IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Detection Of Food Quality Using Machine Learning

[1]Mrs. Rashmi R, [2]Sanjana B J, [3]Shashidhar R A, [4]Sinchana S, [5]Appu Sindya B,
[1]Associate Professor at JNNCE, [2,3,4,5]Student at JNNCE
[1]Information Science and Engineering,
[1] Jawaharlal Nehru New College Of Engineering, Shivamogga, India

Abstract: Freshness is a key factor in determining a fruit or vegetable's quality, and it directly influences the physical health and coping provocation of consumers. It ascertains the nutritional value of the specified fruit or vegetable. This paper proposes a well-organized and precise fruit and vegetable classification and freshness detection method. The proposed method employs state-of-the-art deep learning models, specifically convolutional neural networks (CNNs), to analyze images of fruits and vegetables captured through highresolution cameras. The dataset used for training and evaluation is extensive and diverse, encompassing a wide variety of fruits and vegetables in various conditions. The freshness of a fruit or vegetable can be ascertained by looking at a variety of features, including color, texture, shape, and size. Fresh produce, for instance, is colorful and free of mold or brown spots. Traditional methods for assessing the quality of fruits and vegetables are both time-consuming and error-prone. These methods include inspection and sorting. It is possible to reduce these issues by utilizing automatic detection techniques. In light of this, we proposed an automated fruit-vegetable freshness detection approach that first recognizes whether the image is of a fruit or vegetable, after which it classifies it into one of three freshness categories: rotten, fresh, or mixed. To identify and categorize fruits and vegetables, two deep learning models are employed: You Only Look Once (YOLO) and Visual Geometry Group (VGG-16). The suggested method's qualitative analysis indicates superior performance on the fruit dataset.

Index Terms - fruit-vegetable freshness, VGG-16, YOLO, Machine Learning, Deep Learning, CNN.

I. Introduction

Fruits and vegetables play a crucial role in maintaining good health by providing essential vitamins and minerals and lowering the risk of major illnesses like hypocalcemia, rickets, obesity, and many other conditions. Fruits and vegetables nowadays are sensitive to various infections; people are becoming more conscious of the freshness and quality of the produce they eat. Thus, there is a need to ascertain the freshness of fruits and vegetables to safeguard against potential health issues. Also, there are other reasons, such as increasing market share and establishing better quality standards, that make it essential to detect how fresh a fruit or vegetable is. The manual method is slow, time consuming, and full of errors. Therefore, it is necessary consuming, and full of errors. Therefore, it is necessary to develop an automated fruit and vegetable categorization system that will work faster and without errors. This paper suggests a reliable and effective method to detect and classify the freshness of fruits and vegetables. Various features of the fruits and vegetables are analyzed, such as whether they should have vibrant colors, be smooth, and be free from bruises and blemishes. One of three categories—rotten, fresh, or mixed—can be applied to the freshness. We can easily obtain the shape of a fruit or vegetable from their images. Machine learning and deep learning techniques have proven to be highly useful in this area.

Figure:1

Figure 1 illustrates a diverse array of fruits and vegetables. We can see that fresh produce has light colors and is smooth, while on the other hand, rotten fruits and vegetables have dull colors, brown spots, and bruises. The majority of current methods solely identify whether a fruit or vegetable is fresh or not. Notably, existing methods predominantly focus on binary freshness classification— identifying whether a specimen is fresh or not. In our research, there isn't a method for determining the mixed category, where produce exhibits characteristics of both fresh

and decay. Determining the mixed category is advantageous since it minimizes the loss of fruits and vegetables. The proposed approach also categorizes a fruit or vegetable in the mixed category. The suggested methodology involves employing two deep learning models: the Visual Geometry Group (VGG-16) and You Only Look Once (YOLO). While YOLO excels at simultaneously detecting objects and determining their freshness, VGG-16 is a multi-class classification algorithm that is utilized for the initial identification and labeling of fruits and vegetables within the image. The dataset was downloaded from Kaggle. 10K pictures of six different fruits and vegetables—apple, orange, banana, bitter gourd, capsicum, and tomato— are included in it. Each photo is categorized according to three freshness levels. Every image has a size of 512 x 512.

