ISSN: 2096-3246
Volume 57, Issue 02, JULY, 2025

DATABASE DEVELOPMENT OF DEFECTIVE MANGOES AFTER POST-HARVEST DISEASES

[1] Ms Baig Ishrat Shamsher, [2] Dr.P.D.Deshmukh
Baigishrat786@gmail.com, prapti.research1@gmail.com
[1,2]Department of Computer Science and IT, Dr. G. Y. Pathrikar College, MGM University, Aurangabad, India

Abstract:

Post-harvest diseases in mangoes significantly impact their quality, market value, and supply chain efficiency. Identifying and tracking defective mangoes caused by these diseases is essential for reducing losses and ensuring food safety. This study presents the development of a database system designed to record, monitor, and analyse defective mangoes affected by post-harvest diseases such as anthracnose, stem-end rot, and soft rot.

The database incorporates key features, including disease classification, severity levels, visual symptoms, geographical origins, storage conditions, and time since harvest. Advanced query functionalities allow stakeholderssuch as farmers, distributors, and researchersto access actionable insights and trends, enabling better decision-making in disease management, transportation, and storage protocols.

By integrating data visualization and predictive analytics, the database facilitates early detection of disease patterns, contributing to improved post-harvest handling practices. This initiative aims to enhance the efficiency of mango supply chains, minimize waste, and ensure the delivery of high-quality mangoes to consumers while promoting sustainable agricultural practices.

Keyword: post-harvest, anthracnose, severity levels, detection, decision-making.

Introduction:

A popular tropical fruit throughout the world, mangoes are prized for their economic value, nutritional value, and rich flavor. However, post-harvest illnesses cause a sizable number of mangoes to be lost each year. Mango quality, shelf life, and marketability are all negatively impacted by diseases such as anthracnose, stem-end rot, and bacterial soft rot, which cost farmers and other supply chain participants money. A systematic approach to recognizing, tracking, and examining problems brought on by post-harvest illnesses is necessary for their management. Traditional techniques of monitoring damaged mangoes are generally labor-intensive, error-prone, and unscalable. On the other hand, database systems provide a dependable way to effectively arrange and examine vast amounts of information about faulty mangoes. To improve decision-making, these systems can document illness types, symptoms, severity levels, storage conditions, and geographic sources.

Prasad et al. stated that about 370 million metric tons of fruit are produced annually worldwide, of which 32 million metric tons, or 8%, come from India. Known as the "king of fruits," the mango (Mangifera indica) is one of the most important and ancient fruits in India. With an annual crop of 10.99 million metric tons, India produces 57.18% of the world's mangoes on 1.23 million hectares. Because of centuries of selection from wild plants, the nation is home to more than 500 cultivated mango types. India produces comparatively few mangoes per person, even though it is the world's

largest producer and a major supplier of processed mangoes. Additionally, spoilage at different post-harvest stages is thought to account for 20–22% of India's total fruit yield. China, Pakistan, Thailand, Australia, and the Philippines are some of the other main mango-producing nations. [1]
Using Canonical Discriminant Analysis, **Padda et al**. [4] discovered that firmness, flesh value, and total soluble solids content were the most effective methods for evaluating changes in mango fruit during ripening.Subedi et al. [5] demonstrated a strong correlation between fruit maturity and meat value. Despite being a destructive measurement criterion, flesh colour was shown to be constant among cultivars and is utilized as a maturity index in several producing countries.

Literature Review

Diseases In Mango after Post harvest

Table 1: Disease after post-Harvest [11][12][13]

Aspect	**Anthracnose**	**Stem-end Rot**	**Black Mold Rot**
Causal Agent	*Colletotrichum gloeosporioides*	*Lasiodiplodiatheobromae*, *Phomopsis spp.*	*Aspergillus niger*
Symptoms	- Small, black, sunken spots on the fruit surface. - Lesions enlarge and merge, forming large necrotic patches. - Can also affect the stem and leaves pre-harvest.	Dark brown to black rot beginning at the stem-end. - Internal decay spreads into the flesh. - Advanced cases show softening and discoloration.	- Black, powdery fungal growth on the fruit surface, often around the stem or damaged areas. - The mold can spread quickly in humid conditions.
Mode of Infection	- Latent infections from pre-harvest fungal colonization. - Spores germinate under humid storage conditions.	- Fungi infect the stem during the growing stage. - Latent infections activate during ripening and storage.	- Pathogen enters through damaged skin or stem. - Requires high humidity and warmth to develop.
Factors Favouring Disease	- Warm, humid environments during storage. - Poor field and post-harvest sanitation.	- Mechanical injuries at the stem. - High humidity during storage. - Delayed processing post-harvest.	- Warm, humid conditions. - Damaged or overripe fruits. - Inadequate ventilation during storage.
Economic Impact	- Severe in humid regions, causing up to 50% losses. - Reduces fruit marketability and storage life.	- Major issue during export or prolonged storage. - Causes significant financial losses if untreated.	- Leads to rejection of fruits in the market. - Can spread rapidly, contaminating other fruits.

