Pollution Measurement from Dust Over RoadSide Plant Leaf

REPORT OF PROJECT SUBMITTED FOR PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF

BACHELOR OF TECHNOLOGY
In
INFORMATION TECHNOLOGY

By

Somu Raj

REGISTRATION NO - 181170110234 UNIVERSITY ROLL NO - 11700218017

Sagen Soren

REGISTRATION NO - 181170110213 UNIVERSITY ROLL NO - 11700218038

Angad Kumar

REGISTRATION NO -181170110143 UNIVERSITY ROLL NO - 11700218108

Sheikh Atib

REGISTRATION NO - 181170110226 UNIVERSITY ROLLI NO -11700218025

UNDER THE SUPERVISION OF

Dr. Hiranmoy Roy

Associate Professor



AT
RCC INSTITUTE OF INFORMATION TECHNOLOGY
Affiliated to MAKAUT
CANAL SOUTH ROAD, BELIAGHATA, KOLKATA-700015

RCC INSTITUTE OF INFORMATION TECHNOLOGY

KOLKATA - 700015, INDIA



CERTIFICATE

The report of the Project titled Pollution Measurement from Dust Over RoadSide Plant Leaf submitted by **Somu Raj** (University Roll No.: 11700218017),**Sagen Soren** (University Roll No.: 11700218038), **Angad Kumar**(University Roll No.: 11700218108), **Sheikh Atib** (University Roll No.: 11700218025)of B. Tech. (IT) 8th Semester of 2018) has been prepared under our supervision for the partial fulfillment of the requirements for B Tech (IT) degree in MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY,KOLKATA.

The report is hereby forwarded.

OPTIONAL IN CASE	 Dr.Hiranmoy Roy
[Name of Guide]	Associate Professor
Designation	Information Technology
Institute Name	RCCIIT, Kolkata
(External Supervisor)	(Internal Supervisor)
Countersigned by	
Dr.Indrajit Pan	
•	
Department of IT RCC Institute of Information Technology, Kolkata – 700 015, India	

ACKNOWLEDGEMENT

I express my sincere gratitude to **Dr. Hiranmoy Roy** of Department of Information Technology, RCCIIT and for extending their valuable time for me to take up this problem as a Project.

I am also indebted to **Mr. Soumyadip Dhar**, **Dr. Abhijit Das** for their unconditional help and inspiration.

Last but not the least I would like to express my gratitude to **Dr. Shyantani Maiti** of our department who helped me in their own way whenever needed.

Date: 11th June 2022

Somu Raj Reg.No.:181170110234 Roll No.: 11700218017 B. Tech (IT) – 8th Semester, Session- 2022, RCCIIT

Sagen Soren
Reg.No.:181170110213
Roll No.: 11700218038
B. Tech (IT) – 8th Semester,
Session- 2022, RCCIIT

Angad Kumar Reg.No.:**181170110143** Roll No.: **11700218108** B. Tech (IT) – 8th Semester, Session- 2022, RCCIIT

Sheikh Atib Reg.No.:**181170110226** Roll No.: **11700218025** B. Tech (IT) – 8th Semester, Session- 2022, RCCIIT

RCC INSTITUTE OF INFORMATION TECHNOLOGY KOLKATA – 700015, INDIA



CERTIFICATE of ACCEPTANCE

The report of the Project titled "Pollution Measurement from Dust Over RoadSide Plant Leaf" Submitted by Somu Raj (University Roll No.: 11700218017), Sagen Soren (University Roll No.: 11700218038), Angad Kumar (University Roll No.: 11700218108), Sheikh Atib (University Roll No.: 11700218025) of B. Tech. (IT) 8th Semester of 2018) is hereby recommended to be accepted for the partial fulfillment of the requirements for B Tech (IT) degree in MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, KOLKATA.

Name of the Examiner	Signature with Date
1	
2	
3	
4	

TABLE OF CONTENTS

	Topics	Page No.
1.	Introduction	6
2.	Problem Analysis	7
3.	Review of Literature	8-9
4.	Formulation / Algorithm	10-13
5.	Problem Discussion	14
6.	Implementation Details	15
7.	Implementation of Problem	16-21
8.	Sample Output	22-23
9.	Conclusion / Future Scope of Work	24
10.	Reference	25-26
11	Appendix (Program Code)	27-32

INTRODUCTION

This report will give us an overview of our project Pollution Measurement from Dust over Roadside Plant Leaf. The main objective was to make an accurate assessment of pollution levels, pollution vehicles and develop water sprinkling and techniques to estimate pollution loads in each city. The report contains a description of all the user needs, different concepts that we created in order to meet those user needs, the evaluation process that we followed to evaluate our concepts and techniques that we applied in order to choose our final concept. It also contains a descriptive explanation of our concept and how it will serve our need statement. We have used large datasets (images of different plant leaves) to find accurate pollution measurements. We've used a dataset (plant leaves) from different parts of India. This report also includes the procedure that we followed to complete this project. This heavy concentration of population in a few centers has resulted in the expansion of cities in density as well as in areas.

