SPORTS EVENTS CLASSIFICATION USING COVOLUTIONAL NEURAL NETWORKS

BY

Shahana Shultana
ID: 151-15-5155

AND

Md. Shakil Moharram
ID: 142-15-3662

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

Supervised By

Mr. Nafis Neehal
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Mr. Eshtiak Ahmed
Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

NOVEMBER, 2018

APPROVAL

This Project/internship titled **"Sports Events Classification Using Convolutional Neural Networks"**, submitted by **Shahana Shulatana, ID No: 151-15-5155 & Md. Shakil Moharram, ID No: 142-15-3662** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on November 5, 2018.

BOARD OF EXAMINERS

____

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

____

Narayan Ranjan Chakraborty **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

____

Md. Tarek Habib **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 and co-supervision of **Mr. Eshtiak Ahmed,** Lecturer, Department of CSE, 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:

____

Mr. Nafis Neehal
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

____

Mr. Eshtiak Ahmed
Lecturer
Department of CSE
Daffodil International University

Submitted by:

____

Shahana Shultana
ID: - 151-15-5115
Department of CSE
Daffodil International University

____

Md. Shakil Moharram
ID: - 142-15-3662
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to Almighty Allah for His divine blessing makes us possible to complete the final year research project successfully.

We really grateful and wish our profound our indebtedness to our respected supervisor **Mr. Nafis Neehal**, Lecturer, Department of Computer Science & Engineering, Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Machine Learning*" to carry out this research. His 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 research project successfully.

We would like to express our heartiest gratitude to our honorable department head sir **Prof. Dr. Syed Akhter Hossain**, Professor and Head**,** Department of Computer Science & Engineering, Daffodil International University for his kind help to finish our research project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

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

ABSTRACT

Analysis of different sports data to get valuable insight has become immensely important now-a-days. Profuse application of Artificial Intelligence in different sectors has become a very popular trend as well. However, application of AI in sports analytics is still a new research domain left for exploration. With a view to applying AI in sports analytics, we have deployed Inception V3 and MobileNet which are Google's most popular Convolutional Neural Networks to successfully recognize 5 different sports events from a huge image dataset of these events. We also developed a Convolutional Neural Network model which name is SP-Net and we trained our proposed model with these 5 different sports events. SP-Net correctly predicted the class almost all images during the period of testing and gives a high performance. In terms of performance our proposed model SP-Net surpass Inception v3 and MobileNet both of these models. Besides, Inception v3 and MobileNet also achieved a very high performance in terms of accuracy, precision, recall and f-measure while applied on the target dataset for successful classification.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
Introduction

1.1 Introduction

Sports are a large area of business and it has a heavy market value. For OR (operation research) models we can consider sports as an effective application area. Here we classified several sports properties and activities using classification model. We create five sports classes by downloading the pictures of these classes and install two (Inception v3 and MobileNet) built-in convolutional neural network models and proposed a new convolutional neural network model for classification.

TensorFlow [1] is an example of 2G artificial intelligence and this open source software library is developed by Google. TensorFlow supports several neural networks (CNN, RNN, etc). We just retrained our dataset with Inception v3 and MobileNet architecture to see how much properly they classified. Here we just train the last layer of the neural network. We proposed a new classification model SP-Net where we use CNN and get better performance and accuracy than google built-in model Inception v3 and MobileNet. Images are classified based on the instances and properties they have. Image properties and activities can easily define the types of sports. Hand position, body movement, present or absent of an object are the key factors in this case. If two or more objects are missing in an image, in this case, an image can be classified incorrectly because model will not be able to understand the correct class for this image, like for water polo ball, water, hand position, etc are the key factors.

We used python pillow library for image augmentation. We augmented our images to create differences from the real image so it will be difficult to assigned this image in an appropriate class.

1.2 Motivation

In the current world, sport means a lot. A huge number of sports events arranged regularly around the whole world for entertainment and business purpose. Classify these sports events manually is not effective because it is a matter of time and money. If we can classify these events automatically it can save both of our money and time. Olympic is the world largest sports events and they classify their events manually which is not feasible. Our proposed model could help sports organization for automatic sports events classification.

1.3 Research Questions

1. Is it feasible to classify sports events manually?

1.4 Expected Outcome

- To classify sports events using built-in model and measure their performance.
- To create a CNN model which can perform this classification task very well.
- To make an automated visual system that could classify different sports events.

1.5 Layout of the Report

Chapter one displayed an overview about our work, give an introduction and discuss the motivation, research questions and expected outcome.

- Chapter two will discuss the related works done prior for sports analytics, research summary and challenges.
- Chapter three will discuss the research methodology.
- Chapter four will examine the experimental results and discussion.
- Chapter five will give the future works and conclusion.

CHAPTER 2

Background Study

2.1 Introduction

In this chapter, we will discuss the related works previously done for sports analytics and research summary. In related works section we will talk about other research paper and their working methodology. In research summary part we will present the summary of our related works.

2.2 Related Works

Now-a-days sports are not only the part of entertainment but also the part of international business. Athletic performance is very important for sports. Victor Cordes and Lorne Olfman [2] developed a genetic algorithm functionality to predict the athletic performances. The authors collected summaries of game statics and created feature vector from player performances, used k-fold cross validation for evaluating vectors and then combined an isolated feature subset (genetic algorithm outputs) with the best fitness.

Chan et al [3] described how to find particular types of player like defense, offense, etc in ice hockey. The authors used a clustering technique. Established a relationship between the types of clustered player and performance of the team by used regression model for these clustered. The authors given a tool to assess new deals and the signing of new players which can be use by the team managers and this is an Excel based tool.

Ahmed et al [4] described a method to create an excellent cricket team with best performance and low cost. They used genetic algorithm with several objective and displayed a graph between net bowling average and net batting average and a decision making tool is provided by them to arise as a successful team.

Biao Xu [5] described a Genetic Algorithm Neural Network (GANN) based system which can predict the sport performance. Used genetic algorithm (GA) for feature selection like weather, weight, experience, training time, height, etc and used back-propagation (BP) neural network for prediction. This paper used GANN for the first time to guess the sport performance.

Fister Jr. et al [6] developed a tool which can generate online datasets of sport activity in CSV format and the name of the proposed tool is SportyDataGen. These datasets are already processed so preprocessing is not needed. The authors collected some real data from athletes and they wanted to add more soprts classes in the future for generating more sports activity.

DeSarbo W, Madrigal R [7] described the formation of a particular team, athlete or league based on the choices of sport fan. The authors given a procedure which is based on multidimensional scaling, they collected data from university student and then they performed the segmentation of fans. A few ways provided by the authors that can be used by an organization.

2.3 Research Summary

Table 2.1: Summary of related works

SL	Author	Methodology	Description	Outcome
1.	Victor Cordes, Lorne Olfman	K-fold cross validation, Genetic algorithm	Collected summaries of game statics and created feature vector from player performances.	Developed a genetic algorithm functionality to predict the athletic performances.

2.	Chan TYC, Cho JA, Novati DC	Clustering technique, Regression model	Established a relationship between the types of clustered player and the performance of the team. Given a tool to assess new deals and the signing of new players which can be use by the team managers and this is an Excel based tool.	Find particular types of player like defense, offense.
3.	Faez Ahmed, Kalyanmoy Deb, Abhilash Jindal	Genetic algorithm	They used genetic algorithm with several objective and displayed a graph between net bowling average and net batting average.	A decision making tool is provided by them to arise as a successful team.
4.	Biao Xu	Genetic algorithm (GA), Back-propagation (BP) neural network	Used genetic algorithm (GA) for feature selection like weather, weight, experience, training time, height, etc and used back-propagation (BP) neural network for prediction.	A Genetic Algorithm Neural Network (GANN) based system which can predict the sport performance.
5.	Iztok Fister Jr., Grega Vrbanciˇ c, Lucija Brezo ˇ cnik, Vili	K-means clustering	Developed a tool which can generate online datasets of sport activity in CSV format and these datasets are already processed so	SportyDataGen- a tool which can generate online datasets of sport activity.