Aspect	**Anthracnose**	**Stem-end Rot**	**Black Mold Rot**
Management Strategies	- Pre-harvest fungicide sprays (e.g.,carbendazim). - Hot water treatment post-harvest. - Maintain low humidity and cool temperatures during storage.	- Harvest fruits carefully to avoid stem damage. - Treat fruits with hot water or fungicides (e.g., thiabendazole). - Proper sanitation during handling and storage.	- Avoid mechanical damage during harvest. - Store fruits in cool, dry conditions. - Use fungicidal dips to reduce infection risk.
Prevention Techniques	- Use resistant mango varieties. - Remove infected plant parts pre-harvest. - Ensure proper air circulation in storage.	- De-sap fruits immediately after harvest. - Hot water dips (50–55°C). - Refrigerated storage (10–13°C).	- Avoid storing overripe or damaged fruits. - Ensure good ventilation in storage. - Dry fruits thoroughly after washing.
Image			

One of the most significant postharvest illnesses, anthracnose degrades a variety of fruits, such as bananas, apples, avocados, mangoes, and others.[6]

Ya-qin Tian1et al mentioned in his study the possible anti-disease mechanisms of Metschnikowia pulcherrima yeast as well as its impact on the storage quality of "Tainong" mangoes. According to the findings, M. pulcherrima successfully maintained the quality of mangos by lowering variations in fruit hardness, peel color, and the concentrations of vitamin C, acid, and total soluble solids. It generated antibacterial chemicals, stimulated defense-related enzyme activity (ß-1,3-glucanase and chitinase), and suppressed Colletotrichum gloeosporioides by competing for carbon sources and space. These results demonstrate M. pulcherrima's potential as a biocontrol agent to lessen mango deterioration and spoiling.

Md. Nasir Uddin1 et al. reported that One of the most significant mango diseases in Bangladesh and other humid regions is anthracnose disease, which is brought on by Col. letotrichumgloeosporioides. Anthracnose and stem end rot, which can spread with raindrops, have been observed to cause 25–30% of the entire mango yield to be lost.he also state the estimated loss caused by Anthracnose disease has been reported 60% or higher in the heavy rainy season.[**8**]A direct decrease in the amount or quality of the harvested produce is typically the result of crop losses. According to various reports, the disease incidence in South Africa is 32%, in Costa Rica it is 64.6%, and in conditions that are damp or extremely humid, it can approach 100%. Anthracnose-induced yield loss of 50.28% has been documented in Gondunglegi, Indonesia while 29.6% post-harvest loss has been documented in Himachal Pradesh, India, between 1990 and 1992 . Mango fruits from Hyderabad were found to be 20–30% rotted by Colletotrichum gloeosporioides, according **to Prakash et al. [9]**

On the fruit surface, postharvest anthracnose appears as a circular, brown to black lesion with an

Fig 1: Development of postharvest Diseases

unclear boundary. Large fruit infections typically do not progress to lesions. The fungus first establishes itself in the fruit and then stays inactive or latent until the fruit starts to ripen. On the maturing fruit, dark, depressed circular lesions appear and grow quickly. In the worst situations, they might even completely cover the fruit's surface. From the fruit's base to its distal end, lesions of all sizes might combine to cover large portions of it, usually in the shape of tear stains. [10]

Data Collection:

Alphonso mangoes are among the most costly and superior types. Mangoes are mostly grown in the districts of Sindhudurg, Raigad, and Ratnagiri in western India. In the states of Maharashtra and Gujarat, Alphonso mangoes are frequently referred to as “Hapus.” The finest thing to know is that "the Queen of Mangoes, Kesar," is grown and distributed across the nation in Junagadh. Throughout the year, people travel to Varanasi, one of India's most important pilgrimage sites. This village grows the delectable "Langra" or "Malda" mango kind. Kurukshetra is a major place of cultivation of another famous variety of mango- ‘*Chaunsa*’. The state of Karnataka's most popular mango variety is Badami. Typically grown in southern India, totapuri mangoes are also well-liked in western areas like Gujarat and Maharashtra. Nearly every region of India grows the Neelam mango cultivar. Dasheri is a luscious mango type that is mostly grown in Nepal, Pakistan, and various regions of

North India. These are all varieties of mango found in the market yard in Pune. For database collection, I chose Dasheri and Langra.

We concentrated on three popular mango varieties—Lalbagh, Kesar, and Alphonso—and a **locally grown, non-hybrid variety** of mango to gather data. These groups were selected due to their widespread cultivation, commercial significance, and vulnerability to a few post-harvest diseases. Accurate disease detection and categorization depend on the unique physical traits and ripening patterns displayed by each variety. Our image processing and IoT-based disease diagnosis method is more robust and broadly applicable thanks to the diversified and representative dataset we obtained by gathering data from these categories.