Dust may affect photosynthesis, respiration, transpiration and allow the penetration of phytotoxic gaseous pollutants. Visible injury symptoms may occur and generally there is decreased productivity. Most of the plant communities are affected by dust deposition so that community structure is altered.

Leaf Cells and the cell structure get damaged due to excessive dust deposition. The small pores found in epidermis plant leaves are known as stomata. Stomata are important in case of controlling the vapor exchange and nutrients. Stomata also carry evapotranspiration through which the requirement of water for different functions in the plant body is controlled. Crushing units around selected study areas generate huge dust which directly accumulates on plant leaves. Dust consists of many chemicals as it is emitted by crushing of rocks laid over this area. The contamination of dust particles affects the nutritional quality of plants and vegetables.

Problem Analysis

Our Project is based on Pollution Measurement From Road Side plant Leaf, Using Deep learning tools. We Use Sequential Model for model training.

We will take 400 images for training purposes out of 500 images and the rest 100 images are used for testing the data.

Our second Step of the project is to map the images on graphs to check if the images are trained or not.

Next step is augmentation of images to create a big dataset. We use Adam optimizer to computations, and the accuracy or analysis..

After training the Data We build a sequential model.

This is the approach we take for This projects.

The stages are:

i)Problem Definition

ii)Data Gathering(DataSets)

iii)Model Design

iv)Model Training

v)Model Testing

vi)Deployment

Review of Literature

Patna city shows tropical stunted semi evergreen vegetation patterns which include all types of tropical species indigenous and exotic both. The city has a tree cover distributed throughout the urban-scape. Approximately 380 species of trees are observed in Patna city. A tree census is being conducted by Patna Municipal Corporation. Almost 70% of the tree census was completed up to June 2011. According to the tree census, 23.33 lakhs trees are present in 170 km2 areas.

Impact of air pollution on plant diseases:

Air pollutants directly toxic to the pathogens possibly impair their growth and reproduction and thereby partially or wholly inhibit the diseases. Air pollutants on the other hand, by modifying the host physiology may render it more susceptible to infection and pathogenic damages. Atmospheric pollutants may adversely affect the spore germination, mycelial proliferation, fruiting body formation and spore production by fungi.

Dust induced effects on chlorophyll:

Reductions in Chlorophyll due to dust deposition have also been observed by many coworkers. Dust deposition on leaf surfaces reduces synthesis of chlorophyll-a due to shedding effects. Thus increase in dust deposition and the corresponding chlorophyll degradation can be positively correlated. Alkaline conditions prevailing due to solubility of dust particulates in the cell sap are responsible for chlorophyll degradation and reduced photosynthetic activity. The hydration process of crust formation due to dust deposition on leaf of dusted plants, releases Ca(OH)2 (highly alkaline). This penetrates through stomata and injures the cell beneath and causes partial denaturation of chloroplast and subsequent decrease in pigments in the cells of damaged leaves.

It has been proposed in the above study that there are possibilities of degradation of enzymes responsible for chlorophyll biosynthesis. Alkaline dust deposited on leaf surfaces can also result in leaf chlorosis and death of leaf tissue by combination of thick crust and alkaline toxicity produced in wet weather. Plant cells contain carbohydrate which is useful for many functions of plant anatomy. Chlorophyll is green pigmented material found in green plants. It is also known as chlorophyll a and chlorophyll b. The efficiency of plant functions decides the quantity of plant pigmentation and chlorophyll present in the plant body.

Formulation / Algorithm

Sequential model: Sequence modelling is the process of producing a sequence of values from a set of input values. These input values could be time-series data, which shows how a certain variable, such as demand for a given product, changes over time. The production may be a forecast of demand for future times.

Different Sequential Model

RNN and its Variants Based Models

RNN stands for Recurrent Neural Network and is a Deep Learning and Artificial Neural Network design that is suited for sequential data processing. In Natural Language Processing, RNNs are frequently used (NLP). Because RNNs have internal memory, they are especially useful for machine learning applications that need sequential input. Time series data can also be forecasted using RNNs.