	Podgorelec and Iztok Fister		preprocessing is not needed. Generated clusters of sports activities using k-means clustering.	
6.	DeSarbo W, Madrigal R	Multidimensional scaling, Segmentation	Given a procedure which is based on multidimensional scaling, they collected data from university student and then they performed the segmentation of fans	The formation of a particular team, athlete or league based on the choices of sport fan.

CHAPTER 3

Research Methodology

3.1 Introduction

In this chapter, we will discuss about our working procedures which can be divided into some steps: Data Collection, Data Processing, Model Installation, Train Model and Test Model. Figure 3.1 shows the working steps.

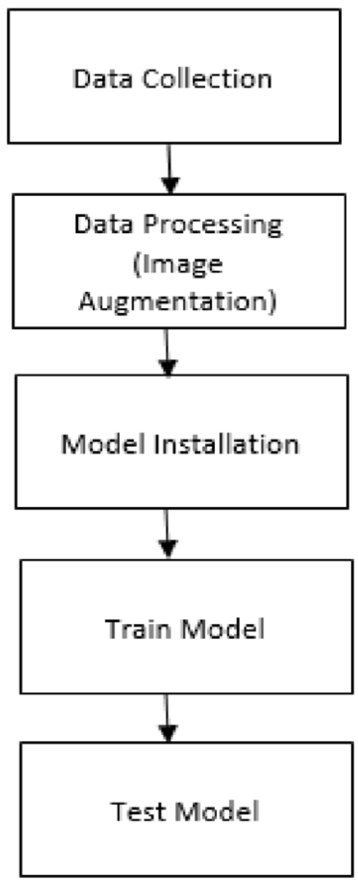

Figure 3.1: Working Steps

3.2 Data Collection

We select five categories/classes of Olympic Sports (Badminton, Basketball, Swimming, Table Tennis, Water Polo). Download image data from google images. For each class we download 100 photos. Total 5 x 100 = 500 images.

3.3 Data Processing

Resize all 500 images to a specific size (200x200). Use python pillow library for resizing. Augment all the images (Rotate +30 degree, Rotate -30 degree, Flip Horizontally, Scale the image 70%, Create a light black shade on each image). So after augmentation, from each original image 5 new images are created. So, now total dataset is like this per class 100 Original Image + 500 Augmented Image = 600 Image per class. As we have 5 classes, total images will be = 600 x 5 = 3000 images.

From this 3000 image (600 image per class), take 20 images from each class (Randomly, the image can be original or can be augmented). So total 20 x 5 = 100 images. Save them in a separate folder. Name it as Test Folder. Download 50 image (10 image per class) from google images and also save them into test folder. Use the rest 2900 images (580 images per class) to train the model. For the validation of SP-Net architecture we download 113 images (Badminton (22), Basketball (18), Swimming (27), Table Tennis (22), Water Polo (24)) and save them in Validation Folder.

3.4 Model Installation

In this work we used two built-in model. At first we install Tensorflow. Then we install Inception v3 model and MobileNet model to train our dataset.

3.4.1 Inception v3

Inception v3 is a built-in model which is developed by google for classification. Inception v3 factorized 7x7 convolutions and it has two parts: Part 1-Feature extraction from input images and Part 2-Classifies images based on their feature. The mean and standard-deviation for all output feature maps of a layer is computed by batch normalization. Multiple 3x3 kernels can be generated from 5x5 kernels and 7x7 kernels. Figure 3.2 shows the decomposition of a 7x7 or 5x5 kernels into multiple 3x3 kernels.

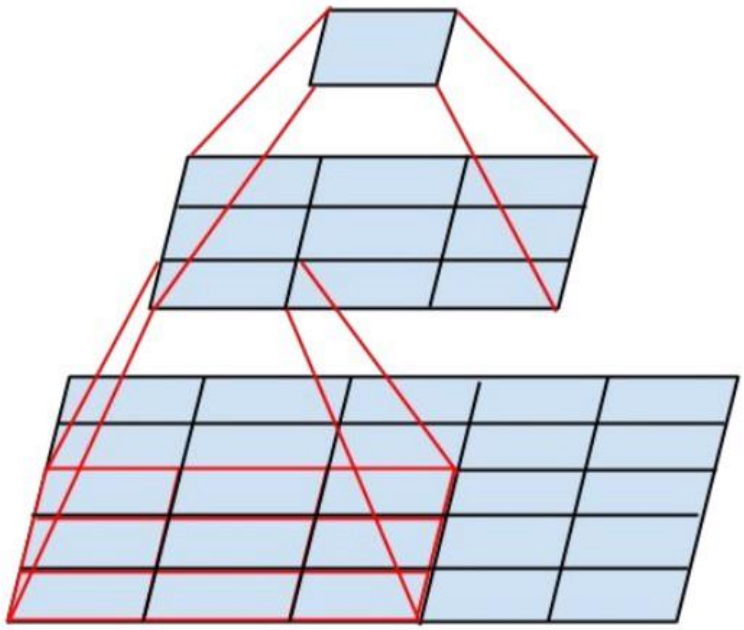

Figure 3.2: Multiple 3x3 kernels generation

3.4.2 MobileNet

In this work we also used MobileNet architecture for image classification. MobileNet is also a built-in model and in this architecture depthwise separable convolutions is used instead of conv3x3. Figure 3.3 shows the overview of MobileNet architecture.

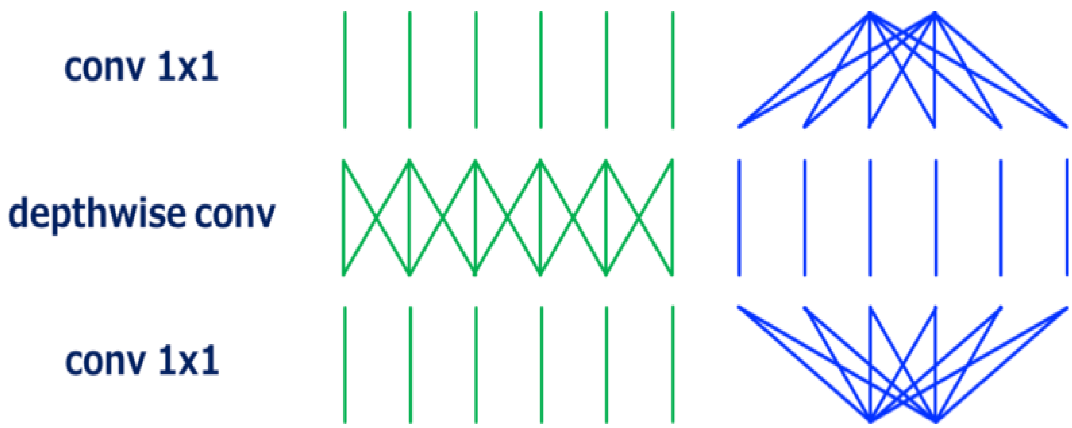

Figure 3.3: Overview of MobileNet Architecture

For MobileNeT architecture the pixel value of input image can be 128,160,192, or 224px. Model size refers to the fractional value of the whole MobileNet model like 0.75, 0.50, 0.2 and in our work we used MobileNet_0.50.

3.4.3 SP-Net

We proposed a new convolutional neural network (CNN) model and the name of our proposed model is SP-Net. In SP-Net we used Convolution2D, MaxPooling2D these two convolutional layers and Dropout, Flatten, Dense, Activation these four non-convolutional layers. SP-Net used a dropout layer with a dropout ratio 0.5, two convolution2D layers where the first conv2D layer used 32 filters, window size 3x3 and the second conv2D layer used 64 filters, window size 2x2, two max pooling layers with a window size of 2x2, three activation layers, one flatten layer and two dense layers. Figure 3.4 shows the model summary of SP-Net architecture.

```
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 200, 200, 32)      896
_________________________________________________________________
activation_1 (Activation)    (None, 200, 200, 32)      0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 100, 100, 32)      0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 100, 100, 64)      8256
_________________________________________________________________
activation_2 (Activation)    (None, 100, 100, 64)      0
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 100, 50, 32)       0
_________________________________________________________________
flatten_1 (Flatten)          (None, 160000)            0
_________________________________________________________________
dense_1 (Dense)              (None, 256)               40960256
_________________________________________________________________
activation_3 (Activation)    (None, 256)               0
_________________________________________________________________
dropout_1 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_2 (Dense)              (None, 5)                 1285
=================================================================
Total params: 40,970,693
Trainable params: 40,970,693
Non-trainable params: 0
_________________________________________________________________
```

Figure 3.4: Model Summary of SP-Net Architecture

3.4.3.1 Convolutional Layer

Convolutional layer is the main part of the CNN and it can be referred as the eyes. This layer explores earmarked feature. Convolution layer must have a kernel size k x k which referred the size of a chunk. For getting an activation map the chunk move over the whole matrix. Filter size refers to the number of filter used by the layer. Just after the convolution layer a pooling layer is added. The purpose of pooling layer is to reduce the size.

