Shot-Net: A Convolutional Neural Network for Classifying Different Cricket Shots

BY

Md. Ferdouse Ahmed Foysal

ID: 151-15-5274

AND

Mohammad Shakirul Islam

ID: 151-15-311

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Nafis Neehal
Lecturer
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

Enamul Karim
Lecturer
Department of Computer Science and Engineering
Faculty of Science and Information Technology
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

NOVEMBER 2018

APPROVAL

This Research titled **"Shot-Net: A Convolutional Neural Network for Classifying Different Cricket Shots"**, submitted by **Md. Ferdouse Ahmed Foysal & Mohammad Shakirul Islam** to the Department of Computer Science and Engineering, Faculty of Science and Information Technology, **Daffodil International University**, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc in Computer Science & Engineering and approved as to its style and contents.

BOARD OF EXAMINERS

Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Sheak Rashed Haider Noori — **Internal Examiner**
Associate Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Md. Zahid Hasan — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Mohammad Shorif Uddin — **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that this research has been done by us under the supervision of **Mr. Nafis Neehal, Lecturer, Department of Computer Science and Engineering and** co-supervision **of Enamul Karim, Lecturer, Department of Computer Science and Engineering**, Faculty of Science and Information Technology, Daffodil International University. We also declare that neither this research nor any part of this research has been submitted elsewhere for the award of any degree.

Supervised by:

Nafis Neehal
Lecturer
Department of CSE
Daffodil International University

Co- Supervised by:

Enamul Karim
Lecturer
Department of CSE
Daffodil International University

Submitted by:

Md. Ferdouse Ahmed Foysal
ID: 151-15-5274
Department of Computer Science and Engineering
Faculty of Science and Information Technology
Daffodil International University

Mohammad Shakirul Islam
ID: 151-15-311
Department of Computer Science and Engineering
Faculty of Science and Information Technology
Daffodil International University

ACKNOWLEDGEMENT

At First we would like to express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year thesis successfully.

We really grateful and wish our profound our indebtedness to **Mr. Nafis Neehal**, Lecturer, **Mr. Enamul Karim**, lecturer, Department of CSE, Faculty of Science and Information Technology, Daffodil International University, Dhaka. Their Deep Knowledge & keen interest with supportive instructions helped us in the field of deep learning based research, finally we completed our work on "A Convolutional Neural Network for Classifying Different Cricket Shots". Their endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Syed Akhter Hossain**, Professor and Head, Department of Computer Science and Engineering, Faculty of Science and Information Technology, DIU, for his valuable support and advice to finish our project and also heartiest thanks to other faculty member and the staff of department of CSE, Daffodil International University.

At last, again we want to thank all the good wishers, friends, family, seniors for all the help and inspirations. This research is a result of hard work and all those inspirations and assistance.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Artificial Intelligence has become the new powerhouse of data analytics in this technological era. With advent of different Machine Learning and Computer Vision algorithms, applying them in data analytics has become a common trend. However, applying Deep Neural Networks in different sport data analyzing tasks and study the performance of these models is yet to be explored. Hence, in this paper, we have proposed a 13 layered Convolutional Neural Network referred as "Shot-Net" in order to classifying six categories of cricket shots, namely Cut Shot, Cover Drive, Straight Drive, Pull Shot, Scoop Shot and Leg Glance Shot. Our proposed model has achieved fairly high accuracy with low cross-entropy rate.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Cricket is one of the most exciting games in the world, batting is the ability of hitting the cricket ball with a cricket bat, and there are different kinds of cricket shots. Batsmen have to accommodate to various conditions when playing on different cricket pitches, especially in different countries therefore, as well as having distinguished physical batting skills, top-level batsmen will have thunder reflex action, excellent decision making and be good strategists [6]. Application of computer vision and machine learning techniques in cricket for different analysis is an emerging domain now. In cricket many technologies are used for visualization and coaching [1, 3]. From recent researches till now satisfactory results for detecting shots are not achieved.

We thought that we can do a deep convolution neural network (CNN) [2] based action detection. Therefore, we are proposing a novel approach to classify different types of cricket shots using Convolutional Neural Network and Deep Learning. An intelligent device can recognize the human corpus parts by feature extraction from the image [4]. Then, enunciated body parts will be recognized [5] using different approaches of action model representation. In machine learning image processing and pattern recognition feature extraction begins from an basic set of measured data and constructs evolved values intended to be informative and non-redundant, facilitating the following learning and generalization moves, and in some cases leading to better human interpretations. Feature extraction is related to dimensional reduction. We proposed a CNN based model where we input images in three convolution layer, three max pooling layer, four dropout layer and two dense layer.

1.2 Objective

Now, this is the time of technological revolution. In today's digital age we are using technology everywhere. In the field of games and sports, technology are used in every milliseconds, every steps and movements of a sportsmen are counted by digital support system. Our main objective is to classify cricket different shots. There are different types of shot like hook, pull, off drive, on drive, cover drive, square cut etc. We want to classify the cricket shots by using deep convolutional network. So we can describes these goals in a list like this:

- Our goal is to study how to classify or recognize different cricket shots.
- To develop a platform that will be able to detect all kind of cricket shots and human poses or activity.
- To visualize some analytical analysis of Human Activity like Cricket Shots classification classified by classifier algorithms.

1.3 Motivation

We were interested to do something different. So we decided earlier that we will do research on Artificial Intelligence (AI) and Machine Learning (ML) field and then we started to search for some ideas. But no idea couldn't satisfied us. As we are sports lover, one day we thought that we can do something on sports. At first we wanted to do something about football and found an idea called "Data-Driven Ghosting using Deep Imitation Learning". But we discovered that it was a high level work and may be tough for us and then we thought that as we are Bangladeshi people and we love cricket so much, then we can do some research on cricket and finally we have reached on an awesome idea and it's called "**Cricket Shot Classification using A Convolutional Neural Network.**" Besides this, we see that today's world is so much focusing on recommendation system. Users expect everything that the better things will be recommended to them by the system. To make a system to be recommendation capable must have the ability to take decision by itself. These made us interested to do such kind of research based work. Our work is fully related with machine learning techniques.

