**Table Of Contents**

|  |  |  |  |
| --- | --- | --- | --- |
| INDEX | | | |
| CHAPTER NO | SECTION NO | TOPIC | PAGE NO |
|  |  | ABSTRACT |  |
| 1. | 1.1  1.2  1.3 | INTRODUCTION  1. Motivation  2. Purpose and Problem Definition  3. Scope |  |
| 2. | 2.1  2.2 | LITERATURE SURVEY  Issues in current system  Benefits of proposed system |  |
| 3. | 3.1  3.2 | METHODOLOGY/APPROACH  Module description  Hardware and software requirements |  |
| 4. |  | Details of Work accomplished/ Project Results |  |
| 5. |  | Benefits of Proposed Work |  |
| 6. |  | Conclusion |  |
| 7. |  | References |  |
| 8. |  | Acknowledgement |  |

**Abstract**

Agriculture contributes significantly to the economical-growth and provides livelihood as well but brings along with it the painstaking process of maintaining the produce. Farmers are often not able to diagnose the type of disease plaguing their crops and thus are unable to administer the correct treatments to the crops. Classification of these diseases would greatly aid the farmers to undertake the required treatments. This model utilizes image classification to distinguish the different types of disease that might manifest themselves in the produce of Eggplant (*Solanum melongena*). The model can classify three types of diseases viz. *Cercospora melongenae, Phomopsis vexans,* Yellow.

**Introduction**

Over the past decade, Computer Vision witnessed tremendous advancement thanks to Machine Learning and Artificial Intelligence. Technological advancements enabled us to feed larger datasets to neural networks and get better results with high accuracy, thus overcoming the "data plateau". Google Lens is a great example of how we have trained machines to not only detect multiple objects but also classify them. In this article, we see how the agricultural industry can benefit from such emerging technologies.

Agriculture is a prominent sector in India, contributing 17-18% to India’s GDP in 2018. This emphasizes the importance of agriculture to the growth of the country. To have a sustained production of crops it is imperative that the crops are healthy and any diseases that might potentially affect them are diagnosed. Classifying the type of disease that might have affected the crop requires expertise which is not readily available. In the absence of which there can be a loss of an entire yield of the crop as these infections can spread to other crops and adversely affect thus rendering them inedible. If the infection is not treated it could even affect the next batch of crops grown in the field. This calls for the need of a model which can classify the disease, once the type of disease is known the treatments can be easily found out. This model is trained to classify the diseases that can be contracted by Eggplant (*Solanum melongena*). The model can classify the image into the following:

1.*Cercospora melongenae*

2. *Phomopsis vexans*

3. Yellow

4. Brinjal

**1. *Cercospora melongenae***

*Cercospora melongenae* is a fungal plant pathogen that causes leaf spot on eggplant. It is a deuteromycete fungus that is primarily confined to eggplant species. This plant pathogen only attacks leaves of eggplants and not the fruit. *Cercospora melongeane* is primarily found in warm climates close to the equator. Symptoms begin to show on the underside of older, lower leaves first due to the proximity to the soil and move upwards; initial symptoms are small circular or oval chlorotic spots on leaves which develop light to dark brown centers but as the lesions expand, they may develop concentric zones. It is categorized as a fungal infection.

**2. *Phomopsis vexans***

The symptoms of *Phomopsis vexans* are circular brown spots with lighter centers on fruits along with that the infested leaves may turn yellow and drop from the plant. This might lead to dark cankers forming on stems. The symptoms visible on the fruit begin as pale sunken areas which are oval in shape, these areas grow bigger and become depressed. The lesions may coalesce to cover all or most of the fruit. It can be categorized as a fungal infection.

3. **Yellow**

Eggplants turning yellow may be indicative of more serious problems if the yellowing is on leaves. Spider mites and lace bugs can cause yellowing when they feed on plant leaves. As the insect populations increase, these damaged leaves may drop or dry up, leading to sunburn on fruits. Yellowing of leaves is often caused by care problems like irregular watering or a lack of nitrogen in the soil. Plants that aren’t getting enough water may initially wilt during midday, yellowing as the water stress increases. Eggplants that develop overall yellowing may need nitrogen.

