How to Cite:
Kanagaraju, P., Aushiq, N. M., & Vanan, R. T. (2022). Disease detection and analysis in fruits using image processing. *International Journal of Health Sciences*, *6*(S8), 1198–1211. https://doi.org/10.53730/ijhs.v6nS8.9879

Disease detection and analysis in fruits using image processing

Dr. P. Kanagaraju
Assistant Professor in Computer Science and Engineering, K. S. Rangasamy College ofTechnology, Tiruchengode-637 215, Namakkal District, Tamil Nadu, India
Corresponding author email: kanagaraju@ksrct.ac.in

N. Mohammed Aushiq
Students of Computer Science and Engineering, K.S.Rangasamy College of Technology,Tiruchengode-637 215, Namakkal District, Tamil Nadu, India
Email: mohammedaushiq2001@gmail.com

R. Tamil Vanan
Students of Computer Science and Engineering, K.S.Rangasamy College of Technology,Tiruchengode-637 215, Namakkal District, Tamil Nadu, India
Email: tamillimat19@gmail.com

***Abstract*---**Fruit diseases are a major problem in the agricultural industry where losses in economic and production are occured. In the existing system, K-means clustering algorithm is used to find whether the fruit is infected or not. Due to low accuracy, it will take more time to show the exact result. In this project, an image processing approach is proposed for identifying apple fruit diseases based on Convolutional Neural Network(CNN). In CNN algorithm, fruit image details are taken by the existing packages in this work. However, it can take a few moments. So, this proposed system can be used to identify fruit diseases quickly and automatically. This proposed approach is composed of the following main steps: getting input image, Image Preprocessing, Identifying affected places, highlighting those affected places, Verifying training set, showing results. Few types of fruit diseases, namely bitter rot, sooty blotch, powdery mildew and fungus images were used for this approach. This approach was tested according to fruit disease type and its stages, such as fresh and affected. The algorithm was used for detecting the disease of the fruit. Images were provided for training, such as fresh apple images, fresh banana images, bitter rot images, sooty blotch images, powdery mildew images and fungus images.. Before the image processing, images were converted to color models, because of finding the most suitable color model for this approach. Local Binary

Pattern was used for feature extraction and Support erosion method was used for creating the model. In this approach, fruit diseases can be identified based on its accuracy better than existing algorithms.

***Keywords*---**image processing, convolutional neural network algorithm, image acquisition, image preprocessing, image segmentation, applying training set.

Introduction

The traditional method for identifying and diagnosing fruit illnesses relies solely on expert observation with the unaided eye. Due to their availability's remote locations, consulting specialists can be costly and time- consuming in some underdeveloped nations. Automatic identification of fruit illnesses is critical in order to detect disease symptoms as soon as they develop on growing fruits. Fruit infections can result in significant productivity and quality losses during harvesting.

To determine what control measures to use next year to avoid losses, it is critical to recognise what is being observed. Apple rot and apple blotch are two frequent apple fruit diseases. Infections of apple rot cause slightly depressed, round brown or black patches which may be bordered by such a crimson haloApple blotches is a fungus disease that causes dark, uneven, or lobed margins on the skin of the apple. Machine vision has been used in the business to automate the size and colour examination of apples. However, due to the natural variety of skin colour in types of fruits, the wide range of fault types, and the presence of diseases, detecting abnormalities remains difficult. Fruit studies can be identified by visible patterns of individual fruit, and that it's essential to measure nutrition and discover disease inside a fruit. Pesticides, fungicides, and chemical applications, for example, can also be used to promote disease control and hence improve quality.

Deep learning, also known as neural networks, is a subset of machine learning that employs a computing model that is heavily influenced by the structure of the brain. Deep learning is already in use in Google search and image search, allowing you to image-search a term such as 'hug.' It's used to deliver Smart Replies to your Gmail inbox.It's in speech and vision. It will soon be used in machine translation, I believe." said Geoffrey Hinton, considered the Godfather of neural networks. Deep Learning models, with their multi-level structures, as shown above, are very helpful in extracting complicated information from input images. Convolutional neural networks are also able to drastically reduce computation time by taking advantage of GPU for computation which many networks fail to utilize.

