Fruit Classification and Quality Prediction using Deep Learning Methods

Shivani R[1], Tanushri[2]

[1]Student, CEG, Anna University, Chennai, India
[2]Student, CEG, Anna University, Chennai, India

---***---

Abstract - *In the agriculture industry, a lot of manpower and resources are wasted in manually classifying the fruits and vegetables and predicting their quality. A traditional method for fruit classification is manual sorting which is time-consuming and labor-intensive and hence requires human presence. Image processing and computer vision can be incorporated to automate these tasks.*

In this paper, we aim to develop an automated system with the help of deep learning models and image processing techniques for accurate predictions. Various deep learning convolutional neural network architectures such as VGG16, Inception V3, ResNet Mobilenet were used to train the model using the transfer learning approach. The model performance and results were compared, and its real-world implementation was studied. Different model performances were evaluated, and Inception V3 seemed to be the most accurate and effective model, with an accuracy of up to 98%. Outcomes of this study show that using the Convolutional Neural Network model; Inception V3 gives promising results compared to traditional machine learning algorithms.

***Key Words*: Fruit Classification, Image Processing, Deep Learning, Inception V3, Quality Prediction**

1. INTRODUCTION

For around 58 percent of India's population, agriculture is their major source of income. Agriculture, together with its linked industries, is indisputably India's primary source of income, particularly in the country's vast rural areas. Agriculture's contribution to GDP has surpassed nearly 20% for the first time in 17 years, making it the only bright point in GDP performance in 2020-21.

India is endowed with vast swaths of fertile land divided into 15 agro-climatic zones, each with its own set of weather conditions, soil types, and crop potential. The varied climate of India assures the availability of a wide range of fresh fruits and vegetables. After China, it is the world's second-largest producer of fruits and vegetables. As a result, the fruit sector plays a critical role in India's economic growth and is an important component of the food processing industry.

Fruits and vegetables are an essential element of a well-balanced diet that help humans stay healthy. Vitamins, minerals, and plant components are abundant in fruits and vegetables. They also contain fiber which lowers cholesterol levels and regulates blood sugar levels. There are numerous fruit and vegetable kinds to choose from, as well as multiple ways to prepare, cook, and serve them. A fruit and vegetable-rich diet can help you avoid cancer, diabetes, and cardiovascular disease.

Unlike digital technology, which has acquired traction, agricultural biotechnology is still developing. Lack of low-cost technology and equitable access has resulted in productivity decline. Despite large-scale agricultural mechanization in some regions of the nation, most agricultural operations are still done by hand using conventional and straightforward tools. Manual sorting is carried out by farmers, which is time-consuming and labor-intensive. It is critical to mechanize agricultural processes to prevent labor waste and make farming more convenient and efficient. Machinery is an essential component of agricultural activities that are efficient and timely. This paper aims to automate the tasks of classification of fruits and vegetables and their quality prediction.

2. LITERATURE REVIEW

A study proposed by Bhavini J. Samajpati aimed to detect and classify apple fruit diseases. Segmentation of infected apple fruit was carried out by the K-mean clustering technique. Different features such as global color histogram(GCH), color coherence vector(CCV) and local binary pattern(LBP) were extracted. Upon testing images with various classifiers, SVM proved to be the most efficient in terms of accuracy and performance.

In [2], the fruit classification system is implemented using the Support Vector Machine classifier. It involved the extraction of statistical and texture features from wavelet transforms.It consisted of three phases: preprocessing, feature extraction and classification phase. In the feature extraction phase, the statistical feature of color and texture feature from wavelet transform were derived.

Winda Astuti[3] proposed an SVM-based fruit classification system that used Fast Fourier Transform as input. The system differentiated the fruit based on their shapes.The results showed that the technique produced

better accuracy than the existing technique based on an artificial neural network. The SVM also required less training time than an artificial neural network..

The convolutional neural network approach was adopted by Zaw Min Khaing [4] to achieve the task of fruit detection and recognition.Alexnet architecture was used to perform classification, and simulation was carried out with the help of MATLAB.

A model by H. Muresan et al. [5] introduces the Fruits-360 dataset, which presently has 49561 photos of 74 different types of fruit. They used deep learning to construct software that can detect fruit from photographs. For recognition, their suggested model uses CNN. Their major goal was to report the findings of a numerical experiment used to train a neural network to recognize fruits. Fruits-360 comprises just single object fruits for both training and testing. Therefore their system does not require any detection. As a result of this flaw, the system is unsuitable for real-time use. Only Fruits360 works better with their recommended model.

3. CONVOLUTIONAL NEURAL NETWORKS

The term "deep learning" refers to a type of machine learning in which data is processed via numerous layers to extract increasingly higher-level characteristics.

It consists of different architectures which comprise artificial neural networks.Convolutional Neural Networks (ConvNet/CNN) are a type of Deep Learning technique used in computer vision.The amount of pre-processing required by a ConvNet is much less than that required by other classification techniques.While in primitive methods, filters are hand-engineered, with enough training, ConvNets could learn these filters/characteristics.CNN employs 2D convolutional layers to combine learned features with input data, making it ideally suited to processing 2D data like photos. Because CNN eliminates the requirement for human feature extraction, the characteristics necessary to categorize the pictures do not need to be selected.

