KLE Society's

KLE Technological University

December 20, 2021

Course Project
On
Data Collection and preprocessing - Category (name board, traffic signal, furnitures)

Machine Learning(17ECSC306)

Submitted by

Name	Roll No	USN
Pradeep Chegur	526	01FE19BCS294
Praveen Thakkannavar	528	01FE19BCS298
Swagat Ingalagaon	529	01FE19BCS299
Naveen Doddamani	534	01FE19BCS305

Team Number:E3

Under The Guidance Of
Mr. Shashidhar V
Mr. Uday Nagraj Kulkarni

SCHOOL OF COMPUTER SCIENCE ENGINEERING
HUBLI-580031
Academic year 2021-2022

Abstract

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so.

The image classification is a classical problem of image processing, computer vision and machine learning fields.Computer Vision is the branch of the science of computers and software systems which can recognize as well as understand images and scenes. Computer Vision is consists of various aspects such as image recognition, object detection, image generation, image super-resolution and many more. Object detection namely name borad detection,traffic signal setection,furniture detection, face detection, vehicle detection, security systems and self-driving cars. In this project, we are using highly accurate object detection-algorithms and methods such as Inception-V4 and CNN models. Using these methods and algorithms, based on deep learning which is also based on machine learning require lots of mathematical and deep learning frameworks understanding by using dependencies such as TensorFlow, pytorch etc, we can detect each and every object in image by the area object in an highlighted rectangular boxes and identify each and every object and assign its tag to the object. This also includes the accuracy of each method for identifying objects.

ACKNOWLEDGEMENTS

The sense of contentment and elation that accompanies the successful completion of our project and its report would be incomplete without mentioning the names of the people who helped us in accomplishing this.

We are indebted to our guides Mr.Shashidhar V and Prof. Uday Kulkarni who was nothing but a constant source of enthusiasm and whose profound guidance, valuable suggestions, and beneficient direction was mainly responsible for us to complete this project.

We take this opportunity to express our deep sense of gratitude and sincere thanks to our HOD Dr. Meena S. M. for her tremendous source of inspiration and help in challenging our effort in the right direction.

Last but not least we like to thank all the staff members, teaching and non-teaching staff for helping us during the project.

Pradeep Chegur

Praveen Thakkannavar

Swagat Ingalagaon

Naveen Doddamani

Contents

1 Introduction **6**
1.1 Overview of the project 6
1.2 Objectives 7
1.3 Literature survey 7
1.3.1 A Review of Content Based Image Classification using Machine Learning Approach 7
1.3.2 Machine Learning Framework for Image Classification 7
1.3.3 Image classification using Deep Learning 8
1.3.4 An analysis of convolution neural network for image classification 8
1.4 Problem definition 9

2 Approach **10**
2.1 Dataset 10
2.2 Methodology 11

3 A Deep Walk Through The Pipeline **12**
3.1 Preprocessing 12
3.2 Model Design 12

4 Results **13**
4.1 Convolutional Neural Network 13
4.2 Inception V3 16
4.3 Comparison 17
4.3.1 Comparison 17
4.4 Application 17
4.4.1 User Interface 17
4.4.2 Predictions 18

5 Conclusion **19**

6 Future Scope **19**

7 References **20**

1 Introduction

1.1 Overview of the project

Classification is a systematic arrangement in groups and categories based on its features. Image classification came into existence for decreasing the gap between the computer vision and human vision by training the computer with the data. The image classification is achieved by differentiating the image into the prescribed category based on the content of the vision.In this paper, we explore the study of image classification using deep learning. Inception-v3 is a convolutional neural network architecture from the Inception family that makes several improvements including using Label Smoothing, Factorized 7 x 7 convolutions, and the use of an auxiliary classifer to propagate label information lower down the network (along with the use of batch normalization for layers in the sidehead). The conventional methods used for image classifying is part and piece of the field of artificial intelligence (AI) formally called as machine learning. The machine learning consists of feature extraction module that extracts the important features such as edges, textures etc and a classification module that classify based on the features extracted. The main limitation of machine learning is, while separating, it can only extract certain set of features on images and unable to extract differentiating features from the training set of data.