3.4.3.2 Non-Convolutional Layer

Non-Convolutional layer mainly used for classification. Collected specific features from convolutional layer and classified them based on these features. Non-Convolutional layer also create the environment to avoid overfitting. Dropout, Dense, Flatten, etc are non-convolutional layers.

In SP-Net we used both types of layer and Figure 3.5 shows the SP-Net architecture.

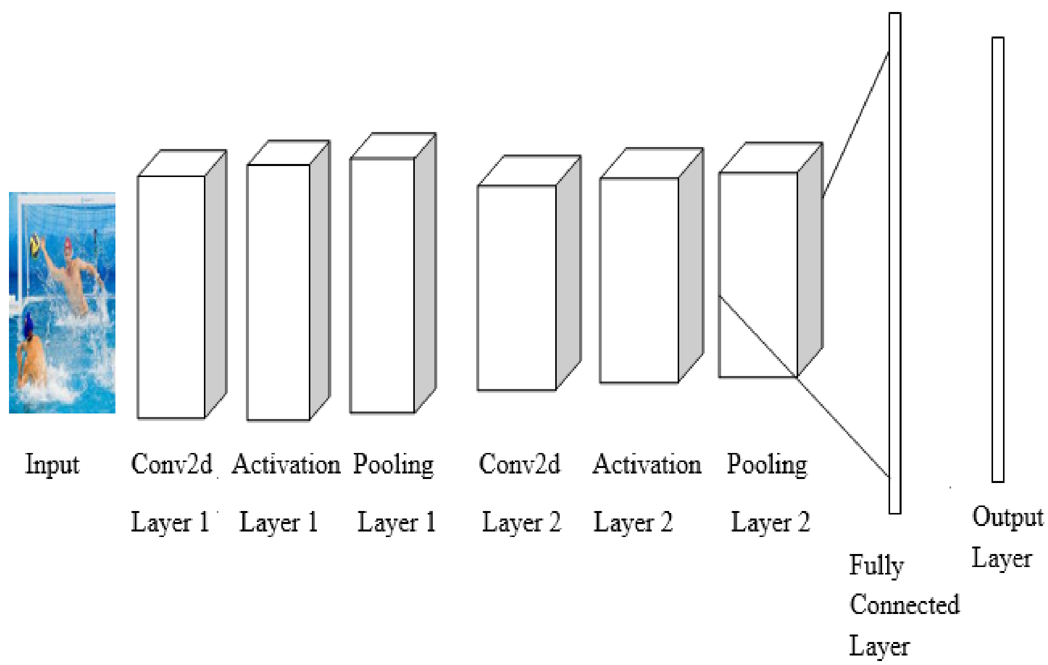

Figure 3.5: SP-Net Architecture

3.5 Train Model

We train our dataset using Inception v3, MobileNet and SP-Net model. Before starting the training, we start - Tensorboard.

3.5.1 Inception v3

Train 2900 images (580 images per class) with Inception v3 model. Validation set created from these 2900 images automatically. The number of training steps is 4000.

From Tensoboard we get the graph of accuracy and cross-entropy. Figure 3.6 shows the accuracy graph and Figure 3.7 shows the cross entropy graph for Inception v3 model.

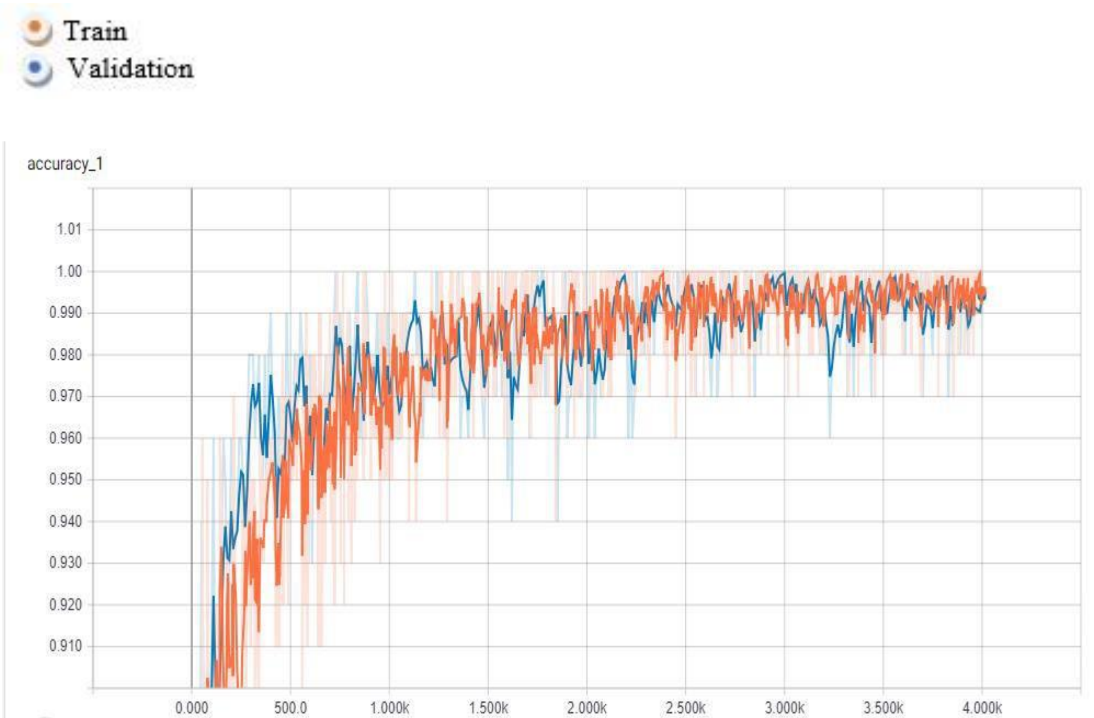

Figure 3.6: Inception v3 (accuracy graph)

From the above graph (Figure 3.6) we can see that the accuracy is very low for very initial step and it’s gradually increases for later steps. At step 4000 the accuracy is so high (almost 100%).

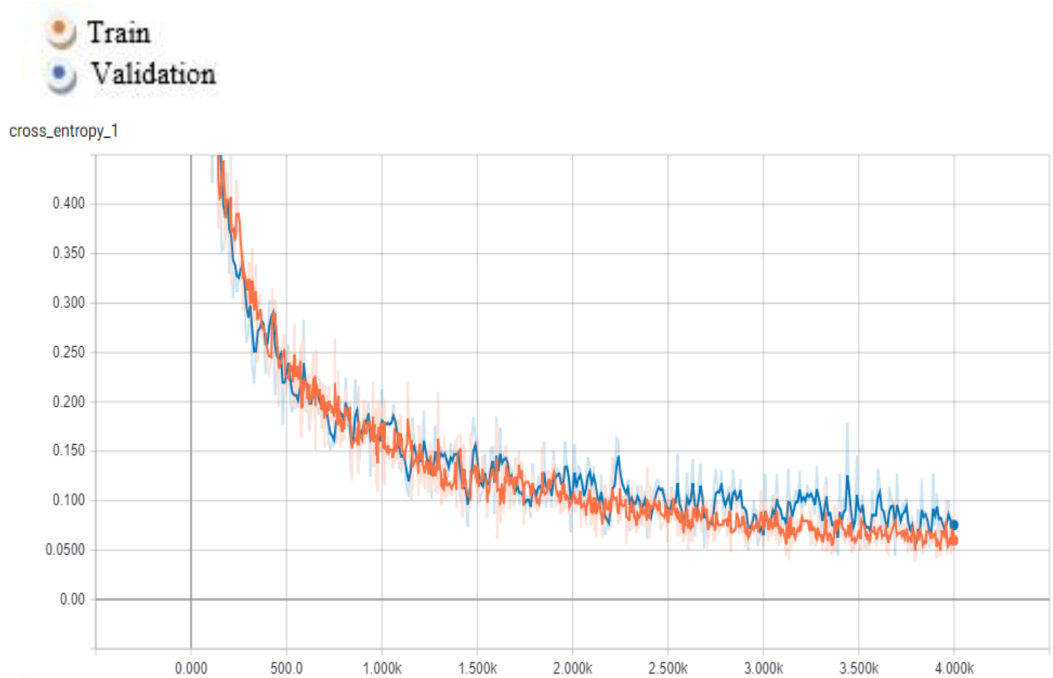

Figure 3.7: Inception v3 (cross-entropy graph)