Convolutional neural networks have a better performance than other types of neural networks with images, voice or audio signal inputs. The convolutional layer, Pooling layer, and Fully-connected (FC) layer are the three primary types of layers.

CNN's most critical component is the convolutional layer. The three components are input data, a filter, and a feature map. A feature detector, sometimes known as a kernel or a filter, detects features and looks for the existence of a feature in the image's receptive fields. This method is known as convolution. A CNN performs a Rectified Linear Unit (ReLU) transformation to the feature map after each convolution process, bringing nonlinearity into the model.

Pooling layers is a dimensionality reduction technique that reduces the number of factors in the input.

The pooling process sweeps a filter across the whole input, but this filter is weightless.

Max pooling is a type of pooling in which the filter picks the pixel with the maximum value to transmit to the output array as it traverses the input. This method is more frequently employed than average pooling.

Average pooling is a type of pooling in which the filter calculates the average value within the receptive field to transmit to the output array as it traverses the input.

Each node in the output layer is connected directly to a node in the preceding layer in the fully-connected layer.

This layer conducts categorization using the characteristics retrieved by the preceding layers and their associated filters. While convolutional and pooling layers often employ ReLu functions, Fully Connected layers typically employ a softmax activation function to accurately categorize inputs, generating a probability between 0 and 1.

4. INCEPTION V3

Inception v3 focuses on reducing computational power consumption with modifications to prior Inception designs. This concept was first proposed in the 2015 article Rethinking the Inception Architecture for Computer Vision. Different model architectures such as VGG16, ResNet 50, MobileNet, Inception v3 were implemented using the Datasets, and their performances were evaluated. Inception Networks have been demonstrated to be more computationally efficient, both in terms of the number of parameters created and the cost incurred (memory, loss and accuracy). An inception network is a type of deep neural network with an architectural design consisting of repeating components referred to as Inception modules.

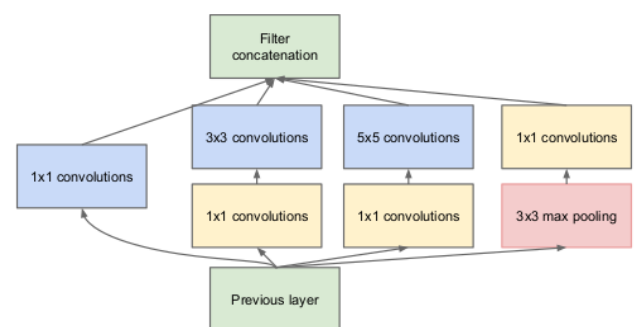The inception module has a Convolution layer (1x1,3x3,5x5), Max pooling layer, Concatenation layer.

Fig-1:Inception Module

4.1 1 x 1 Convolution

The purpose of 1 x 1 convolution is to reduce the dimensions of data passing through the network,

This gives the added advantage of increasing the network's breadth and depth.

It acts as a channel collapser, allowing subsequent procedures to be performed on a two-dimensional picture rather than a costly three-dimensional one, significantly reducing the number of operations required while maintaining color recognition.

4.2 3 x 3 and 5 x 5 Convolutions

As a consequence of the varied convolutional filter sizes of 3 x 3 and 5 x 5 Convolutions, the network will be able to learn distinct spatial patterns at different scales.Inception network has the luxury of leveraging different filter sizes within their convolutional layers.Within a convnet, different convolution filter sizes learn spatial patterns and detect features at varying scales.3x3 and 5x5 learn spatial patterns across all input's dimensional components (height, width and depth).

4.3 Concatenation layer

The Inception module comprises a concatenation layer that combines all outputs and feature maps from the convolutional filters into a single object to form the module's output. It accepts a list of tensors with identical shapes except for the concatenation axis as input and gives a single tensor that is the concatenation of all inputs.

5. DATASET

The primary concept of machine learning is collecting image datasets for training and model building. The dataset "Fruit and Vegetable Image Recognition"(from Kaggle) was used for the task of Fruit and Vegetable Classification. It consists of images of fruits including Apple, Banana, Watermelon and Vegetables, including Eggplant, Onion, Bell pepper. It contains three folders: train (100 images each), test (10 images each), validation (10 images each). For the prediction of the quality of fruits, the dataset "IEEEFRUITSDATA_train&test" was used.The Dataset consists of 12,050 images comprising fruits such as Apple, Banana, Guava, Orange, Lime and pomegranate categorized based on their quality as Good or Bad.

6. METHODOLOGY

The Inception v3 model with pre-trained weights was imported and loaded for training the classifiers using transfer learning.

Transfer Learning is a method where we use a pre-trained model. This Inception-v3 model has been trained on a vast dataset, and we transferred weights accumulated during hundreds of hours of training on many high-performance GPUs. We then set the parameters in the pre-trained model to be untrainable, optimizing only the parameters in the succeeding dense layers during training.