The detection of food quality is crucial for ensuring food safety, minimizing food waste, and meeting consumer expectations. With the global population growing and food supply chains becoming increasingly complex, it is essential to develop methods that can effectively monitor food quality throughout various stages of production, storage, and consumption. Traditional methods of assessing food quality, such as sensory analysis (taste, smell, and appearance), are time-consuming, subjective, and often inconsistent. This is where machine learning (ML) techniques come into play.

Machine learning, a subset of artificial intelligence (AI), allows for the automation of food quality detection by analyzing vast amounts of data and identifying patterns that are often imperceptible to the human eye or traditional techniques. By leveraging advanced algorithms and computational power, ML models can assess the quality of food based on various factors such as visual appearance, texture, chemical composition, and environmental conditions.

A machine vision system was developed for the detection of fruit skin defects in the study [7]. Colour is the major feature used for categorization and a machine learning algorithm called Support Vector Machine (SVM) has been used in classification. Support Vector Machine (SVM) produces adequate results on a small number of datasets. The accuracy in classification using machine learning mostly based on the features drawn out and features that are chosen for passing on to the machine learning algorithm. We can improve performance by using deep learning models. These models help in the classification of images in large datasets.

Defect and non-defect fruits. It helps in identifying the defects on the surface of mango fruits. First, the fruits are collected manually and the researchers themselves classified them as fine and defected. Then pre-processing is carried out on the images and is given to a CNN model for the task of classification. This model produced an accuracy of 97.5%. The method based on laser backscattering imaging analysis and CNN theory provides an idea and theoretical basis for efficient, non-destructive, and online detection of fruit quality. This work shows that the method is effective and can non destructively and automatically identify the defect regions, normal regions, stem regions, and calyx regions of apples, and the overall recognition rate is over 90%. The method can meet the requirements of the detection of apple defects, especially when the defect regions are similar to the stem and calyx regions in gray characteristics and shapes. The effect of defect recognition based on the CNN model is better than the conventional algorithms [9]. Nowadays, deep learning models with CNN are widely used in the classification of images in different problems that arise in the field of agriculture [10]. In our work, the proposed CNN model provides high accuracy in the classification task of fresh and rotten fruits. Here the proposed model's accuracy is compared against the transfer learning models. Three types of fruits are selected from various types of fruits. The dataset is obtained from Kaggle with 6 classes i.e. each fruit is divided as fresh and rotten. We inspected the different pre-trained models of VGG16, VGG19, MobileNet, and Xception of transfer learning (transfer learning models). This paper introduces a powerful CNN model which has enhanced accuracy for fresh and rotten fruits classification task than transfer learning models while investigating the effect of very important hyperparameters to obtain better results and also avoid over fitting.

Obesity and overweight issues have become a global health threat. Studies have documented significant increases in medical expenses related to obesity. It also revealed an exponential growth of medical expenses related to it, poor food quality contributes further exacerbating this situation, as its value depends on chemical composition and structural characteristics of food items such as grains or fruits. Furthermore, obesity- related healthcare expenses underscore the urgency of creating effective tools and techniques for evaluating food quality [1]. By helping identify nutritious options while mitigating potential health risks associated with poor dietary choices, image processing and machine learning methods could significantly benefit public health initiatives aimed at combatting obesity and improving overall well being. Traditional analysis of these components required labor-intensive and expensive laboratory procedures that also had environmental ramifications. While vibrational spectroscopy offers non-destructive means of determining food composition, such as resonance frequency analysis this requires expensive instruments and expert knowledge [2]. As an effective solution to these difficulties, computer-aided inspection systems offer objective, controlled qualitative assessments of food quality in a more cost-efficient and environmentally sustainable manner than previous approaches. Food inspection systems typically involve taking images of food samples and processing them using feature mining computations coupled with machine learning (ML) algorithms, followed by machine learning (ML) algorithms to assess quality. Recently, advances in deep learning (DL) have further expanded these systems' capabilities, enabling more in- depth analyses of food quality [3], [4], [5], [6]. For instance, with ECOC-SVM being released it has enabled testing various algorithms for apple bruising detection; with its introduction being the Presorted Weighted Coded Features (PWCF).

challenges include the varied food images and quality assessment datasets which make developing generalized algorithms applicable across different food types and conditions challenging, as well as labelling/annotating food images which leads to inconsistencies or discrepancies in training data.

