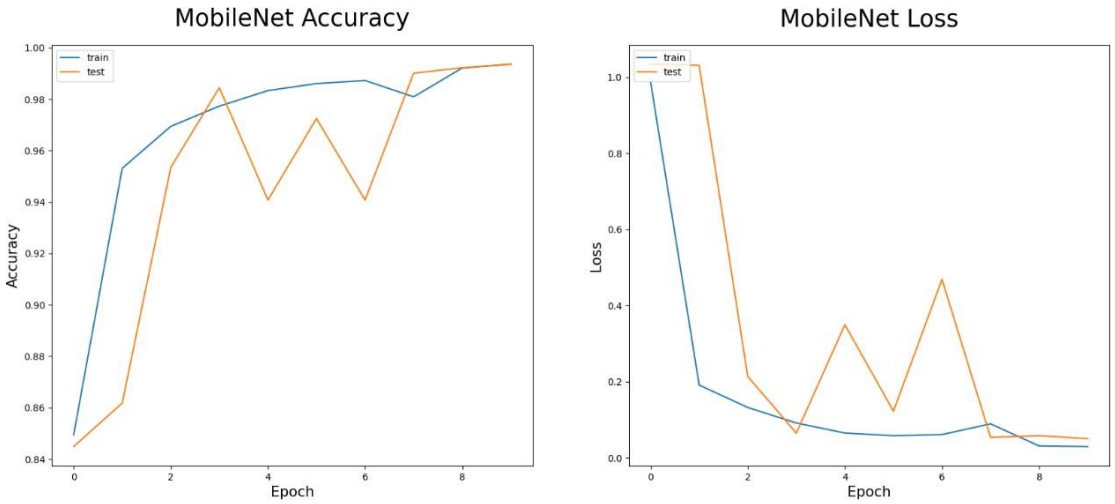
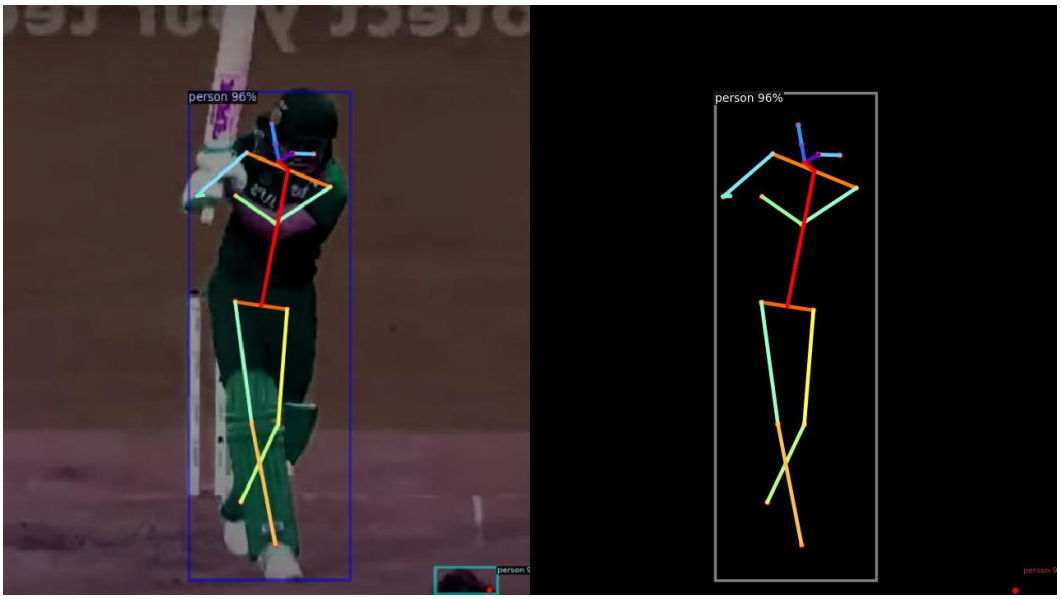


Figure 5.15. MobileNet accuracy and loss plot

5.2 Case 2: Gesture Detection Analysis

The pursuit of gesture detection using the Detectron2 framework has ushered in a paradigm shift in the understanding and analysis of human gestures. The implementation involved a multi-faceted process, encompassing data preparation, model configuration with Detectron2, and a detailed analysis of gesture instances through visualization and kinematics assessments.