1.4 Rational of the study

There is no doubt that there are thousands of works done on Image processing or object classification domain. But there are only a few works done on Cricket shot classification. So our work is a new approach using different algorithms and simulation. To develop more efficient classifier application in the field of games and sports we give out best effort to develop our own model.

Image processing is an advance approach it can be split into different categories: one of these is Image Compression another image enhancement and the last is the restoration, and measurement extraction. It helps to reduce the amount of memory which is needed to store a digital image. Image can be defected. By digitization process and by faults image can be defected .Defected image can be corrected using Image Enhancement techniques.

1.5 Research Questions

It was so challenging for us to complete this work. In order to have a realistic, efficient and accurate response to the problem, the researchers wishes to propose following questions to express this feelings and outcomes this problem.

- Can we collect row image data for deep learning research?
- Is it possible to pre-process the row data using deep learning approach?
- Is it possible to improve cricket coaching system using this approach?
- How cricket game will be benefited by this approach?

1.6 Expected Outcome

In this section there is some points given that points was our min expected outcome. Expected outcome of this research based project is to build an algorithm or making a complete efficient procedure that will categorize cricket shots with respect to the built model of trained dataset.

- Cricket shots can be classified.
- A cricket team can be benefited by this.
- They can find the weakness of a batsman and solve it etc.

1.6 Layout of the Report

Chapter one have demonstrated an introduction to the project with objective, motivation, research questions, and expected outcome, this section describes the whole layout of this report.

Chapter two provides the discussion on what already done in this domain before. Then the later section of this second chapter shows the scope arisen from their limitation of this field. And very last, the root obstacles or challenges of this research are explained.

Chapter three describes the theoretical discussion on this research work. To discuss the theoretical part of the research, this chapter elaborates the statistical methods of this work. Besides, this chapter shows the procedural approaches of the CNN and Machine Learning classifier. And in the last section of this chapter, to validate the model as well as to show the accuracy label of the classifier, confusion matrix analysis is being presented.

Chapter four provides the experimental results, performance evaluation and result discussion. Some experimental pictures are presents in this chapter to make realize the project.

Chapter five discussed with summery of the study, future work and conclusion. t. This chapter is responsible to show the whole project report adhering to recommendation. The chapter is closed by showing the limitations of our works that can be the future scope of others who want to work in this field.

CHAPTER 2
Background Study

2.1 Introduction

In this section, we will discuss related works, research summary and challenges about this research. In related works section, we will discuss other research paper and their works, their methods, and accuracy which are related to our work. In research summary section we will give the summary of our related works. In challenges section, we will discuss how we increased the accuracy level.

2.2 Related Works

Several studies published on Cricket since last decade. The Hawk-Eye [1] by H. Collins and R. Evans is studied about advanced system of coaching for cricket. In 2015 a research paper was published on cricket shot classification based on motion vector by a group of Bangladeshi researcher. For action recognition, they use 3D MACH to classify the shots and to detect cricket shots they define 8 classes of angle ranges [7]. In 2016 K. Dixit and A. Balakrishnan from Stanford University published a report on Deep Learning using CNN's for Ball-by-Ball Outcome Classification in Sports [8]. They compare the performance of three different Convolutional Neural Network architectures, inspired by literature on activity recognition in videos.

In 2010 B. Yao and Li Fei-Fei published a paper on Human-Object Interaction Activities by modeling mutual context of human pose [9]. In their research paper they described a new random field model to encode the mutual context of objects and human poses in human-object interaction activities. They cast the structure learning problem as a model learning task, of which the structural connectivity between the object, the overall human pose, and different parts of the body are estimated using a structure search approach and new max-margin algorithm used for estimating the parameters of the model.

In another research paper, N. Batra, H. Gupta, N. Yadav, A. Gupta and A. Yadav proposed a multi-valued automated decision whether a ball is no-ball or wide ball [6]. Presenting game specific concept selection and event selection criteria. Another cricket shot classification using computer vision proposed by E Chowdhury and Abu U Jihan divided the approach into four phases of identifying batsman's hand stroke direction, tracking, detection of a collision of bat and ball and detection of human pose and skeleton joints [10]. M H Kolekar and K Palaniappan form a semantic video analysis based on low level image features and high-level knowledge for cricket video sequences encoded in hierarchical classification [12]. S A Angadi and Vilas Naik's shot boundary detection technique based on local color moments in YCbCr. They said that this color similarity can become a measure of finding difference between consecutive frames in a video [11]. In [13-19], different approaches of moving object detection, and sports analysis works has been shown.

2.3 Research Summary

Deep learning is a technique for implementing Machine Learning. It is made of artificial neural networks. Neural networks work as similar as our brain. CNN that means Convolutional Neural Network is one of the strongest networks in deep learning. It is an artificial neural network, which is also known as feed-forward ANN. In a "feed-forward" network information flows right through the networks.

Yann LeCun was the inventor of CNN. Inspired from human processes he made it. Actually CNN works like biological visual cortex. CNN is one of the most successful models in image classification. CNN's classification accuracy is better than any other traditional image classification algorithms. In CNN we don't have to do feature selection, but in other image classification algorithms, we have to do it. There are different types of layers that are used in CNN.

Convolution layer has a moving filter or kernel which passes through the image. Generally it passes through a 2D matrix (representation of image) and take a certain portion and applies dot multiplication and stores it in another matrix.

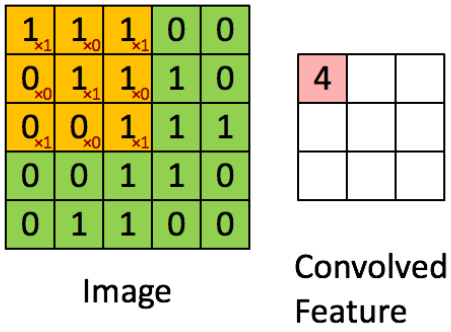

Fig 2.3.1: Convolution of a filter over 2D image

Dimension of the output matrix can be calculated by an equation. We can see an equation bellow where,

n_{out} - Output dimension

n_{in} - Input dimension

f - Window size

s- Stride

$$n_{out} = floor\left(\frac{n_{in}-f}{s}\right) + 1$$

The above equation is used to find Output of the dimension.

