[1]*Deepali Joshi

[2]Harshali Patil

[3]Varsha Bendre

[4]Jayashree Katti

Classification of Mango Leaf Disease using Machine Learning to Enhance Yield Productivity

Journal of Electrical Systems

Abstract: - Fruits provide critical nourishment to the human body. Proper Care and upkeep are essential for the fruit to be healthy. Lack of maintenance, infections, spot, fungus all cause significant production and profit losses. Mango is a seasonal and famous fruit which is consumed all over the world. It is a delicate fruit that is susceptible to disease that reduce the quality as well as quantity. Manual illness or infection inspection is a time-consuming and labour-intensive technique that necessitates a large amount of resources and therefore it is inefficient. On the other hand, automatic inspection provides various advantages such as less time consuming, less labour and also the number of resources required are less. Image classification techniques and algorithms can be used to distinguish between infected and healthy mangoes, decreasing losses.

Keywords: Mango, Mango Leaves, Mango disease, Production, ANN, KNN, Machine Learning

I. INTRODUCTION

India is the world's largest producer of fruits, vegetables, flowers, grains, and pulses. Mango, known as the king of all fruits, is a famous and well-liked fruit by people of all ages all over the world. In fact the entire mango tree offers numerous benefits, such as it has anti-microbial qualities, it releases large amounts of oxygen. Mango fruit has various characteristics such as it is high in fiber, vitamins, antioxidants, and is good for the intestines. Despite all of its benefits, the mango tree and fruit are extremely delicate and susceptible to illness. There are various diseases that infect several parts of the mango tree such as fruit, leaves, and flowers. First and foremost, it is difficult to detect the disease, and if the mango tree or fruit is contaminated with any disease, the yield suffers in terms of quality and quantity.

Once the tree or its parts are infected with diseases the quantity and the quality of the yield is affected and the farmer loses the profit. There are various methods through which diseases can be detected. The farmers can utilize the manual or automatic disease detection. If the farmer follows the manual disease detection method it requires time, skilled labourers who can accurately detect the diseases and is a tedious process. On the other hand if the automatic method is followed then it will be accurate, less time consuming and will require less number of resources.

II. LITERATURE SURVEY

The plant disease detection and prediction literature survey comprises reviewing numerous studies. The datasets were subjected to a variety of algorithms by the authors. Some of the authors worked with real-time data, while others used data obtained from web resources.

Rajeev Singh [4], the authors discuss fungal illnesses, specifically Anthracnose disease. CNN and Alexnet are used to identify illness. The mango leaves utilized in the dataset were obtained from the GBPUAT field location.

Siddharth Singh Chouhan [5], the authors concentrate on establishing an effective approach for diagnosing Anthracnose and its symptoms. The MCNN algorithm is presented for classifying Mango leaves afflicted with the fungal disease Anthracnose. The dataset under consideration contains 1070 photos of mango leaves. The proposed model's accuracy is 97.13%.

[1 1] * Department of Computer Engineering, Thakur College of Engineering & Technology, Maharashtra, India, deepali.joshi@thakureducation.org

[2] Department of Computer Engineering, Thakur College of Engineering & Technology, Maharashtra, India, harshali.patil@thakureducation.org

[3] Department of E&TC, Pimpri Chinchwad College of Engineering, Maharashtra, India, varsha.bendre@pccoepune.org

[4] Department of Information Technology, Pimpri Chinchwad College of Engineering, Maharashtra, India, jayashree.katti@pccoepune.org

Swathi R [7], bacterial canker, Scab, and Powdery Mildew are three diseases on which the authors have worked. The Alexnet model has been trained and tested to determine whether the leaves are diseased or healthy. The dataset under consideration consists of mango leaves. The proposed model has an 89% accuracy.