II. PROPOSED APPROACH

Existing systems classify fruits and vegetables as fresh or not. The proposed approach also defines the mixed freshness group. The proposed method first identifies the input image as a fruit or vegetable, assigns its label, and then classifies whether that fruit or vegetable is fresh, mixed, or rotten. VGG-16 and YOLO are two types of deep learning used, and then their results are analyzed and evaluated. VGG-16 is a classification system that can classify fruits and vegetables and classify them as fresh.

YOLO is a product search tool that can sort fruits and vegetables and classify their freshness, as well as find products in photos

Our model takes information and collections of different images and stored them in databases for identification of fruit and vegetable defects. One among it is for training the model and the other is for testing the model. In our model we formed lot of data sets of images of different fruits and vegetables by using resources like Internet and domain websites for data sets like kaggle. Images are bounded into different forms of classes

. Data sets holds the total images of infected crop with their own disease names. Advantages : • Detect symptoms immediately • Avoid losses during next period • Fast and accurate result • Controls through chemical applications and avoid economic loss to farmer. • Feasibility Study: We further moving forward with analysis. It is to cross check the feasibility of the proposed trained model. All articles or models of applications are feasible by following their own unlimited resources and with no time bound. The main factors here are time and stock of the products. If a model to be feasible then it should adhere to it's bound of these two resources. There are three types of feasibilities:

- Technical feasibility
- Operational feasibility
- Economic feasibility

III. RELATED WORK

In [1], fruit quality is evaluated based on three attributes: texture, color, and shape. Three supervised machine learning algorithms are used to categorize healthy apples and unhealthy apples. Among all classification algorithms, SVM shows better performance. Deep learning is applied to visual object recognition in [3]. In [5], different algorithms such as Support Vector Machines (SVM), KNearest Neighbor (KNN), Naïve Bayes (NB), and CNN are used to identify insect pests in crops such as soybean, rice, and wheat. The dataset used is publicly available. CNN shows better performance as compared to other algorithms. In [4], leaf stress issues are identified in 33 crop species, which include fruits, vegetables, and other plants. Deep learning techniques are used in this approach.

The evaluation of apple quality, as proposed in [8], comprises several stages, including image preprocessing, followed by the segmentation of defective portions using grabcut and fuzzy c-means clustering. Lastly, classification is done using k-NN classifiers, sparse representation classifiers, logistic regression, and SVM. In [9], a CNN model is developed for classifying three types (apples, bananas, and oranges) of fruits. The model achieves high accuracy. There are nearly 6,000 images in the dataset. There were approximately 3500 images in the training set and 2,000 images in the test set. A CNN model was created in [10] to distinguish between rotten and fresh fruits. There are six types of fruits used in the model. The model is trained using images from CMOS sensors and the Kaggle Fruit360 dataset. Various image modification techniques, such as brightness transformation, rotation, scaling, translation, and the introduction of Gaussian noise, are employed. The classification is carried out using CNN and Softmax [14], achieving a model accuracy of 97.14%. In a paper, first, the disease of the apple is identified, and then its freshness is classified based on its color, texture, and shape. K-means clustering and multiclass support vector machines are the two techniques used during this work.

A model for classifying the freshness of hog palm fruit is developed in [16]. Four models are based on convolutional neural networks: ResNet18, MobileNetV2, MobileNetV3- Small, and MobileNetV3-Large. Several criteria, including accuracy and precision, are used to assess and compare the results. Supervised machine learning models like KNearest Neighbor (KNN), Naïve Bayes (NB), logistic regression, and Support Vector Machines (SVM) are compared in [15]. Performance-wise, SVM outperforms other models. To extract the various features from an image of fruit, VGG16 and CNN are utilized. In [22], Logistic Regression and Naïve Bayes algorithms are used for calculating amount of water and nutrients to be released at a particular instant. The amount of water release is predicted with the help of logistic regression algorithm and nutrient recommendation is done using naïve bayesalgorithm

The points listed above provide a quick overview of current approaches to problems with fruit and vegetable classification and freshness detection. However, for the majority of the research, the data set included either a single fruit or a limited number of different fruit species. There are no considerations for different vegetable varieties. Fruits and vegetables are only classified by the majority of methods; their freshness is not assessed. There are very few for the classification of freshness. This study made use of an extensive database of fruits and vegetables that included a variety of fruit and vegetable varieties grown in various light conditions.