For our problem statement, the last layer -The softmax layer, was modified to have the required number of classes as the original model was trained on the ImageNet dataset. It significantly decreases training time and needs far less data to improve performance.

6.1 Training Parameters

The loss metric used was Categorical cross-entropy , a loss function used in multi-class classification tasks.

Adam optimizer was implemented, a stochastic gradient descent replacement technique for training deep learning models. Adam combines the finest qualities of the AdaGrad and RMSProp methods to provide a sparse gradient optimization technique for noisy problems.

The activation function utilized in the hidden layer substantially impacts how well the network model learns the training data. The type of predictions the model may produce is determined by the activation function used in the output layer. The type of predictions that the model makes is heavily influenced. Hidden layer: Rectified Linear Activation (ReLU) was used because it performs better than other activation functions and overcomes problems such as the vanishing gradient problem. ReLU is represented by the formula $y=max(0,x)$. The ReLU activation function changes the negative values of the neuron to 0, and positive values remain as it is.

In the output layer, the activation function depends on the type of prediction problem. For multilabel or multiclass such as fruit prediction SoftMax function is used.

6.2 Image Augmentation

Image augmentation is the technique of altering the existing data to create some more data for the model training process. Different Augmentation techniques such as Image rotation, Image flipping, and Image shifting have been used. The ImageDataGenerator class allows you to instantiate generators of augmented image batches (and their labels) via. flow(data, labels) or .flow_from_directory(directory).

7. RESULT

The Fruit and Vegetable Classification model was trained for 50 epochs. The Validation loss incurred is 0.4914, and the accuracy achieved was 0.9832.

The Fruit Quality Prediction model was trained for 10 epochs. The Validation loss incurred is 0.2925, and the accuracy achieved was 0.9851.
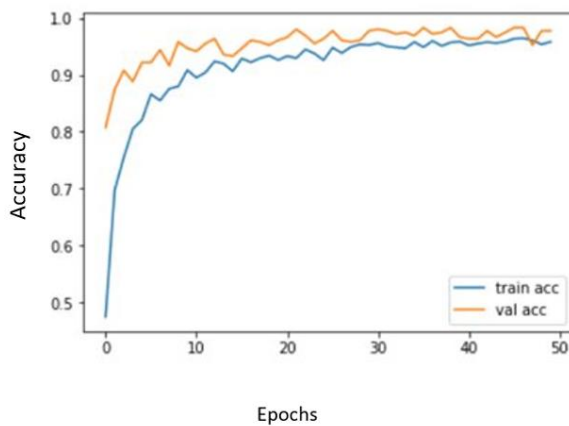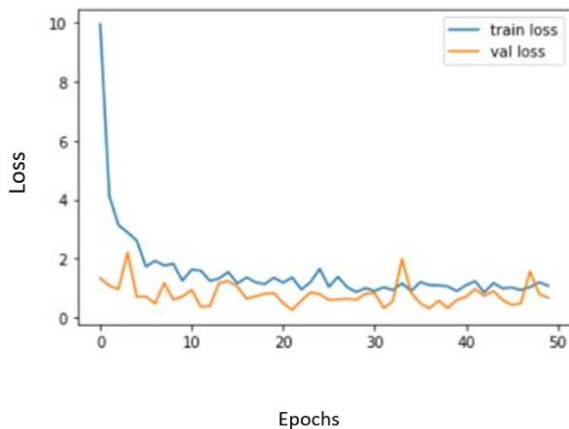
Chart-1: Model Performance Loss

Chart-2: Model Performance Loss

8. CONCLUSION

This paper focuses on developing an automated system for Fruit and Vegetable Classification and their quality prediction. Image processing methods and several CNN models were experimented with, and transfer learning was utilized. Inception V3 exhibited the best results and showed an accuracy of 0.98. Automation technology enables produce to reach customers faster, more freshly, and sustainably. Increased productivity due to automation increases the output and pace of production, lowering consumer costs. Routine, manual tasks may be mechanized using robot technology, lowering labor costs and reducing the number of workers required in a farm business facing a labor crisis.

ACKNOWLEDGEMENT

REFERENCES

[1] Bhavini J. and Sheshang D., "A Survey on Apple Fruit Diseases Detection and Classification", International Journal of Computer Applications (0975 – 8887) Volume 130, No.13, November 2015

[2] R Shantha and V Gomathy, "Fruit Classification using Statistical Features in SVM Classifier," IEEE 2018 4th International Conference on Electrical Energy Systems (ICEES), doi:10.1109/ICEES.2018.8442331

[3] Winda Astuti, Satrio Dewanto, Khristian Edi Nugroho Soebandrija and Sofyan Tan, "Automatic fruit classification using support vector machines: a comparison with artificial neural network"," ICEED 2018,doi:10.1088/1755-1315/195/1/012047

[4] Zaw Min Khaing, Ye Naung and Phyo Hylam Htut3, "Development of Control System for Fruit Classification Based on Convolutional Neural Network ",2018 IEEE 978-1-5386-4340-2/

[5] H. Muresan and M. Oltean, "Fruit recognition from images using deep learning," Acta Universitatis Sapientiae, Informatica 10(1), 26–42 (2018).