4. **Brinjal**

This is the category that implies that the plant is not affected by any of the diseases mentioned above and signifies that the plant is healthy and needs no further treatment.

**Scope**

This model pertains to the classification of diseases in a crop. Though the model classifies the type of disease in the crop it can only the diseases that are seen in Eggplant (*Solanum melongena*). Due to a limited dataset, data augmentation has been implemented to increase the number of images available for a class for training the model. The model is able to classify the diseases from the types of *Cercospora melongenae, Phomopsis vexans,* Yellow.

**LITERATURE SURVEY**

**ISSUES AND CHALLENGES IN DISEASE PREDICTION**

**SYSTEM ON INDIAN BRINJAL CROP:**

Brinjal is very common tropical vegetables cultivated and grown in India. It is known by different names like Begun (Bengali), vange (Marathi), Kathiri (Tamil), baingan (Hindi), badane (Kannada), Vashuthana (Malayalam), waangum (Kashmiri), baigan (Oriya), venkaya (Telugu), ringna (Gujarati) and Peethabhala (Sanskrit) A large number of famers cultivate different types of Brinjal in their shape; size and color. These Brinjal fruits are used in vegetable curries and a variety of dishes are prepared out of Brinjal. The main important issues and challenges on Indian Brinjal plant which we need to focus and work more on it are: Monitoring Brinjal plant growth periodically Predicting plant disease based on variations on leaf and on Brinjal both well in advance Recommending appropriate solution for predicted disease in reliable way Increasing the quality and hygienic production required to serve the society.

**BENEFITS:**

In this direction, our research focus is to resolve some of the issues related to Machine Learning and Image Processing so that it is useful to sense the Brinjal plant periodically and is to get the current conditions of Brinjal crop in reliable way. Also this information will be helpful to get info about prediction decease related data & suggestion to use appropriate pesticides or organic materials for the same.

In specific, the research objectives are as follows.

* Intelligent prediction of disease on Brinjal crop based on symptoms of leaf and Brinjal in reliable way
* Recommending appropriate solution in terms of pesticides or organic material for the same disease predicted earlier designing of human friendly system which monitors plants growth by using machine learning techniques. To publish the work for information dissemination to the computer or agriculture researchers and different kind of industries

**Theory**

One of the most challenging aspects of training any image classifier by neural networks is the amount of data it requires. Conventionally, thousands of positive training set images are to be present if you want your model to be accurate and with their individual bounding boxes files(.XML). This is where Transfer Learning comes into the picture. In essence, transfer learning can be signified as how pre-trained models can be used to solve new problems. In this case the crop disease classification problem can be solved by using Tensorflow's pre-trained models available at <https://www.tensorflow.org/resources/models-datasets>.

These models are nothing but convolutional neural networks trained on millions of images varying in multiple classes. Convolutional neural networks aka ConvNets can be visualised as having

1.) Input Layer

2.) Hidden Layer

3.) Output Layer

Input layer is where we input our dataset (with labels) and for each training example we keep modifying the parameters of hidden layer unless we get the correct corresponding output (or label) at the output layer. The modification of hidden layer is basically changing "weights" of the activation function present in each node.

