Applied Deep learning approaches on canker effected leaves to enhance the detection of the disease using Image Embedding and Machine learning Techniques

K Badri Narayanan[1*], Devatha Krishna Sai[2], Korrapati Akhil Chowdary[3], Srinivasa Reddy K[4]

[1,3,4]School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
[2]School of Electronics and Communication Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

Abstract

Canker, a disease that causes considerable financial losses in the agricultural business, is a small deep lesion that is visible on the leaves of many plants, especially citrus/apple trees. Canker detection is critical for limiting its spread and minimizing harm. To address this issue, we describe a computer vision-based technique that detects Canker in citrus leaves using image embedding and machine learning (ML) algorithms. The major steps in our proposed model include image embedding, and machine learning model training and testing. We started with preprocessing and then used image embedding techniques like Inception V3 and VGG 16 to turn the ROIs into feature vectors that retained the relevant information associated with Canker leaf disease, using the feature vectors acquired from the embedding stage, we then train and evaluate various ML models such as support vector machines (SVM), Gradient Boosting, neural network, and K Nearest Neighbor. Our experimental results utilizing a citrus leaf picture dataset show that the proposed strategy works. With Inception V3 as the image embedder and neural network machine learning model we have obtained an accuracy of 95.6% which suggests that our approach is effective in canker identification. Our method skips traditional image processing techniques that rely on by hand features and produces results equivalent to cutting-edge methods that use deep learning models. Finally, our proposed method provides a dependable and efficient method for detecting Canker in leaves. Farmers and agricultural specialists can benefit greatly from early illness diagnosis and quick intervention to avoid disease spread as adoption of such methods can significantly reduce the losses incurred by farmers and improve the quality of agricultural produce.

Keywords: Canker, Convolutional Neural Network, Machine Learning, Image Embedding, InceptionV3

Received on 01 December 2023, accepted on 01 March 2024, published on 08 March 2024

doi: 10.4108/eetiot.5346

*Corresponding author. Email: badrinarayanan78@gmail.com

1. Introduction

Canker poses a serious threat to the production of oranges, apples, and stone fruits in many countries. The disease manifests as spots on the leaf surface, which can merge and cause extensive areas of dead tissue. It can be caused by different types of fungi or bacteria. Symptoms of canker disease can be observed on leaves, stems, or fruit [3]. The pathogens invade the plant's bark, branches, and leaves, leading to lesions or open sores.