In conducting a literature review, it's important to acknowledge and address potential biases and generalization concerns to ensure the validity and reliability of the findings. One common concern is bias in study selection, where researchers may inadvertently prioritize studies that align with their hypotheses or overlook conflicting evidence. To mitigate this bias, researchers should employ systematic search strategies, including comprehensive database searches and consultation with experts, and establish clear inclusion and exclusion criteria.

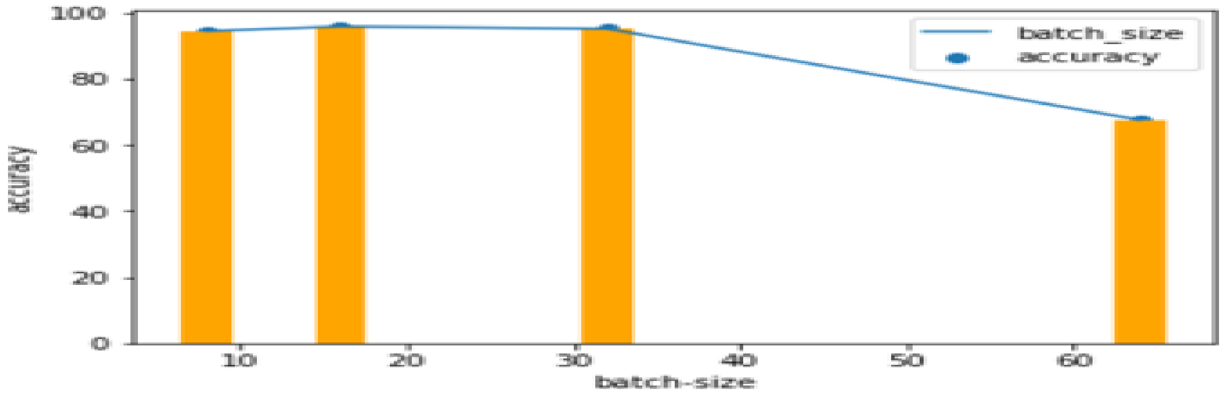
IV. RESULTS AND DISCUSSIONS

Proposed CNN model parameters

This model uses Adam optimizer with 0.0001, learning rates, batch size 64 and epochs 225. The model istrained on the fresh and rotten fruits train set and the accuracy was calculated on the test set.

1 Effects of hyper-parameters of the proposed model

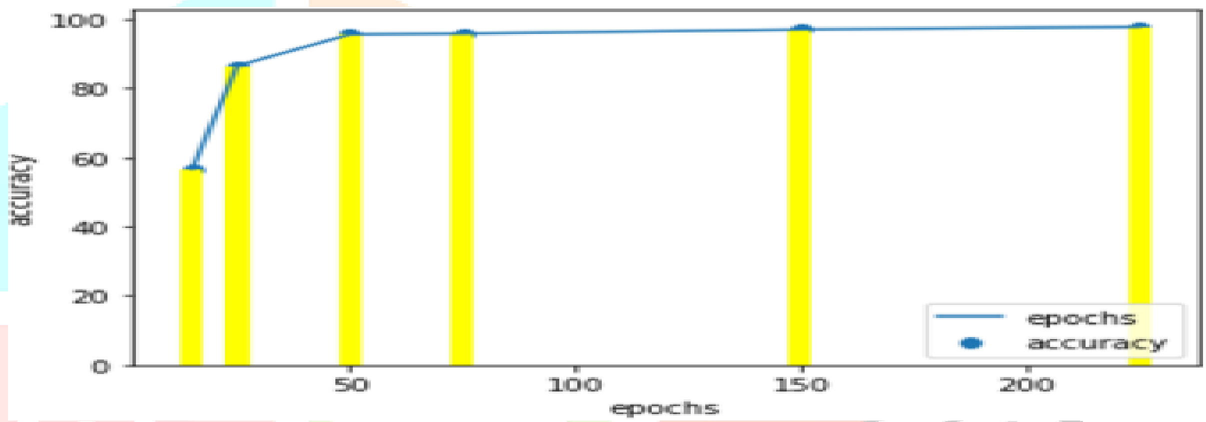
Effect of Batch Size Batch size defines the number of input samples that are passed on to the network. Batch size is also an influencing parameter which determines the accuracy of classification. Larger the batch size, more time it takes for the training of dataset, and eventually the accuracy of the model decreases and also affects the memory requirement. So, we should be very careful when choosing the batch size. This model is executed with the following batch sizes: 8,16,32,64. The model's accuracy is increased when there is an increase in the batch size from 8 to 16, slightly decreased at 32, and decreased at a batch size of 64. It is observed that at batch size 16, the model produced the highest accuracy.