A different task that can be achieved using RNN areas,

One-to-one

With one input and one output, this is the classic feed-forward neural network architecture.

One-to-many

This is referred to as image captioning. We have one fixed-size image as input, and the output can be words or phrases of varying lengths.

Many-to-many

model=build model()

This paradigm is suitable for machine translation, such as that seen on Google Translate. The input could be a variable-length English sentence, and the output could be a variable-length English sentence in a different language. On a frame-by-frame basis, the last many to many models can be utilized for video classification.

We create a *Sequential model* and add layers one at a time until we are happy with our network architecture.

In order to train a Sequential model, we first have to configure our model using model.compile() with the following arguments:

```
#train model

#Sequential Model

def build_model():
    model=Sequential()
```

```
#layer-1
model.add(ZeroPadding2D((1),input shape=train set.image shape))
model.add(Conv2D(filters=64, kernel size=(3,3), activation='relu'))
model.add(MaxPooling2D(pool size=(2,2),strides=(2,2)))
#Layer-2
model.add(ZeroPadding2D(1))
model.add(Activation('relu'))
model.add(Conv2D(filters=128, kernel size=(3,3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2,2),strides=(2,2)))
#Layer-3
model.add(ZeroPadding2D(1))
model.add(Activation('relu'))
model.add(Conv2D(filters=256, kernel size=(3,3), activation='relu'))
model.add(MaxPooling2D(pool size=(2,2),strides=(2,2)))
```

```
#Layer-4
  model.add(ZeroPadding2D(1))
 model.add(Conv2D(filters=512, kernel size=(3,3), activation='relu'))
 model.add(MaxPooling2D(pool size=(2,2),strides=(2,2)))
  #Flatten the data
 model.add(Flatten())
  #FC
  model.add(Dense(units=256,activation='relu'))
  #Dropout
  model.add(Dropout(0.20))
  #output layer
  model.add(Dense(units=10, activation='softmax'))
  adam optimizer=Adam(learning rate=0.0001)
model.compile(loss='categorical_crossentropy',metrics=['accuracy'],optimizer=adam
optimizer)
  return model
```

Problem Discussion

Dust is a major concern in urban areas .It has been estimated that about 30 million tons of dust enter the atmosphere each year worldwide. In fact, dust when inhaled can increase breathing problems, damage lung tissue, and aggravate existing health problems. In addition to health concerns, dust generated from various activities can reduce visibility, resulting in accidents. By the use of plant leaves we can detect pollution in that area . Measure Pollution level so plants that Absorb Pollutants can be planted in a particular locality. Here we first click an image of a leaf which then scans it and runs the program which undergoes various processes and gives out a result of how much dust is there in the locality. If there is dust on leaves, we can build a mechanism that can sprinkle water to wash the plants.

Implementation Details

Create Datasets(picture of leaves)

- i) Summer Season (150 photos)
- ii) Rainy Season(50 photos)
- iii) Winter Season(300 photos)

Coding part progress:

- i)Convolutional Neural Network
- ii)Data augmentation
- iii)Pooling(maximum)
- iv)Padding
- v)Activation layer
- vi)Stride
- vii)Train sequential model
- viii)Find accuracy

Implementation of Problem

The following steps that we used to solve the problem are following:

- Step 1:Firstly Create a Datasets of 500 photos.
- Step 2:Import Library Such as Numpy,Os,matplotlib,pandas for calculation and plotting image on graph.
- Step 3:For training of Model Import keras library.
- Step 4:Use Adam optimisers for computations, and the accuracy or analysis.
- Step 5:After Compiling the code of the built model(Sequential Model).
- Step 6:Once the training is completed and analyzed for a fixed number of epochs, we can move ahead to the next step of evaluation and making predictions.
- Step 7: The deployment Stage is the final Stage of the model.

1. Defining Your Architecture —

Deep learning is one of the most preferable methods to solve complex tasks like image classification or segmentation, face recognition, object detection, chatbots, and so much more. But, with each of these projects, every deep learning model goes through five fixed stages to accomplish the task at hand. The first and most significant step of building your deep learning model is to define the network and architecture successfully. Depending on the type of tasks that are being performed, we prefer to use certain types of architecture.

Usually, Convolutional Neural Networks (CNNs) or ConvNets are preferred for computer vision tasks such as image segmentation, image classification, facial recognition, and other similar projects.