Pooling layer generally sits next to convolution layer. It mainly used to reduce memory and for fast computation. It reduces the volume. Max pooling is one of the most used layers in CNN. It sets a kernel and finds the max number from the matrix.

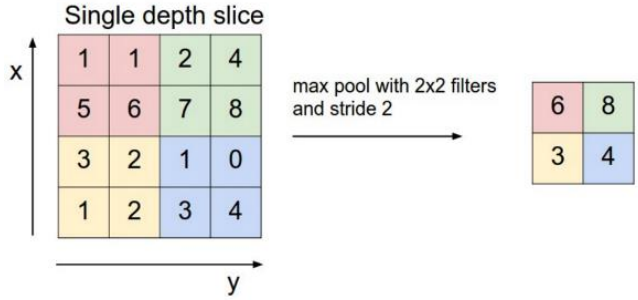

Fig 2.3.2: Max pooling

Fully connected layer gets 2D or 3D array as input from previous layer and converts the 2D or 3D array into 1D array.

The output layer of a convolutional neural network shows the probability of the classes. It is calculated by "Softmax" function. The equation of calculating the probability is given bellow.

$$\sigma\ (x_j) = \frac{e^{x_j}}{\Sigma_i e^{x_i}}$$

2.4 Challenges

The main challenges of this work is collecting and processing the dataset, dealing with the data set was too hard. To clean and normalize we used several steps and methods. After all training with many layers with different size of epoch took long time in our machine, so getting the final output we waited so much with keeping patience. There was not another dataset or resources regarding this paper domain. There was not enough work done before so we have to start from our own motivation.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

In this section we are going to elaborate the workflow of our novel approach to classify cricket shots. There are some key point like data collection, processing, proposed model also described with relevant equation, graph, table and description. Own developed CNN based model applied and own dataset used in this work. The chapter is being closed by giving the explanation of our project's statistical theories and besides, giving the clear concept of the implementation requirements.

3.2 Research Subject and Instrumentation

Research subject can be called as research area that was reviewed and studied for clearing concepts. Not only for implementation but also for design model, collecting data, implement or process data and training the model. On the other section is Instrumentation that is which technology and method we used. We used windows platform, python language with many packages like numpy, pandas, skit learn, matplotlib etc. Anaconda application was used for all the training and testing process, Anaconda is a free and open-source distribution of the Python and R programming languages for data science and machine learning applications.

3.3 Workflow

This research have few stages of workflow such as data collection, data processing data resize and augmentation, model selection etc.

Stage 1 - Data Collection: We collected data from web and created our own data-set by processing those raw data. Collecting data was so challenging, there is not a single dataset available in this domain.

Stage 2 - Data Processing: All data have been processed class by class after collection from various sources. There are lots of data having noise and errors. We manually process those data first then implement the selective dataset to the next step.

Stage 3 - Data Resize and Augmentation: After processing class by class data have been augmented and resized. For training purpose we had to go through data augmentation and resize. Augmented data give some overfitting that's why we done only a few and most important augmentation.

Stage 4 - Model Selection: To train and validate our data for better accuracy we choose out model. There are hundreds of convolutional neural networks. To get better accuracy with our machine configuration we implement few model and finally one model was selected for final training and testing process.

Stage 5 - Performance Evaluation: In this section, all the results have been discussed with graph. After training and testing those process gave us few accuracy graph with validation loss and accuracy. We also calculated the confusion matrix and a table for showing the precision, recall and f1 measure.

Stage 6 - Conclusion and Future Work: In this section there will be a conclusion and future work map.

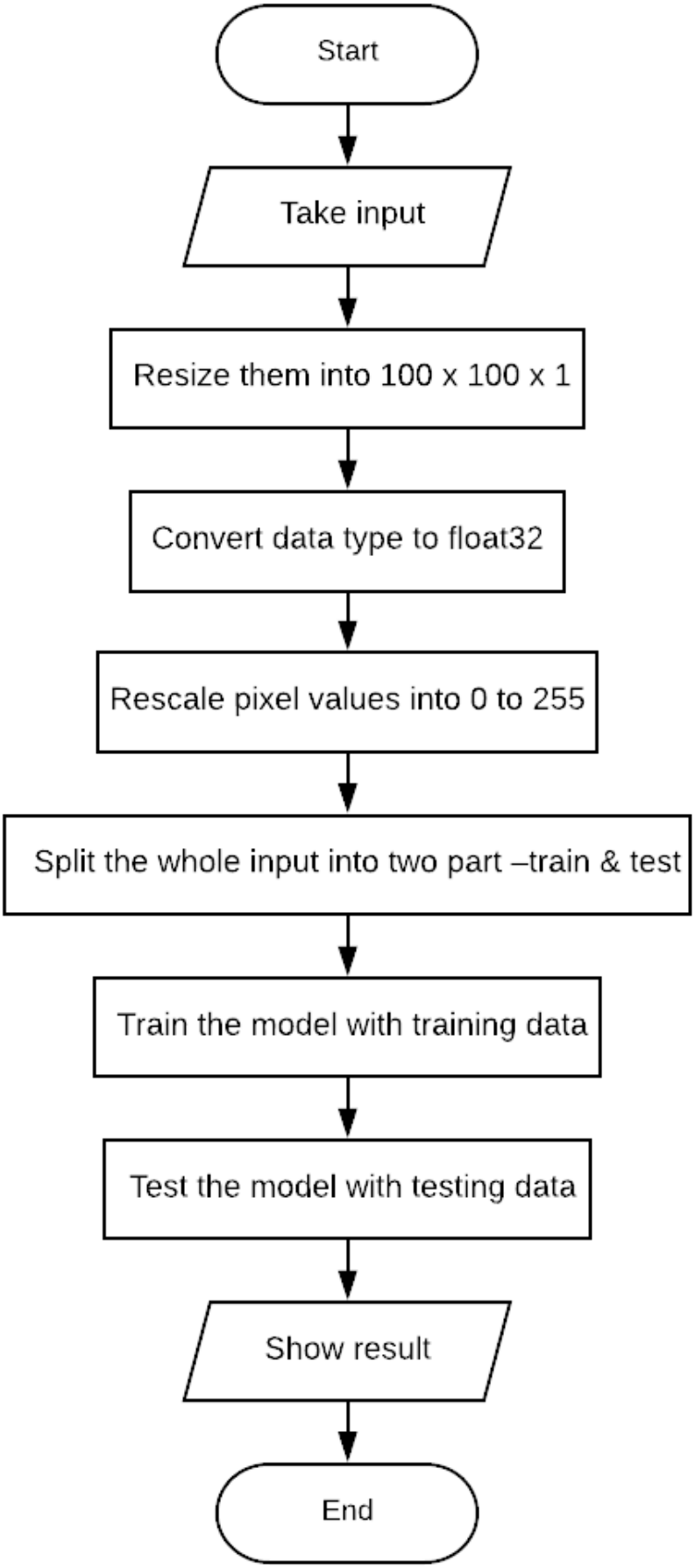

Fig 3.3: Workflow of our approach.