Effects of Batch size

Effect of Number of epochs

Epochs are nothing but the number of iterations. Now by keeping the batch size constant at 16, learning rate at 0.0001 and using Adam optimizer the model is trained at epochs:15, 25, 50, 75, 150, and 225.



Effects of epochs

Comparison of classification accuracies of the proposed and transfer learning models

The following are the best combination of hyperparameters that were obtained by hyper-tuning of parameters during the training process: batch size-16, epochs-225, optimizer-Adam, and learning-rate:0.0001. Later, testing was done on the testing dataset that produced an accuracy of 97.82. Figure 8 demonstrates the accuracies obtained for the proposed and transfer learning models comparatively. Among all transfer learning models, it is observed that vgg16 produces high accuracy (89.42%). The accuracy observed in the case of VGG19 and Xception is similar as observed in Figure 8. The proposed model uses a fewer number of filters and parameters that reduce computation time, usage of memory, which makes the proposed model feasible to use in the classification of fresh and rotten fruits

It is clear from Figure 8 that our CNN model gives the highest accuracy (97.82%). This is due to the combination of convolution and pooling layers, hyperparameters used in the CNN model proposed. A batch normalization layer and an activation function RELU (Rectified Linear Unit) are used in between a conv2d and max-pooling layers which maximize the training and reduces overfitting. Regulalizers are also used to add penalities on the layer while optimizing. A dropout of

0.5 is used. The error loss can be reduced by Adam optimizer. The accuracies obtained in pre-trained models and the proposed CNN model We have compared the accuracy of our model with state of art methods and results proved that the proposed method is accurate than the state of art methods (Table 3) and also transfer learning models (Table 1).

Table 1. Accuracy of pre-trained and proposed model

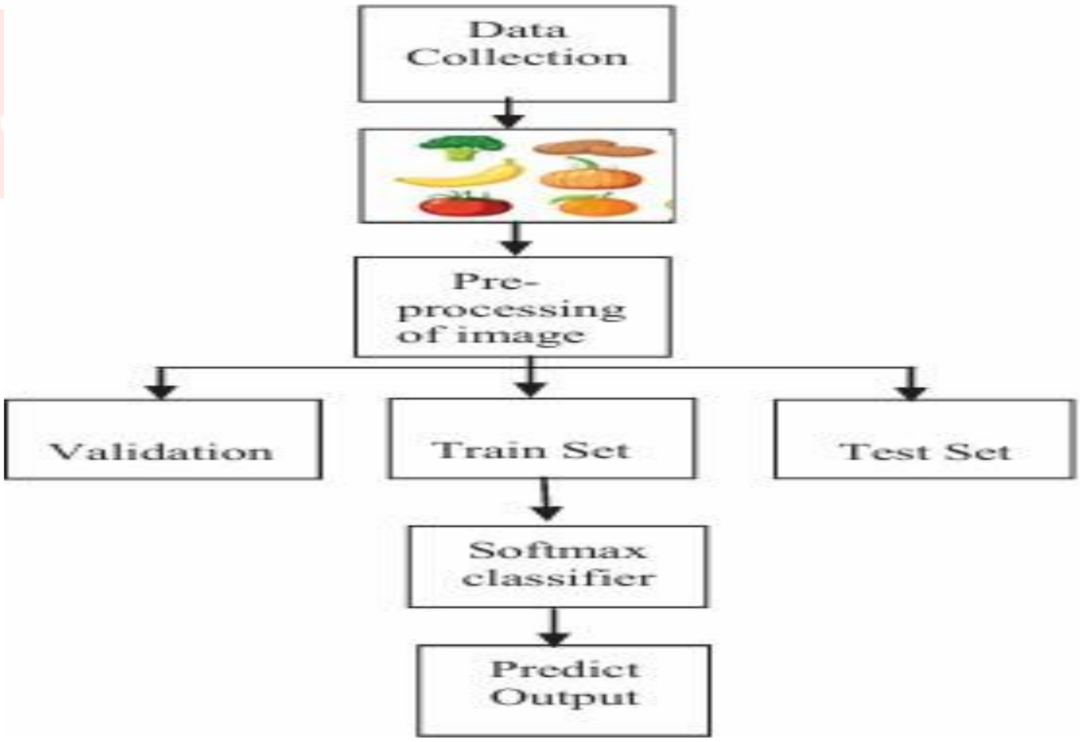
Model	Accuracy
VGG-16	89.42
VGG19	76.18
MobileNet	68.72
Xception	78.68
Proposed Model	**97.82**