With billions of smartphones around the globe, wouldn't it be great if the smartphone could be turned into a image identification tool, recognizing the object from the image of it is pretty challenging. As a part of our course project, our aim is to develop learning model that can accurately recognize the category based on an image.

1.2 Objectives

1) To detect whether the given image is traffic signal,name board or furniture.
2) To build a model for detection of class of the object.
3) Build an application using the trained model to predict the class of a given image.
4) To incorporate the application in the edge device such as mobile.

1.3 Literature survey

1.3.1 A Review of Content Based Image Classification using Machine Learning Approach

Three classification methods are used to classify the images using machine learning technique .KNN, Decision Tree and support Vector machine. Evaluated performance of the algorithm using a general-purpose image database containing 500 JPEG images with size of 256*256 or 256*384 pixels from COREL photo gallery. Tested the performance of the model based on recall, precision and Accuracy. In terms of accuracy, precision, and recall SVM produce better result than other with a percentage of 97.14 96.11 94.10 respectively.

1.3.2 Machine Learning Framework for Image Classification

This paper is concerned with the different techniques and algorithms used in machine learning framework for image classification. Presented machine learning state-of-the-art applied to computer vision. Introduced the Bag of Features paradigm and highlighted the SURF as its technique for image features extraction and description. Experimentation's results proved that using SURF local feature extractor method and a SVM

(cubic SVM) training classifier performs best average accuracy. Measurements show that the image classification process performs better with SVM. It's reported that the Cubic SVM yields average accuracy which reaches 90The KNN techniques offer an average accuracy around 65Among the ensemble classifier trainers (2 last tested algorithms) the bagged trees achieves the best accuracy.

1.3.3 Image classification using Deep Learning

Four test images sea anemone, barometer, stethoscope and radio interferometer are chosen from the AlexNet database for testing and validation of image classification using deep learning. The convolutional neural network is used in AlexNet architecture for classification purpose. From the experiments, it is observed that the images are classified correctly even for the portion of the test images and shows the effectiveness of deep learning algorithm.

1.3.4 An analysis of convolution neural network for image classification

This paper presents an empirical analysis of the performance of popular convolutionl neural networks (CNNs) for identifying objects in real time video feeds. The main aim of our work is to understand the performance of the networks for static as well as live video feeds. The first step for the following is to perform transfer learning on the networks with image data sets. This is followed by checking the prediction rate of the same object on static images and real-time video feeds.Third important criteria for evaluating the performance was to check whether prediction accuracy varies across all CNNs chosen for the study. The main purpose of this paper was to find out the accuracy of the different networks

on same datasets and evaluating the consistency of prediction by each of these CNN.The results suggested that trained networks with transfer learning performed better than existing ones and showed higher rates of accuracy. The results also showed that use of more the number of layers, more will be the training and therefore, higher the rate of accuracy in prediction will be achieved.

1.4 Problem definition

From the given dataset of three category namely name board, traffic signals, furniture's and given an image of a object, our task is to predict the category that it belongs to.

2 Approach

2.1 Dataset

The dataset consists of about 500 images of different class like Name boards, Traffic signals and Furniture collected under controlled environmental conditions.We analyze 500 images of all three classes. Each category has set of images and we have predict to which category belongs to.