3.4 Data Collection Procedure

We have made a dataset of 3600 images. The dataset have 6 classes of cricket shots, each class contains 600 images. The classes are Cut shot, Cover drive, Straight drive, Pull shot, Leg glance shot and Scoop shot.

We took 80% image of the dataset that means 2880 images to train the model and 20% image of the dataset that means 720 images for testing. In the train dataset, each class contains 480 images and in the test dataset, each class contains 120 images.

All the data were collected and processed by us, and we named it Shot-Net.

Fig 3.4: A small part of our dataset

3.5 Data Processing

Data processing system has two steps one is data augmentation another one is data preparation. When we deal with the row data, the success mostly depends on the pre-processed data. The more efficiently data will be pre-processed; the outcome will be more accurate. In one word, it is the beginning challenge for such kind of research based work.

3.5.1 Data Augmentation

We artificially expanded the dataset to avoid overfitting. It adds value to base data by adding information derived from internal and external sources within an enterprise. It helps to increase the amount of relevant data in the dataset. We augmented the main data in 5 different methods, these methods given below:

- Rotate left -30 degree
- Rotate right +30 degree
- Shear by a certain amount
- Adding Salt and Pepper noise
- Shadding

3.5.2 Data Preparation

All the images of our dataset have different heights and widths. Since our model requires a fixed pixel for all images, we resize our data into 100 x 100 pixels.

We have also converted the images into grayscale. Since we have not better GPU in our computer that we have used to train the model .So we used grayscale images to train the model.

3.6 Proposed Methodology

We proposed our own CNN model, which have 13 layers .There are three convolutional layers:

- The first layer have 32-3 x 3 filters and ‘linear’ as activation function.
- The second layer have 64-3 x 3 filters and ‘linear’ as activation function.
- The third layer have 128-3 x 3 filters and ‘linear’ as activation function.

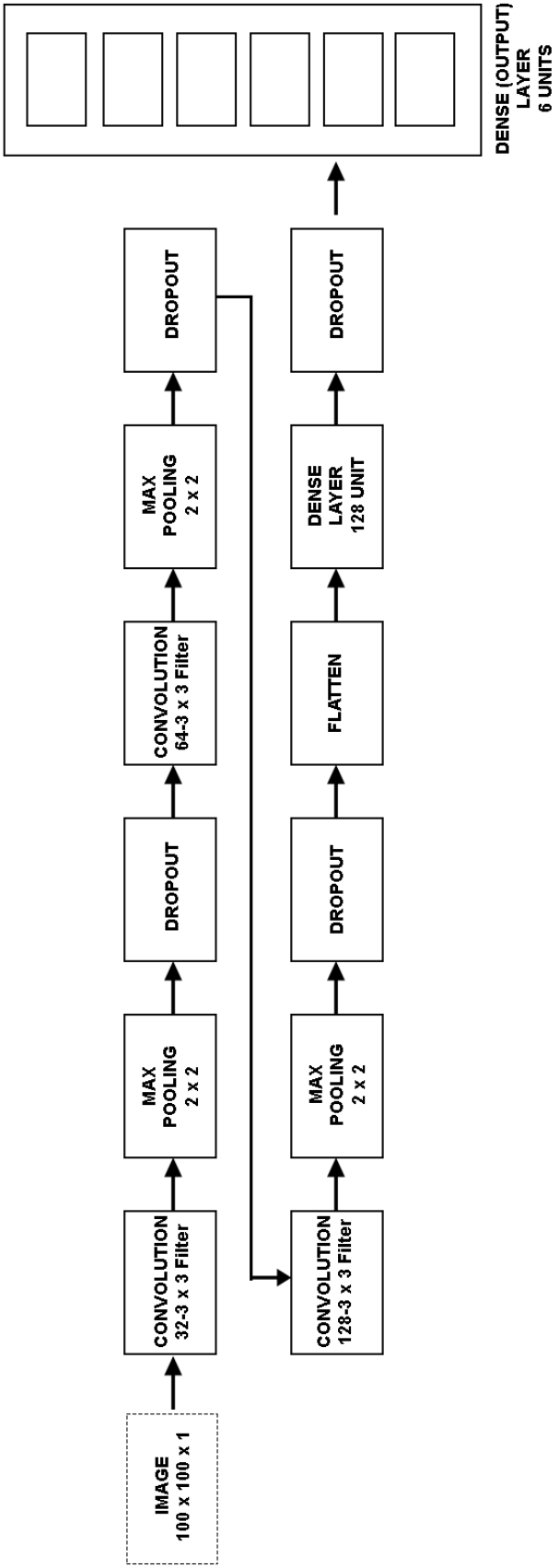

Fig 3.6: Architecture of our model.

In addition, there are three max-pooling layers each of size 2 x 2. There are four dropout layer with parameter 0.20. We have a flatten layer, in the model. Lastly there are two dense layer , where in one we used ‘linear ’ as activation function and in the other we used ‘soft-max’ as activation function. We used soft-max activation function to get the probability of each class. Let’s have a brief discussion about these layers below.