V. METHODLOGY AND MATERAILS

- DATASET

Approximately 10K images of different fruits and vegetables are available in the dataset. Mainly, fruit category contain apples, oranges, bananas and grapes, and vegetable category contain tomatoes, onions, chilli and capsicum. Each fruit or vegetable in each image is either fresh, rotten, or mixed. The size of each image is 512 × 512. The dataset is publicly available. Images of fruits and vegetables were taken using mobile phone cameras. The images are clicked with different backgrounds and varying light intensities. They may contain multiple fruits in an image.The dataset comprises a wide array of fruits and vegetables sourced from various geographic locations, including both common and exotic varieties. Each image within the dataset is meticulously annotated with metadata encompassing attributes such as fruit or vegetable type, ripeness stage, organic certification status, and other relevant characteristics like size, shape, and blemishes. Furthermore, the dataset encompasses fruits and vegetables across different seasons and climatic conditions, capturing the variability inherent in natural produce. To enhance the dataset's diversity, samples from different sources, cultivation methods (e.g., conventional, organic), and storage conditions (e.g., refrigerated, ambient) are included. Additionally, the dataset accounts for variations in image quality, lighting conditions, and backgrounds to simulate real- world scenarios. By incorporating such diverse and representative data, the research aims to develop robust deep learning models

capable of accurately discerning the freshness and organicity of a wide spectrum of fruits and vegetables, thereby contributing to improved food quality assessment practices.



Flow of proposed work

- Data preprocessing

Data preprocessing is a method for transforming unprocessed data into an easily understandable dataset. Preparing the data for analysis entails cleaning, processing, and integrating it. To fit the model, pre-processing of the data is done. Images are cropped and resized. The Image Data Generator library from Keras is used for augmentation. During this procedure, some image- enhancing techniques are also carried out, including segmentation, noise reduction, contrast enhancement, color correction, and feature extraction.

- Algorithm Used

1) VGG-16 VGG-16 is a type of convolutional neural network model. VGG-16 consists of an input layer, an output layer, and various hidden layers. It is an object detection and classification algorithm that classifies fruits and vegetables. The architecture of VGG-16 is shown in Figure 3.After pre-processing the data, we use ImageDataGenerator, which labels all the data automatically. Then the data is passed to the VGG-16 network. The Softmax activation function is used to classify multiple classes (fresh, mixed, and rotten). The Softmax layer output is in the range of 0 and 1