Image classification using CNN is most effective. First and foremost, we need a set of images. In this case, we take images of beauty and pharmacy products, as our initial training data set. The most common image data input parameters are the number of images, image dimensions, number of channels, and number of levels per pixel. There are some illnesses that spread to the tree's branches,

leaves, and other parts, including the twigs. Apple scab, apple rot, and apple blotch are three common illnesses that affect apples. Apple scabs are grey or brown corky patches on the skin. Infections of apple rot cause slightly depressed, round brown or black patches with a crimson halo. Appleblotch is a fungal disease and appears on the surface of the

fruit as dark, irregular or lobed edges. Visual inspection of apples is already automated in the industry by machine vision with respect to size and color. There are various approaches available such as spectroscopic and imaging technology, applied to achieve better plant disease control and management. The increased amount of commercialization of agricultural farms are always on the lookout to reduce manpower in whatever way possible without affecting productivity. Utilizing automatic harvesters would greatly streamline the procedure and is a specific consideration. Fruit detecting systems are mostly used in robotic harvesting. But from the other hand, the technology can be specifically designed to be suited for various applications including disease detection, maturity detection, tree yield monitoring, and other comparable processes. Fruit varieties are being exported all over the world as cold storage facilities and transportation improve.

It becomes the necessity of maintaining the highest level export quality which is mainly carried out by visual checking by experts. Given how far away the farms are, this is both expensive and time-consuming. Precision agriculture enables farmers to supply sufficient and cost- effective information and control technology as a result of advancements and disclosures in numerous industries. Systematizing agricultural inputs, increasing profits, and minimising environmental harm are the goals. As a result, in this paper, a strategy for detecting and classifying fruit illnesses is provided and experimentally confirmed. This system analyses input in the form of an image of fruit and determines whether it is contaminated or not. The technique which helps the farmers to identify disease properly by using this proposed work.

Fruit industry is one of the major drivers to grow the economy of the country. There is a possibility of erroneous sorting and packaging of fruits due to manual inspection and lack of knowledge of quality evaluation. The farmers are under pressure for demand of rapid supply due to shortage of skilled workers and rising labor costs. In such a scenario, automation can reduce the costs by promoting production efficiency.In Agricultural image processing significant research has been done for identification of fruits and detection & quantification of diseases. Most of the previous works are based on C-Mean, K-Mean and KNN for identification and quality analysis of fruits. In this paper an automatic system is proposed, which is less time consuming and cost effective for farmers to identify the type of fruit and grade it according to appearance of defects and disease.

This research considers five types of fruits i.e. apple, mango, orange, pomegranate and tomato with two common appeared diseases i.e. fruit rot & anthracnose. The Symptoms of anthracnose seem at first as little, roundabout, marginally depressed sores on the outside of aging organic products. The spots fast spread, become severely discouraged, and develop a water-splashed appearance directly beneath the skin of the foods developed from the ground rings framed in the focal

point of the injuries. The organic product spoils malady advances with slow increment of sore, causing a vile decay, and microscopic organisms may overflow the epidermis and split the influenced region. Side effects can go from shallow flecking to depressed sores, dark colored dark, dry.

Literature Review

[1] Recently, many people have done research for detecting fruit and vegetable diseases using image processing and deep learning. According to the research paper, authors had used image processing technology to identify the pomegranate diseases. Image preprocessing was the first step of the methodology. Image resizing was done under the image preprocessing. Because a digital camera had been used to capture the images in this study. The size of those images was very large and took more time to process. So all images were resized to 300 x 300 PX. Morphology, color and CCV features were used for feature extraction. K-means clustering technique was used for partitioning the training dataset according to their features. After the clustering, SVM was used for classification to identify the image as infected or non- infected. An intent search technique was provided to find the user's intention. The best result was obtained using morphology feature extraction. Experimental evaluation of this approach was effective and 82% accurate to identify pomegranate disease.