These diseases have a detrimental impact on fruit yield and overall plant health and are often challenging to control [15]. The spread of these diseases through contaminated tools, infected plant material, wind, and rain makes it difficult to predict and detect them early. To address this issue, advanced techniques like deep learning approaches, particularly those based on convolutional neural networks (CNN), are being utilized for early detection and management of canker leaf disease. These techniques enhance the accuracy of detecting affected leaves. Accurate diagnosis and management of cankers are crucial for maintaining plant health and productivity. The agriculture industry has witnessed increased yields due to recent technological advancements [10]. However, without accurate disease diagnosis, implementing effective control measures in a timely manner is not possible [15]. Plant leaf diseases, such as canker, pose a severe threat to various crops and agricultural productivity. Traditional categorization methods, such as visual inspection and laboratory tests, have limitations, including subjectivity and time consumption. Deep learning techniques, particularly those based on CNN, are now widely used to classify plant diseases [4][5][12]. Image processing is a popular method for identifying and categorizing plant leaf diseases [15]. Deep learning tools, including image embedding and machine learning algorithms, show great potential in the detection and control of canker leaf disease. By extracting features from images of canker-infected leaves using image embedding techniques and feeding them into deep learning and machine learning models, the accuracy of detection and classification of canker disease in plants can be improved. These advancements eliminate the need for manual detection and diagnosis, which are labor-intensive and time-consuming. In this study, we employed convolutional neural networks, a type of deep learning model, for image embedding to extract features and accurately classify different types of canker disease affecting plants. Machine learning models such as SVM, Gradient Boosting, Neural Network, and KNN were utilized to achieve accurate detection of canker disease. The objective of these disease-specific models is to efficiently and accurately identify various diseases that can affect single or multiple crops, including canker disease in plants. Machine learning approaches have proven successful in automating disease detection by recognizing plant illnesses [16]. Ghosh et al.'s 2023 study focuses on "Water Quality Assessment Through Predictive Machine Learning",[17] highlighting the use of machine learning for analyzing and predicting water quality parameters. In "Unraveling the Heterogeneity of Lower-Grade Gliomas," Rahat, Ghosh,[18] and colleagues (2023) delve into deep learning-assisted segmentation and genomic analysis of brain MR images, offering new insights into this medical condition. Potato Leaf Disease Recognition and Prediction using Convolutional Neural Networks," by Ghosh, Rahat, and team [19] showcases the application of convolutional neural networks in accurately identifying diseases in potato leaves. Mandava, Vinta, Ghosh, and Rahat's 2023 research presents "An All-Inclusive Machine Learning and Deep Learning Method [20] for Forecasting Cardiovascular Disease in Bangladeshi Population", integrating advanced AI techniques for health predictions. The 2023 study [21] by Mandava et al., titled "Identification and Categorization of Yellow Rust Infection in Wheat through Deep Learning Techniques", applies deep learning methods to detect and categorize wheat infections effectively. Khasim, Rahat, Ghosh,[22] and colleagues' 2023 article, "Using Deep Learning and Machine Learning: Real-Time Discernment and Diagnostics of Rice-Leaf Diseases in Bangladesh", explores AI-based solutions for diagnosing rice-leaf diseases. Deciphering Microorganisms through Intelligent Image Recognition", authored by Khasim, Ghosh, Rahat [23] and others in 2023, discusses the use of machine learning and deep learning in identifying microorganisms through advanced image recognition techniques. The 2023 study by Mohanty, Ghosh, Rahat, and Reddy, "Advanced Deep [24] Learning Models for Corn Leaf Disease Classification", focuses on the application of deep learning in classifying diseases in corn leaves based on a field study. Alenezi and team's 2021 research, "Block-Greedy and CNN[25] Based Underwater Image Dehazing for Novel Depth Estimation and Optimal Ambient Light", investigates novel CNN-based methods for enhancing underwater image clarity and depth estimation.

These advancements in deep learning techniques, incorporating image processing and machine learning algorithms, hold great potential for the detection and control of canker leaf disease in agriculture. They enable the tracking of disease severity and progression over time, 3 while reducing the need for manual detection and diagnosis, which is time-consuming and labor-intensive.

2. Related Work

Plant leaf diseases have a significant impact on the quantity and quality of agricultural products. These diseases can be identified by observing symptoms on leaves, stems, or fruit. Canker leaf disease, for example, leads to leaf spots, defoliation, stunted growth, reduced crop yields, and inferior fruit quality. Severe cases can even result in plant death, incurring substantial costs for farmers and the agricultural sector. Traditional categorization methods, such as visual observation and laboratory tests, have limitations in terms of time consumption and subjectivity. To address these challenges, machine learning and deep learning techniques, particularly those based on convolutional neural networks (CNN), have gained popularity for plant disease classification.

Several studies have compared the effectiveness of machine learning and deep learning methods for detecting and classifying various plant diseases. In one such study, machine learning algorithms such as Support Vector Machine (SVM), Random Forest (RF), and Stochastic Gradient Descent (SGD), along with deep learning models like Inception-v3, VGG-16, and VGG-19, were evaluated for citrus plant disease detection. The results showed that deep learning methods, specifically Inception-v3 and VGG-16, outperformed machine learning methods in terms of disease identification accuracy. Other studies also reported high accuracy rates for disease detection using deep learning architectures such as mobilenetv2, deep CNN with 9 layers, and CNN combined with SVM. These approaches achieved accuracy rates ranging from 95% to 96.46%.