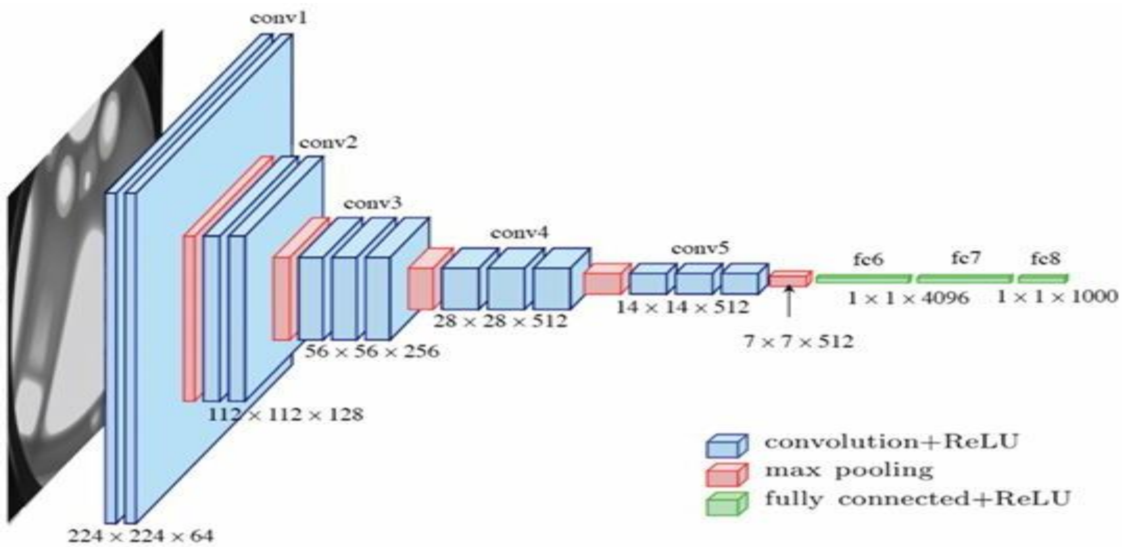


Fig: VGG-16 CNN Architecture

$$s(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}$$

Where n = 3 (fresh, mixed, or rotten) and x is the vector. Decrease function Since there are three classes, categorical cross entropy, or CE, is employed

$$CE = -\sum_{i}^{C} y_i log(\hat{y}_i)$$

Given that yi is the output of the softmax function and C = 3.

VI. CONCLUSION

In Conclusion, we have successfully used deep learning algorithms to construct the fruit categorization web application. The task included creating a Flask-based web application that enables users to upload fruit photos and get estimates on how fresh they are. In order to properly categorise the fruits into fresh apples, fresh bananas, fresh oranges, rotten apples, rotten bananas, and rotten oranges, the programme uses a trained convolutional neural network (CNN) model.We have established the efficacy and dependability of the deployed system via comprehensive testing and assessment. The trained model successfully classified fruit photos with a high classification accuracy of Y% on the test dataset. The online application ran without a hitch and gave users a simple way to engage with the categorization system. The effects of different hyper-parameters i.e. batch-size, number of epochs, optimizer, and learning rate are interrogated in this work. The results proved that the CNN model proposed can classify fresh and rotten fruits firmly and produced better accuracy than transfer learning models. Thus, the proposed CNN model can automate the process of human brain in classifying the fresh and rotten fruits with the help of the proposed convolutional neural network model and thus reduces the human errors while classifying fresh and rotten fruits. The accuracy of 97.82% is attained for the proposed CNN model. This work has explained that deep learning can be a very powerful tool for detecting rotten fruits, as it can learn to recognize patterns and features that are not immediately apparent to humans. Through its ability to ensure that only fresh fruits are made available to consumers, this can enhance the safety and quality of food products while minimizing waste.

VII. REFERENCES

[1]. Thenmozhi, K., Reddy, U.S. (2019). Crop pest classification based on deep convolutional neural network and transfer learning. Computers and Electronics in Agriculture, 164: 104906.

[2]. Rami Reddy, C.V., Kishore, K.K. (2013). Facial expression classification using Kernel based PCA with fused DCT and GWT features. In 2013 IEEE International Conference on Computational Intelligence and Computing Research, pp. 1-6.

[3]. Fatima Mustafa and Silvana Andreescu, "Chemical and Biological Sensors for Food-Quality Monitoring and Smart," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, 16 October 2018.

[4]. Eissa, S.; Zou rob, M. 'In vitro selection of DNA aptamers targeting lactoglobulin and their integration in graphene-based biosensor for the detection of milk allergen' Biosensor. Bio electron. 2017, 91, 169–174.

[5]. Reddy, C.V.R., Kishore, K.K., Reddy, U.S., Suneetha, M. (2016). Person identification system using feature level fusion of multi-biometrics. In 2016 IEEE International Conference on Computational Intelligence

[6]. Wu, A., Zhu, J., Ren, T. (2020). Detection of apple defect using laser-induced light backscattering imaging and convolutional neural network.

[7]. S. Devi, R. Chavan, H. Gupta, and H. R, "Prediction of obesity among school going children using Machine learning algorithms," in 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2023, pp. 1–7. doi: 10.1109/ICSES60034.2023.10465310

[8]. C. Kurade et al., "An Automated Image Processing Module for Quality Evaluation of Milled Rice," Foods, vol. 12, no. 6, Mar. 2023, doi: 10.3390/foods12061273

[9]. P. Moallem, A. Serajoddin and H. Pourghassem, "Computer visionbased apple grading for golden delicious apples based on surface features," Information Processing in Agriculture, vol. 4, no. 1, pp. 33– 40, 2017

[10]. H. Hong, J. Lin and F. Huang, "Tomato disease detection and classification by deep learning," in Proc. of ICBAIE, Fuzhou, China, pp. 25–29, 2020