3.6.1 Convolutional Layer

CNNs have wide applications in image and video recognition, recommender systems and natural language processing. Convolutional neural networks. Sounds like a weird combination of biology and math with a little CS sprinkled in, but these networks have been some of the most influential innovations in the field of computer vision. 2012 was the first year that neural nets grew to prominence as Alex Krizhevsky used them to win that year’s ImageNet competition (basically, the annual Olympics of computer vision), dropping the classification error record from 26% to 15%, an astounding improvement at the time. When a computer sees an image (takes an image as input), it will see an array of pixel values. Depending on the resolution and size of the image, it will see a 32 x 32 x 3 array of numbers (The 3 refers to RGB values). In a traditional convolutional neural network architecture, there are other layers that are interspersed between these conv layers. I’d strongly encourage those interested to read up on them and understand their function and effects, but in a general sense, they provide nonlinearities and preservation of dimension that help to improve the robustness of the network and control overfitting.

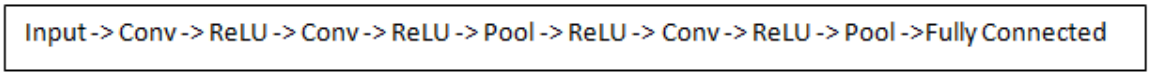

Fig 3.6.1: A classic CNN architecture

3.6.2 Feature extraction

Convolution is one of the main building blocks of a CNN. The term convolution refers to the mathematical combination of two functions to produce a third function. It merges two sets of information. In the case of a CNN, the convolution is performed on the input data with the use of a filter or kernel to then produce a feature map. We execute a convolution by sliding the filter over the input. At every location, a matrix multiplication is performed and sums the result onto the feature map.

3.6.3 Max pooling

Max Pooling is a down sampling strategy in Convolutional Neural Networks. The objective is to down-sample an input representation reducing its dimensionality and allowing for assumptions to be made about features contained in the sub-regions binned. It is basically used to reduce the size of the image because the larger number of pixels contribute to more parameters which can involve large chunks of data. Thus we need less parameters such that a CNN can still identify the image. Max pooling is discarded 75% of the activations and controlling overfitting.

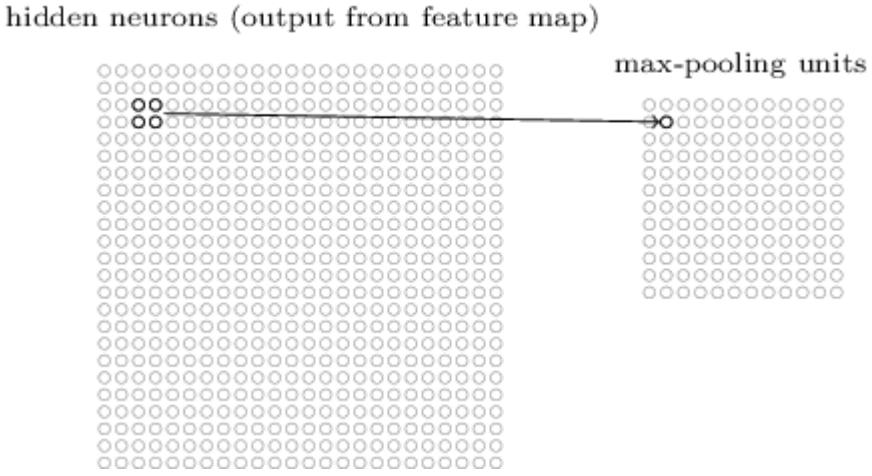

Fig 3.6.3: A pooling unit outputs the maximum activation.

3.6.4 Dense Layer

A dense layer is just another name of fully connected layer. Similar operations take place in dense layer where every neuron is connected with each other. It is also called dense because it represents a dense connection of dense neurons. A dense layer has weights associated with every neuron pair and with unique values. Different types of function like softmax activation function, SVM, and many others are used here for high-level reasoning in the neural network. But in our model, we stick used softmax for classification. After several convolution and pooling layers, we get some high-level features as input. These input images features are used as classifying to explore various classes. But when we combine convolution layer`s features and polling layer`s features it gives the better result of classifications. In Fully Connected layers summation of output probabilities is 1.One Conv layer share weights with other Conv layers. It is very difficult to attach all nodes with a softmax layer that's why we use a fully-connected layer to increase the efficiency of classification in our model.

3.6.5 Soft-max

Let us consider a classification model to classify with n classes. This model takes input datasets and an algorithm and produces a score of each class.the softmax activation function converts from score to the probability between 0 to 1.the summation of all probabilities is 1.we used this function to the final layer of convolutional neural networks to classify the classes.this function is produced multiple class from an input array.the probability distribution of softmax function is:

$$\sigma(x_j) = \frac{e^{x_j}}{\sum_{j=1}^{n} e^{x_i}}$$

Where i=1, 2, 3 ……. n and j=1, 2, 3 …….n

3.6.6 Dropout layer

Dropout is a technique used to improve over-fit on neural networks. The data structure of a neural network is a directed graph where each node represents a bias, while each edge represents a weight. If there are unit y with bias b, and in edges e1,e2,e3 with weights w1,w2,w3, then, when the signals that comes into y along e1,e2,e3 are x1,x2,x3 respectively, the output of y is w1x1+w2x2+w3x3+b. As we already know, the deeper the network is, the more parameter it has. For example, VGGNet from ImageNet competition 2014, has some 148 million parameters. That's a lot. With that many parameters, the network could easily over fit, especially with small dataset. The CNN has to work within a robust environment hence dropout becomes necessary. Dropout is basically chosen between 0.2 to 0.8. Dropout removes the neurons randomly based on the parameters provided by the user like 0.4 etc.

3.6.7 Flatten Layer

The reason to use this is to need to insert this data into an artificial neural network later on. Fully connected layers don't have a local limitation like convolutional layers (which only observe some local part of an image by using convolutional filters). This means it can combine all the found local features of the previous convolutional layers. Each feature map channel in the output of a CNN layer is a "flattened" 2D array created by adding the results of multiple 2D kernels (one for each channel in the input layer).

3.7 Training the Model

We used Adam optimizer to compile our model. We used 80 % of training dataset to train and 20 % is used for validation. Training dataset has 2880 images so that we used 2304 images to train and 576 images to validate. We used a batch size of 64 .We have trained the network for 40 epochs.