Methodology:

We captured an image of the mango using a Raspberry Pi device integrated with a camera, controlled by Python. The captured image is passed through an image processing pipeline that utilizes OpenCV to analyze the mango's portion in the image. The image is then converted to grayscale and transformed into a NumPy array for further processing. After conversion, the image is analysed using a Simple Blob Detector with an area filter to detect blobs that fall within specific criteria. Once the blobs are identified, the application evaluates their color range, the area they cover, their size, and assesses the impact of the blob, such as identifying potential diseases and affected areas. The user is notified of results, including a percentage of the affected area. The detailed process mechanism is outlined below:

Pre-processing techniques:

- **Methodology:**

We captured an image of the mango using a Raspberry Pi device integrated with a camera, controlled by Python. The captured image is passed through an image processing pipeline that utilizes OpenCV to analyze the mango's portion in the image. The image is then converted to grayscale and transformed into a NumPy array for further processing. After conversion, the image is analyzed using a Simple Blob Detector with an area filter to detect blobs that fall within specific criteria. Once the blobs are detected, the application calculates their color range, the area they are in, their size, and the effect of the blob, such as detecting possible diseases and infected areas. The user is notified of results, such as a percentage of the infected area. The whole process mechanism is explained below: Pre-processing techniques:

Resizing

When resizing, we may lose some information in the image, but it will not make the image unusable. It is essential to examine the image preview and determine which part of the image is retained by the resizing function. In this resizing process, we will preserve the aspect ratio of the image by computing the width and height within a function that can reshape different images to the desired aspect ratio, typically 256x256.

The resizing method is very handy when you have thousands of images that you want to standardize in terms of size. It is essential to trial different methods of resizing images, as the choice of resize method can significantly impact the result of the image. For example, opting for trilinear, bilinear, or nearest neighbor can affect the overall good and bad features of the image. In OpenCV resize () function is available to resize the image.

This function also provides options for different interpolation methods, such as nearest-neighbors, bilinear, and cubic interpolation. You can select the interpolation method you want based on the trade-off between speed and quality. Resizing images uniformly in this way means that the dataset will remain uniform for machine learning algorithms to be able to run and learn from the dataset.

Grayscale: Changing color photographs to grayscale can lessen the amount of data/image complexity, making image interpretation and machine learning less computationally demanding for certain techniques and methods. In many cases, the color part of an image can be considered extraneous. That can be true when we work on tasks related to structure, texture, and other geometric shapes. In image processing, if we can simplify the image from three colors (RGB) to one color (a single intensity value), we often improve the processing time, and ignore the color is often necessary. The transformation from RGB to grayscale works in OpenCV's cv2.cvtColor () function. OpenCV's cv2.cvtColor () function can convert an image from a color version to a grayscale version, as it will take the RGB channels and convert them to one intensity value to represent varying shades of gray. Consequently, the image becomes simpler, requiring less processing power while maintaining the key characteristics required for analysis.

Noise reduction: In machine learning models for image analysis and processing, noise is an unwanted aspect of images that can negatively affect performance, as it creates inconsistencies that can interfere with the extraction of features and measures of interest. Various methods can be used to reduce noise (e.g., filtering, blurring, and smoothing) in images to improve the quality of the images. One of the more common methods involves using blurring techniques to reduce high-frequency noise and smooth the image. OpenCV has multiple functions to apply blurring to images, including the Gaussian Blur () and medianBlur () functions. The Gaussian Blur () function uses a Gaussian kernel on the image and blurs the image by averaging each pixel with its neighbours based on a weighted approach that reduces noise while still keeping some degree of edge definition. On the other hand, the medianBlur () function replaces the value of each pixel with the median of the neighbouring pixels; one major benefit of the median blur is the peaceful removal of salt-and-pepper noise from the images. Thus, the filtering and blurring techniques can enhance the quality of the images when undertaking image analysis tasks.

Normalization: In image processing, one typically wants to translate the pixel intensity values to a consistent and appropriate scale, usually between 0 and 1. This is done to ensure that machine learning will work on the image effectively. Normalization of the pixel intensities helps standardize datasets so that extreme intensity values do not have a negative effect on learning, as well as not creating difficulty learning the important aspects of the image. The key idea is that normalization scales the pixel intensities and will allow our model to learn relative scale instead of engaging in learning from the overall range of pixel values. There are many ways to apply normalization to images and one of the easiest methods is to use the normalize () function provided by the scikitimage library. The normalize () function automatically translates and rescales the pixel intensities to the desired scale, and the common one is between 0 and 1, but you can use others too. This step is an important preprocessing step since using a consistent scale for data presented to your machine learning models will help with convergence and performance, both while training.