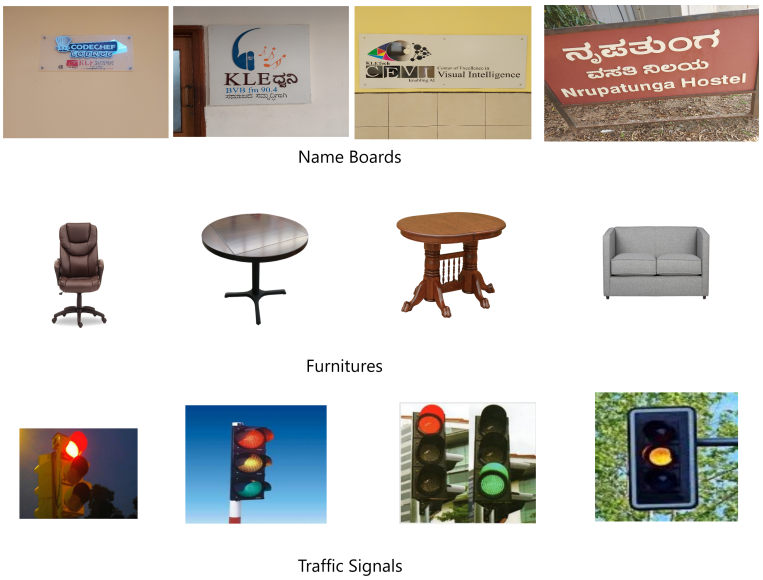

Figure 1: Images of a three different category

Figure.1: shows one example each from every crop-disease pair from the Plant Village dataset.In all the approaches described in this paper, we resize the images to 520 x 520 pixels, and we perform both the model optimization and predictions on these downscaled images.The dataset contains a total of 3 category.

2.2 Methodology

Our Methodology consists of 3 phases, Preprocessing, Model Training, App Building. The images in the dataset are Preprocessed so that the model will be robust and will be able to classify even in realtime.
Also the height and width of the images is kept 520 itself. The images are split into train and test datasets. In second phase i.e. model training the model is trained on the preprocessed data the model with best test accuracy is selected. These model weights are saved and used for building the app in the third phase.

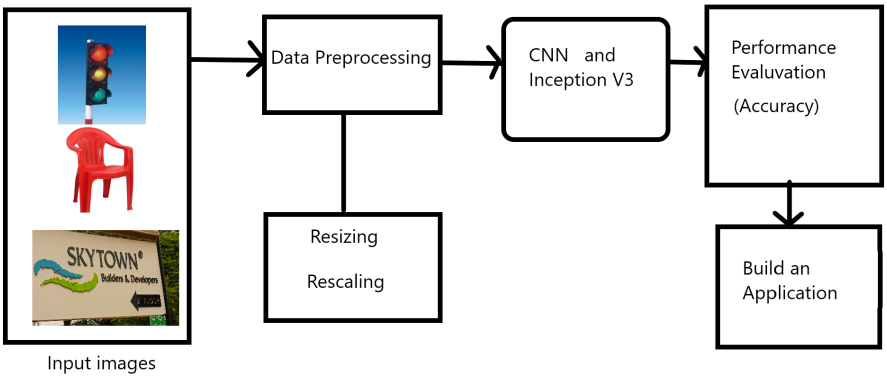

Figure 2: Methodology

3 A Deep Walk Through The Pipeline

3.1 Preprocessing

All the images were converted to an array. The images were then converted to a default size(520*520) and all the pixel intensities of the image were scaled down from [0,519] to the range [0,1].The data was then split into training(80%) and testing(20%) data.

3.2 Model Design

Our Model consists of 5 convolution blocks followed by relu activation functions also batch normalization blocks. The first Convolution block has 32 output filters with kernel size (3,3) and activation function applied is relu. MaxPooling is done with pool size (3,3) after Batch normalization and ends the block with dropout layer with p = 0.25. The Second Convolution block has convolution2d layer with 64 filters followed by relu activation and batch normalization. The third Convolution block contains convolution2d layers 128 filters with kernel size (3,3) and relu activation and batch Normalization and Maxpooling with pool size (3,3) and dropout as p = 0.25. The fourth convolution block has convolution2d layer with 128 filters, relu activation and batch Normalization.The fifth convolution block has convolution2d layer with 256 filters, relu activation and batch Normalization.After the 5 convolution blocks the tensors are flattened followed by a fully connected layer of 1024 units and a dropout layer with p = 0.5 and a fully connected layer with number of units equal to number of classes and softmax activation function.