**Methodology**

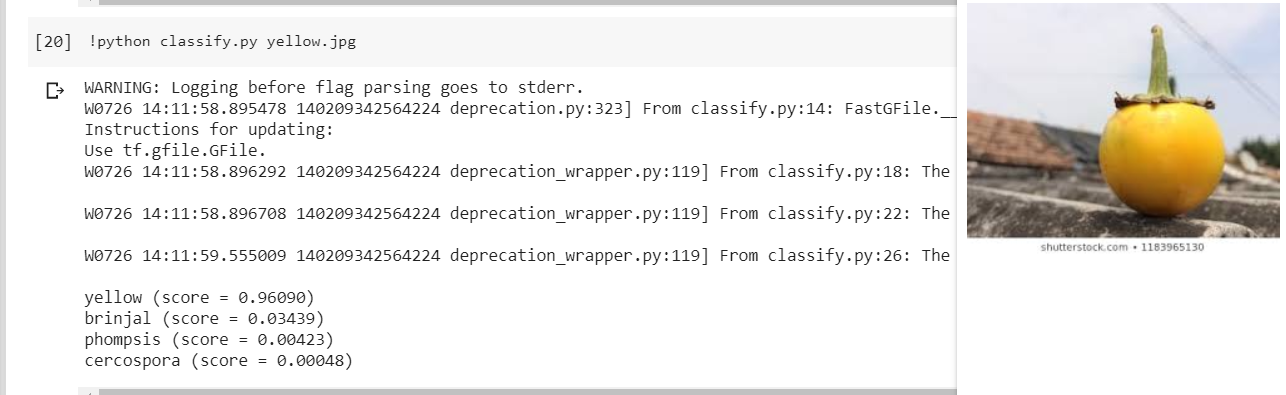
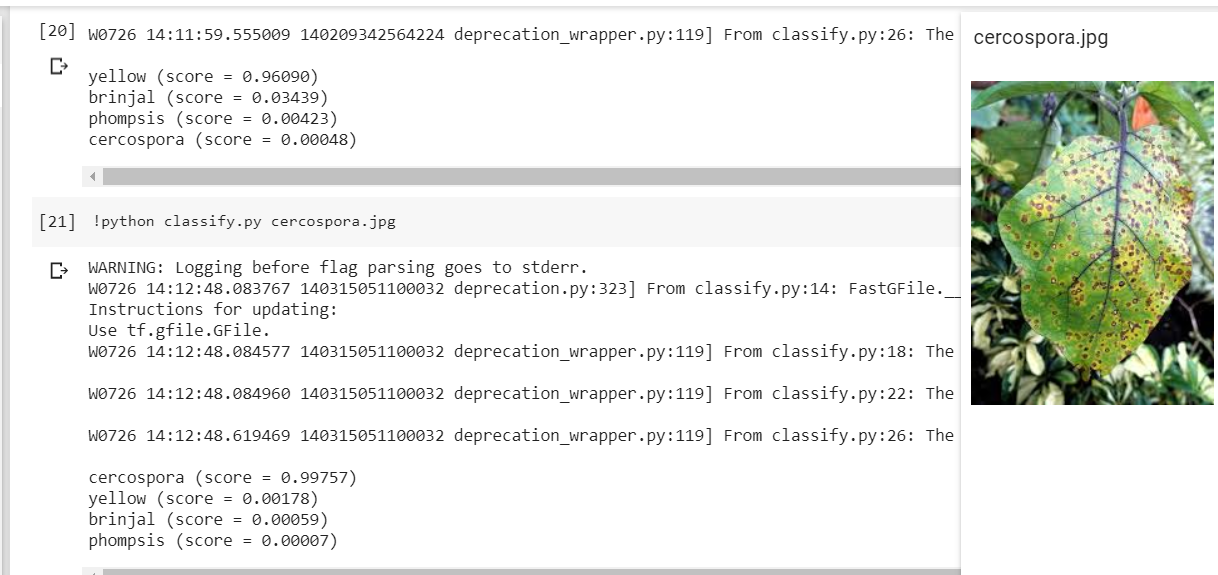
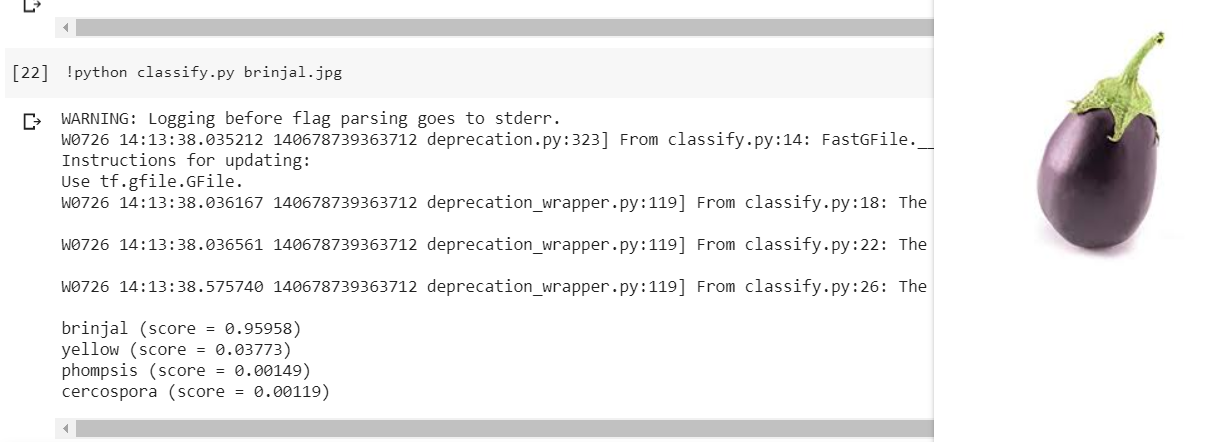
In this article, we use the Inception V3 model, 1st runner up for image classification in ILSVRC (ImageNet Large Scale Visual Recognition Competition). This model was trained on 1.2 million training images, 50,000 validation images and 100,000 testing images. Another challenging aspect of deep learning is the hardware it requires to train models. Enter Google Colab (https://colab.research.google.com). Gone are the days where you would wait for days for your model to "bake". Google Colab is a FREE Juptyer notebook environment provided by Google with GPUs and TPUs.