Binarization: Binarization is the conversion of grayscale images into black and white images using a simple thresholding technique. It reduces the grayscale image into two-pixel values, one representing the foreground (white usually) and the other the background (black usually) or two classes. Binarization is helpful in many other image-processing processes, such as object detection, document analysis, and others. In OpenCV, the binarization of an image is accomplished through a threshold () function. It takes in an image and a threshold value, and the function segregates all pixel values greater than the threshold into the maximum value (usually white) and the rest below the threshold into the minimum value (usually black). This separation allows for effective separation of features to analyse specific objects or areas of interest in depending on your definition of the focus area.

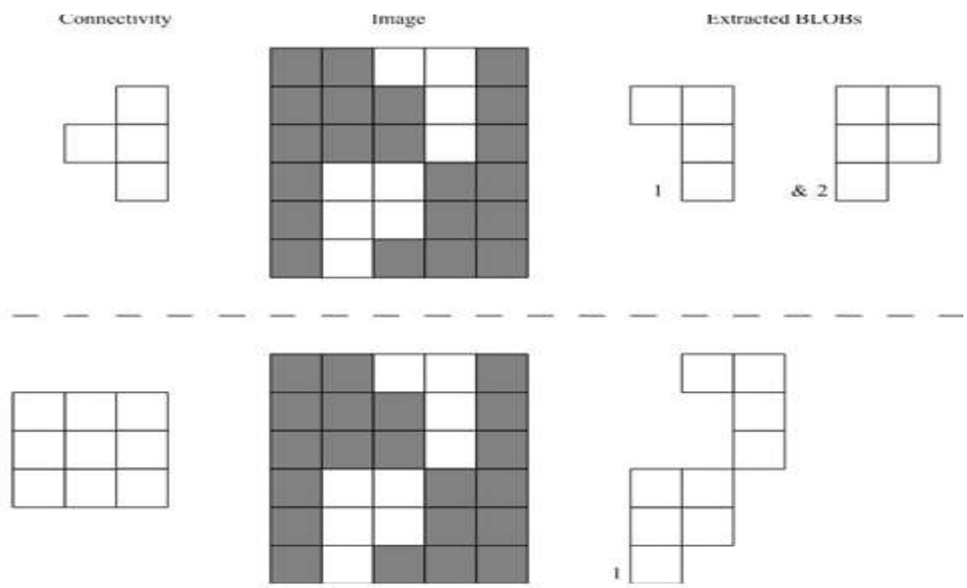
• **Contrast Enhancement:** Histogram equalization is a technique to enhance the contrast of a poorly contrasted image by stretching the intensity value of the pixel values to be as distributed across the entire intensity possible range. This change tends to highlight the details in the most dark and bright areas of the image, improving the clarity and definitiveness of the image. Oftentimes images may be poorly contrasted and as such, it is making it difficult for the machine learning model to understand the key features of the image that are important. Histogram equalization is a great solution that will spread the most common intensity values out across the possible range improving their contrast and importance as image details.
OpenCV uses the equalizeHist() method to create histogram equalization. This function applies to grayscale images and manipulates the pixel intensity distribution of a grayscale image, which will improve the contrast of certain areas that were previously difficult to see. The end result will be a better equilibrium between light and dark areas for contrast, allowing features to be seen more readily by the models.
We can improve your image data significantly by performing a pre-processing step that allows you to use a number of techniques, including resizing the image, converting to grayscale, denoising the image, normalizing the image, binarizing the image, and histogram equalization. With these steps, you have prepared the data in a way that is conducive to use with your machine learning algorithms and will benefit the effectiveness and efficiency of your computer vision application. These techniques take unprocessed or raw images and prepare them in a somewhat standard, clean, informative format, and work within the constraints of your specific problem.

BLOB (Binary Large Object):
BLOB analysis is a basic image processing technique that looks at unique features of an object, and is the image segmentation technique used in this study, D. Therefore, BLOB analysis can be used to identify and isolate the size and shape of the object, compared to its surroundings.

The BLOB Features: The next step after finding the various BLOBS is classifying the different BLOBS. The first step in the BLOB extraction process is separating each BLOB by the various traits and suggested features. Extraction of features transforms each BLOB into representative figures, where only pertinent data is considered, and the remainder is ignored. The first step is to remove all BLOBs that are attached to an image's edge because, often, there is no information regarding anything that is not in the picture. The number of pixels that make up a BLOB is its number. Since this function is utilized to choose the BLOB size, BLOBs that are too large or little can be disregarded. To display the identified BLOB, draw a convex hull, box, or bounding circle. The bounding circle is used in this study to display the identified image.