4 Results

4.1 Convolutional Neural Network

```
Model: "sequential_3"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 rescaling_2 (Rescaling)     (None, 520, 520, 3)       0

 conv2d_13 (Conv2D)          (None, 520, 520, 32)      896

 batch_normalization_10 (Bat  (None, 520, 520, 32)     128
 chNormalization)

 max_pooling2d_13 (MaxPoolin  (None, 260, 260, 32)     0
 g2D)

 conv2d_14 (Conv2D)          (None, 260, 260, 64)      18496

 dropout_8 (Dropout)         (None, 260, 260, 64)      0

 batch_normalization_11 (Bat  (None, 260, 260, 64)     256
 chNormalization)

 max_pooling2d_14 (MaxPoolin  (None, 130, 130, 64)     0
 g2D)

 conv2d_15 (Conv2D)          (None, 130, 130, 64)      36928

 batch_normalization_12 (Bat  (None, 130, 130, 64)     256
 chNormalization)
```

```
conv2d_16 (Conv2D)           (None, 65, 65, 128)       73856

dropout_9 (Dropout)          (None, 65, 65, 128)       0

batch_normalization_13 (Bat  (None, 65, 65, 128)       512
chNormalization)

max_pooling2d_16 (MaxPoolin  (None, 33, 33, 128)       0
g2D)

conv2d_17 (Conv2D)           (None, 33, 33, 256)       295168

dropout_10 (Dropout)         (None, 33, 33, 256)       0

batch_normalization_14 (Bat  (None, 33, 33, 256)       1024
chNormalization)

max_pooling2d_17 (MaxPoolin  (None, 17, 17, 256)       0
g2D)

flatten_3 (Flatten)          (None, 73984)             0

dense_7 (Dense)              (None, 128)               9470080

dropout_11 (Dropout)         (None, 128)               0

dense_8 (Dense)              (None, 2)                 258

dense_9 (Dense)              (None, 3)                 9

=================================================================
Total params: 9,897,867
Trainable params: 9,896,779
Non-trainable params: 1,088
```

Total params: 9,897,867
Trainable params: 9,896,779
Non-trainable params: 1,088
Training accuracy: 73.25
Testing accuracy : 65.66