(a) **(b)**

Figure 5.16. **(a)** Illustrates a player executing a shot with kinematics, while **(b)** focuses solely on kinematics.

The preliminary stages encompassed a rigorous data preparation process, wherein gesture images were meticulously sourced from the designated path. Each subfolder within the repository denoted a distinct type of gesture. The images underwent systematic reading and organization, culminating in the creation of lists paired with corresponding labels. This foundational step ensured that the dataset was well-organized and ready for integration into the Detectron2 pipeline. The code randomly selected five gesture images from the dataset and utilized the trained Keypoint RCNN model to make predictions. These predictions were visually represented using the Detectron2 Visualizer module, which overlays keypoint information on the original images. The resulting visualizations provided a tangible representation of the model's ability to identify and locate keypoints within the human gesture, offering insights into its accuracy and precision. Understanding the model's performance in different scenarios, lighting conditions, and gesture variations is essential for ensuring robustness and reliability in real-world applications. The validation process in our specific case is centered around CricNet, a dedicated Deep Learning Network designed to elevate the analysis of cricketing gestures. The significance of visual analysis extends beyond mere model validation; it holds practical implications across diverse domains. In our context, the validation pertains to CricNet's proficiency in accurately identifying and localizing keypoints within dynamic cricketing gestures. This precision is paramount for applications in human-computer interaction, sign language recognition, and virtual reality interfaces. The overarching goal is to assess CricNet's suitability and effectiveness in various contexts related to cricketing gesture analysis.

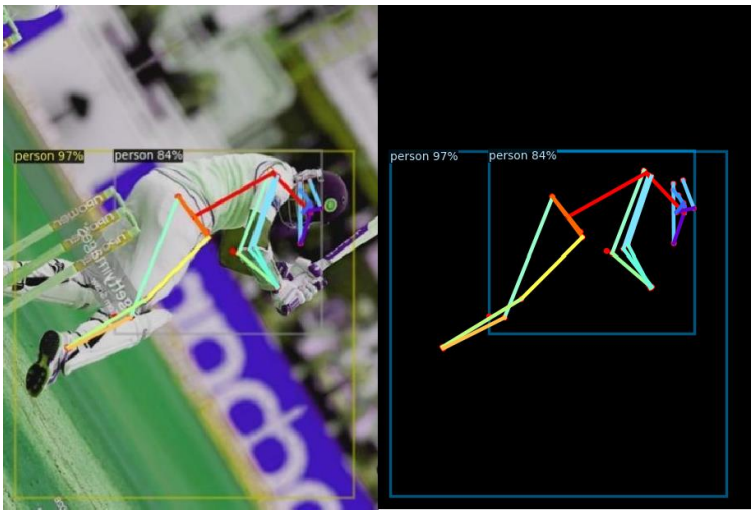


Figure 5.17. Detection of the Angle between the lines for better analysis using shot1

Angle between elbow, hand, and wrist: 5.24 degrees
Angle between shoulder, elbow, and hand: 109.61 degrees

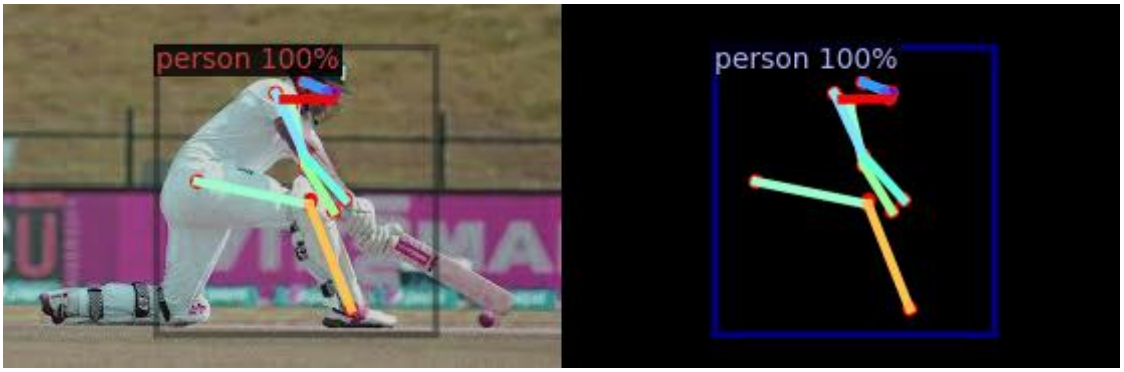


Figure 5.18. Detection of the Angle between the lines for better analysis using shot2