In this graph (Figure 3.7) we see the cross-entropy loss or log loss value is high (>0.40) for very initial step and it's gradually decreases for later steps. At step 4000 the log loss value is so low (0.05).

So, Inception v3 performed this classification task very well.

3.5.2 MobileNet

We train 2900 images (580 images per class) with MobileNet model. Validation set created from these 2900 images automatically. The number of training steps is 500, the pixel value of input images is 224px and the model size is 0.50.

From Tensoboard we get the graph of accuracy and cross-entropy. Figure 3.8 shows the accuracy graph and Figure 3.9 shows the cross entropy graph for MobileNet_0.50_224 model.

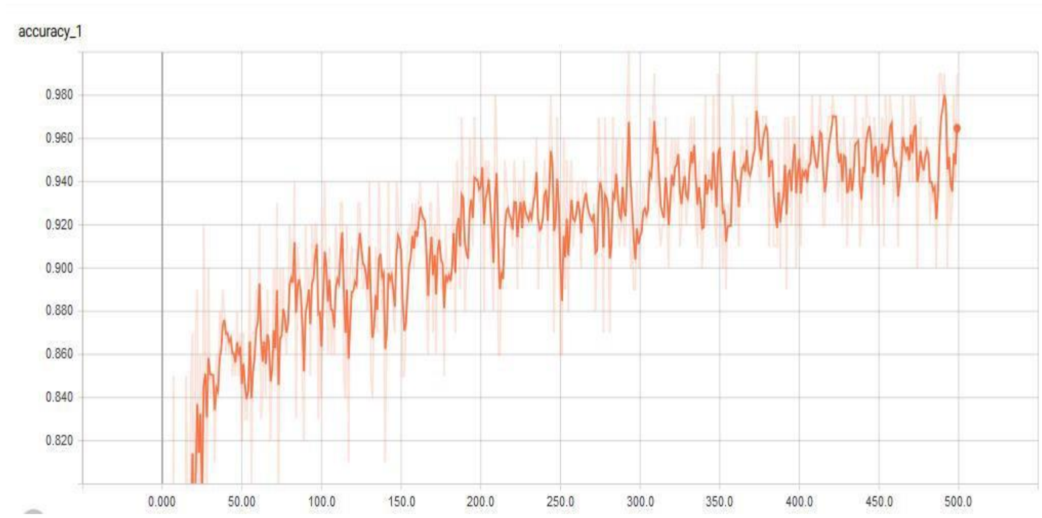

Figure 3.8: MobileNet (accuracy graph)

Here (Figure 3.8) we can see that the accuracy is very low for very initial step and it's gradually increases for later steps. At step 500 the accuracy is so high (almost 98%).

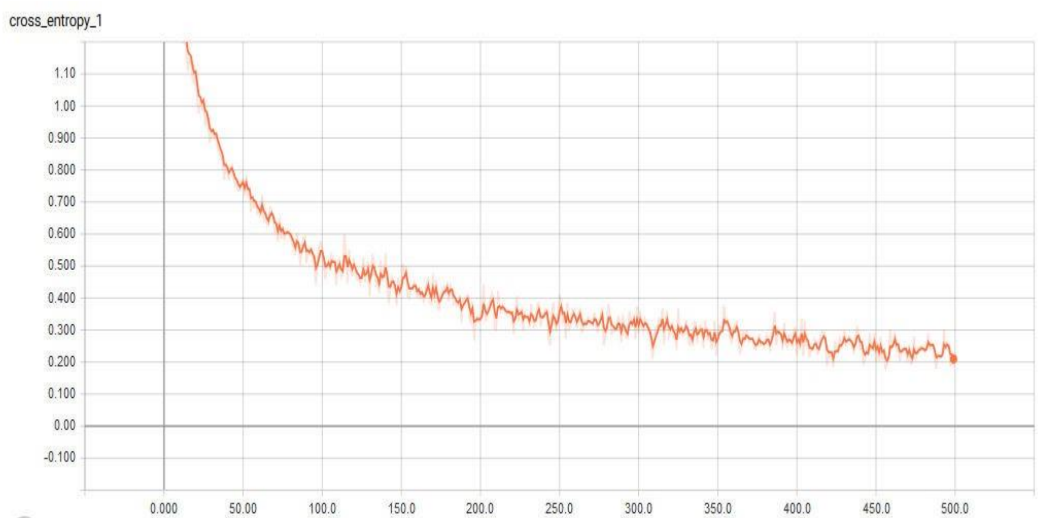

Figure 3.9: MobileNet (cross-entropy graph)

Here (Figure 3.9) the value of cross-entropy loss or log loss is high (> 1.10) for very initial step and it's gradually decreases for later steps. At step 500 the log loss value is so low (0.20).

So, the performance of MobileNet architecture is so good for this classification task.

3.5.3 SP-Net

We train our proposed model SP-Net with 2900 images (580 images per class). For validation we used 113 images (Badminton (22), Basketball (18), Swimming (27), Table Tennis (22), Water Polo (24)). The number of validation steps is 1000, the number of epoch is 30. We get 97% training accuracy and 98% validation accuracy.

We get the graph of accuracy and cross-entropy or loss for both train and validation. Figure 3.10 shows the train accuracy graph, Figure 3.11 shows the train cross entropy graph, Figure 3.12 shows the validation accuracy graph and Figure 3.13 shows the validation cross entropy graph for SP-Net model.

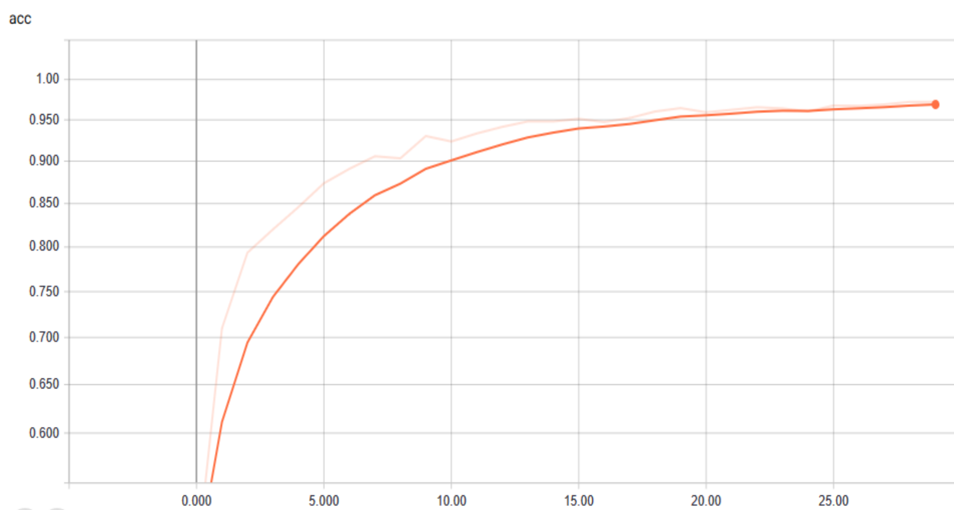

Figure 3.10: SP-Net (train accuracy graph)

From the above graph (Figure 3.10) we can see that the accuracy is very low for very initial step/epochs and it's gradually increases for later steps/epochs and for very last step/epochs the accuracy is so high (almost 98%).