Ram Prasad [8], author have worked on Anthracnose and Scab. The input is gathered from several sources and is divided into three categories: Anthracnose, Scab, and Healthy. The dataset is divided into 80%-20% training data and 20% testing data. The implementation employs a total of 20 epochs. CNN was utilized as the method to distinguish between healthy and diseased mangoes. The suggested model's accuracy is 91.8% for 64 batch sizes, 95.6% for 32 batch sizes, and 94.3% for 16 batch sizes.

Ly Van Tran [9], Authors have worked on Anthracnose, Powdery Mildew, and Gall Mildew. VGG, Alexnet, ResNet, and ResNet 50 are algorithms employed. The dataset considered consists of leaves. The model is implemented using Adaptive Particle-Grey Wolf metaheuristic. The strategy is better than deep learning methods like VGG, AlexNet, and ResNet-50, which were upgraded considering transfer learning, with scores of 89.41%, 78.64%, 79.92%, and 84.88%, respectively.

Yu-Lun Hsu [10], the authors address the topic of fruit grading utilizing CNN algorithms using various models like Mask R-CNN, AlexNet, VGGs, ResNets. Mango fruit visuals make up the dataset.

Sunidhi A [12], the writers concentrate on the Anthracnose disease. CNN is the algorithm employed. The mango leaf dataset under consideration was acquired in real time from various states like Karnataka, Maharashtra, New Delhi. The approach has a classification accuracy of roughly 96.16%.

Parinya Chantrasi [13], the authors develop a model which will detect Anthracnose, with CNN algorithm serving as the system's major core. Mango fruit was utilized as the dataset. The primary goal is to separate the sick mangoes. The method is 70% accurate.

Rajneet Kaur [14], Anthracnose, Red Rust, and Powdery Mildew are among illnesses studied. The authors suggested a model based on CNNs for detection, classification of leaf diseases. dataset under consideration consists of leaves. The CNN is trained to identify, characterize leaf diseases. The CNN-based model is 90.36% accurate.

Md Nadeem [15], the writers give detailed descriptions of numerous diseases that afflict mango trees. PlantVillage, DigiPathos, Indian leaves, and Indonesian mango leaves are among datasets which are analyzed. The symptoms are stated together with an illustration, making it simple to grasp. CNN, CNN-SVM, and Resnet 50 are the algorithms employed, with CNN providing the best accuracy. The dataset under consideration consists of mango leaves.

Idy Diop [16], has contributed to a review study in which the author concentrates on the CNN algorithm. Mango leaf and fruit are included in the dataset. The method is classified as traditional ML-based solutions, DL-based solutions. The types of diagnosis: automatic, manual. It indicates a full overview of how the two methods function and the differences between them.

TABLE 1: LITERATURE REVIEW SUMMARY

Paper Tile	**Algorithms, Mango Disease**	**Findings**	**Gaps**
"Mango Leaf Diseases Detection using Deep Learning", IJKBCS	CNN Anthracnose, Red Rust, Powdery Mildew.	The CNN model achieves an accuracy of 90.36%.	Further research is required on various parts of the plant like fruit, flowers.
"MangoLeafBD: A Comprehensive Image Dataset to Classify Diseased and Healthy Mango Leaves", Research Gate	CNN, CNN-SVM, Resnet 50 Anthracnose, Bacterial Canker, Cutting weevil, Die Back, Sooty Mould, Powdery Mildew.	CNN algorithm provides highest accuracy	Further research is required on various parts of the plant like fruit, flowers.