While Recurrent Neural Networks (RNNs) and Long Short Term Memory (LSTMs) are preferable for natural language processing and problems related to text data. In this step, you can also decide the type of model building structure of the entire deep learning architecture. The three main steps for performing this are the Sequential Models, Functional API, or a custom architecture defined by the user. We will discuss each of these methods in further detail in a future article.

2. Compiling Your Model —

With the preferred architecture constructed, we will move on to the next step of the compilation of the built model. The compile step is usually a single line of code in the TensorFlow deep learning framework and can be achieved with the model.compile() function. The requirement of the compilation step in deep learning is to configure the model for the fitting/training process to be successfully completed. It is during the compile

step that some of the critical components of the training procedure is defined for the evaluation procedure. To mention a few of the necessary parameters, we have the loss, optimizer, and the metrics to be assigned in the following step.

The kind of loss is determined by the type of problem we are encountering and what we need to solve. The optimizers are usually adam, or similar optimizers for computations, and the metrics can be accuracy or any other user-defined metrics for analysis.

3. Fit The Model —

After defining the overall architecture and compiling your model, the next logical step is to fit the model on the training dataset. The fit function trains the model for a fixed number of epochs (iterations on a dataset). The important parameters like the number of epochs for training, the input and output data, validation data, and many others are decided with the help of this function. It is used for computing and calculating these essential features.

The fitting step must be continuously evaluated while the training procedure is underway. It is essential to make sure the model being trained is performing well with improving accuracies as well as a reduction in the overall loss.

It is equally important to consider that the model is not being overfitted in any manner.

For this, continuous evaluation with a tool like Tensorboard must be used for analyzing

the various graphs and understanding if these models are being overfitted by any chance. Once the training is completed and analyzed for a fixed number of epochs, we can move ahead to the next step of the evaluation and making predictions with the trained model.

4. Evaluating And Making Predictions —

Evaluation of deep learning models is an extremely significant step to check out if your article is working out as desired. There are chances that the deep learning model that you built might not perform well in real-world applications. Therefore, evaluation of your deep learning models becomes critical.

One main method of evaluating your deep learning models is to ensure that the predictions made by your model on the test data that is split at the beginning of the pre-processing step are considered for the purposes of validating the effectiveness of the trained model.

Apart from the test data, the model must be tested with variable data and random tests as well to see its effectiveness on un-trained data and if its efficiency of performance matches the required commitments.

To explain with a simple example — let us assume that we built a simple face recognition model. Consider the model you have trained with the images and try to evaluate these images with various faces both on the test data and a real-time video reel as well to make sure the trained model is performing well.

5. Deploying The Model —

The deployment stage is the final stage of any model constructed.

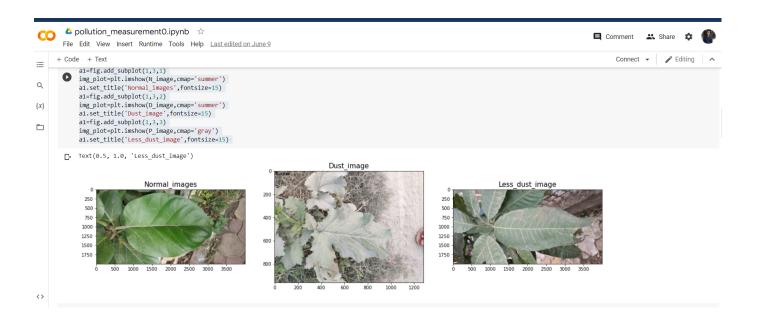
Once you have successfully completed building your model, this is an optional step if you want to keep it with yourself or deploy it so that you can target a wider audience.

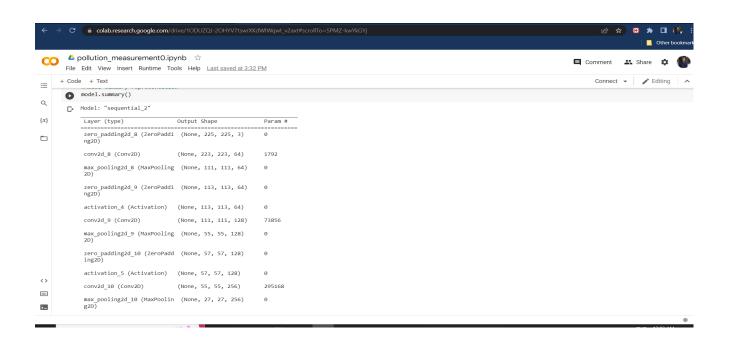
The methods of deployment vary from deploying it as an application that can be transferred across, or by using the AWS cloud platform provided by amazon for deployment, or by making use of an embedded system.