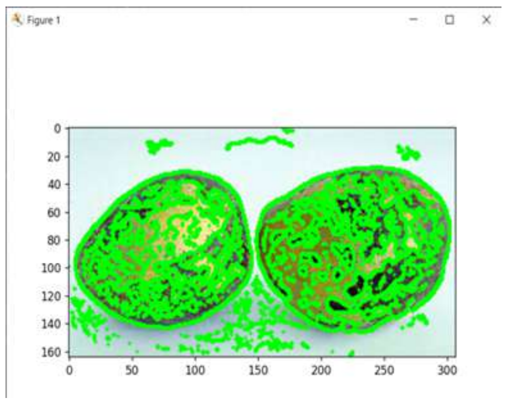
A BLOB's bounding box, which is the smallest rectangle inside a BLOB, is used to eliminate BLOBs that are too large or small from an image. Going over each pixel for a BLOB and identifying the four pixels with the lowest x, maximum x, minimum y, and maximum y values results in the rectangle. The width, xmax xmin, and height, ymax ymin, are used to draw the bounding box. The ROI (Region of Interest) is another name for the bounding box. A BLOB's bounding box ratio is calculated by dividing its height by its width.

Figure 2: 4 Connectivity and 8 Connectivity



Original Image **Gray Scale Image with Filter noise and blob detection**

Fig3: BLOB Detection Using OpenCV



BLOB Detected

The ratio of a blob's area to the area of the bounding circle is known as its **compactness**, and it's used to identify compact blobs from non-compact ones. A blob's compactness is determined by

Area of a BLOB

Compactness = width · height

A binary image's centre of mass is the average of the binary object's x and y positions, which serves as the image's centre of balance. It can be described as a point whose x and y values are provided by

$$x_c = \frac{1}{N}\sum_{i=1}^{N} xi \quad , \quad y_c = \frac{1}{N}\sum_{i=1}^{N} yi$$

A quick approximation of the centre of mass is the centre of the bounding box, which is determined by

$$x_{bb} = \frac{Xmin\ +Xmax}{2} \quad , \quad y_{bb} = \frac{Ymin+Ymax}{2}$$

The length of the counter inside a BLOB, which is determined by counting the number of pixels encountered while scanning along an object's counter, is known as the BLOB's perimeter.

Database Creation:

After identification of each BLOB (Binary Large Object) as separate image. data base has been updated with the pixel values which are extracted from the BLOB image.

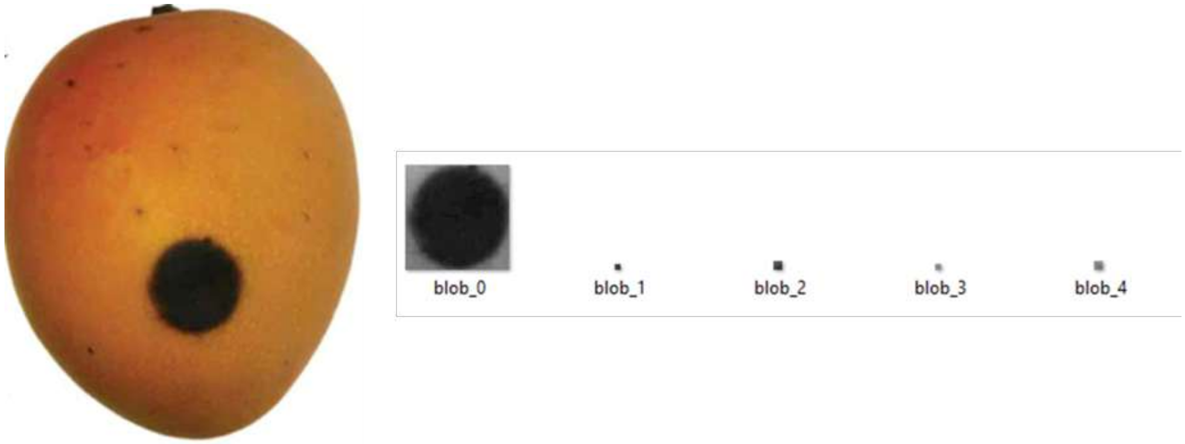


Fig 4: After BLOB detection each blob as separate Image

The width, XmaxXmin, and height, YmaxYmin, are used for ROI (Region of Interest).

DiseaseRecognisationMaster

ImagePath	BlobName	XCoOrdinat	YCoOrdinat	Size	MaxValue	MinValue	DiseaseID
D:\testImage\IshratMango\Set1\1.jpeg	blob_0	1381	2959	3	69	61	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_1	2310	2709	4	62	55	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_547	2098	370	3	126	120	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_536	916	184	3	125	117	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_552	1995	889	5	153	146	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_551	2011	589	4	147	135	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_550	1743	702	4	150	130	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_549	686	51	4	139	126	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_553	1435	23	8	162	135	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_548	1503	73	4	134	122	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_535	947	409	3	121	113	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_546	2296	478	4	136	116	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_544	2129	524	4	138	125	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_543	2746	659	4	136	121	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_542	2385	1843	4	140	126	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_541	2462	2360	5	139	126	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_540	2049	2472	5	135	123	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_539	1952	2535	4	134	121	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_538	1595	80	5	130	103	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_537	1715	86	4	124	113	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_545	1800	513	13	157	110	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_533	1260	468	3	122	118	2
D:\testImage\IshratMango\Set1\1.jpeg	blob_525	2057	835	3	124	118	2

Fig 5: Screenshot of Database

Data is the most important component of any intelligent agricultural monitoring system, especially one that aims to identify post-harvest mango illnesses. The analytical core is made up of image processing and machine learning, but the SQL-based database offers an organized, safe, and query able framework for handling and storing this data.