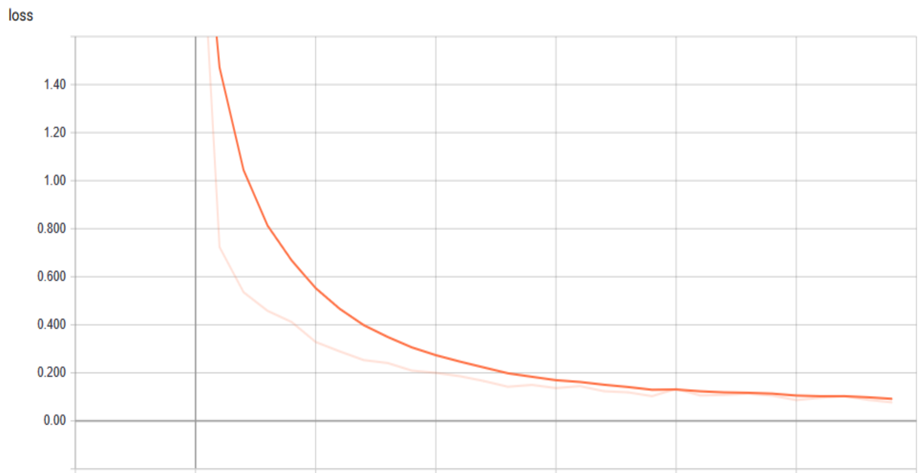

Figure 3.11: SP-Net (train cross-entropy graph)

From the above graph (Figure 3.11) we can see that the value of cross-entropy loss is high (>1.40) for very initial step/epochs and it's gradually decreases for later steps/epochs and for very last step/epochs log loss value is so low (0.10).

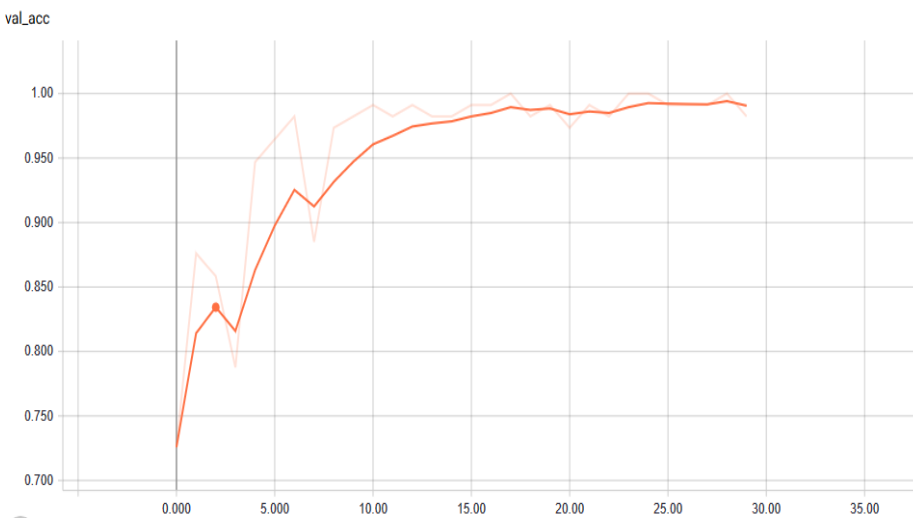

Figure 3.12: SP-Net (validation accuracy graph)

From the above graph (Figure 3.12) we can see that the accuracy is very low for very initial step/epochs and it's gradually increases for later steps/epochs and for very last step/epochs the accuracy is so high (almost 100%).

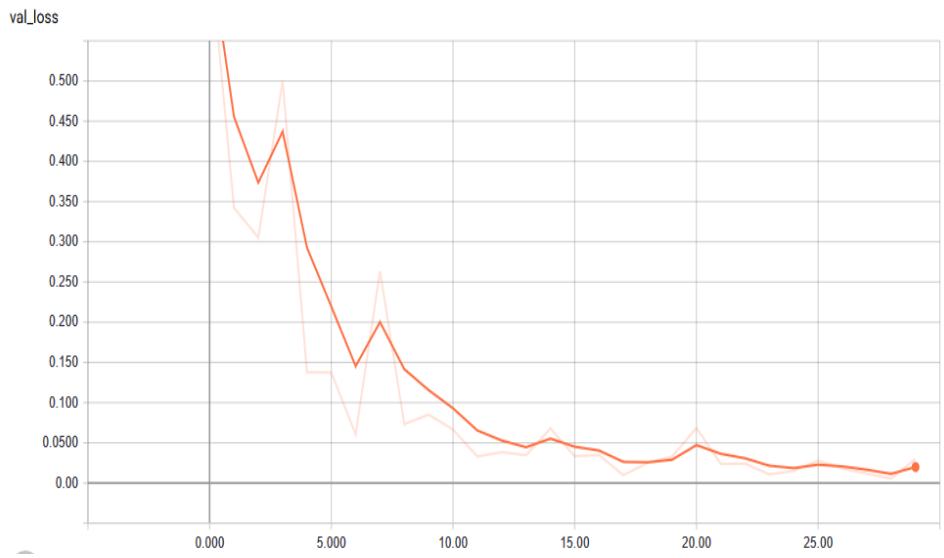

Figure 3.13: SP-Net (validation cross-entropy graph)

Here (Figure 3.13) the cross-entropy loss or log loss value is high (>0.50) for very initial step/epochs and it's gradually decreases for later steps/epochs and for very last step/epochs log loss value is so low (<0.05).

So, the performance of SP-Net architecture is so good for this classification task.

3.6 Test Model

We retrain Inception v3 and MobileNet architecture with a train dataset of 2900 images and train our proposed model SP-Net with the same dataset and then we test these three architecture with our test dataset which contains 150 images (30 images per class).

3.6.1 Inception v3

After completion the testing of Inception v3 with 150 images we get 94% test accuracy. The number of correctly classified images is 129 and incorrectly classified images is 21.

3.6.2 MobileNet

After testing with 150 images we get 90% test accuracy for MobileNet architecture. The number of correctly classified images is 113 and incorrectly classified images is 37.

3.6.3 SP-Net

For our proposed architecture SP-Net we get 98% test accuracy after the completion of testing with 150 images. Only 8 images are incorrectly classified and the number of correctly classified images is 142.

CHAPTER 4

Experimental Results and Discussion

4.1 Introduction

In this section, we are talking about our test results and performance evaluation. We test these three architectures (Inception v3, MobileNet and SP-Net) with 150 images (30 images per class) and get good test accuracy for all these three models. In Experimental Results section, we will discuss the confusion matrix, test accuracy and performance measure. In performance comparison section we will compare the performance of these three classification models. In receiver operating characteristic (ROC) curve section we will show the ROC curve for these three models and describe ROC AUC. In summary section we will give the overall summary of our experimental results.

4.2 Experimental Results

After testing each model with 150 images we get the confusion matrix of these three models. Confusion matrix of Inception v3 which we have shown in table [4.1]. Confusion matrix of MobileNet which we have shown in table [4.2]. Confusion matrix of SP-Net which we have shown in table [4.3]. Confusion matrix is also known as performance matrix because it gives the performance measure of any classification model and we get the performance measure of Inception v3, MobileNet and SP-Net from these three performance matrix. TP (true positive), FP (false positive), FN (false negative) and TN (true negative) these four terms can easily define a confusion matrix.

TP (true positive): when our predicted class and actual class both are true.

FP (false positive): when our predicted class is true and our actual class is false.

FN (false negative): when our predicted class is false and our actual class is true.

TN (true negative): when our predicted class and actual class both are false.