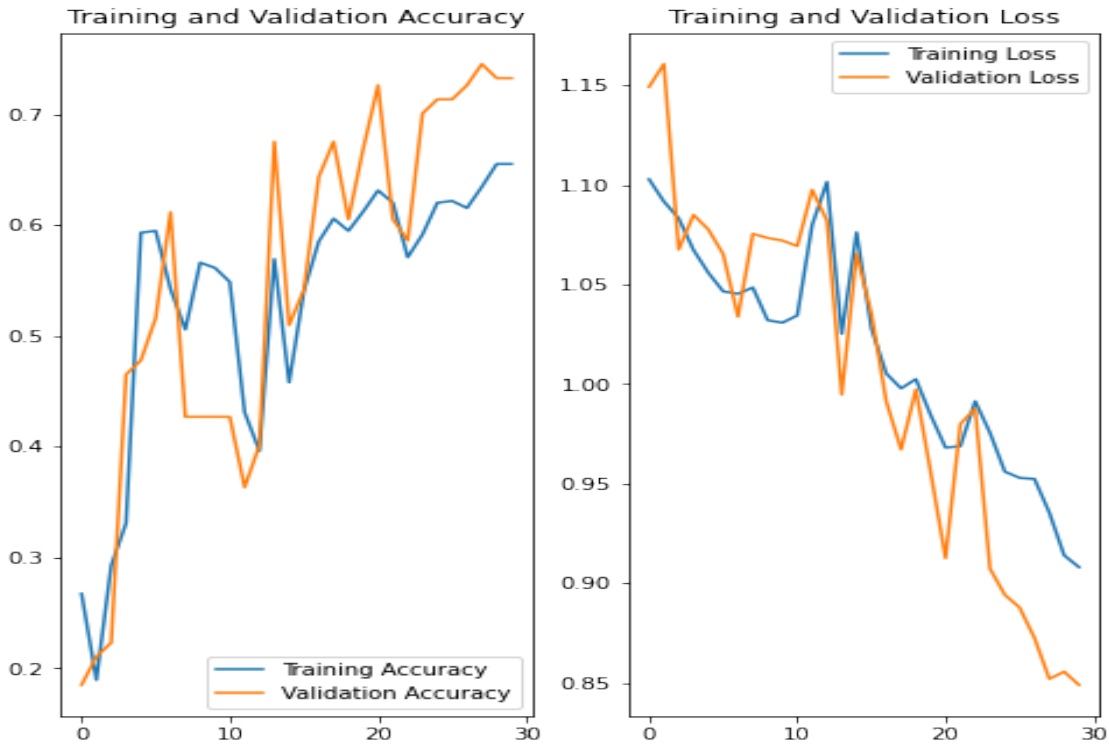

Figure 3: Epoch vs Loss

This is identified by a validation loss that is lower than the training loss. In this case, it indicates that the validation dataset may be easier for the model to predict than the training dataset.

4.2 Inception V3

Training accuracy: 99.03
Testing accuracy : 95.68

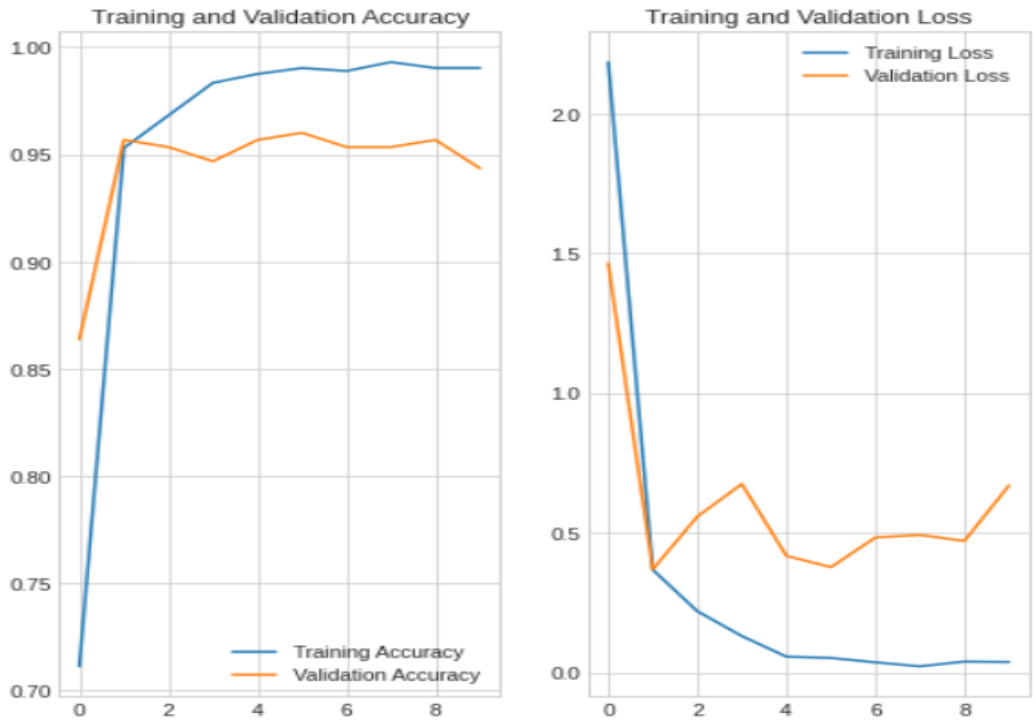

Figure 4: Epoch vs Loss

This good fit is identified by a training and validation loss that decreases to a point of stability with a minimal gap between the two final loss values.

