

A Fruit Recognition System based on Modern Deep Learning Technique

Swapnil Srivastava^{#1}, Tripti Singh^{#2}, Sakshi Sharma^{#3}, Anil Verma^{#4}

^{1, 2, 3, 4}Department of Electronics and Communication Engineering, Raj Kumar Goel Institute of Technology, Ghaziabad
Dr. A.P.J. Abdul Kalam Technical University, Lucknow

Abstract:- The technology popularly used in this innovative time is Computer vision for the process of fruit recognition. In comparison to other machine learning (ML) algorithms, such networks like deep neural networks (DNN), provide results to identify fruits in images. And there results are promising. As for now, in order to identify fruits, different DNN-based classifications of algorithms are used. However, there issue in ordering to recognizing fruits has yet to be addressed. And it is due to similarities in size, shape and other features. This section discusses this in brief. The deep learning is used for recognizing fruits and its other applications. This section will also provide a concise and precise explanation of convolution neural networks (CNNs). Moreover, it also throws light on the EfficientNet architecture to recognize fruit by using the Fruit 360 dataset. The results show that the proposed model is 95% more accurate.

Keywords: Deep Learning, DNN, CNN.

I. INTRODUCTION

Taking rapid advancement of Human Race into Consideration, a legitimate concern may be given to the consumed food. Various techniques have been systematically used over the past times for fruit recognition by using Computer vision technology. The most notable and highlighted application is the use of DNN in identifying, classifying, and differentiating among different varieties of fruits from a dataset of pictures showing how they outperform any other algorithms. Moving ahead, DNNs are used in a huge set of application to provide optimally concluding solutions for the problems faced in multiple and complex domains such as large dataset analysis, image analysis, forecasting, prediction, speech recognition, and marketing.

Today, the most commonly used artificial neural network (ANN) type across multiple domains is CNN. CNNs are basically taken in use for classification of 2-D input images/pictures and to recognize the objects based on convolution and pooling layers. There are three layers in ANN architecture, starting with the input layer, then the hidden layer and finally ending with the output layer. Every layer is made up of neurons. The input tin each of the layer's neurons consist of the summations of output of neurons from the previous layer. The differentiation of output is with the target values done on the cost function. This is very important as the more accurately recognizing fruits is done, more is paramount importance in the yield mapping field. In this study, an optimal scheme has been introduced for differentiating variety of fruits by using a dataset, that is accessible and it also simulates current time predictions by using EfficientNet.

The general structure of the paper/study is as such:

Section 1 deals with the applications of DL in identifying fruits.

Section 2 would elaborate the topic of CNNs. (The architecture of EfficientNet will be highlighted in more detail in the following paragraph).

Section 3 would explain the reasons for selecting the dataset.

Section 4 would present the prediction of outcome and then the results and will have the discussion of the potential for future improvements.

II. DEEP LEARNING ALGORITHMS:-

The Deep Learning is basically the sub-field of Machine Learning, as sub-field of Artificial Intelligence. It is a group/collection of techniques that models high-level of abstractions in data. A computer-based statistically derived model scans, understands and learns from images, sounds, or texts to conduct the analysis. Such models might attain a State Of Art accuracies and sometimes even exceed human level's performance. These models are trained by using a group of labelled data and neural network architectures that has many layers of accuracy.

As the concept of deep learning was firstly came forward in 1980s, the idea subsequently started becoming popular mainly of these two reasons. It requires a great amount of labeled data and significant computing power. In the recent decades, the number of deep learning applications have experienced an excessive increase, along with, image classification, natural language processing, information retrieval, and so on. Deep learning, further, could be divided into two parts and could be understood individually: deep and learning. Learning deals with the previous understanding information and then making an inner depiction of the subject that can be used to act by the agent. As there is the internal depiction which is a compact presentation to form a summary of the data. Machine Learning provides many functions and various techniques to learn automatically from the available information, and this concept of learning from the information is used to forecast and promote projections in the future. Artificially designed Neural Network (ANN) gets the basic inspiration from the human brain and is the most commonly deployed algorithm in the case of Machine Learning. It consists of integrated processing units named as neurons. ANN is comprised of input, hidden, and output layer. Input layer takes an input, for example, an image and passes it to the hidden layer, and then the output layer gives output - the maximum probability that what object in an image. We can have a multiple of hidden layers for more complexity of functions.

A. Convolutional neural networks:-

Taking visual analysis into account, CNN is a neural network from the DNN class, which is widely used. CNNs are normally viewed as feed-forward neural networks (FFNN) which can identify, classify, and recognize any features present in an image. The first CNN is known as LeNet. It was created by Yann LeCun in 1988, a member of the Artificial intelligence research group at Facebook. In CNNs, the network input consists of image pixel that contain values of different weights depending on the feature that needs to be extracted as defined in the hidden layer. In an input image, CNNs are made up of fully connecting layers in order to recognize different items, despite the pooling and convolution layer.

The convolution operations are performed into a CNN classifier taking pixels in an image. It consists of 4 of the most commonly used layers. The first one being the

convolution layer, that is tasked with convolving the pixels in an image with a chosen kernel/Harris to extract or different features. The 2nd one is the ReLU layer, which stated an activation function, that can be a sigmoid. The image is then passed several times between the convolutions, and ReLU layers. It is here, where all the negative pixel are converted to zero and trends and attributes of an image are analyzed. The third layer is known as the Pooling layer, and the only purpose of this layer is to do the transformation of the image into the required dimension instead of blurring it. It is for that purpose, the pooling layer encompasses different kernels to identify the sharp edges and to detect different contours in an image. The image is then transformed into a 1-D linear matrix. The last layer is fully connected. It is used to identify the images and classify them with accuracy. A typical architecture of the CNN is shown in Figure 1.