"Classification of Mango Leaves Infected by Fungal Disease Anthracnose Using Deep Learning", ICCMC 2021,, IEEE Xplore Part No: CFP21K25-ART, 2021.	CNN Anthracnose	The algorithm proposed by the authors provides classification accuracy of about 96.16%.	It is important to address various diseases that infect the plant/ tree. Further research is required on different parts of the plant such as fruit, flowers.
"Deep Learning Precision Farming: Grapes and Mangoes Leaf Disease Detection by Transfer Learning", Global Transitions Proceedings 2(2021) 535-544, KeAi, Chinese Roots Global Impact	Alexnet Bacterial canker, Scab, Powdery Mildew	The proposed model provides accuracy of 89%.	Further research is required on different parts of the plant such as fruit, flowers.
"Detection and Classification of Diseased Mangoes", Research Gate, November 2020, DOI: 10.1109/ICOSICA49951.2020.9243277	CNN Anthracnose, Scab	The accuracy of the methodology with batch size 64 is 91.8%, for batch size 33 is 95.6%, for batch size 16 is 94.3%.	Further research is required on different parts of the plant such as fruit, flowers.
"Multilayer Convolutional Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease", IEEE Access, March 2019, DOP: March 27,2019	CNN, MCNN Anthracnose	The model achieves accuracy of 97.13% .	It is important to address various diseases that infect the plant/ tree.
"Early Disease Classification of Mango Leaves using Feed-Forward Neural Network and Hybrid Metaheuristic Feature Selection", IEEE Access, October 12,2020	ANN, Alexnet Anthracnose	transfer learning values 89.41% vs 78.64%, 79.92%, and 84.88%.	Further research is required on different parts of the plant such as fruit, flowers.
"Mango Diseases Classification Solutions Using Machine Learning or Deep Learning: A Review", Journal of Computer and Communications, 2022,10,16-28, Scientific Research Publishing	CNN Describes Mango diseases.	The author is providing detailed description about the algorithm	It The disease that infect plant should be addressed.
"Machine Learning Algorithm Development for detection of Mango infected by Anthracnose Disease", 2021 IEEE, The 6th International Conference on Digital Arts, Media and Technology (DAMT) and 4th ECTI Northern Section Conference on Electrical, Electronics, Computer and	CNN Anthracnose	The methodology achieves accuracy greater than 70%.	Further research is required on different parts of the plant such as fruit, flowers.

Telecommunications Engineering (NCON)			
"Deep Learning for Automatic Quality Grading of Mangoes: Methods and Insights", 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA), ©2020 IEEE	Grading of mango Fruit	The issues of grading mango in different categories are identified.	Further research is required on different parts of plant such as leaf, flowers.
"A Comparative study of CNN and Alexnet for Detection of Disease in Potato and Mango Leaf", 2019 2nd International Conference on ICICT, 2019 IEEE	Anthracnose	The issues of anthracnose disease are identified.	It is important to address various diseases that infect the plant/ tree.
"Hybrid approach for Anthracnose detection using intensity and size features", 2016 IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development	Anthracnose	Detection of affected areas using the proposed hybrid approach.	Further research is required on different parts of the plant such as leaf, flowers.

III. METHODOLOGY

The proposed methodology consists of numerous steps such as the flow, acquisition of data, and the proposed model. The detailed description is as follows.

A. Process Flow

1) Kaggle dataset is used to collect on field data that contains diseased mango as well as healthy mango leaves from two classes.

2) Preprocessing of leaf pictures includes cropping, contrast improvement with equalization of Histogram.

3) Image Labelling.

4) The dataset is segregated as: training as well as testing dataset.

5) The algorithms are trained using a training dataset and the testing will be done through test dataset.

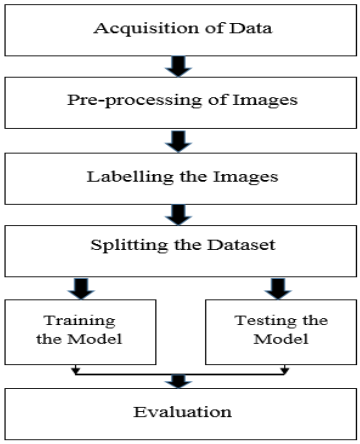

Figure 1 Block Diagram of Proposed Model

B. Acquisition of Dataset

Kaggle Dataset is used, comprises image data. The categories include: healthy, unhealthy leaf images. There are eight classes and consists of 500 photos in each category.

A 3500 picture dataset was utilized, with 80% used for training and 20% used for validation. The images given below are the sample dataset.