A strong tool for creating, modifying, and retrieving data from relational databases like MySQL, PostgreSQL, SQL Server, or SQLite is SQL (Structured Query Language). SQL is essential to this project's management of sensor readings obtained by IoT devices as well as image information.

Fig 6: Raspberry PI-3

Fig 7: ELP 180-degree fisheye Raspberry Pi 1080P H. 264 microphone PC Web USB security camera

Raspberry Pi is a credit-card-sized single-board computer designed and manufactured by the Raspberry Pi Foundation in the United Kingdom. Raspberry Pi has an ARMv6 700 MHz single-core processor, a Video Core IV GPU, and 512MB of RAM. It uses an SD card for its operating system and data storage. The Raspberry Pi officially supports Raspbian, a lightweight Linux OS based on Debian.

I have analysed the Raspberry Pi architecture and decided to install Windows OS because of the following reasons.

Supports all types of Codes, Vast peripheral support, can be used as a Portable Computer

Primary Data Collection for Post-Harvest Mango Disease Detection

Primary data collection is an essential first step in creating a strong and trustworthy system for identifying post-harvest illnesses in mangoes. The quality and volume of the input data that a machine learning model gets are critical to its success. To ensure that the dataset is typical of real-world situations and has accurate annotations for a variety of diseases, our strategy in this context focuses on collecting actual image data of mango fruits straight from the post-harvest environment.

Image Acquisition Assisted by IoT

The installation of IoT-enabled video systems in key locations throughout mango handling and storage facilities marks the start of the data collection process. These high-definition cameras, including the Kreo Owl 4K, the ELP 1080p USB camera, and others, can take close-up pictures of individual mango fruits. The system's capability to rotate and scan multiple angles of mangoes ensures that all visible signs of decay or infection are effectively captured. IoT technology enables automation, data transfer, and remote monitoring. Cameras are attached to hardware modules

(Techno tics and Arduino) designed to scan mango conveyor belts or trays. The layout allows for scale and continuous data collection with little human interaction. The scanned images are processed by tagging or other processing and transferred in real-time to a local server or cloud storage using the images collected.

Visibility of Disease and Image Quality

To ensure signs of disease are seen, high-resolution imaging is crucial. By correctly adjusting variables, such as light, camera view, background/contrast, and sharpness, key signs of illness are emphasized. To be able to see the symptoms of post-harvest diseases, photos need to show Mold patches, black spots, discoloration, or shrivelled skin clearly. Good imaging conditions help improve the accuracy of subsequent classification tasks, as well as minimize the burden of complex preprocessing.

Labelling and Annotation of Diseases

Each image has been labelled according to professional confirmation or prior laboratory diagnosis and assigned the appropriate location for the diseased category types considered in this study. The following is a list of the major post-harvest diseases this project examines:

Anthracnose - Characterized by black, sunken spots that grow into necrotic areas and coalesce with one another.

Stem-End Rot - This is defined by the darkening around the stem end of the mango and often includes softening and internal rot.

Black Mold Rot - This is black, powdery fungal growth from Aspergillus Niger and is normally found at or near fissures or wounds.

Aspergillus Niger is the pathogen responsible for black Mold rot. The disease is a black, powdery fungal growth that usually appears at or near fissures or wounds, often sporulating profusely in these areas.

Brown to black circular lesions with mouldy patches, with or without concentric rings, are representative of Alternaria Rot.

These labels help keep an orderly structure to the dataset as well as help supervised machine-learning algorithms. All the images have been archived using a consistent naming convention and also organized in folders distinguished by disease name.

Challenges in Primary Data Collection

The first critical step toward developing a reliable approach to identifying postharvest diseases of mangoes encompasses the utilization of primary data. In this step, several challenges must be properly navigated to ensure that the final dataset is accurate, adequately diverse, and usable to a machine learner. These challenges include technical, contextual, operational, and annotation challenges.

1. **Variability in the Environment**

Another major challenge is the variability in surroundings while taking images. Little to no control over lighting in warehouses, including other storage facilities, can lead to transitions in illumination, colour balance, and shadows. This may conceal or imbalance disease symptoms, such as a discolouration or Mold growth, which would affect the accuracy of disease detection models.

Variations in temperature and humidity may also affect the development and visibility of post-harvest diseases. For instance, a disease like black Mold may be more visible in humid scenarios versus dry ones; timing is key with imaging.

2. **Camera Adjustment and Quality**

The quality and calibration of IoT-enabled cameras used in the field will influence how useful the data collected can be. Images with low resolution or poor focus may miss some of the subtle signs of illness, while very early on (De Silva et al., 2020). So, it will be very important to properly adjust how the camera captures the mangoes not only in position, angle, and distance, but also by training the image to obtain sharp, full images of every fruit.
Additionally, if automated scanning platforms or a camera rotation device are not able to stay in sync, they can easily introduce motion blur or alignment issues and further degrade image quality.

3. **Fruit Overlapping and Positioning**

It's another challenge getting the mangoes placed for photos. When mangoes are very close, or overlap one another, fully revealing each fruit to discern any diseases and label each fruit becomes difficult. Reliable and accurate data collection involves taking single-fruit photos, taken at equal intervals, with minimum background noise.

4. **Disease Similarities and Overlap**

There are some disease symptoms which are challenging to categorize due to visual similarities or concurrent symptoms. For example, the early manifestation of anthracnose disease symptoms can look like bruising or stem-end rot. In these cases, mislabelling could lead to confusion when training the model and a decreased classification accuracy. This will require an expert to be involved in the assistance of labelling, multispectral imaging, or other sensor input may be needed for distinction.

5. Unbalanced Distribution of Datasets

Class imbalance in the dataset is a consistent issue. In the dataset, certain classes will contain a disproportionate number of photographs; this is likely due to the more common occurrence of certain diseases than others. This class imbalance likely causes poor identification rates for uncommon diseases, such as Alternaria Rot, because it creates an artificial bias towards the majority class in the machine learning model.

To address this issue, data augmentation methods, such as flipping, rotating, and brightness modifications, are commonly used to artificially inflate the sample size of the neglected classes.

6. Complexity of Manual Annotation

Manually labelling photos according to disease symptoms with effort and expertise can be a challenge. While mechanical damage and biological disease symptoms can appear similar, skilled annotators need skilled annotators to differentiate them. Human error, fatigue, and limited topic expertise can also pose a threat to the reliability of the dataset through inconsistent labelling. As part of the validation process, one strategy employed to minimize this risk is inter-annotator agreement analysis.

7. Transmission and Storage of Data

IoT systems that capture large numbers of high-resolution images on an ongoing basis use a considerable amount of bandwidth and storage allocation for transfer. Handling large amounts of data over a longer time period, managing datasets requires secure, scalable, and reliable data storage with either local servers or cloud-based systems. There is an additional risk of data loss or corruption during transfer that needs to be minimized with redundancy and backup plans, as well as robust communication protocols.

Discussion And Conclusion:

The need for creative post-harvest management methods is underscored by the growing need for quality assurance in agricultural products, especially fruits like mangoes. One major issue in the supply chain is post-harvest losses brought on by illnesses, including Alternaria Rot, Stem-End Rot, Black Mold Rot, and Anthracnose. In order to effectively identify and handle faulty mangoes during the post-harvest phase, this study suggested the development and deployment of a progressive database system combined with image processing, Internet of Things-enabled data collecting, and machine learning technologies. The system's objective is to serve as an end-to-end intelligent platform that facilitates data-driven decision-making throughout the whole agricultural value chain, not just to detect illness.

This system is built on a structured intelligent database of mango images and related metadata like timestamp, storage conditions, sensor data, and disease diagnosis. Predictive modelling and advanced analytics with machine learning approaches, including convolutional neural networks (CNNs), are based on this database. The system can achieve exact illness identification through systematic image capturing, processing, labelling, and classification. The machine learning techniques are improving the ability of the system to detect mango diseases based on visual descriptors.

Traditional manual inspection methods are greatly enhanced due to the availability of IoT devices to collect the environmental and image-related data involved in the inspection. IoT cameras installed in the warehouse or cold storage are capturing images of mangoes from multiple angles automatically, increasing the consistency of the data collection while decreasing labour costs. This real-time data collection approach supports operational inferences by performing image collection in real-time operational environments, which improves the model's generalization performance in operational settings. The technology is intended for use in internal warehouses as well as a component of a larger smart agriculture ecosystem, which allows data from many sources to be combined, examined, and used for forecasting, market optimization, and policymaking.

However, the study also ran into a few practical and technological issues with the database management and data collection procedures. One of the main difficulties that had to be addressed was ensuring the reliability and consistency of the annotated data. As certain post-harvest diseases share similar characteristics, expert agricultural knowledge was essential to properly categorize each of the photos. To enhance the accuracy and reliability of labelling, we built a quality assurance workflow that contained expert review, inter-annotator agreement, and iteration-facilitating feedback loops. We utilized controlled lighting arrangements and camera calibration processes to try to stabilize lighting and image quality in the variable situation of warehouse environments.

There were a few accommodations to consider related to the database management aspect, such as managing large amounts of photo data, establishing rapid indexing and retrieval, and building in privacy protections for agricultural-related data. The system needed scalable storage approaches that could cope with constant incoming data without losing a level of efficiency, given that IoT cameras were taking thousands of photos each day. The cloud storage was used with local buffering and syncing to help it remain accessible and efficient.