[2] The authors presented the image processing based approach for fruit disease detection. First, read input image and transformed it from RGB to L*a*b color space. Because the color information in the L*a*b color space is stored in only two channels. Input images were partitioned into four segments using the K-means cluster in this research. Because of the empirical observations it was found that using 3 or 4 clusters yields good segmentation results. GCH, LBP, CCV and CLBP were used for feature extraction. More accurate results could be taken using the CLBP feature extraction technique. K-means clustering was used for segmentation. Those segmented images were extracted to label each pixel in the image. SVM algorithm was used for training and classification of fruit disease. Authors used apples as a test case and evaluated the classification model for three types of apple diseases, which were apple rot, apple blotch and apple scab. The accuracy of this approach was achieved by up to 93%.

[3] The author had used SVM classification for identifying and classifying the grape leaf diseases. Grape leaf images were taken using a digital camera and those were used for both training and testing the system. Collected images included the leaves infected by Powdery Mildew and Downy Mildew. Removing background noise and resizing to 300*300 PX to improve the image quality were done under the image preprocessing. Gaussian filtering had been used to remove noise in the image. K- means clustering was used for segmenting an image into three groups. Features were extracted based on both color and texture for taking accurate disease information. Finally, the classification model was used to detect the leaf disease. LSVM was used in this research for the classification of leaf diseases. This system could detect and classify the examined disease successfully. The accuracy of this system was 88.89%.

[4] Authors had used image processing technology for identifying the leaf

diseases. First authors selected the plants, which were affected by the disease and then took the snapshot of the diseased leaf. Contrast enhancement and converting RGB to HIS was done under the image preprocessing step. K-means clustering algorithm was used to cluster the object based on the feature of leaf into k number of groups. The SVM algorithm had been used in this system for classification purposes. SVM is a statistical learning-based solver. Finally, when entering a diseased leaf image to a system, the system was able to detect the leaf disease successfully.

[5] Authors had presented the image processing based system to identify pomegranate fruit diseases. This fruit is mainly affected by Bacterial Blight, Anthracnose and Alternaria. After capturing the disease images, image resizing, filtering, segmentation, morphological features were used to preprocess the images. Image segmentation isthe process of dividing the image into multiple parts. color-based segmentation was used in this research, such as clustering, YCbCr, RGB, L*a*b and HSV. However, the best performance in terms of segmentation error was achieved by the HSV and YCbCr. Morphology, texture and color features were extracted for classification purposes. HIS color model and color histogram techniques had been used to color feature extraction. Under the morphology feature extraction, boundary extraction was used to identify the region and shape. The deteriorated photos were removed from the original image in order to obtain a shape from the healthy fruit image. The Gabor filter was employed to extract textural features. Fruits with and without diseases were distinguished using the Minimum Distance Classifier following picture training and testing (MDC).

Proposed Methodology

There are five phases in this methodology. Those are Image Acquisition, Image Preprocessing, Image Segmentation, Applying training dataset and Experimental results.

Image Acquisition

In this phase, the sample images are collected, which are required to train the classifier algorithm and build the classifier model. Yellowish or Reddish fruit varieties were selected to take sample images. Because the yellowish variety is widely cultivated on our site. Healthy and diseased fruit images were taken by using mobile phone digital cameras and used for both training and testing the classifier algorithm. Images were taken in different angles, under the different environmental and lighting conditions. The standard JPG format was used to store these images. In this study, images were collected from farms in different regions. Fruits infected by scab disease and woodiness virus that had been included in collected images.