Figure 2 Diseased Leaf

Figure. 3 Healthy Leaf

C. Proposed Model

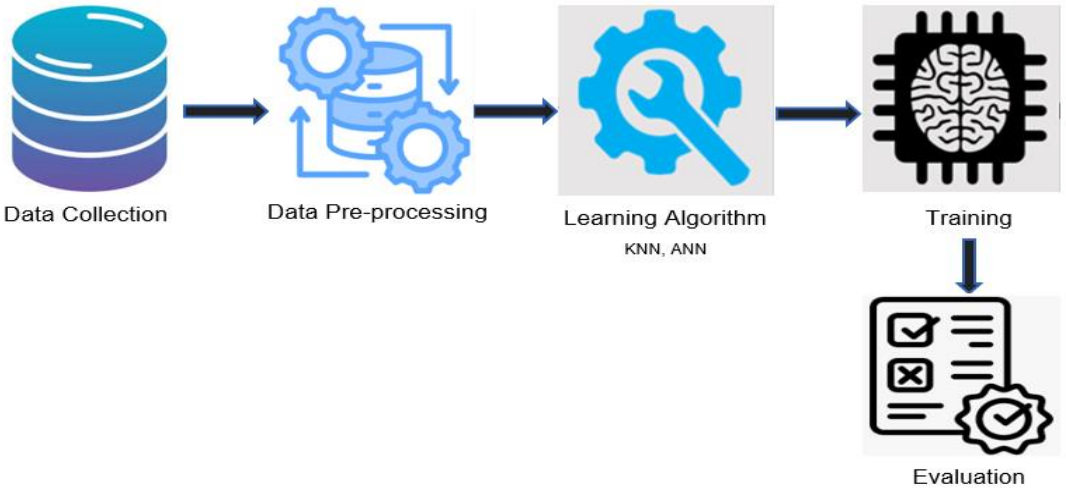

Figure. 4 Proposed Model

The proposed model comprises various steps such as data collection, data pre-processing, learning algorithms, training and evaluation. Each of the step plays a crucial role in the working of proposed method.

1) Data Collection

The kaggle dataset of images is used. It consists of healthy and unhealthy leaf images with eight groups. There are 500 photos in each category[18].

2) Data Pre-processing

It refers to processing the data in raw format and transforming into information. The phase is crucial for development of model.

3) Learning Algorithms

KNN and ANN are the learning algorithms employed in the suggested method for detecting mango leaf disease.

4) Training

A dataset of 3500 photos was utilized for training, with 80% used for validation.

5) Evaluation

The evaluation process comprises of evaluating the data using the various models.

D. Artificial Neural Networks

At foundation of Deep Learning, the advanced application of Machine Learning techniques, are ANN. ANNs are flexible, scalable, versatile, making them ideal for managing large datasets and incredibly and extremely difficult Machine Learning tasks such as classifying images, recognizing speech, recommending videos, or analysing consumer sentiment. The structure of ANN algorithm is similar to human brain and therefore the term "artificial neural network". ANN features neurons which are interconnected to each at various network levels.

ANN Layers are as follows:

1) Input Layer:

The input layer will accept inputs in different formats from the programmer.

2) Hidden Layer:

It is visible between input, output layer. It does the computations required to get patterns and features.

3) It is identifies the output once the input goes through various alterations in hidden layer.

Advantages of ANN:

1) Parallel processing:

ANN have a numerical value that allows them to carry out multiple tasks simultaneously.

2) Ability to work with incomplete information:

After ANN training, the information may still provide output with insufficient data. However due to missing data the performance of the algorithm is affected.

Disadvantages:

1) Network structure

There is no set formula for figuring out how artificial neural networks should be structured. Trial and error method is used to achieve the right network structure.

2) Hardware dependence

Due to the topology of artificial neural networks, parallel processors are required. As a result, the equipment's realization is dependent.

E. K-Nearest Neighbors Algorithm

It assigns new data that will resemble the existent categories based on the presumption and the new and old data are compared. The classification of new data depends on its similarity to existing data. The new data can be precisely segregated into an suitable category using K-NN. KNN algorithm saves information during its training phase, and when new data is received it will be categorized as per its similarity to existing data.