Let's take the example of brinjal crop. We'll see the three major diseases in this article:

1. Sunburnt Brinjal:- 155 images
2. Phomopsis blight:- 487 images
3. Cercospora leaf spot:- 374 images

For the healthy brinjal crop dataset is collected by employing Web Scraping and 60-70 images are collected. Through Image Augmentation i.e., angular rotation, noise addition, horizontal and vertical flip, the dataset is increased for the model's requirements. This was necessary as dataset for the diseased brinjal types were scarce and not recorded anywhere. The model is trained for 2000 training steps. On Google Colab's GPU (Tesla K80) it takes around 5-10 minutes for the training to complete. Accuracy measure for each class is as follows:-

1. Healthy Brinjal:- 95.95%
2. Phomopsis blight :- 95.57%
3. Cercospora leaf spot :- 99.75%
4. Sunburnt Brinjal :- 96.09%



Tensorflow also provides libraries to deploy your model on various other devices like Android, iOS, Raspberry Pi etc. which is Tensorflow Lite. A mobile app can be devised to enable farmers to better understand the types of crop diseases; how to control and treat them.

The second attempted method for this project, Haar cascading is an approach where by using a lot of positive and negative images a 'cascading function' is trained. Positives are where we gather 1000 images of the relevant data i.e. images with brinjal crop disease variations in this case Phomopsis blight , Cercospora leaf spot, Sunburnt Brinjal as well as Healthy Brinjal; and the negatives where we gather 1000 images of non-relevant data. The cascading function is trained based on these positives and negatives to make up a cascade (.xml) file.

Next big part of the brinjal crop detection to detect diseases specifications ROI-Region of Interest pooling is used which provides fixed-size feature maps from uniform inputs. ROI pooling takes every ROI from the input and takes a section of input feature map which corresponds to that ROI and converts that feature-map section into a fixed dimension map. The output fixed dimension of the ROI pooling solely depends on the layer parameters.

Reasons as to prefer the use of Tensorflow method in this topic over Haar cascading –

* **(.XML) File**: The creation of (.xml) file for the cascade classifier for the Haar classifier model is the most complex task as compared to Tensorflow model. Here, the composition for the images to be classified in positives and negatives is a very important step. It have been observed that for this article where not only the detection but also the N layers of ROI’s (Region of Interest) for differentiation amongst the diseases are very likely to be affected to a very high degree by the selection of the samples to be taken as positives, which indirectly affects the execution accuracy.