Image Preprocessing

After the image acquisition, image processing was done to improve the image quality. All original fruit images were stored in one folder. Those images were

named as we like our wish to take any value of numbers. Only horizontal images were rotated by 90 degrees and resized by 200x300 pixels.Vertical images were resized by 200x300 pixels and when the width and height of the image are same,those images were resized to 250x250 pixels.When the image size is too large,the processing task takes more time.After that,one of the noise reduction methods was used to remove the noises from images and increase the sharpness of image.Literally preprocessed images were saved in folder.

Image Segmentation

The third phase of the methodology is image segmentation. As the first step, all preprocessed images were converted into L*a*b, HSV, Gray color models and kept one in the original way (RGB). Because identifying suitable color models for preprocessing is one of the outcomes of this research. After that, the image was converted to binary format. This format values were clustered using the CNN algorithm. According to the algorithm used an image segmentation was done.

Converting to Hue Saturation Value

found=0 img1=cv2.imread(source) folder='E:/FruitRot/apple fresh' for filename in os.listdir(folder):
img2 = cv2.imread(os.path.join(folder,filename))if img2 is not None:
img1_hsv=cv2.cvtColor(img1,
cv2.COLOR_BGR2HSV)
img2_hsv=cv2.cvtColor(img2,cv2.COLOR_BGR2HSV)

Applying Training Set

The fourth phase of the methodology is applying training set images. The segmented outputs were done, which were created using feature extraction. However, three image sets were created to do experiments. Preparation of those image sets is discussed here. Field expertise support was taken for the categorization of images and each image was selected from the categorized sets of an image randomly.

Calculating and normalizing

hist_img1 = cv2.calcHist([img1_hsv], [0,1], None,[180,256], [0,180,0,256])
cv2.normalize(hist_img1, hist_img1, alpha=0, beta=1,
norm_type=cv2.NORM_MINMAX);
hist_img2 = cv2.calcHist([img2_hsv], [0,1], None,[180,256], [0,180,0,256])
cv2.normalize(hist_img2, hist_img2, alpha=0, beta=1,
norm_type=cv2.NORM_MINMAX);

Experimental Results

After applying the training set images, base folders were used for identifying fruit disease according to its name. These files are called the “Classes dataset”. Another way is counting the number of affected places to identify fruit diseases according to its stage. This method is called an alternative method. Every training

and testing time, rows of training files were shuffled randomly to increase the accuracy of the model. Each training file was verified and tested five times and accuracy was taken. Average of those accuracies was taken as the accuracy of each model. Using this image dataset, types of diseases were found.

find the metric value

```
metric_val = cv2.compareHist(hist_img1, hist_img2,
cv2.HISTCMP_BHATTACHARYYA)
if metric_val == 0.0:
print("Fruit") cv2.imshow("Fruit",img1)found=found+1
```

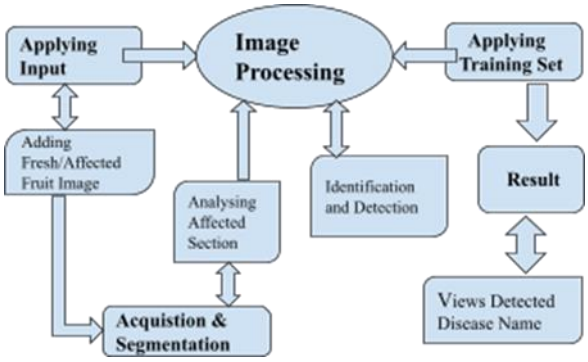

Figure 1. Block Diagram of Disease Detection and Analysis in Fruits Using Image Processing

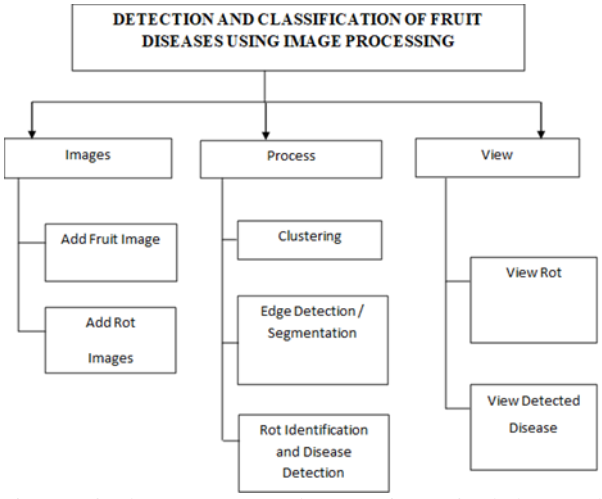

Figure 2. Flow chart of Disease Detection and Analysis in Fruits using Image Processing