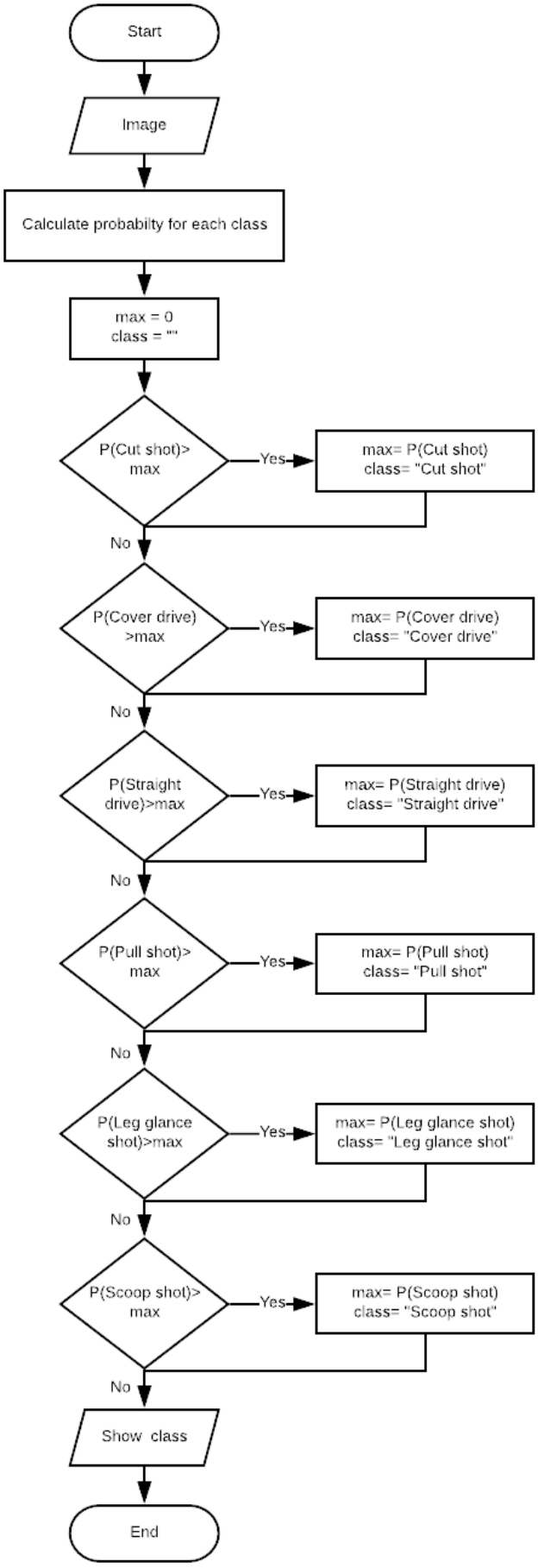

Fig 3.7: Training procedure of the model

3.8 Implementation Requirments

After the proper analysis on all necessary statistical or theoretical concepts and methods, a list of requirement has been generated that must be required for such a work of image Classification. The probable necessary things are:

Hardware/Software Requirements

- ✓ Operating System (Windows 7 or above)
- ✓ Hard Disk (minimum 500 GB)
- ✓ Ram (Minimum 4 GB)

Developing Tools

- ✓ Python Environment
- ✓ Spyder (Anaconda3)

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

In this section we described the construction process of cricket shot classification model. The overall process of the model divided into few steps like dataset collection, data preparation, data augmentation, data resize, proposed model description and finally training procedure of the model.

4.2 Performance Evaluation

Training accuracy is usually the accuracy when the model is applied on the training data. When the model is applied on a randomly-selected images from different class, is known as validation accuracy. Fig shows a graph which contains training and validation accuracy of our model.

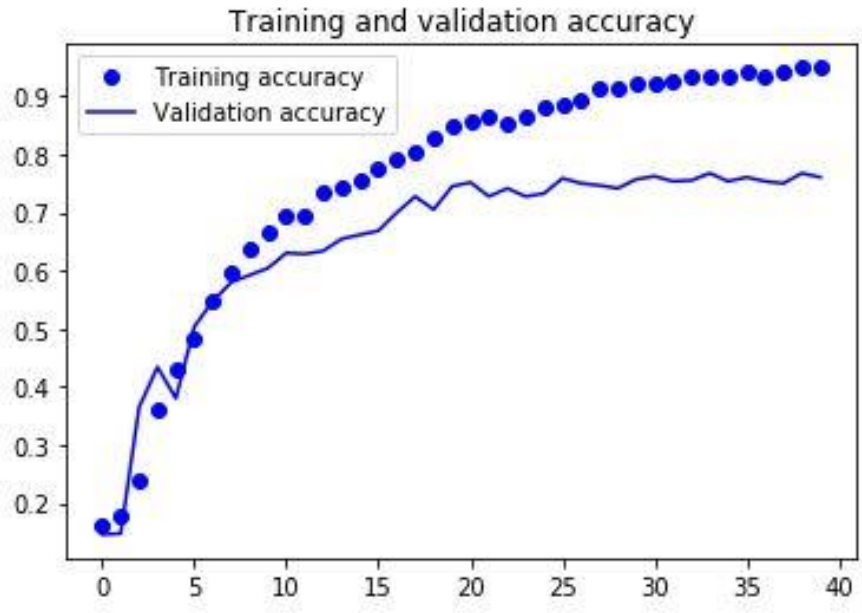

Fig 4.2.1: Training and validation accuracy

Training loss is the error on the training set of data. Validation loss is the error after running the validation set of data through the trained network. Fig shows a graph which contains training and validation loss of our model.