Angle between elbow, hand, and wrist: 43.68 degrees
Angle between shoulder, elbow, and hand: 161.78 degrees

Figure 5.19. Detection of the Angle between the lines for better analysis using shot3

Angle between elbow, hand, and wrist: 50.57 degrees
Angle between shoulder, elbow, and hand: 15.51 degrees

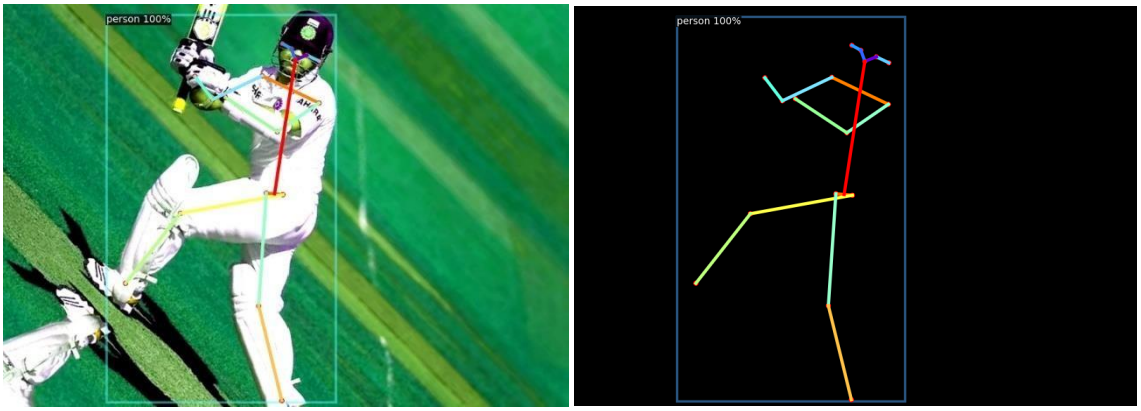


Figure 5.20. Detection of the Angle between the lines for better analysis using shot4

Angle between elbow, hand, and wrist: 26.00 degrees
Angle between shoulder, elbow, and hand: 112.03 degrees

The testing of gesture angle recognition is essential to assesing the level of accuracy of the algorithm within the framework of cricket analytics. The algorithms understanding of cricket specifically gesture movements can be measured via calculating degrees of angles among important points indicating joints such as shoulder and elbow. Furthermore, the visual representation exhibits reliable angle identification, that is crucial for improving the player performance. Additionally, this study highlights the real-world significance and possible challenges of cricket analytics. However, it directs potential advances tailored for this distinctive field. The illustrations supporting the written material not just demonstrate different viewpoints or players movements, instead they additionally assist in broadening the investigation in the larger context of cricket-specific occurances.

5.3 Discussion of the model performances

5.3.1 Case 1 – Shot Classification

InceptionV3 emerges as the best-performing model in the analysis of cricket shot classification, exhibiting exceptional accuracy, precision, recall, and F1 Score, all standing at an impressive 99.58%. This model's outstanding performance is attributed to its efficient use of inception modules with multiple filter sizes, enabling it to capture intricate features within gesture images. The high precision indicates a minimal rate of false positives, showcasing the model's proficiency in accurately classifying different cricket shots. Simultaneously, the high recall suggests that InceptionV3 effectively captures a comprehensive range of relevant instances, making it particularly adept at recognizing diverse gestures. The F1 Score, representing the harmonic mean of precision and recall, further emphasizes the model's balanced performance. InceptionV3's success underscores its suitability for applications demanding both precision and robust recognition of various cricket shots. In conclusion, InceptionV3 stands out as the optimal choice among the models analyzed, demonstrating superior capabilities in cricket shot

classification. Its architectural efficiency, as evidenced by its outstanding metrics, positions it as a top-performing solution for tasks requiring nuanced gesture recognition in the dynamic realm of sports analytics.