Input Design

The method of changing user-generated inputs into a format that computers can interpret is known as input design. One of the most costly aspects of running a

computerized system is input design, which is frequently the main issue with a system. Imaging is a technique used to apply certain operations to an image in order to produce an improved image or to draw out some relevant information from it. It is a form of signal processing where a picture serves as the input and the output could be another image or features related to the original image. A system's fault input design and mechanism are typically the source of a significant number of issues. Each step of the input design process needs to be carefully considered. The design of the input should be made the input as the over to the numerous networks in the reliable area that should be passed as the installation in the remote network.

It has the following constraints in the input database.

- All the files from the disk should be acquired by data.
- It is suitable for more available data clearance and made available.
- The menu of design should be understandable and it is in the right format.

The system takes input from the users, processes it and produces an output. Input design is a link that ties the information system into the world of its users. The system should be user-friendly to gain appropriate information to the user. The decisions made during the input design are the project gives the low time consumption to make the sensitive application simple. When applying the project it provides the low man- power attrition with the reasonable output. The amount of funds that the company can spend into the research and development of the system is limited. Inputdata of a system may not necessarily be raw data captured in the system from scratch. These can also be the output of another system or subsystem. Theexpenditures must be justified. Thus, the developed system as well within the budget and this was achieved because Most of the technologies used are freely available. Only the customized products had to be purchased.The goal of designing the input data is to make data entry as easy and error free as possible.In the project, the forms are designed with easy-to-use options such as selecting the master records through a dropdown list in transaction forms. The coding is being done such that proper validations are made to get the perfect input. No error inputs are accepted. The end users need not to give the id themselves.

Example images of training set

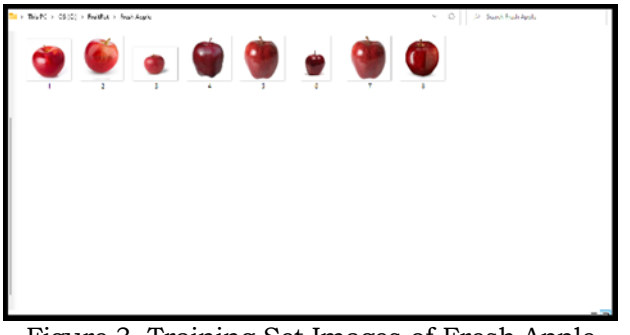

Figure 3. Training Set Images of Fresh Apple

Apples are an incredibly nutritious fruit that offers multiple health benefits. An apple is an edible fruit produced by an apple tree. Apple trees are cultivated worldwide and are the most widely grown species in the genus Malus. The tree originated in Central Asia, where its wild ancestor, Malus sieversii, is still found today.

Figure 4. Training Set Images of Fresh Banana

Banana is a long yellow fruit, found on the market in groups of three to twenty fruits, similar to a triangular cucumber, elongated and normally yellow. An elongated tropical fruit usually conical with soft pulp enclosed in a soft usually yellow skin, a bunch of bananas sliding on a banana peel.