Advantages:

1) Simple implementation

2) It is robust to noise in data

3) Effective when the training data is large.

Disadvantages:

1) The value of K should be identified which may be complex in some cases.

2) High computation cost due to distance calculations between the data points

IV. RESULTS AND DISCUSSION

The "Mango Leaf Disease Dataset" dataset from kaggle is used for the implementation of the proposed model. It has over 4000 photos of mango leaf. There are eight classifications, with healthy leaves and diseased leaves. The size of dataset is approximately 108MB. Implementation of KNN, ANN is done. The following are the algorithm results:

A. K-Nearest Neighbors Algorithm

The figure indicates the results that are achieved for KNN algorithm. Parameters such as precision, recall, f1-score and support are considered. The achieved accuracy is 66%.

Classification Report

	precision	recall	f1-score	support
0	0.78	0.57	0.66	102
1	1.00	0.59	0.74	117
2	0.71	0.51	0.60	101
3	0.69	0.93	0.79	87
4	0.81	0.68	0.74	99
5	0.64	0.57	0.60	91
6	0.84	0.51	0.64	96
7	0.68	0.93	0.78	107
micro avg	0.75	0.66	0.70	800
macro avg	0.77	0.66	0.69	800
weighted avg	0.78	0.66	0.70	800
samples avg	0.66	0.66	0.66	800

Figure 5 KNN Classification Report

The macro average and weighted average accuracy are given in the graph below. The observed values are 0.77, 0.66, 0.69 and 0.78, 0.66, 0.66.

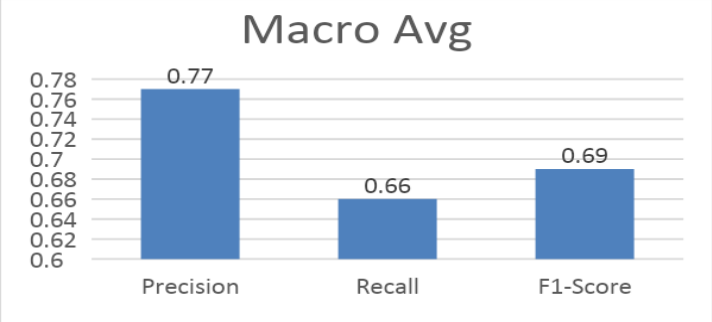

Figure 6 KNN Macro Avg

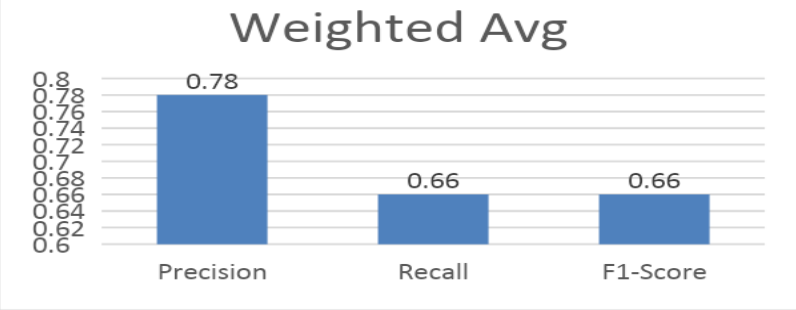

Figure. 7 KNN Weighted Avg

B. Artificial Neural Networks

The results of ANN algorithm is given below. The parameters considered are precision, recall, f1-score, support. The accuracy achieved by ANN algorithm is 62%.

Classification Report

	precision	recall	f1-score	support
0	0.61	0.42	0.50	102
1	0.99	0.72	0.83	117
2	0.65	0.80	0.72	101
3	0.82	0.54	0.65	87
4	0.25	0.02	0.04	99
5	0.29	0.78	0.43	91
6	0.89	0.75	0.81	96
7	0.71	0.88	0.79	107
accuracy			0.62	800
macro avg	0.65	0.61	0.60	800
weighted avg	0.66	0.62	0.60	800

Figure 8 Classification Report of ANN

The macro average and weighted average accuracy values are given below. It is observed that the macro average values are 0.65, 0.61, 0.60 and the weighted average values are 0.66, 0.62, 0.60.