Table 4.1: Confusion Matrix of Inception v3

	Predicted					
		badminton	**basketball**	**swimming**	**table tennis**	**water polo**
Actual	**badminton**	25	0	0	5	0
	basketball	1	28	0	1	0
	swimming	0	1	24	0	5
	table tennis	4	0	0	26	0
	water polo	0	0	4	0	26

Here (Table 4.1), for class badminton TP=25, FP=5, FN=5 and TN=115, for class basketball TP=28, FP=1, FN=2 and TN=119, for class swimming TP=24, FP=4, FN=6 and TN=116, for class table tennis TP=26, FP=6, FN=4 and TN=114, for class water polo TP=26, FP=5, FN=4 and TN=115. So, the total number of correctly classified instances is 25+28+24+26+26=129.

Table 4.2: Confusion Matrix of MobileNet

		Predicted				
		badminton	**basketball**	**swimming**	**table tennis**	**water polo**
Actual	**badminton**	22	2	0	6	0
	basketball	0	26	4	0	0
	swimming	0	1	27	1	1
	table tennis	9	0	1	20	0
	water polo	0	1	11	0	18

Here (Table 4.2), for class badminton TP=22, FP=9, FN=8 and TN=111, for class basketball TP=26, FP=14, FN=4 and TN=116, for class swimming TP=27, FP=16, FN=3 and TN=104, for class table tennis TP=20, FP=7, FN=10 and TN=113, for class water polo TP=18, FP=1, FN=12 and TN=119. So, the total number of correctly classified instances is 22+26+27+20+18=113.

Table 4.3: Confusion Matrix of SP-Net

	Predicted					
		badminton	**basketball**	**swimming**	**table tennis**	**water polo**
Actual	**badminton**	28	2	0	0	0
	basketball	0	30	0	0	0
	swimming	0	0	28	0	2
	table tennis	0	0	0	30	0
	water polo	0	0	4	0	26

Here (Table 4.3), for class badminton TP=28, FP=0, FN=2 and TN=120, for class basketball TP=30, FP=2, FN=0 and TN=118, for class swimming TP=28, FP=4, FN=2 and TN=116, for class table tennis TP=30, FP=0, FN=0 and TN=120, for class water polo TP=26, FP=2, FN=4 and TN=118. So, the total number of correctly classified instances is 28+30+28+30+26=142.

Accuracy: accuracy is a term which is directly related to the performance and it is calculated as the total number of correctly predicted instances divided by the total number of instances present in a dataset.

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

After the completion of testing with 150 images we get 94% test accuracy for Inception v3, macro average precision 0.862, macro average recall 0.860 and macro average F1 score 0.860. From these 5 classes, get the best accuracy for basketball 98%, then for water polo 94% and 93% accuracy for rest three classes. Figure 4.1 shows the test accuracy, precision, recall and F-measure of the classes using Inception v3 model.

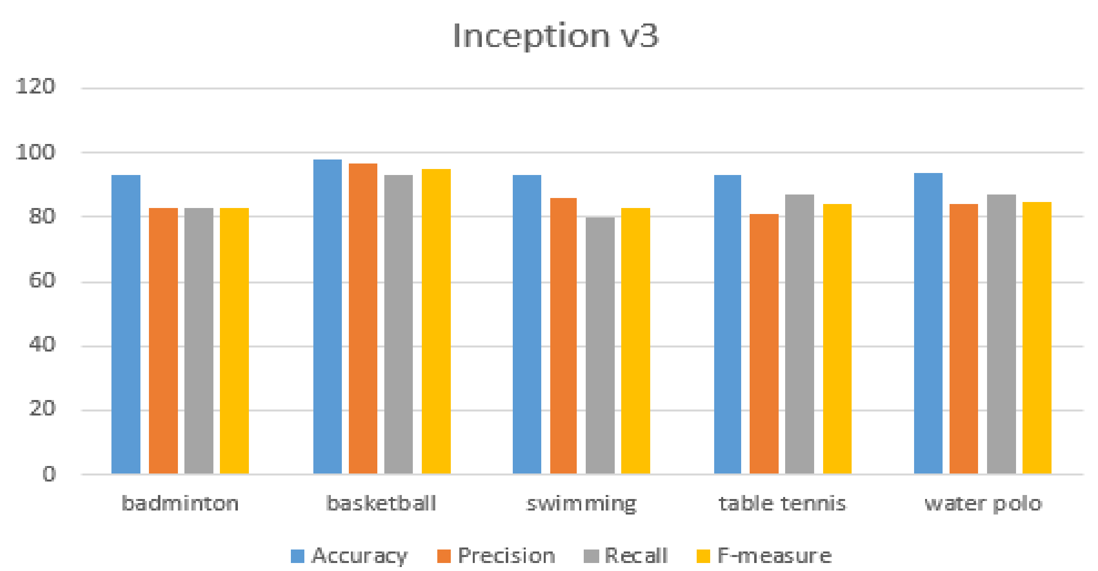

Figure 4.1: Inception v3 (Accuracy, Precision, Recall, F-measure)

We Get 90% test accuracy for MobileNet, macro average precision 0.779, macro average recall 0.753 and macro average F1 score 0.753. From these 5 classes, get the best accuracy for basketball 95%, then for water polo 91%, for badminton and table tennis 89% and for swimming 87%. Figure 4.2 shows the test accuracy, precision, recall and F-measure of the classes using MobileNet model.

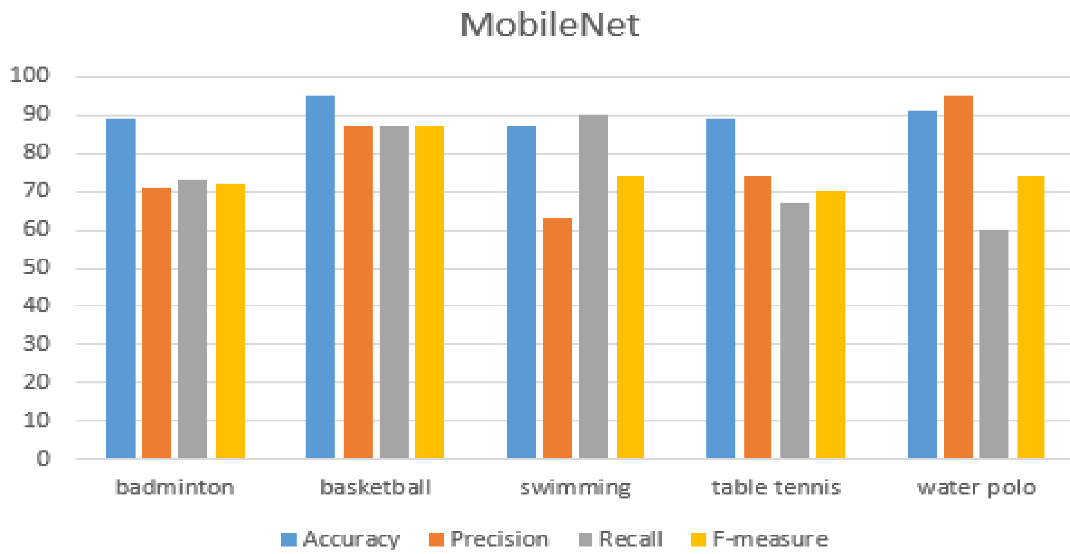

Figure 4.2: MobileNet (Accuracy, Precision, Recall, F-measure)

We Get 98% test accuracy for SP-Net architecture, macro average precision 0.95, macro average recall 0.95 and macro average F1 score 0.95. From these 5 classes, get the best accuracy for table tennis 100%, then for badminton and basketball 99% and for water polo and swimming 96%. Figure 4.3 shows the test accuracy, precision, recall and F-measure of the classes using SP-Net model.