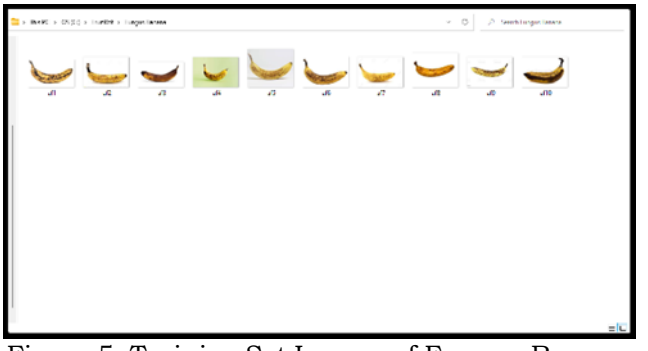

Figure 5. Training Set Images of Fungus Banana

Fungus banana, popularly known as Fusarium disease, is a lethal fungal disease caused by the soil- borne Fusarium oxysporum. This fungus also causes trouble for melons and tomato.

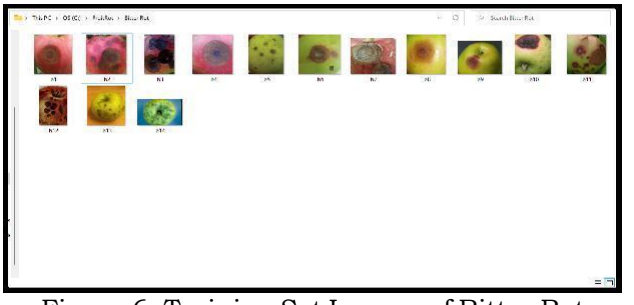

Figure 6. Training Set Images of Bitter Rot

Bitter rot is a common fruit rotting disease of apple (and pear) that occurs in all states where apples and pears are grown. Bitter rot on apple and pear fruit is caused by the pathogenic fungi in the Colletotrichum genus, specifically those species in the acutatum and gloeosporioides species complexes.

Figure 7. Training Set Images of Sooty Blotch

Sooty Bloch usually appears in August or September and is sometimes called the Summer Diseases of Apples. Sooty Bloch and FlySpec is a descriptive term for the state of dark pigment spots and smudges caused by several different fungi that affect the fruit, including apples, pears, persimmon, banana, papaya and many other cultivated tree and vine crops.

Figure 8. Training Set Images of Powdery MildewPowdery Mildew, caused by the fungus Podosphaera leucotricha,attacks bud,blossoms,leaves,new shoots,and fruit of wild and cultivated apples and crabapples.

Output Design

Output design generally refers to the results and information that are generated by the system for many end- users; it should be understandable with the enhanced format. The Output of the software is used to make the remote installation of the new software in the system and it awakens the immediate alert to the system that should be enhanced as the input to the system. Output is the main reason for developing the system and the basis on which they evaluate the usefulness of the application. Computer output is the most important direct source of information to the user output design deals with form design. Efficient output design should improve the interfacing with the user. The term output applies to any information produced by an information system in terms of display.

When analysts design system output, they Identify the specific output that is needed to meet the requirements of the end user. Previewing the output reports by the user is extremely important because the user is the ultimate judge of the quality of the output and, in turn, the success of the system.When designing output, system analysis accomplishes more things like, to determine what applications, websites or documents are blocked or allowed. Allowing should be in various options. The output is designed in such a way that it is attractive, convenient and informative.Forms are designed in VB.NET with various features, which make the console output more pleasing.

As the outputs are the most important sources of information to the users, better design should improve the system's relationships with users and also will help in decision-making. Form design elaborates the way output is presented and the layout available for capturing information.

Table shows the performance between K-means Clustering Algorithm and CNN Algorithm

	K-means Clustering	CNN
Accuracy	Accuracy is Low	Accuracy isHigh
Time Taken	Time taken is High	Time taken is Low
Disease Name	Name of the Disease not Found	Name of the Disease Found
Pixel Applicable	Applicable for Low Pixel Only	Applicable for both High andLow Pixel