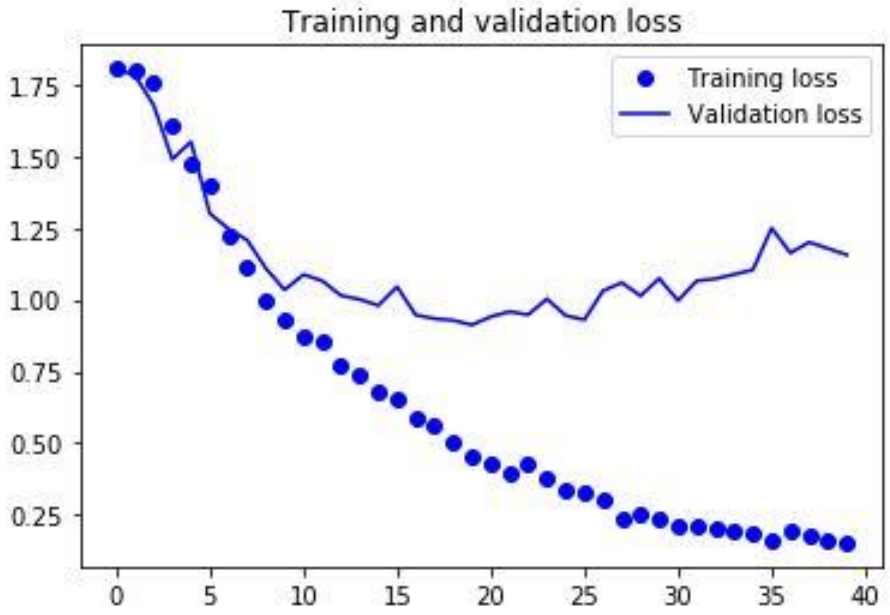

Fig 4.2.2: Training and validation loss

4.3 Result Discussion

We calculated Precision, Recall and F1-score from test dataset containing 840 images. From the classification report we can see Precession average is 0.80, Recall average is 0.79 and average F1-score is 0.79. So it can be said that the performance of our classifier is pretty good. From table of Classification report we can see that the classifier achieved a decent accuracy, which is 80%.

Table 4.3: Classification report

Class	Precession	Recall	F1-score
Cut shot	0.69	0.76	0.72
Cover shot	0.74	0.78	0.76
Straight drive	0.78	0.83	0.81
Pull shot	0.89	0.77	0.83
Leg glance shot	0.79	0.88	0.83
Scoop shot	0.88	0.72	0.79
Avg.	0.80	0.79	0.79

We calculated Normalized confusion matrix from confusion matrix, fig.1 shows the confusion matrix without normalization and fig.2 Show the normalized confusion matrix.

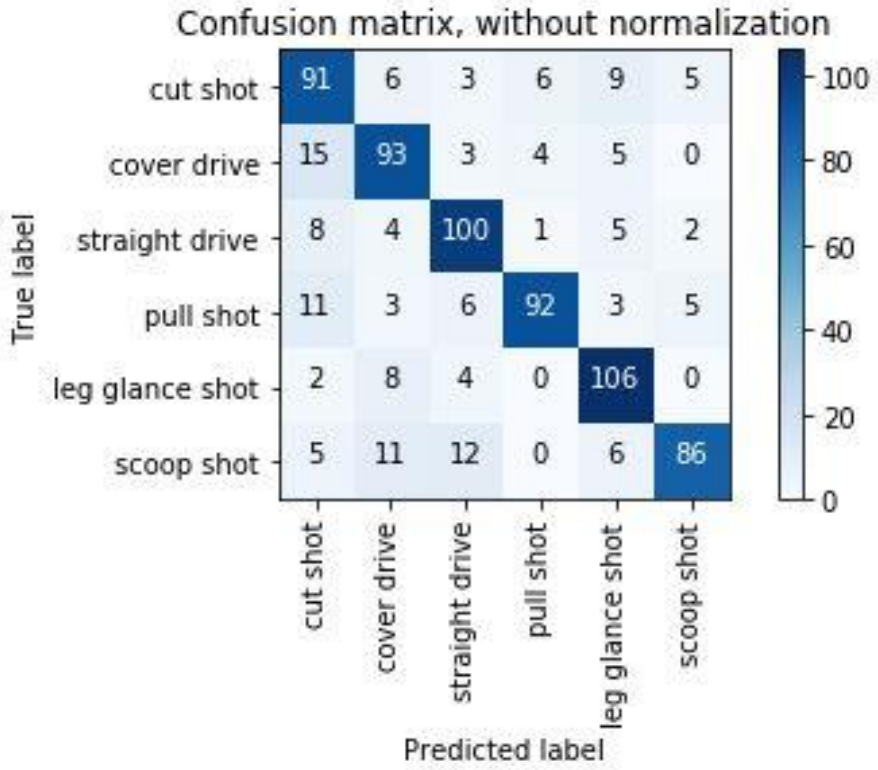

Fig 4.3.1: Confusion matrix

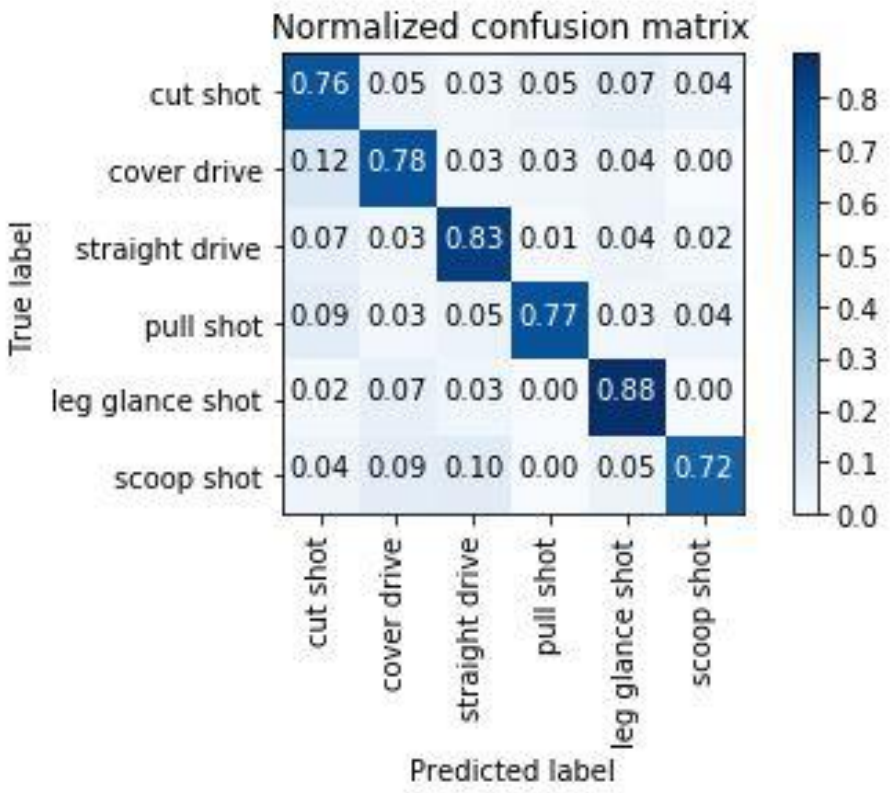

Fig 4.3.2: Normalized confusion matrix

Precision: In the field of information retrieval, precision is the fraction of retrieved documents that are relevant to the query:

$$\text{precision} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{retrieved documents}\}|}$$