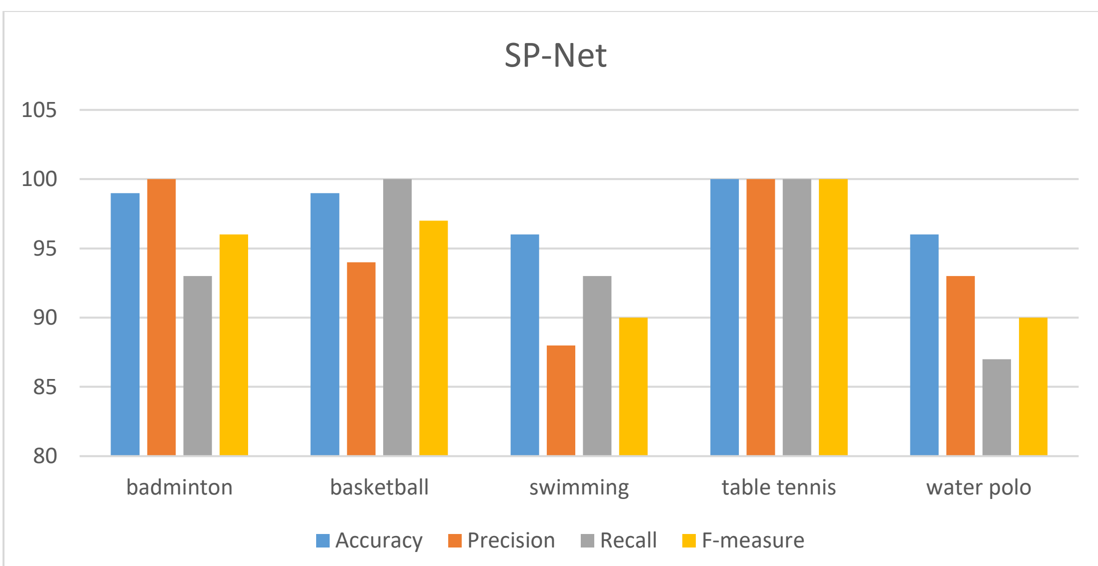

Figure 4.3: SP-Net (Accuracy, Precision, Recall, F-measure)

4.3 Performance Comparison

We can compare the macro average accuracy, precision, recall, F-measure of Inception v3, MobileNet and SP-Net and from this comparison it is clear that the performance evaluation of SP-Net architecture is better than Inception v3 and MobileNet architecture. The performance of Inception v3 model is also so good and it has better performance than MobileNet. Figure 4.4 shows the Performance Comparison between Inception v3, MobileNet and SP-Net.

For evaluating the test performance receiver operating characteristic (ROC) curve of Inception v3, MobileNet and SP-Net architectures are generated.

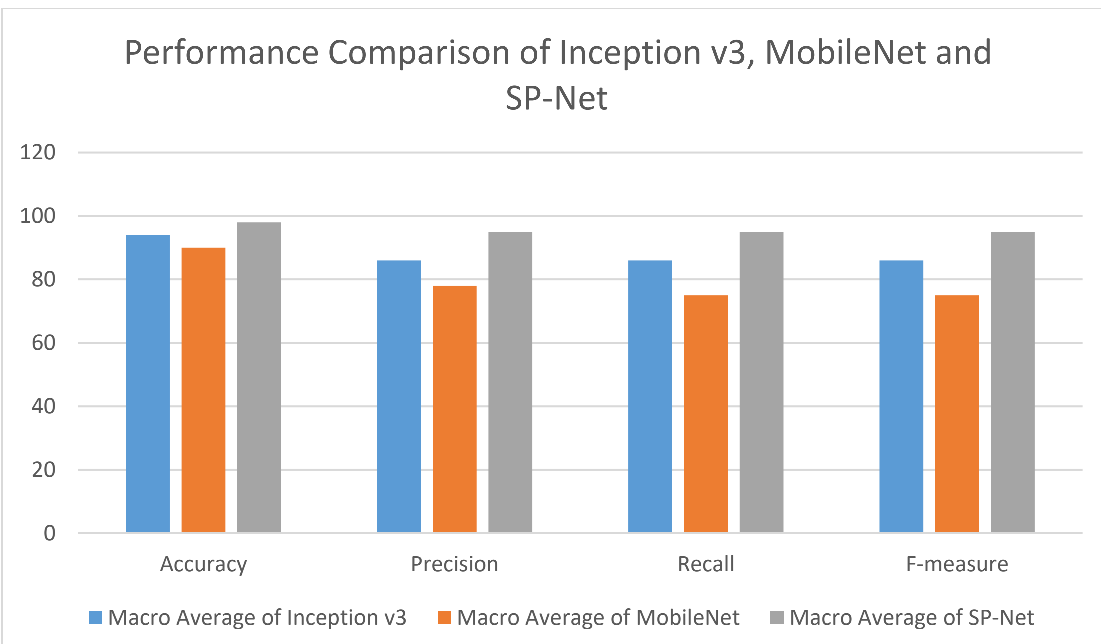

Figure 4.4: Performance Comparison between these three models

4.4 Receiver Operating Characteristic (ROC) Curve

For evaluating the test performance receiver operating characteristic (ROC) curve of Inception v3, MobileNet and SP-Net architectures are generated.

Receiver operating characteristic (ROC) curve shows the FPR (false positive rate) along X-axis and TPR (true positive rate) along Y-axis which can give the test performance measurement.

FPR (false positive rate): False positive rate is equivalent to 1-specificity. We can calculate the FPR by the given equation:

$$False\ positive\ rate\ (FPR) = \frac{FP}{FP + TN}$$

TPR (true positive rate): True positive rate is equal to sensitivity or recall. We can calculate the TPR by the given equation:

$$True\ positive\ rate\ (TPR) = \frac{TP}{TP + FN}$$

ROC AUC refers to the area under the curve. We can measure the accuracy by the area which is under the ROC curve. Figure 4.5 shows the ROC curve of Inception v3, Figure 4.6 shows the ROC curve of MobileNet, Figure 4.7 shows the ROC curve of SP-Net.

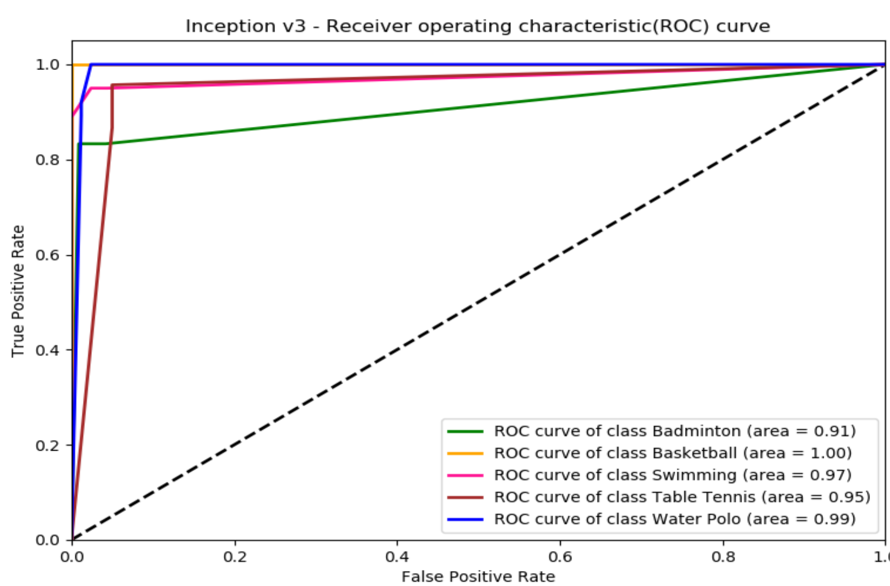

Figure 4.5: ROC curve of Inception v3

ROC curve of Inception v3 (Figure 4.5) provides 91% area for badminton, 100% area for basketball, 97% area for swimming, 95% area for table tennis and 99% area for water polo.