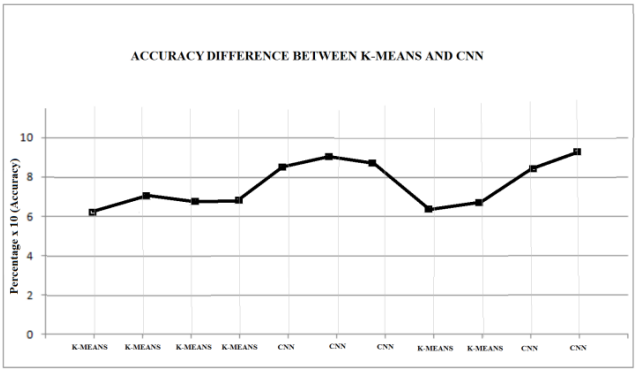

Figure 9. Shows the accuracy between K-means Clustering Algorithm and CNN Algorithm

Sample Outputs

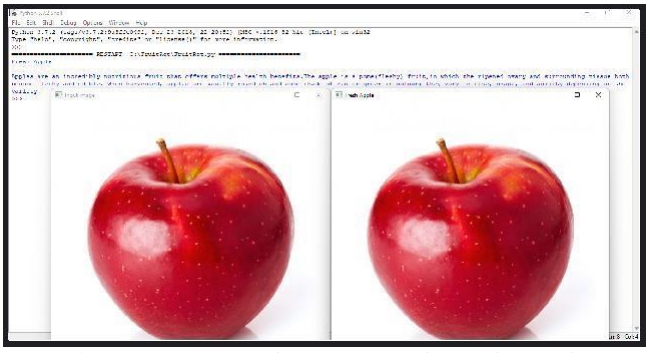

Figure 10. Sample Output of Fresh Apple

The apple is a pome(fleshy) fruit,in which the ripened ovary and surrounding tissue both become fleshy and edible. An apple is an edible fruit produced by an apple tree.

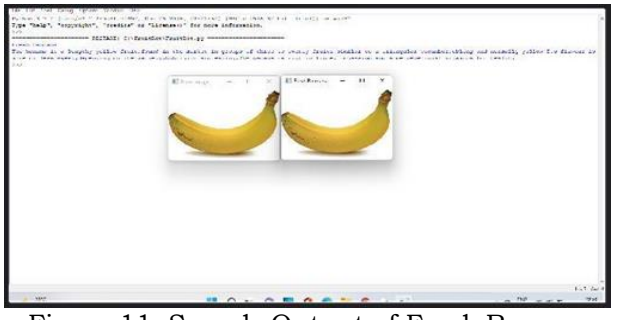

Figure 11. Sample Output of Fresh Banana

Banana's flavor is more or less sweet,depending on the variety.Nutrition and eating.The banana is rich in fiber, potassium and some beneficial vitamins for health. An elongated usually tapering tropical fruit with soft pulpy flesh enclosed in a soft.

Figure 12. Sample Output of Fungus

Fungus is the first disease of bananas to have spread globally in the first half of the 20th century. Fungus banana, popularly known as Fusarium disease, is a lethal fungal disease caused by the soil-borne Fusarium oxysporum.

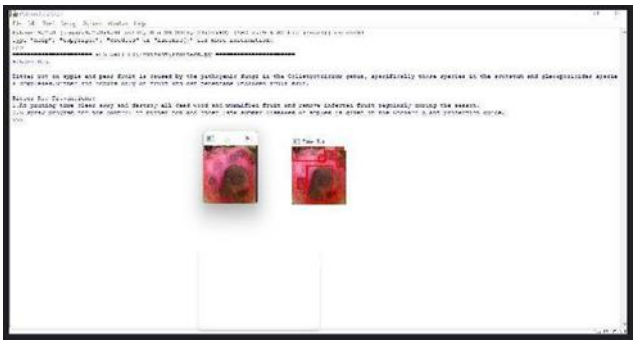

Figure 13. Sample Output of Bitter Rot

Bitter rot occurs only on fruit and can penetrate unbroken fruit skin. Bitter rot is caused by the pathogenic fungi in the Colletotrichum genus. Bitter rot appears on young fruit as small, circular brown lesions. Lesions expand rapidly and radially under wet and warm conditions.