In addition, let’s take an example –

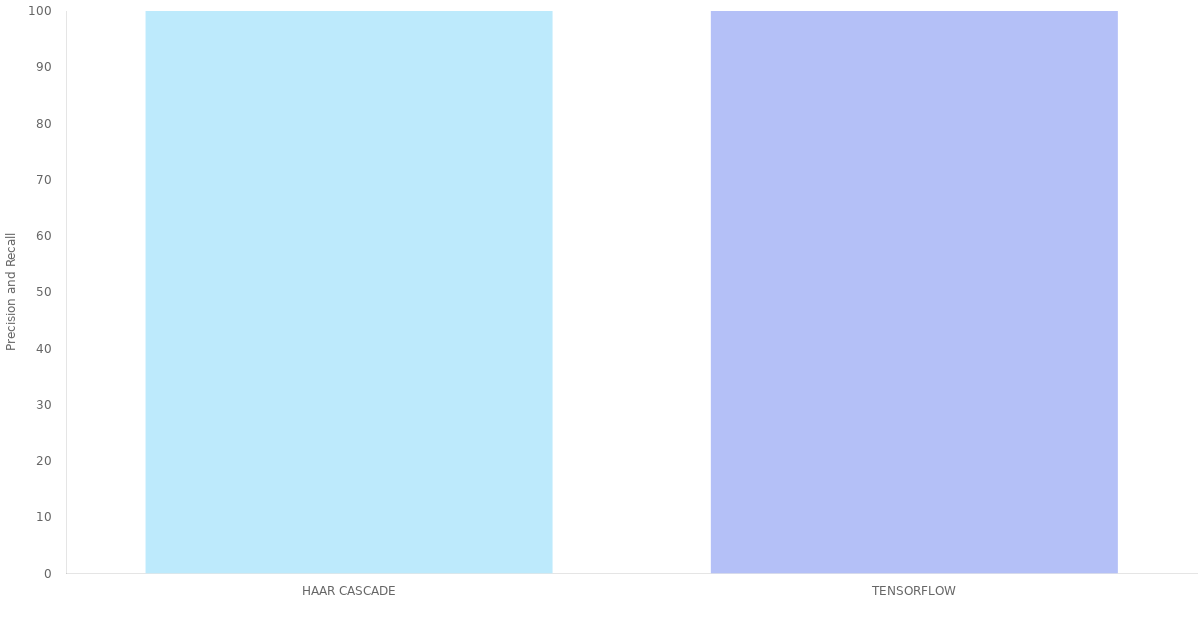
For training a set of 100 positive and 100 negative samples for creating a (.xml) file it make take around half an hour to a whole hour, which follows as for 1000 samples it takes a very high processing time to create one (.xml) cascade classifier file. Since, the accuracy is dependent on (.xml) file takes a great amount of time to be created will become an inconvenience when trying to improving the accuracy.

* **Processing Time:** Compared to Tensorflow, Haar classifier is been noted to tend to take more processing time by a large margin.
* **Scalability:** The Haar-like features as mentioned earlier cannot properly represent edge features for shape outlines. So if the task is somehow rich then we need a much richer set of features and a Tensorflow (CNN) is one such system capable of learning the relevant features that are necessary in order to increase performance. As in Tensorflow one will register an improvement with increased data and scale of the Tensorflow. This makes the Tensorflow scalable while the Haar cascade classifier is not that scalable because it is limited by what Haar-like features can do.
* **Consistency with tilting objects:** Tensorflow have a great consistency with tilting objects while Haar classifier are non-consistent in this case.