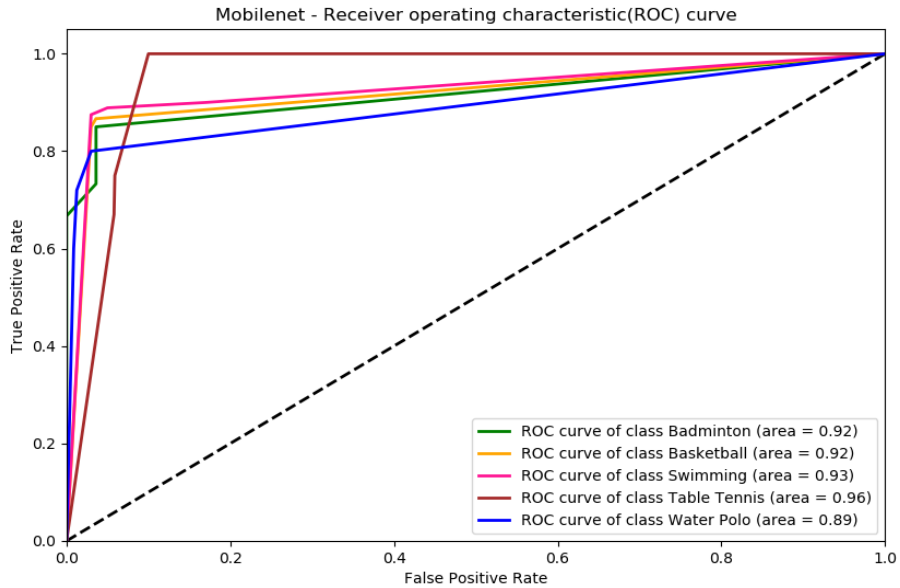

Figure 4.6: ROC curve of MobileNet

ROC curve of MobileNet (Figure 4.6) provides 92% area for badminton, 92% area for basketball, 93% area for swimming, 96% area for table tennis and 89% area for water polo.

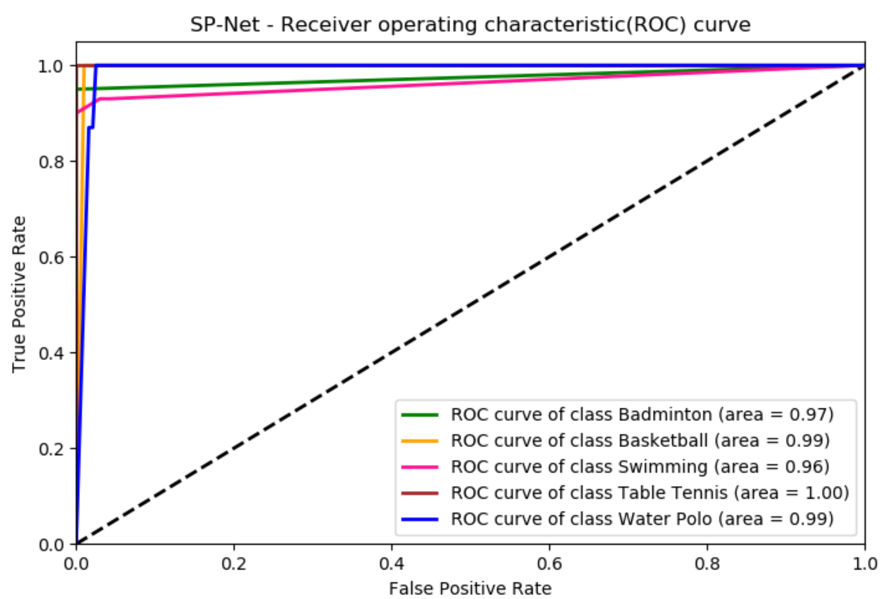

Figure 4.7: ROC curve of SP-Net

ROC curve of SP-Net (Figure 4.7) provides 97% area for badminton, 99% area for basketball, 96% area for swimming, 100% area for table tennis and 99% area for water polo.

Among these 5 classes SP-Net gives more ROC AUC (area under the curve) for class badminton and table tennis. For water polo SP-Net and Inception v3 both architecture gives 99% area which is greater than the area provided by MobileNet architecture (89%). Inception v3 provides more ROC AUC (area under the curve) for class basketball and swimming. On the other hand, the ROC AUC (area under the curve) of MobileNet of these five classes can't beat the ROC AUC of Inception v3 and SP-Net.Inception v3 gives better accuracy for 2 classes, SP-Net gives better accuracy for 2 classes and for one class Inception v3 and SP-Net provides the same area. So the correctly classification rate of Inception v3 and SP-Net is greater than the correctly classification rate of MobileNet and between SP-Net and Inception v3 the classification performance of SP-Net is best.

4.5 Summary

If we look at our experimental results, we can see that the test accuracy of SP-Net architecture is better than the test accuracy of Inception v3 and MobileNet architecture. From the graph of performance comparison (Figure 4.4) we can say that SP-Net provides the best result for each performance measure and between Inception v3 and MobileNet architecture, Inception v3 performed well. Finally, from the ROC curve evaluation we get the more ROC AUC (area under the curve) for SP-Net, then for Inception v3 and then for MobileNet architecture.

CHAPTER 5

Future Works and Conclusion

5.1 Future Works

In this paper, we classify only five olympic sports categories but a lot of sports events are arranged by olympic game so increasing the number of classes is important. Our proposed model gives us 98% test accuracy but if we increment the number of images in our dataset and use the high resolution images in dataset this accuracy will be increased. Therefore, our aim is to increase the number of classes and added more images in our dataset.

5.2 Conclusion

Throughout the paper, Inception v3, MobileNet, SP-Net all these three architectures are classified dataset appropriately. Only few cases these models are predicted wrong classes. But among these three classification model SP-Net gives the best result, it's correctly classified 142 images out of 150 and we get a test accuracy 98% where Inception v3 correctly classified 129 images out of 150 and we get a test accuracy 94% and MobileNet correctly classified 113 images out of 150 and we get a test accuracy 90%. So we can say that our proposed SP-Net architecture is more accurate than Inception v3 and MobileNet. Contrariwise, in ROC curve SP-Net gives more AUC for all classes, only for two (basketball and swimming) classes the ROC AUC of Inception v3 beat SP-Net. However, all these three (Inception v3, MobileNet and SP-Net) models are good for classification.

REFERENCES

[1] Martín Abadi, Ashish Agarwal, et al, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems", CoRR abs/1603.04467, 2016

[2] Victor Cordes, Lorne Olfman, "Sports Analytics: Predicting Athletic Performance with a Genetic Algorithm", Twenty-second Americas Conference on Information Systems, San Diego, 2016.

[3] Chan TYC, Cho JA, Novati DC, "Quantifying the contribution of NHL player types to team performance", Interfaces , 42(2), pp. 131–145, 2012.

[4] Faez Ahmed, Kalyanmoy Deb, Abhilash Jindal, "Multi-objective optimization and decision making approaches to cricket team selection", Applied Soft Computing, 13 (1), pp. 402–414, 2013.

[5] Biao Xu, "Prediction of Sports Performance based on Genetic Algorithm and Artificial Neural Network", International Journal of Digital Content Technology and its Applications(JDCTA), 6(22), pp. 141-149, 2012.

[6] Iztok Fister Jr., Grega Vrbanciˇ c, Lucija Brezo ˇ cnik, Vili Podgorelec and Iztok Fister, "SportyDataGen: An Online Generator of Endurance Sports Activity Collections", 29th CECIIS, Varaždin, Croatia, 2018.

[7] DeSarbo W, Madrigal R, "Exploring the demand aspects of sports consumption and fan avidity", Interfaces, 42(2), pp. 199–212, 2012.

[8] BJ Coleman, "Identifying the "players" in sports analytics research", Interfaces, 42(2), pp. 109–118, 2012.

[9] Learn about TensorFlow, available at << https://www.tensorflow.org/ >>, last accessed 07-09-2018 at 01:07am.

[10] Learn about Confusion Matrix, available at << https://medium.com/greyatom/performance-metrics-forclassification-problems-in-machine-learning-part-i-b085d432082b >>, last accessed 13-09-2018 at 10:22am.

[11] Learn about MobileNet, available at << https://medium.com/@sumit.arora/training-a-neural-network-using-mobilenets-in-tensorflow-for-image-classification-on-android-14f2792f64c1 >>, last accessed 15-09-2018 at 9:23pm.

[12] Learn about Inception v3, available at << https://becominghuman.ai/transfer-learning-retraininginception-v3-for-custom-image-classification-2820f653c557 >>, last accessed 29-09-2018 at 12:03am.

[13] Learn about keras, available at << https://towardsdatascience.com/building-a-deep-learning-model-using-keras-1548ca149d37 >>, last accessed 12-10-2018 at 07:57pm.