5.3.2 Case 2 – The Gesture and Angles Analysis

The visual analysis of gesture instances, facilitated by the Detectron2 model, adds a layer of interpretability and assurance to the model's predictions. The ability to visually inspect and validate the localization of key points contributes to building trust in the model's performance. Moving forward, this insight into the visual aspects of gesture detection opens avenues for refining the model, addressing potential challenges, and expanding its applications in real- world scenarios.

Chapter 6: Conclusion and Future Work

In conclusion, the purpose of this research is to address the long-standing difficulty of effectively identifying and categorizing cricket shots by strategically integrating deep learning approaches in a unique manner. The invention of CricNet, a novel system for automated cricket shot categorization, demonstrates the possibility of utilizing player posture photos to improve the comprehension and recognition of a variety of shot types among cricket players. Our study not only aims to improve upon the methods that are currently used in computer vision, but it also intends to be a pioneer in the use of deep learning techniques within the framework of cricket analytics. An all-encompassing framework for assessment has been made available because of the specific study objectives, which range from determining the efficiency of player posture photographs in shot categorization to investigating the function of deep learning models, in particular Convolutional Neural Networks (CNNs).

The CricNet system that has been suggested has the potential to revolutionize the processes of training cricket players, analyzing player performance, and developing multimedia applications. Our model attempts to improve the accuracy and reliability of cricket shot detection by including posture information. This will allow it to transcend the capabilities of standard image-based classification approaches. The angles between player joints, as analyzed by CricNet, provide a sophisticated knowledge of the movement and motion of the cricketer. This insight contributes greatly to the model's ability to differentiate between various strokes. The real-world usability of the proposed approach in professional cricketing contexts will be a key point as we move further with this project. In conclusion, this research not only makes a significant contribution to the development of a cutting-edge model for the automatic identification of cricket shots, but it also sets the framework for future breakthroughs in the use of deep learning and computer vision in the field of sports technology.

In future we will work towards the image reconstruction based on the real time angles changed. This will help in comparing the actual shot being played and the reconstruction of the image. This will give a better judgement for the batter to play a shot and understand the area of improvement.

References

Analytical Vidhya. Cricket Shot Analysis: A Complete Guide with Python. Retrieved November 25, 2023 : https://www.analyticsvidhya.com/blog/2021/06/analyze-cricket-data-with-python-a-hands-on-guide/

Cao, Zhe, GinesHidalgo , TomasSimon , Shih-EnWei, andYaserSheikh "OpenPose: Realtime multi-person 2D pose estimation using Part Affinity Fields." IEEE transactions on pattern analysis and machine intelligence 43.1 (2021): 172-186.

Francois Chollet. Xception: Deep Learning with Depthwise Separable Convolutions. CVPR 2017. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1251-1258

G James, D Witten, T Hastie, R Tibshirani.,2013. An introduction to statistical learning (Vol. 112, p. 18). New York: springer.

Goodfellow Y Bengio, A Courville., 2016. Deep learning (Vol. 1, No. 2). Cambridge: MIT press.

Hobday, H., Cole, J.H., Stanyard, R.A., Daws, R.E., Giampietro, V., O'Daly, O., Leech, R. and Váša, F., 2022. Estimating Brain Age Using UK Biobank Imaging Data with Deep Learning.

Hotz, N. What is CRISP DM? Data Science Process Alliance 2023.

Howard et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications 2017.

Huang et al. Densely Connected Convolutional Networks. CVPR 2017.

Iqbal, Saad, et al. "Cricket action recognition: A review." Pattern Recognition Letters 138 (2020): 21-30.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li and Li Fei-Fei.,Imagenet: A large-scale hierarchical image database. CVPR, 2009.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun Deep residual learning for image recognition. CVPR, 2016.

Khan, Sirajum, et al. "Classifying cricket shots in broadcast videos." Proceedings of the 23rd ACM international conference on Multimedia. 2015.