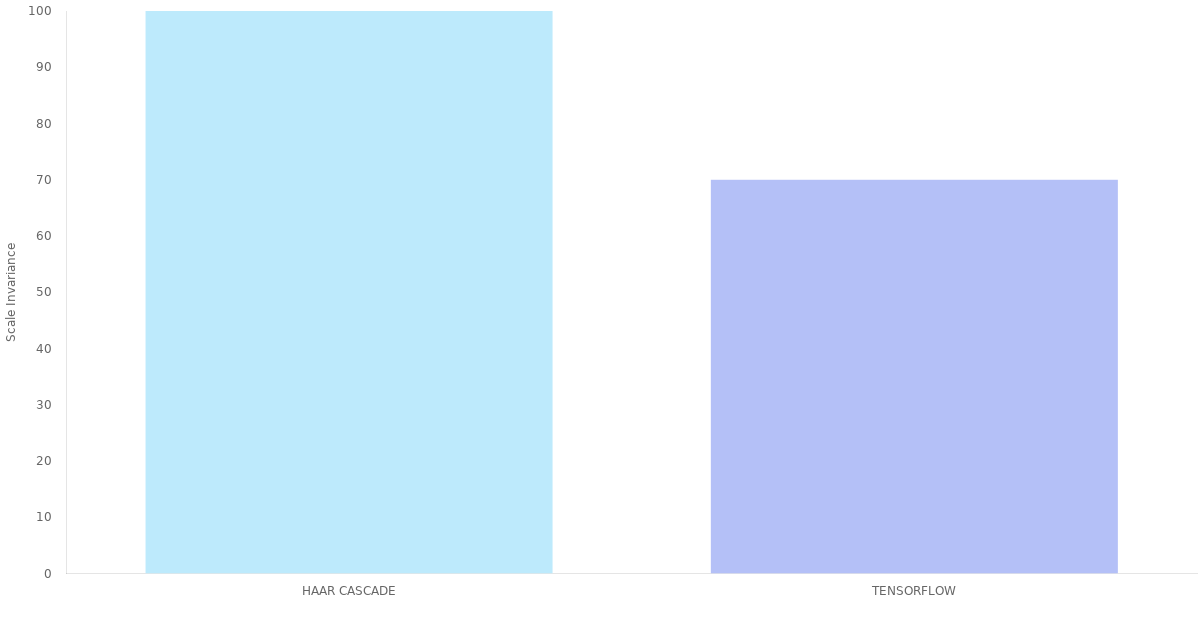
**Characteristics of Haar Cascade**

**& Tensorflow**

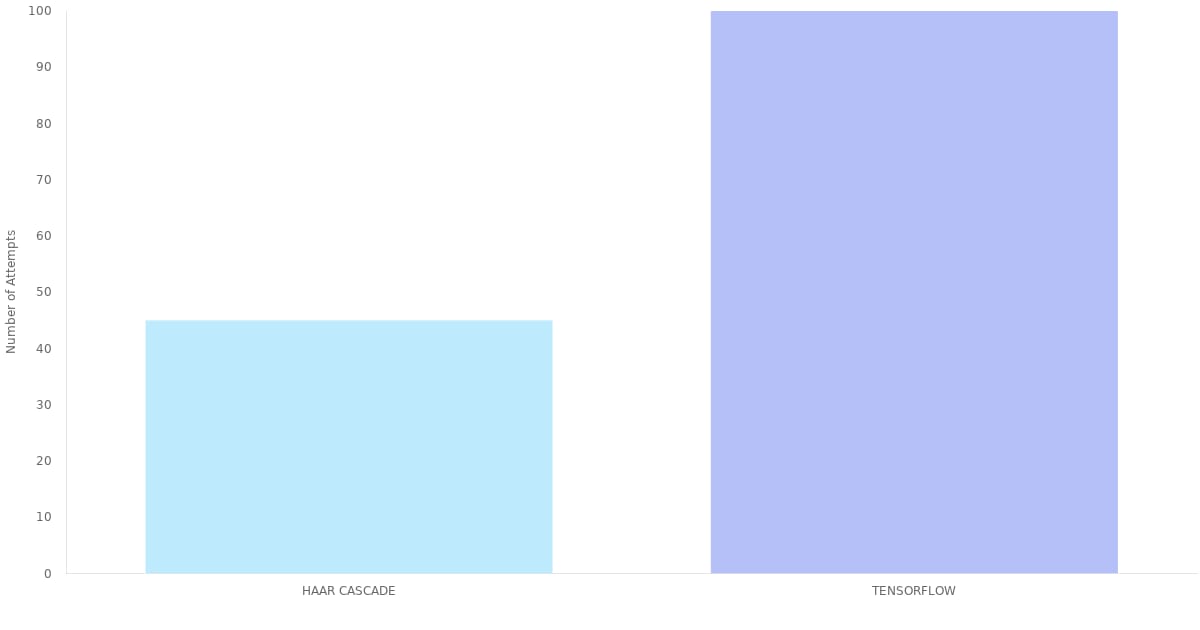
**Precision & Recall:**



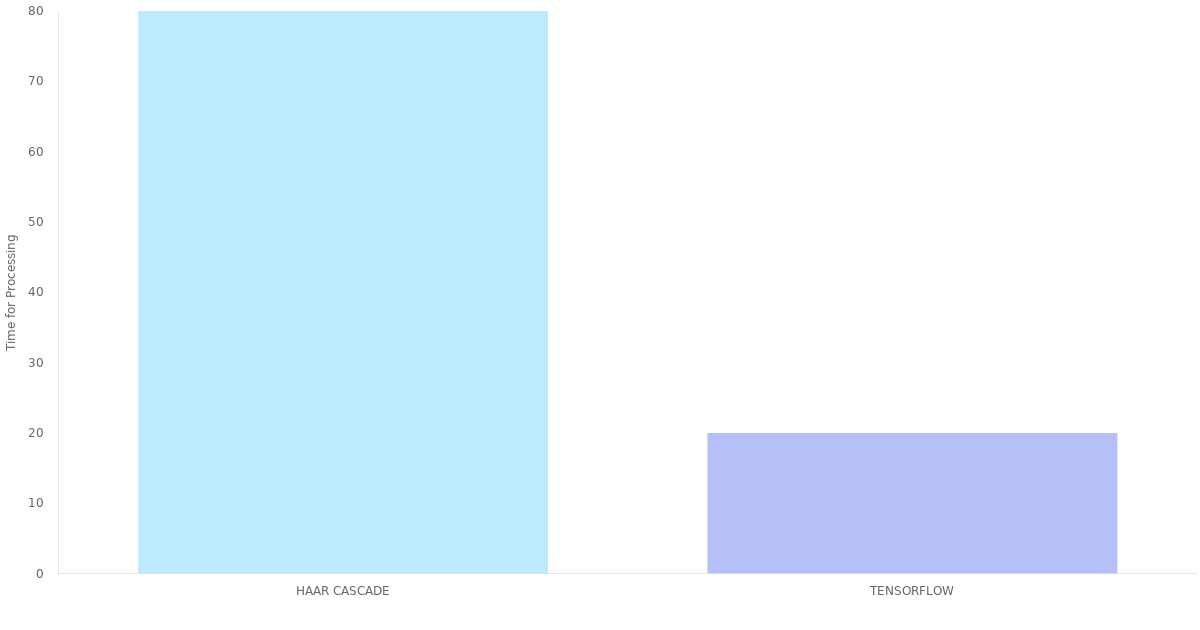
**Scale Invariance:**

****

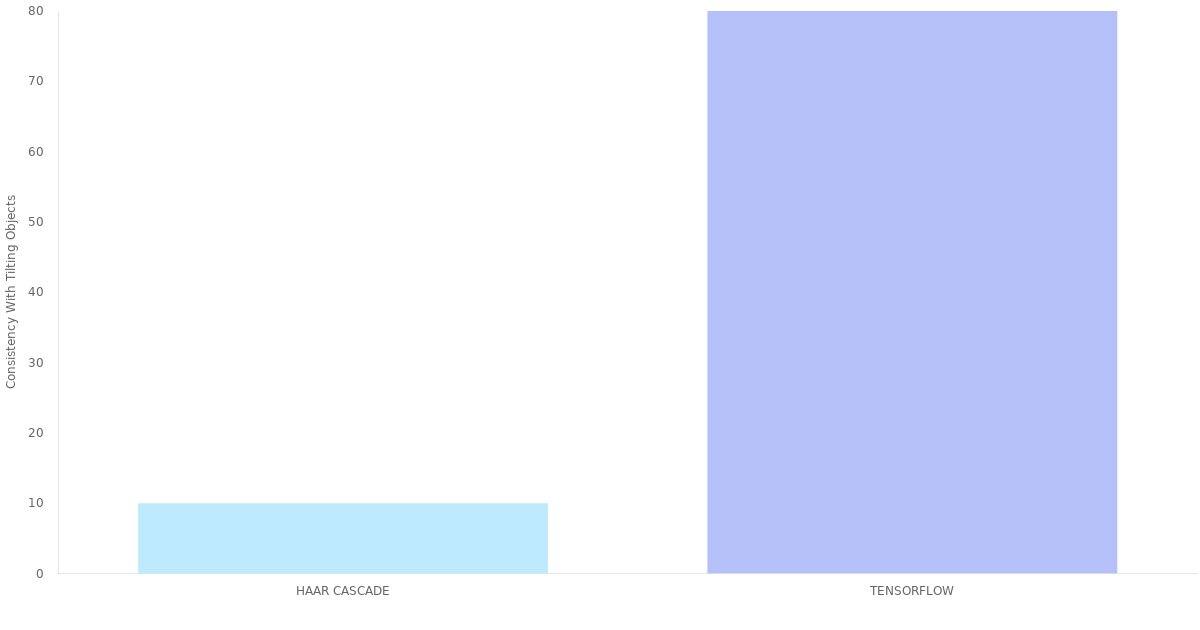
**Number of Attempts:**

****

**Processing Time:**

****

#### **Consistency with Tilting Objects:**

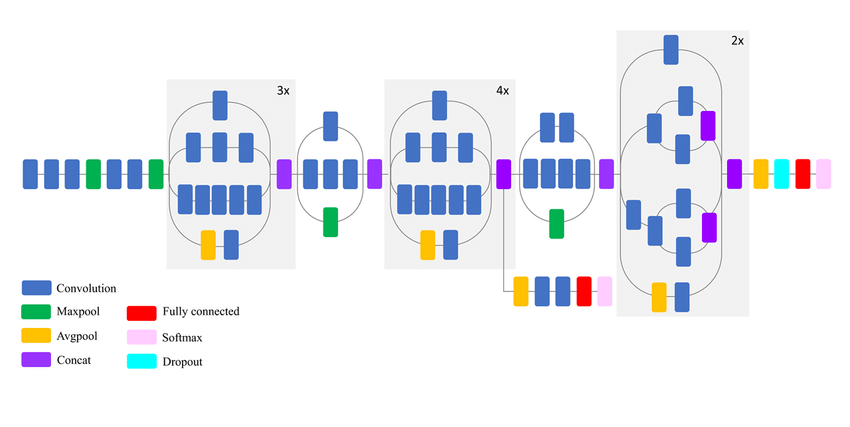
****

**Summing up:**

Tensorflow Model outperform Haar cascade classifiers in a number of cases. It is more than just reasonable to choose Tensorflow Model if you have enough time for training and your objects don’t scale much.