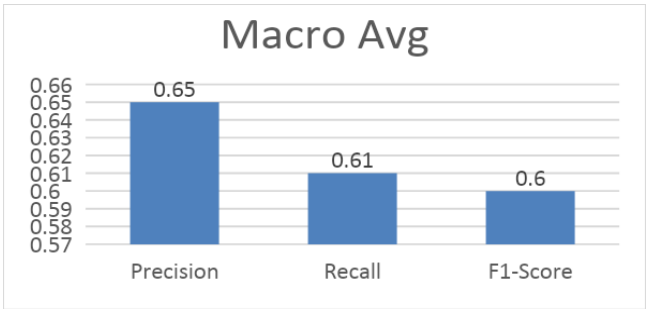

Figure. 9 ANN Macro Avg

Figure. 10 ANN Weighted Avg

The accuracy achieved by KNN, ANN algorithms is given below:

Table 2: Accuracy Comparison for ANN, KNN

Sr.No	Algorithm	Accuracy %
1	ANN	62%
2	KNN	66%

The figure given below depicts the results obtained from KNN, ANN.

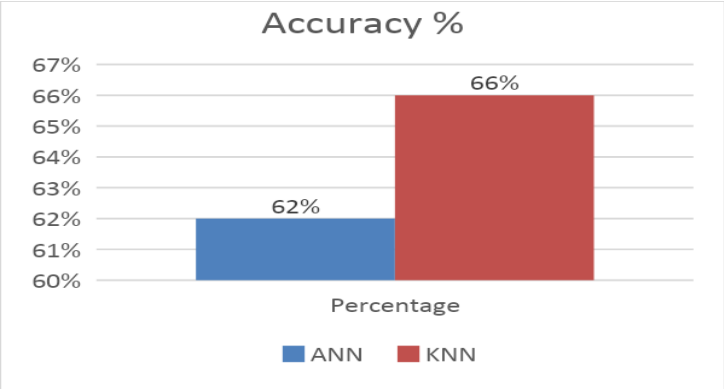

Figure. 11 ANN, KNN Accuracy %

V. CONCLUSION

The mango leaves dataset from kaggle is explored in this research. There are total eight types of leaves, which include seven sick leaves and one healthy leaf. To detect infected leaves, algorithms such as KNN, ANN are used. Training and testing datasets taken into account is 80%-20%. The KNN has a 66% accuracy, whereas the ANN has a 62% accuracy.

REFERENCES

[1] Peter Flach, "Machine Learning: The Art and Science of Algorithms that Make Sense of Data", Cambridge University Press, 20-Sept-2012

[2] K.Swetha, Veena V, Sadhana.G.D, R Priyatharshini, "Hybrid approach for Anthracnose detection using intensity and size features", 2016 IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development, 978-1-5090-0615-1/16

[3] Andreas C. Müller, Sarah Guido, "Introduction to Machine Learning with Python", O'Reilly Media, Inc., September 2016, ISBN: 9781449369897

[4] Sunanya Arya, Rajeev Singh, "A Comparitive study of CNN and Alexnet for Detection of Disease in Potato and Mango Leaf", 2019 2nd International Conference on ICICT, 978-1-7281-1772-0 2019 IEEE

[5] Uday Pratap Singh , Siddharth Singh Chouhan , Sukirty Jain , Sanjeev Jain, "Multilayer Convolutional Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease", IEEE Access, Volume 07, DOI:10.1109/ACCESS.2019.2907383, March 2019

[6] P Rama Koteswara Rao, K Swathi, "Mango Plant Disease Detection using Modified Multi Support Vector Machine Algorithm", PJAEE, 17 (7) 2020