4.3 Comparison

4.3.1 Comparison

Model Name	Accuracy
CNN	65.66
Inception V3	95.68

By comparing the accuracy of the two proposed models we got high accuracy for Inception V3 model with the accuracy of 95.68.

4.4 Application

4.4.1 User Interface

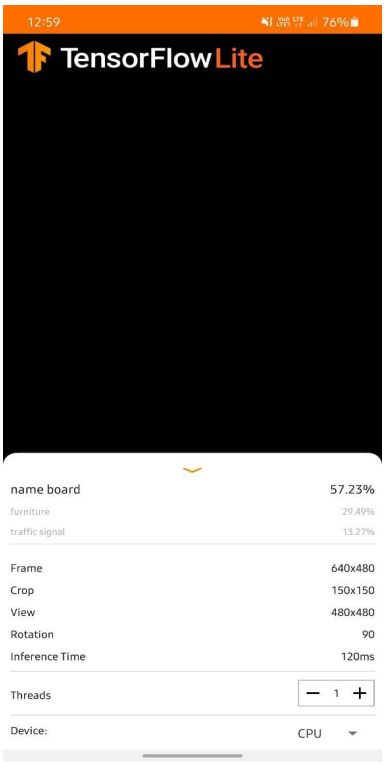

Figure 5: Main Screen of Application

4.4.2 Predictions

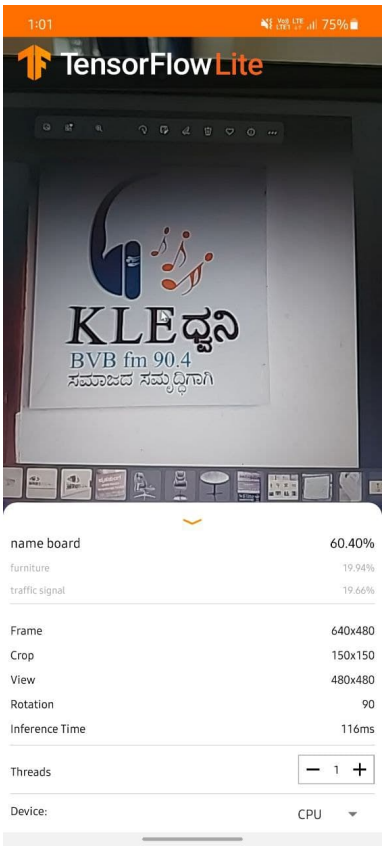

Figure 6: Result

The Predicted class label is nameboard.

5 Conclusion

With respect to the context of proposed learning model to detect the category of an image it belongs to, we tried with inception V3 and CNN model and it was trained to:

- To predict the category label of the given image

We explored different learning models used for image detection. The three category namely Nameboard, Traffic signal, Furniture on which proposed method is tested.To improve recognition rate in classification process some other pre-trained models and hybrid algorithms can also be used.

6 Future Scope

- The images in Dataset are taken in noiseless setups and not in real conditions, this motivates to build a public dataset which captures the images in real life conditions which would enable us to achieve a better efficacy in solving this problem.
- Few more techniques like image segmentation can be tried along with different pre-trained models to help in detection of images with high accuracy and faster.

7 References

[1]- Kumar, Sandeep; Khan, Zeeshan; Jain, Anurag."A Review of Content Based Image Classification using Machine Learning Approach."Vol. 2, Iss. 3, (Sep 2012): 55-60.

[2]- S. Loussaief and A. Abdelkrim, "Machine learning framework for image classification." 2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), 2016, pp. 58-61, doi: 10.1109/SETIT.2016.7939841.

[3]- Manoj krishna1, M Neelima, M Harshali, M Venu Gopala Rao."Image classification using Deep learning." March 2018:International Journal of Engineering Technology 7(2.7):614 DOI:10.14419/ijet.v7i2.7.10892.

[4]- Neha Sharma,Vibhor Jain, Anju Mishra."An analysis of convolution neural network for image classification." International Conference on Computational Intelligence and Data Science (ICCIDS 2018).