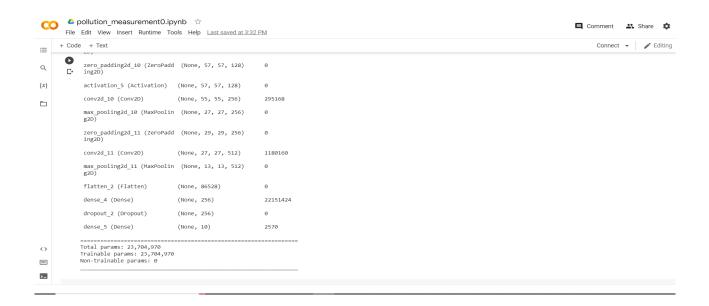
If you want to deploy something like a security camera, then you can consider using an embedded device like the raspberry pi alongside a camera module for performing this function. Embedded systems with AI are common ways of deploying your IoT projects.

You can also deploy these deep learning models on your website after it is built with either flask, Django, or any other similar framework. Another way to effectively deploy your models is by developing an android or iOS app for smartphone users to reach a wider range of audiences.

Sample Output







Conclusion / Future Scope of Work

The study shows that dust generated from the different industries causes damage to the Mango plant. Leaves are the most suitable organ of the plant to hold the dust on its upper and lower surface. Dust holding capacity of any leaf determined by its size, orientation and arrangement. Simple leaves hold more dust than compound leaves. Dust has physical and chemical properties, it may produce necrotic spotting. In presence of moisture, it solidifies into a hard adherent crust, which can damage plant tissue and restrict growth .Mango is sensitive to dust pollution thus, dust pollution decreases the economic and nutritional value of the crop. Plants improve air quality through several mechanisms: they absorb carbon dioxide and release oxygen through photosynthesis, they increase humidity by transpiring water vapor through microscopic leaf pores, and they can passively absorb pollutants on the external surfaces of leaves and on the plant root-soil system. All type of pollution has their kind of negative impact on our environment. The lives of humans and animals get impacted due to this. It is our responsibility to take various initiatives to protect nature. We need to fight against pollution to take steps towards a better tomorrow.

Our natural environment makes human life possible, and our cultural environment helps define who we are. It is therefore essential that our population and economic growth are environmentally sustainable.

While the effects of air pollution on materials, vegetation, and animals can be measured, health effects on humans can only be estimated from epidemiological evidence. Most of the evidence comes from occupational exposure to much higher concentrations of pollutants than the general public is exposed to.

Reference

Andrew Ng, Richard Socher 2014. RECURSIVE DEEP LEARNING FOR NATURAL LANGUAGE PROCESSING AND COMPUTER VISION

Akhiezer, Naum I., and Glazman, Izrail M. 1993. Theory of Linear Operators in Hilbert Space. Dover Publications.

Alpaydin, Ethem. 2010. Introduction to Machine Learning. MIT Press

Simonyan, Karen and Andrew Zisserman. 2014. "Very Deep Convolutional Networks for Large-Scale Image Recognition." CoRR abs/1409.1. Retrieved (http://arxiv.org/abs/1409.1556).

Sun, Yi, Xiaogang Wang, and Xiaoou Tang. 2013. "Deep Convolutional Network Cascade for Facial Point Detection." Pp. 3476–83 in Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on.

P. Duygulu, K. Barnard, N. de Freitas, and D. Forsyth. 2002. Object recognition as machine translation. In ECCV.

Yunus M, Ahmed KJ. Use of epidermal traits of plants in pollution monitoring. National Seminar on Environmental Pollution: A Status Review. National Productivity Council, India; 1979. p. 1–10.

Prajapati, S.K., and Tripathi, B.D. Seasonal variation of leaf dust accumulation and pigment content in plant species exposed to urban particulates pollution. *Journal of Environmental Quality*, 37: 865-870 (2008)

Prusty, B.A., Mishra, P.C., and Azeez, P.A. Dust accumulation and leaf pigment content in vegetation near the national highway at Sambalpur Orissa, India. *Ecotoxicological Environmental Safety*. 60: 228-235 (2003)

Walker, D.A., and Everett, K.R. Road dust and its environmental impact on alaskan taiga and tundra. *Arctic and Alpine Research*, 19(4): 479-489 (1987)