Precision is used with recall, the percent of all relevant documents that is returned by the search. The two measures are sometimes used together in the F1 Score (or f-measure) to provide a single measurement for a system. Note that the meaning and usage of "precision" in the field of information retrieval differs from the definition of accuracy and precision within other branches of science and technology.

Recall: Recall is the piece of relevant instances that have been retrieved over the total amount of relevant instances. High recall means that an algorithm returned most of the relevant result.

$$Recall = \frac{tp}{tp + fn}$$

F-measure: f-score is a measure of test's accuracy by considering both precision and recall. it is a harmonic average of precision and recall.

$$F - score = 2 * \frac{precision * recall}{precision + recall}$$

Accuracy: accuracy refers to the familiarity of the measured value to a known value.

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

CHAPTER 5

CONCLUSION, RECOMMENDATION AND FUTURE WORKS

5.1 Introduction

It has no doubt that there are lots of research works on. Cricket shot recognition and classification are one of the important domain of cricket game. In a variety of applications, it has become important. Now a days there are many kind of technology used in cricket, so this approach will invent a new technology what is our main goal to find out something for games and sports arena.

5.2 Conclution

In this paper, we provide an approach of cricket shot classification approach by our CNN model. We used three convolution layer, three max polling layer, four dropout layer, one flatten and two dense layer. We use dropout layers to reduce overfitting. The result we've achieved is really promising. Hopefully this approach will be pursued and developed in future as part of further contributions in cricket.

5.3 Future Works

In our proposed method we can classify different kinds of cricket shots, we have used convolution neural networks to build a model for our Shot-Net data. Our future goal is to make a better neural network to get a better accuracy. We have a plan to do a 3D depth image based classification by deep learning. Where we will use MS Kinect or Intel RealSense, We will use different types of algorithm and that will select the efficient one.

REFERENCES

[1] H. Collins and R. Evans, “Hawkeye second edition public understanding revised 07 clean,” 2008

[2] A. Krizhevsky, Ilya Sutskever, G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks”.

[3] H. Collins and R. Evans, “hawkeye-final-submission,” 2012.

[4] J. Shotton, T. Sharp, and A. Kipman, “Real-time human poserecognition in parts from single depth images,” Communications of the 2013. [Online] Available: http://dl.acm.org/citation.cfm?id=2398381.

[5] M. Andriluka, S. Roth, and B. Schiele, “Pictorial structures revisited: People detection and articulated pose estimation,” Computer Vision and Pattern, pp. 1014–1021, 2009.

[6] Batra, Nikhil & Gupta, Harsh &Yadav, Nakul& Gupta, Anshika & Yadav, Amita. (2014). Implementation of augmented reality in cricket for ball tracking and automated decision making for no ball. 316-321. 10.1109/ICACCI.2014.6968378.

[7] Wikipedia, Different Cricket Shots, https://en.wikipedia.org/wiki/Batting (cricket).

[8] Kalpit Dixit and Anusha Balakrishnan, Deep Learning using CNN's for Ball-by-Ball Outcome Classification in Sports", report submission on the course of Convolutional Neural Networks for Visual Recognition, Stanford University,2016.

[9] B. Yao, Li Fei-Fei, “Modeling Mutual Context of Object and Human Pose in Human-Object Interaction Activities”, 978-1-4244-6985-7/10/$26.00 ©2010 IEEE.

[10] Chowdhury, AZM Ehtesham, and Abu Umair Jihan. “Classification of Cricket Shots Using Computer Vision.” (2014).

[11] S. A. Angadi, Vilas Naik, “A Shot Boundary Detection Technique Based On Local Color Moments In YCbCr Color Space”, CS & IT-CSCP 2012.

[12] M. H. Kolekar, K. Palaniappan, and S. Sengupta, "Semantic Event Detection and Classification in Cricket Video Sequence," 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing, Bhubaneswar, 2008, pp. 382389.

[13] M. Shakirul Islam, F. Ahmed Foysal, N. Neehal, E. Karim, S. A. Hossain, “InceptB: A CNN Based Classification Approach for Recognizing Traditional Bengali Games”, ICACC-2018.

[14] Hitesh A Patel, Darshak G Thakore, “Moving Object Tracking Using Kalman Filter”, IJCSMC, Vol. 2, Issue. 4, April 2013, pg.326 – 332.

[15] G. Zhu, Q. Huang, C. Xu, L. Xing, W. Gao, and H. Yao. Human behavior analysis for highlight ranking in broadcast racket sports video. in IEEE Transactions on Multimedia, 9(6), 2007.

[16] Andrew Simonyan, Karen; Zisserman. Very deep convolutional networks for large-scale image recognition.

[17] Rock, R & Als, Adrian & Gibbs, P & Hunte, Carlos. (2012). The 5 the Umpire: Cricket"s Edge Detection System.

[18] D.a. Forsyth and V.O. Brien, “Computer Vision second edition,”Computer Vision: A Modern Approach (2003): 88-101.

[19] D. G. Lowe, “Distinctive image features from scaleinvariant keypoints,” International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

APPENDIX

To complete the project we faced so many problem, first one was to determine the methodological approach for our project. It was not traditional work it was a research based project, more over there were not much work done before on this area. So we could not get that much help from anywhere. Another problem was that, collection of data, it was big challenge for us. There was no dataset available on this kind of sports, that's why we collect our own data and developed a best fit model. Working with this kind odd data is so interesting.

Plagiarism Report: