Identification of Papaya Disease Using a Deep Learning Approach

PREETI GUPTA[1], OMPRAKASH SINHA[2]

[1]*M. Tech, Computer Technology, Columbia Institute of Engineering &Technology, Raipur*

[2]*Asst. Professor, Dept. of CSE, Columbia Institute of Engineering &Technology, Raipur*

Abstract - Papaya disease poses a significant threat to farmers worldwide, leading to substantial losses annually. Recognizing the urgency of mitigating these losses, researchers have increasingly focused on developing papaya disease recognition systems. However, farmers often lack awareness of detection techniques, resulting in disease identification only after the papayas are already affected, leading to wasted crops and financial losses. Consequently, many farmers are hesitant to continue papaya cultivation. To address this issue, we conducted research leveraging deep learning technology for papaya disease detection and classification

Index Terms— Papaya Diseases, Deep Learning, Keras, Classification.

I. INTRODUCTION

1. Papaya (Carica papaya L.) is a tropical fruit native to Central America and southern Mexico, now cultivated in many tropical regions around the world. It is a rich source of vitamins A, C, and E, as well as papain, a digestive enzyme. Papaya production plays a significant role in the economies of many developing countries, and it is also gaining popularity in developed countries due to its nutritional value and delicious taste. However, papaya cultivation faces numerous challenges, with various diseases significantly impacting yield and fruit quality. Papaya Ringspot Virus (PRSV), Papaya Black Spot (PBS), and fungal diseases like Anthracnose and Phytophthora are some of the most common and destructive diseases affecting papaya crops. These diseases can cause devastating losses, leading to reduced fruit production, poor fruit quality, and even complete crop failure. Early and accurate identification of papaya diseases is crucial for effective disease management strategies. Traditional methods of disease identification rely on visual inspection by farmers. However, this approach has limitations. Firstly, it is subjective and depends on the experience and expertise of the farmer. Inexperienced farmers might struggle to identify diseases in their early stages, leading to delayed intervention and greater yield losses. Secondly, visual inspection is time-consuming, especially for large farms. Finally, some diseases can exhibit subtle symptoms that are difficult to detect with the naked eye.

These limitations highlight the need for more objective, automated, and reliable techniques for papaya disease identification. This thesis explores the application of deep learning, a subfield of machine learning, for achieving this goal. Deep learning algorithms have the potential to revolutionize the field of plant disease detection by offering accurate, efficient, and user-friendly tools for farmers. Furthermore, with the growing concerns about food security and sustainable agriculture, the need for effective disease management strategies in papaya cultivation becomes even more pressing. Deep learning algorithms have shown promise in various agricultural applications and could significantly enhance disease detection and management in papaya crops.

II. RELATED WORK

A: Inspiration

The inspiration for this project stems from the critical need for a more efficient and accurate method for papaya disease identification. Traditional methods, while providing a baseline for disease detection, are limited by subjectivity, time constraints, and the potential for human error.

Here are some of the specific reasons that motivated this project: -

- Economic Impact of Papaya Diseases: Papaya diseases cause significant economic losses for farmers due to reduced yield and fruit quality. Early and accurate disease detection is crucial for implementing effective disease management strategies and minimizing these losses.

- Limited Expertise for Disease Diagnosis: Farmers, particularly smallholder farmers in developing countries, may lack the necessary expertise to accurately diagnose papaya diseases, especially in their early stages. A deep learning-based tool can offer a user-friendly and reliable alternative.
- Time Constraints in Large Farms: Large papaya farms require efficient methods for disease detection to cover vast areas. Deep learning models can automate the disease identification process, saving time and resources.
- Potential for Early Disease Detection: Deep learning models have the potential to detect diseases in their early stages, based on subtle visual cues that might be missed by human inspection. Early detection is critical for implementing timely interventions and preventing the spread of diseases.
- Advancements in Deep Learning Technology: Recent advancements in deep learning architectures and image classification techniques have opened new avenues for developing robust and accurate disease detection models.

B: Papaya Disease Recognition

Traditional methods of papaya disease identification rely on visual inspection by farmers. This method is subjective and depends on the farmer's experience and knowledge of papaya diseases. An inexperienced farmer might struggle to differentiate between healthy and diseased plants, especially in the early stages of infection when symptoms are mild. Furthermore, visual inspection is a time-consuming process, especially for large farms with vast areas to cover. Additionally, some papaya diseases can exhibit subtle visual symptoms that are difficult for the human eye to detect. These limitations can lead to delayed intervention, missed diagnoses, and significant crop losses. Deep learning offers a promising alternative to traditional methods of papaya disease identification. Deep learning algorithms can be trained to automatically identify diseases from images of papaya plants. This approach offers several advantages over visual inspection. Firstly, deep learning models are objective and consistent in their analysis. They can be trained on a large dataset of labeled images, encompassing various papaya diseases and their different stages of progression. This training enables the model to learn the characteristic visual patterns associated with each disease, allowing for accurate identification even in the early stages. Secondly, deep learning models are fast and efficient. They can analyze images rapidly, making them suitable for large-scale disease detection in commercial papaya farms.

C: Machine Learning Techniques

Supervised learning algorithms are fundamental in machine learning, particularly for tasks such as papaya disease prediction. In supervised learning, a labeled dataset is essential, comprising data points (papaya images) each linked to a corresponding label (healthy or diseased). Through analysis of these labeled examples, the machine learning model discerns patterns and connections between features extracted from the images and the presence of specific diseases. These acquired patterns and connections essentially encode the knowledge necessary to differentiate between healthy and diseased papaya. Subsequently, the trained model applies these learned relationships to predict the disease status of new, unseen papaya images. This process enhances disease detection efficiency, aiding in timely interventions and mitigating yield losses. By leveraging supervised learning, farmers can adopt proactive measures to manage papaya diseases effectively, thereby safeguarding crop yields and livelihoods. Through the integration of advanced technologies like supervised learning, the agricultural sector can adapt to evolving challenges and contribute to global food production resilience.

D: Convolutional Neural Networks

This project focuses on a specific type of machine learning technique called a Convolutional Neural Network (CNN). CNNs are particularly well-suited for image recognition tasks due to their unique architecture. Unlike traditional machine learning algorithms that require manual feature extraction, CNNs can automatically learn these features directly from the image data. This is achieved through convolutional layers that apply filters to the image, extracting low-level features like edges and textures. Subsequent pooling layers then reduce the dimensionality of the data while preserving important features. By stacking multiple convolutional and pooling layers, CNNs can learn increasingly complex hierarchical features, ultimately leading to accurate image classification. CNNs have revolutionized the field of computer vision and are a powerful tool for tasks like object detection, image segmentation, and, as this project demonstrates, disease identification in plants and fruits.

III. LITERATURE REVIEW

Deep learning approaches, particularly Convolutional Neural Networks (CNNs), have emerged as powerful tools for automated disease identification in various crops. CNNs are able to learn complex patterns from image data, making them well-suited for tasks like image classification, which is essential for disease identification. By analyzing large datasets of papaya leaf and fruit images with disease symptoms, CNN models can be trained to distinguish between healthy and diseased samples, as well as differentiate between various papaya diseases.

Md. Sagar Hossen, Imdadul Haque, Md. Saif Islam, Md. Tanvir Ahmad, Md. Jannati Nime and Md. Ashiqul Islam in January 2021 proposed A deep Learning Based Classification of Papaya disease Recognition. On this matter they have performed a research by the advancement of deep learning technology to detect and classify the papaya disease. They have used CNN model according to Keras API. That model reliable with fully connected where classification is completed and all the process is a deep learning based. It took the fixed size of image 200x200 RGB image as input. And their accuracy is ~92%.

Jairo Lucas de Moraes * [ORCID] , Jorcy de Oliveira Neto, Claudine Badue, Thiago Oliveira- Santos and Alberto F. de Souza in May 2023 proposed a Papaya Fruit Disease Detector and Classifier Using CNNs and Convolutional Block Attention Modules. The authors proposed a novel image dataset comprising 23,158 examples divided into nine classes of papaya fruit diseases, and a robust papaya fruit disease detector called Yolo-Papaya based on the YoloV7 detector with the implementation of a convolutional block attention module (CBAM) attention mechanism. This detector achieved an overall mAP (mean average precision) of 86.2%, with a performance of over 98% in classes such as "healthy fruits" and "Phytophthora blight". The proposed detector and dataset can be used in practical applications for fruit quality control and are consolidated as a robust benchmark for the task of papaya fruit disease detection. The image dataset and all source code used in this study are available to the academic community on the project page, enabling reproducibility of the study and advancement of research in this domain.

Rashidul Husan Hardoy and Mosammat Rokeya Anwar Tuli in November 2021 prosed A Deep Ensemble Approach for Recognition of Papaya Diseases using EfficientNet Models. In this projet they have made An efficient diagnosis approach for papaya diseases is enormously desired to control and prevent the spread of diseases. At first, using a dataset of 138980 images of affected and healthy leaves and fruits of papaya which was generated with image augmentation techniques from 13898 collected images, eight models of EfficientNet between B0 and B7 were trained via transfer learning technique to recognize eight diseases. Afterward, finetuned versions of the three best-performing models were selected for ensemble learning such as EfficientNet B5, B7, and B6, which achieved 98.13%, 96.93%, and 96.87% accuracy under the test set of 6931 images, respectively. The deep ensemble model showed more effective recognition performance than single models, and test accuracy increased by 1.61%. The experimental result demonstrates that the proposed ensemble model can recognize papaya diseases more efficiently than single models of EfficientNet.

VI. PROPOSED METHODOLOGY

A: Deep Learning Model Architecture

The core of our methodology revolves around the design and implementation of a deep learning architecture tailored to papaya disease identification. Inspired by state-of-the-art convolutional neural network (CNN) architectures such as ResNet, DenseNet, and Inception, our model comprises multiple layers of convolutional, pooling, and fully connected units. The architecture was optimized to capture intricate spatial features and hierarchical representations present in papaya disease images, thereby facilitating accurate classification and diagnosis.

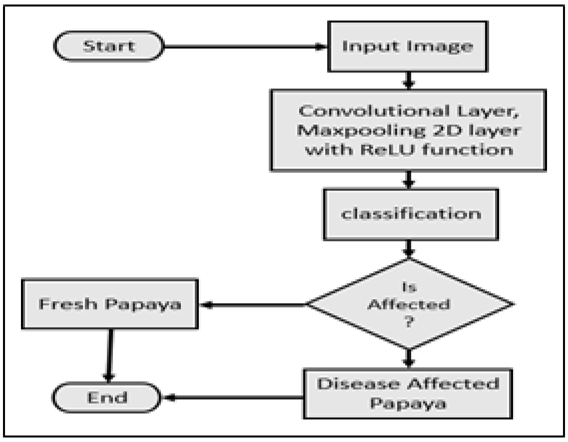

Fig: Papaya disease classification model

B: Preprocessing for CNN Image Classification

Preprocessing plays a pivotal role in the success of convolutional neural network (CNN) models for image classification tasks. These initial steps are essential for ensuring that the dataset is in a suitable format for training and that the model can effectively learn from the provided data. By meticulously preparing the dataset, researchers and practitioners can enhance the model's ability to generalize well to unseen examples, ultimately leading to more accurate and reliable classification results. The preprocessing pipeline typically involves various operations, including data loading, normalization, and augmentation, each aimed at optimizing the dataset for training.

C: Reading Training and Test Data

Reading training and test data is the initial step in the preprocessing pipeline for building convolutional neural network (CNN) models for image classification. This process involves accessing and loading the dataset into memory, enabling subsequent manipulation and analysis. Typically, datasets consist of a collection of images, each associated with one or more labels indicating their respective classes or categories. For supervised learning tasks like image classification, the dataset is divided into two subsets: the training set, used to train the model, and the test set, used to evaluate its performance on unseen data.

D: Representation of images

Representation of images is a critical preprocessing step in preparing data for convolutional neural network (CNN) models in image classification tasks. In this step, raw image data is converted into a format suitable for processing by the model, typically numerical arrays. Each image in the dataset is represented as a multidimensional array, where each dimension corresponds to a specific aspect of the image, such as width, height, and color channels (e.g., red, green, blue). This transformation enables the model to interpret and analyze the visual information contained within the images.

E: Feature extraction

Feature extraction is a crucial step in the machine learning pipeline, particularly in tasks involving image classification such as papaya disease identification. In essence, feature extraction involves transforming raw input data, such as images, into a format that is more conducive to model training and interpretation. In the context of papaya disease classification, feature extraction entails extracting meaningful patterns or features from images of papaya leaves. These features may include textures, shapes, edges, or other visual attributes that are indicative of different disease types. Feature extraction techniques aim to capture the salient characteristics of the input images in a compact and informative representation, facilitating subsequent classification by the model.

CONCLUSION

This study deals with the problem of classification of Papaya Disease Recognition. The advance scope of work has been spelt out and contribution made in this thesis has been summarized. This proposed system recognizes Papaya disease Recognition. The strategy utilizes combination of various machine learning approaches. For improvement of older algorithm this thesis utilizes feature extraction with data augmentation technique.

FUTURE RECOMMENDATIONS

In this proposed work, we implemented many of the concepts. Currently We didn't get the Best and correct dataset, So, in future if we get the correct and best dataset then, In future we would like to carry out the work with the correct and best dataset and increase our models accuracy and under fitting will be minimized.

REFERENCES

[1] S. J. Banarase and S. D. Shirbahadurkar, "Papaya Diseases Detection Using GLCM

Feature Extraction and Hyperparatuning of Machine Learning Approach BT - Proceedings of Third International Conference on Sustainable Expert Systems," in Springer, 2023, pp. 145–158.

[2] J. L. de Moraes, J. de Oliveira Neto, C. Badue, T. Oliveira-Santos, and A. F. de Souza, "Yolo-Papaya: A Papaya Fruit Disease Detector and Classifier Using CNNs and Convolutional Block Attention Modules," Electronics (Switzerland), vol. 12, no. 10, 2023, doi: 10.3390/electronics12102202

[3] U. Premchand et al., "Survey, Detection, Characterization of Papaya Ringspot Virus from Southern India and Management of Papaya Ringspot Disease," Pathogens, vol. 12, no. 6, p. 824, 2023, doi: 10.3390/pathogens12060824.

[4] Md. Sagar Hossen, Md. Saif Islam, Md. Jannati Nime, Imdadul Haque, Md. Tanvir Ahmed, Md. Ashiqul Islam, Deep Learning Based Classification of Papaya Disease Recognition, Jan 2021

[5] Rashidul Hasan Hridoy, Mosammat Rokeya Anwar Tuli, A Deep Ensemble Approach for Recognition of Papaya Diseases using EfficientNet Models, November 21

[6] Jairo Lucas de Moraes * [ORCID] , Jorcy de Oliveira Neto, Claudine Badue, Thiago Oliveira- Santos and Alberto F. de Souza, Yolo-Papaya: A Papaya Fruit Disease Detector and Classifier Using CNNs and Convolutional Block Attention Modules, 12 May 2023

[7] Veeraballi, R. K., Nagugari, M. S., Annavarapu, C. S. R., & Gownipuram, E. V. (2018, December). Deep Learning Based Approach for Classification and Detection of Papaya Leaf Diseases. In International Conference on Intelligent Systems Design and Applications (pp. 291- 302). Springer, Cham 6. Behera, S. K., Rath, A. K., & Sethy, P. K. (2020).

[8] Kawaljit kaur, Chetan Marwaha, "Analysis of Diseases in Fruits using Image Processing Techniques", International Conference on Trends in Electronics and Informatics ICEI 2017, 978- 1-5090-4257-9/17/$31.00

[9] Nikhitha M, Roopa Sri S, Uma Maheswari B, "Fruit Recognition and Grade of Disease Detection using Inception V3 Model", Proceedings of the Third International Conference on Electronics Communication and Aerospace Technology [ICECA 2019] IEEE Conference Record # 45616; IEEE Xplore ISBN: 978-1-7281-0167-5.

[10] J. A. Bacus and N. B. Linsangan, "Detection and Identification with Analysis of Carica papaya Leaf Using Android," Journal of Advances in Information Technology, vol. 13, no. 2, pp. 162–166, 2022, doi: 10.12720/jait.13.2.162-166.

[11] M. T. Habib, A. Majumder, A. Z. M. Jakaria, M. Akter, M. S. Uddin, and F. Ahmed, "Machine vision-based papaya disease recognition," Journal of King Saud University - Computer and Information Sciences, vol. 32, no. 3, pp. 300–309, 2020, doi: 10.1016/j.jksuci.2018.06.006.

[12] S. K. Behera, A. K. Rath, and P. K. Sethy, "Maturity status classification of papaya fruits based on machine learning and transfer learning approach," Information Processing in Agriculture, vol. 8, no. 2, pp. 244–250, 2021, doi: 10.1016/j.inpa.2020.05.003.

[13] A. K. Azad, L. Amin, and N. M. Sidik, "Gene Technology for Papaya Ringspot Virus Disease Management," The Scientific World Journal various, vol. 2014, 2014.

[14] M. A. Islam, M. S. Islam, M. S. Hossen, M. U. Emon, M. S. Keya, and A. Habib, "Machine Learning based Image Classification of Papaya Disease Recognition," Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2020, no. November, pp. 1353–1360, 2020, doi: 10.1109/ICECA49313.2020.9297570.

[15] Papaya Disease Dataset 1: https://universe.roboflow.com/sidd-b/disease-b/browse?queryText=&pageSize=50&startingIndex=0&browseQuery=true

[16] Papaya Disease Dataset 2: https://github.com/imdadulhaque1/papaya

[17] Papaya Fruit Diseases Detection (Binary Classification) using ML Techniques by Jhaji Bhaskar: https://github.com/Jhajibhaskar/Papaya-Fruit-Disease-Detection-using-ML-Techniques?tab=readme-ov-file

[18] Papaya-Fruit-Disease-Detection By Abhishek Kirtey: https://github.com/abhishektirkey/Papaya-Fruit-Disease-Detection

[19] React JS Library: https://react.dev/

[20] Tailwind CSS Framework Guide: https://tailwindcss.com/docs/installation/framework-guides

[21] ViteJS Build Tool Guide: https://vitejs.dev/guide/

[22] Machine Learning Guidance Using Krish

[23] Naik's Playlist: https://www.youtube.com/channel/UCNU_lfiiWBdtULKOw6X0Dig

[24] S. J. Banarase and S. D. Shirbahadurkar, "Papaya

[25] Diseases Detection Using GLCM Feature

[26] Extraction and Hyperparatuning of Machine

[27] S. J. Banarase and S. D. Shirbahadurkar, "Papaya

[28] Diseases Detection Using GLCM Feature

[29] Extraction and Hyperparatuning of Machine

[30] S. J. Banarase and S. D. Shirbahadurkar, "Papaya

[31] Diseases Detection Using GLCM Feature

[32] Extraction and Hyperparatuning of Machine

[33] S. J. Banarase and S. D. Shirbahadurkar, "Papaya

[34] Diseases Detection Using GLCM Feature

[35] Extraction and Hyperparatu