**Conclusion**

Image classifiers while previously required tremendous hardware support and data can now easily be concocted and deployed by

1. ConvNets for multi-class image classification.
2. Tensorflow and Tensorflow Lite for a robust framework to deploy models to mobile devices.



1. Google Colab for training ML/AI models terminating the need of GPU and CPU intensive hardware and seamless cloud deployment.



1. Image Augmentation for dataset manipulation and multiplication.



Emerging technologies mentioned in this project enables developers to create more use-cases and implement them in real-time to solve problems and immensely benefit society without computational impediments.

**CHAPTER No.**

**[chapeter title]**

**CHAPTER-7**

**References**

**Papers and websites referred:**

1. ["The Electronic Mall"](http://www.gsbrown.org/compuserve/electronic-mall-1984-04/). GS Brown. 30 April 2010*.* Retrieved 17 June2013.
2. Burgess, S; Sellitto, C; Karanasios, S (2009),Effective Web Presence Solutions for Small Businesses: Strategies and Successful Implementation, IGI Global
3. ["India Weighs FDI In E-Commerce"](http://blogs.wsj.com/indiarealtime/2014/01/08/india-weighs-fdi-in-e-commerce/?KEYWORDS=e-commerce). The Wall Street Journal India.
4. <https://en.wikipedia.org/wiki/E-commerce>
5. <https://www.shopify.in/tour/ecommerce-website>

**Acknowledgement**