Spatt, P.D., and Miller, C. Growth conditions and vitality of *Sphagnum* in a tundra community along the alaska pipeline haul road. *Arctic*, 34(1): 48-54 (1981)

Appendix (Program Code)

#import library
import numpy as np
import os
import matplotlib.pyplot as plt
from PIL import Image
import pandas as pd
#Deep Learning or CNN
import keras
from keras.models import Sequential
from keras.layers import Conv2D,MaxPooling2D,Flatten,Dense,ZeroPadding2D,Dropout,Activation
#Image Augmentation and Regeneration
from keras.preprocessing.image import ImageDataGenerator,load_img
#print directory folder of datasets
print(os.listdir('/content/drive/MvDrive/Datasets'))

```
#Train & Test folder
train_folder='/content/drive/MyDrive/Datasets/train_data/'
test_folder='/content/drive/MyDrive/Datasets/test_data'
train_n=train_folder+'normal_image/'
train_d=train_folder+'Dust_image/'
train_l=train_folder+'Less_Dust_image/'
#Load and display the images belongs to these folders
D_image=Image.open(train_d+'dust.jpg')
N_image=Image.open(train_n+'normal.jpg')
P_image=Image.open(train_I+'LESS.jpg')
#Plot the images
fig=plt.figure(figsize=(20,10))
a1=fig.add_subplot(1,3,1)
img_plot=plt.imshow(N_image,cmap='summer')
a1.set_title('Normal_images',fontsize=15)
a1=fig.add_subplot(1,3,2)
img_plot=plt.imshow(D_image,cmap='summer')
a1.set_title('Dust_image',fontsize=15)
a1=fig.add_subplot(1,3,3)
img_plot=plt.imshow(P_image,cmap='gray')
a1.set_title('Less_dust_image',fontsize=15)
28
```

```
train_folder='/content/drive/MyDrive/Datasets/train_data'
test_folder='/content/drive/MyDrive/Datasets/test_data'
val folder='/content/drive/MyDrive/Datasets/val data'
#Data Augmentation:- Create some new data
train_datagen=ImageDataGenerator(rescale=1./255,shear_range=0.2,zoom_range=0.2,horizontal_flip=
True, rotation range=40, width shift range=0.4, height shift range=0.2)
test datagen=ImageDataGenerator(rescale=1./255,shear range=0.2,zoom range=0.2,rotation range=4
0,width_shift_range=0.2,height_shift_range=0.2,)
train set=train datagen.flow from directory(train folder,target size=(226,226),batch size=32,class
mode='categorical')
test_set=test_datagen.flow_from_directory(test_folder,target_size=(226,226),batch_size=32,class_mo
de='categorical')
val_set=test_datagen.flow_from_directory(val_folder,target_size=(226,226),batch_size=32,class_mode
='categorical')
```

from tensorflow.keras.optimizers import Adam

```
#train model
#Sequential Model
def build_model():
 model=Sequential()
 #layer-1
 model.add(ZeroPadding2D((1),input_shape=train_set.image_shape))
 model.add(Conv2D(filters=64,kernel_size=(3,3),activation='relu'))
 model.add(MaxPooling2D(pool_size=(2,2),strides=(2,2)))
 #Layer-2
 model.add(ZeroPadding2D(1))
 model.add(Activation('relu'))
 model.add(Conv2D(filters=128,kernel_size=(3,3),activation='relu'))
 model.add(MaxPooling2D(pool_size=(2,2),strides=(2,2)))
```

```
#Layer-3
model.add(ZeroPadding2D(1))
model.add(Activation('relu'))
model.add(Conv2D(filters=256,kernel_size=(3,3),activation='relu'))
model.add(MaxPooling2D(pool_size=(2,2),strides=(2,2)))
#Layer-4
model.add(ZeroPadding2D(1))
model.add(Conv2D(filters=512,kernel_size=(3,3),activation='relu'))
model.add(MaxPooling2D(pool_size=(2,2),strides=(2,2)))
#Flatten the data
model.add(Flatten())
#FC
model.add(Dense(units=256,activation='relu'))
#Dropout
model.add(Dropout(0.20))
```

```
#output layer
  model.add(Dense(units=10,activation='softmax'))
 adam_optimizer=Adam(learning_rate=0.0001)
 model.compile(loss='categorical_crossentropy',metrics=['accuracy'],optimizer=adam_optimizer)
  return model
model=build_model()
#Model summary representation
model.summary()
#history of model
history=model.fit_generator(epochs=1,shuffle=True,generator=train_set,steps_per_epoch=300)
np.round(model.predict(test_set))
```