KristinaHost , MarinaIvasic-Kos (2022).An overview of Human Action Recognition in sports based on Computer Vision[online].Croatia.ElsevierLtd

Kumar, S., & Das, P. (2018). "Biomechanical Analysis of Stress Points in Cricket Shots for Injury Prevention." Journal of Sports Medicine and Health Science 14(2): 201-215.

Liu, Jun, et al. "An end-to-end network for pose detection of team sports players in images." IEEE Access 7 (2019): 175767-175776.

Long et al., Learning transferable features with deep adaptation networks. ICML 2015.

Lu, Haiping, et al. "Augmented skeleton-based action recognition using multi-scale deep CNN." Proceedings of the 24th ACM international conference on Multimedia. 2016.

M.Jagadeesh , Sooraj Nikam.P , Sudharsan.V , Vishnu.S , Rithesh S and Sagar Y (2023). Shot Analysis of Batsmen in Cricket Matches using Transfer Learning Techniques [online] 7th International Conference. Chennai: IEEE

Madhukar, A., et al. "Advancements in Pose Estimation for Cricket Shots." Journal of Sports Technology 12.3 (2020): 45-58.

Moon, Gyeongsik, et al. "InterHand2. 6M: A dataset and baseline for 3D interacting hand pose estimation from a single RGB image." Proceedings of the European Conference on Computer Vision (ECCV). 2020.

Oquab et al., Learning and transferring mid-level image representations using convolutional neural networks. CVPR 2014.

Peng et al., VisDA: The Visual Domain Adaptation Challenge 2017.

Russakovsky et al. Imagenet large scale visual recognition challenge. IJCV, 2015.

Saini, Jasbir, et al. "Cricket video annotation based on action recognition by normal human actions." Proceedings of the First International Conference on Intelligent Computing and Communication. Springer, Singapore, 2017.

Sampath et al., "Big data analytics for identifying batting and bowling leadership qualities in cricket." Procedia computer science 112 (2017): 1817-1826.

Sharma and Aloimonos. "A multi-sensory system for analyzing cricket shots." MultiMedia Modeling. Springer, Cham, 2016. 197-208.

Sharma and Gupta. "Deep Learning Approaches for Cricket Shot Classification Based on Player Poses." International Journal of Sports Science & Coaching 16(4) (2021).

Simonyan et al. Very deep convolutional networks for large-scale image recognition. arXiv 1409.1556, 2014

Simonyan et al., Very deep convolutional networks for large-scale image recognition. ICLR 2015.

Smith, J., et al. "Biomechanical Analysis of Cricket Shots: A Pose-Based Approach." Journal of Sports Biomechanics 21(2) (2019): 112-128.

Subhash Conv2D function of Keras and its use in CNNs 2023.

Sun et al., Return of frustratingly easy domain adaptation. AAAI 2016.

Szegedy et al. Rethinking the Inception Architecture for Computer Vision. CVPR 2016.

Szegedy et al., Rethinking the inception architecture for computer vision. CVPR 2016.

Tan et al. Efficientnet: Rethinking model scaling for convolutional neural networks. ICML, 2019.

Tan et al., A survey on deep transfer learning. ICANN 2018.

Vats, Kuldeep, et al. "Action recognition using enhanced motion vectors and deep residual network." The Visual Computer 35.6 (2019): 1-15.

Weiss et al., A survey of transfer learning. JMLR 2016.

WM Kouw, M Loog, An introduction to domain adaptation and transfer learning 2018: :1812.11806

Wu et al., Fang, F., Bai, X., & Tegabi, G. (2019). Detectron2.

Y LeCun, B Boser, JS Denker, D Henderson, RE Howard, W Hubbard, LD Jackel. Backpropagation applied to handwritten zip code recognition. Neural computation, 1989.

Yan, S., Xiong, Y., & Lin, D., "Spatial temporal graph convolutional networks for skeleton-based action recognition." Proceedings of the AAAI conference on artificial intelligence. Vol. 32. No. 1. 2018.

Ye, Qixiang, et al. "A survey on deep learning-based action recognition in images and videos." Pattern Recognition 112 (2021): 107584.

Yosinski et al., How transferable are features in deep neural networks? NIPS 2014.?