Implementing user access control, encrypting important metadata, and adhering to pertinent agricultural and digital data rules all helped to allay security and privacy concerns. By taking these precautions, data integrity and confidentiality were preserved while being securely shared between platforms and stakeholders.

Additionally, the training data's quality and variability were critical to machine learning model performance. Biased model performance could occur with an unbalanced dataset dominated by a handful of diseases while other diseases were rare. To alleviate this disparity in sample amount, the study applied image data augmentation in the form of flipping, rotating and adding brightness to the underrepresented classes to create a proportionate training sample size.

In summary, the development of this progressive database model marks an enormous advancement for smart technologies in the post-harvest management of agriculture. By combining IoT-based image collection, intelligent labelling, and predictive machine learning models with a scalable database design, the system provides a holistic approach to accept the challenge of mango rot caused by post-harvest diseases.

The results validate the feasibility and effectiveness of this strategy, despite implementation challenges. The platform not only addresses a specific agricultural challenge, but it also sets a precedent for using technology to build responsible, effective, and sustainable agricultural systems. Ultimately, this project advances the notion of a digital agriculture revolution where data is the most meaningful resource for ensuring food quality, reducing waste, and positioning all stakeholders, from farm to table.

References:

1.Prasad, R., Ranjan, K. R., & Sinha, A. K. (2006). AMRAPALIKA: An expert system for the diagnosis of pests, diseases, and disorders in Indian mango. *Knowledge-based systems*, *19*(1), 9-21.
2.Ullah, M. A., Kiloes, A. M., Aziz, A. A., & Joyce, D. C. (2024). Impact of factors contributing to internal disorders of mango (Mangifera indica L.) fruit—A systematic literature review. *Scientia Horticulturae*, *331*, 113150.
3.Musale, S. S., & Patil, P. M. (2014). Database development of defective and healthy alphonso mangoes. *Int J Adv Agric Environ Eng*, *1*(1), 155-160.
4. Padda S.M., do Amarante C. V. T., Garcia R. M., Slaughter D. C., and Mitcham E. M, "Methods to analyzephysico-chemical changes during mango ripening: A multivariate approach," Postharvest Biology and Technology, vol 62, pp 267-274, 2011. http://dx.doi.org/10.1016/j.postharvbio.2011.06.002
5. Subedi P.P., Walsh K.B. and Owens G, "Prediction of mango eating quality at harvest using short-wave near infrared spectrometry," Postharvest Biology and Technology, vol 43, pp 326-334, 2007. http://dx.doi.org/10.1016/j.postharvbio.2006.09.012
6. Campos-Martı´nezA,Velazquez-del Valle MG, Flores-Mocte zuma HE, Suarez-Rodrı´guez R, Ramı´rez-TrujiLLo JA, Hernan dez-Lauzardo AN. Antagonistic yeasts with potential to control Colletotrichum gloeosporioides (Penz.) Penz. and Sacc. and Colletotrichum acutatum J.H. Simmonds on avocado fruits. Crop Prot. 89: 101–104 (2016).
7. Ya-qin Tian1 • Wen Li2 • Zi-tao Jiang2 • Min-min Jing1 • Yuan-zhi Shao The preservation effect of Metschnikowia pulcherrima yeast on anthracnose of postharvest mango fruits and the possible mechanism Food Sci Biotechnol DOI 10.1007/s10068-017-0213-0
8. Md. Nasir Uddin1*, Shoaibe Hossain Talukder Shefat1, Mansura Afroz1 and Nusrat Jahan Moon Management of Anthracnose Disease of Mango Caused by Colletotrichum gloeosporioides: A Review ACTA SCIENTIFIC AGRICULTURE (ISSN: 2581-365X) Volume 2 Issue 10 October 2018
9. Prakash P., et al. "Elucidation of some in vitro exo-petrolytic enzymes by mango anthracnose fungus Colletotrichum gloeosporioides Penz". Archiv für Phytopathologie und Pflanzenschutz 25 (1989): 197-199.
10. Kumlachew Alemu Department of plant science, Assosa University, P.O.Box 18, Assosa, "Ethiopia Dynamics and Management of Major Postharvest Fungal Diseases of Mango Fruits"

Journal of Biology, Agriculture and Healthcare ISSN 2224-3208 (Paper) ISSN 2225-093X (Online) Vol.4, No.27, 2014
11.Adapted from multiple sources including:
Ploetz, R. C. (2003). *Diseases of mango*. In R. C. Ploetz (Ed.), **Diseases of tropical fruit crops** (pp. 327–363). CABI Publishing.
12.Prusky, D., & Lichter, A. (2007). Activation of latent infections by postharvest pathogens during transition from the quiescent to necrotrophic stage. *Postharvest Biology and Technology*, 44(2), 125–130.
13. Johnson, G. I., & Hofman, P. J. (2009). *Postharvest diseases of horticultural produce: Management strategies* (Vol. 1). ACIAR.