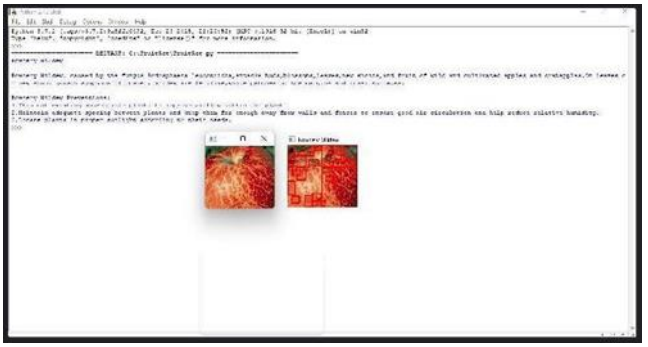

Figure 14. Sample Output of Powdery Mildew

Powdery mold is a fungal disorder that influences a huge variety of plants. Powdery mold sicknesses are because of many special species of fungi withinside the order Erysiphales. Powdery mold is one of the simpler plant sicknesses to identify, as its signs are pretty distinctive.On leaves of recent shoot boom signs of powdery mold are felt like,white patches at the margins and decrease surfaces. Sooty blotch causes dark smudges on the surface of the apple.Both fungi are limited to the surface of the fruit and do not cause rot.

Conclusion

An image processing-based solution is proposed and evaluated in this project for the detection and classification of fruit diseases. The proposed approach is composed of mainly three steps. In the first step image segmentation is performed using convolutional neural network technique. In the second step affected places are found. In the third step training and classification are performed. With the

advent of fast and cheap machines, digital image processing has become a very highly demanded field of study and practice. It provides solutions to various real-life applications in an economical way. Image processing is a way to carry out a few operations on an photo, with the intention to get an more advantageous photo or to extract a few beneficial statistics from it. It is a kind of sign processing wherein enter is an photo and output can be photo or characteristics related to that photo. It would also promote Indian Farmers to do smart farming which helps to take time to time decisions which alsosave time and reduce loss of fruit due to diseases. The leading objective of our project is to enhance the value of fruit disease detection.

Scope for future development

Future of this project can be easily updated. To achieve the benefits that are expected from the user must understand the overall system and they must be able to carry out their specific tasks effectively. The successful implementation depends upon the right people at the right time. The application becomes useful if the below enhancements are made in future.

- ➢ If the application is designed as a web service, it can be integrated in many websites.
- ➢ More accuracy can be detected using various machine learning algorithms.

The application is developed such that above said enhancements can be integrated with current modules.

References

[1] Bhange, M. & Hingoliwala, H. A. Pomegranate Disease Detection Using Image Processing. India, Elsevier B.V, 2019

[2] Dubey, S. R. & Jalal. A. S. Adapted Approach for Fruit Disease Identification using Images. India, International Journal of Computer Vision and Image Processing, 2018

[3] Padol, P. B. SVM Classifier Based Grape Leaf Disease Detection. India, Conference on Advances in Signal Processing (CASP), 2017

[4] Sujatha, R., Kumar, Y. S., Akhil, G, U. Leaf disease detection using image processing. Vellore, Journal of Chemical and Pharmaceutical Sciences, 2017

[5] Khot, S. T., Supriya, P., Gitanjali, M., & Vidya, L. Pomegranate Disease Detection Using Image Processing Techniques. Pune, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2016

[6] Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2021). Health and treatment of diabetes mellitus. *International Journal of Health Sciences*, *5*(1), i-v. https://doi.org/10.53730/ijhs.v5n1.2864

[7] Fattakhov, N., Normatova, S., Madaminov, S., Tilyakhodzhaeva, G., & Abdulkhakimov, A. (2021). Hirudotherapy as an effective method for treatment of migraine - a disease of unknown etiology. *International Journal of Health & Medical Sciences*, *4*(2), 232-237. https://doi.org/10.31295/ijhms.v4n2.1714