[7] Sanath Rao, U. ,Swathi, R. ,Sanjana, V. ,Arpitha, L. ,Chandrasekhar, K. , Chinmayi ,Naik, Pramod Kumar, "Deep Learning Precision Farming: Grapes and Mangoes Leaf Disease Detection by Transfer Learning", Global Transitions Proceedings 2(2021) 535-544, KeAi, Chinese Roots Global Impact

[8] Akshay Koushik, Ritanya B Bharadwaj, Ram Prasad, Sana Habeeb,"Detection and Classification of Diseased Mangoes", ICOSICA, Research Gate, November 2020, DOI: 10.1109/ICOSICA49951.2020.9243277

[9] Tan Nhat Pham , Ly Van Tran, and Son Vu Truong Dao, "Early Disease Classification of Mango Leaves using Feed-Forward Neural Network and Hybrid Metaheuristic Feature Selection", IEEE Access, October 12,2020, DOI:10.1109/ACCESS.2020.3031914

[10] Shih-Lun Wu, Hsiao-Yen Tung, and Yu-Lun Hsu," Deep Learning for Automatic Quality Grading of Mangoes: Methods and Insights", 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA), 978-1-7281-8470-8/20/$31.00 ©2020 IEEE DOI 10.1109/ICMLA51294.2020.00076

[11] Balasubramanyam meda, S Kalyan, "Early Disease Classification of Mango Leaves using Feed- forward Neural Network Model", IJTE, Volume XIII, Issue III, 2021

[12] Pankaj Kumar,Sunidhi Ashtekar,S. S Jayakrishna, K P Bharath, P. T Vanathi, M Rajesh Kumar,"Classification of Mango Leaves Infected by Fungal Disease Anthracnose Using Deep Learning:, ICCMC 2021, IEEE Xplore Part No: CFP21K25-ART, DOI:10.1109/ICCMC51019.2021.9418383

[13] Suwit Wongsila, Parinya Chantrasri , Pradom Sureephong,"Machine Learning Algorithm Development for detection of Mango infected by Anthracnose Disease", The 6th International Conference on Digital Arts, Media and Technology (DAMT) and 4th ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (NCON), 978-1-6654-1569-9/21/$31.00 ©2021 IEEE

[14] Amisha Sharma, Rajneet Kaur Bijral , Jatinder Manhas, Vinod Sharma, "Mango Leaf Diseases Detection using Deep Learning", International Journal of Knowledge Based Computer Systems, IJKBCS, Volume 10, Issue 1, June 2022, 40-44

[15] Sarder Iftekhar Ahmed , Muhammad Ibrahim , Md Nadim , Md Mizanur Rahman , Maria Mehjabin Shejunti , Taskeed Jabid , Md Sawkat Ali ,"MangoLeafBD: A Comprehensive Image Dataset to Classify Diseased and Healthy Mango Leaves", Research Gate, Aug 2022, Elsevier

[16] Dembe Faye, Idy Diop, Doudou Dione, "Mango Diseases Classification Solutions Using Machine Learning or Deep Learning: A Review", Journal of Computer and Communications, 2022,10,16-28, Scientific Research Publishing, ISSN Online: 2327-5227, ISSN Print: 2327-5219

[17] Israa Mohammed Hassoon, "Classification and Diseases Identification of Mango Based on Artificial Intelligence: A Review", Journal of Al-Qadisiyah for Computer Science and Mathematics Vol. 14(4) 2022, pp Comp. 39–52

[18]] Kaggle Dataset

[19] Asha Rani K P, Gowrishankar S "ConvNeXt-based Mango Leaf Disease Detection: Differentiating Pathogens and Pests for Improved Accuracy", International Journal of Advanced Computer Science and Applications (IJACSA), Volume 14 Issue 6, 2023.

[20] Sapna B Kulkarni, J Keerthi, "Mango Leaf Disease Detection Using Deep Learning", IJRASET, ISSN: 2321-9653, Volume 11 Issue VIII Aug 2023