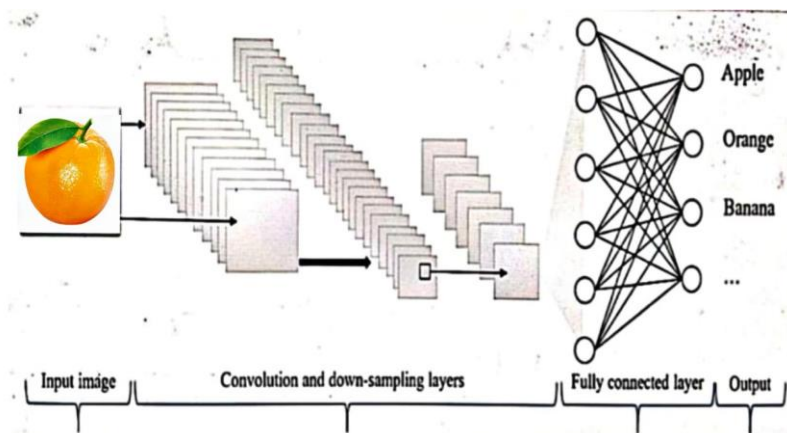


Fig.1 A convolutional neural network

B. EfficientNet:-

It uses pre-trained convolution neural networks as for performing image related functions in form of a base network. Such base networks are efficiently capable of to learn from a wide range of dataset hence more particular models with restricted training data can be created more. Networks as such, are useful for functions like classification of images and facial recognition so as to provide benefits for more precise and efficient models. The conventional

arbitrarily defined scaling as a process nevertheless provides functional results.

EfficientNet initially performs a grid search for the base network in order to determine the relationships between the different scaling dimensions of the network while watching both model size and availability of computational resources. In most of the situations, the initial testing shows a higher amount of precision and velocity. The following table demonstrates the summary of the architecture of EfficientNet-B0:

TABLE 1. EfficientNet-B0 architecture

Stage i	Operator F_i	Resolution $H_i \times W_i$	Channels C_i	Layers L_i
1	Conv 3x3 MBConv _i ,	224*224	32	1
2	k3x3 MBConv6, k3x3	112*112	16	1
3	MBConv6, k3x3	112*112	24	2
4	MBConv6, k3x3	56 x 56	40	2
5	MBConv6, k3x3	28 x 28	80	3
6	MBConv6, k3x3	14 x 14	112	3
7	MBConv6, k3x3 Conv1x1	14 x 14	192	4
8	& Pooling & FC	7x7	320	1
9		7x7	1280	1

III. DATASET

For the work of training and testing, the images were chosen from the fruits 360 dataset, which are available on Kaggle. This dataset contained 77917 different types of fruits

pictures belonging to 103 categories. These pictures were received by registering the fruits. In due process, a motor revolves them and then produces frames.

A white paper is placed in front of the fruits and is used as a background. Due to the disparity in the lighting, a flood-fill algorithm is applied to extract the fruit from the background. After removing such backgrounds, all the fruits are resized to 100x100 pixels of standard RGB pictures. From the

fruits-360 dataset, we had selected 5548 pictures from 8 different categories. We had used 4161 images (75%) to create the training set and the rest 1387 images (25%) for testing the model. Table 2 shows the 8 categories of fruits that we used for analysis.

TABLE 2. Using categories of fruits

Name of fruits	Number of training images	Number of testing images
Apple Red	492	164
Banana	489	163
Lemon	492	164
Mango	489	163
Orange	480	160
Pear	492	164
Tomato	738	246
Pineapple	489	163

IV. EXPERIMENTAL RESULTS

In this module, we have applied EfficientNet-b0 on Fruit Dataset to implement the better classification performance of the network. From Fruits 360 dataset, we have taken 5534 images from 7 different categories of fruit: 74% of the images are used for training, and 24 % are used for testing it. The network is trained for 33 epochs with a batch of 18. We compared our model with the present state models and

results were different. The accuracy of the implemented model was 93.67 %. The comparison of the implemented model with the models shows that the results of our model are good and promising to use in real-world applications. This sort of accuracy and precision will work to boost the machine's efficiency in fruit recognition more accurately. As a prototype, a program was developed in Python with PyQt library in a Visual Studio environment. The appearance of the program is shown in Figure 2.



Fig 2. Main window of the program

V. CONCLUSION

This paper elaborates a fruits recognition classifier system based on the EfficientNet algorithm. The recognition rate has improved throughout the several experiment. Among all the tried cases, the model has achieved the best test accuracy of 97% in case 3 from 10 to 14 epochs and best accuracy of 95.79% at epoch 12. This type of accuracy will stimulate the overall performance of the system more adequately in fruits recognition. In the future, our plan is to improve recognition system by updating its functions to process and recognize more variety of different fruit images.

REFERENCES

- [1] Rocha A, Hauagge D C, Wainer J, Goldenstein S 2010 Automatic fruit and vegetable classification from images Comput. Electron 70 96-104
- [2] I Sa, Z Ge, F Dayoub, B Upcroft T. Perez, and C McCool 2016 Deepfruits: A fruit detection system using deep neural networks Sensors 16(8) 1222
- [3] L Den, G.M Hinton . New classification of deep neural network learning and for speech recognition and its application.
- [4] Y LeCun and Y Bengio Convolutional networks for images, speech, and time series," The handbook of brain theory and neural networks 3361(10) 1995
- [5] M. Tan and Q. V. The Model Scaling for Convolutional Neural Networks preprint .
- [6] H. Muresan, M. Oltean 2018 Fruit recognition from images using deep learning, Proceeding of the Acta Univ. Sapientiae, Informatica 10(1) 26-42