Different image processing and machine learning techniques have been proposed for the automated identification of diseases in specific crops. For example, in grapevine disease identification, approaches utilizing grab cut segmentation, K-means clustering, and SVM achieved accuracies of 93% and 88.89%. In rice leaf disease detection, machine learning methods including KNN, Decision Tree, Naive Bayes, and Logistic Regression were employed, with the decision tree method achieving an accuracy above 97%. CNN and SVM classifiers were

combined for the classification of major rice leaf diseases, yielding accuracy rates of 91.37%. Similarly, in potato leaf disease categorization, a deep learning model achieved an accuracy of 97.2%. The combination of EfficientNetB0 and InceptionV3 demonstrated the highest accuracy of 91.14% in classifying pear leaf diseases. For cotton leaf disease identification, an adaptive neurofuzzy inference system achieved an accuracy of 85% using pattern recognition and Hu's moments.
Overall, these studies highlight the effectiveness of machine learning and deep learning techniques for accurate and efficient detection and classification of plant leaf diseases, providing valuable tools for farmers and the agricultural industry.

3. Methodology

This study was conducted using a dataset downloaded from Mendeley Dataport which consisted of 163 images of canker infected leaves and 446 images of leaves infected with various other diseases like black spot, greening, scab, melanose and healthy leaves as well. Sample images of the considered dataset are shown in Fig.1. The method followed majorly includes Image Embedding, to which the image dataset is given as input, here we used two deep learning based image embedders namely Inception V3 and VGG16, upon image embedding we obtain the features, these features are now given to various machine learning models as input with a train tesr split of 70% and 30%, and 5 fold cross validation, here we used SVM, Gradient Boosting, Neural Network and KNN, which then detect the Canker leaf disease. Fig.2. shows the block diagram of the method followed.

Fig.1. Canker and Non-Canker sample image

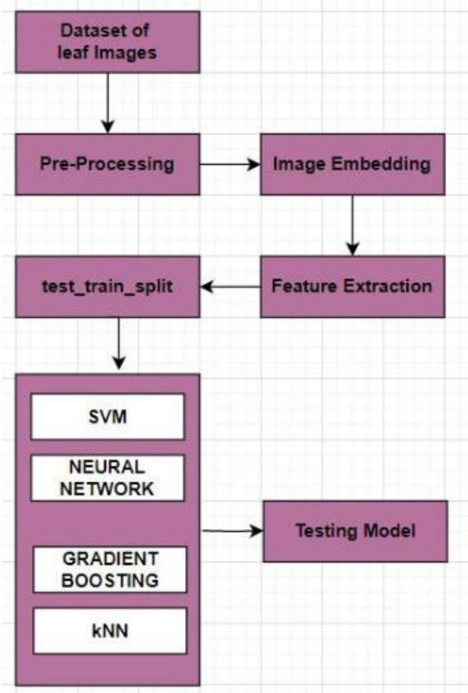

Fig.2. Methodology block diagram

3.1 Image Embedding

Image embedding is a technique used to create a high-dimensional vector representation of an image, which can be applied in various computer vision applications. This approach has gained popularity due to its ability to efficiently encode visual information. Typically, a pre-trained (CNN) such as VGG or Inception is employed to generate these image embeddings. These networks are trained on image datasets to identify and extract information from images. The output of the layer just prior to the final softmax classification layer is used as the image embedding. This output represents the image as a high-dimensional vector, where each dimension corresponds to a learned feature of the CNN. Compared to using raw images, image embeddings offer several advantages. They are more compact and efficient, making them particularly useful in large-scale computer vision applications as they can be easily stored and transmitted. Image embeddings are also more robust to changes in the input image, such as variations in size, rotation, or lighting conditions. This allows the CNN to extract important features from the image that are resilient to such changes. Moreover, image embeddings can be compared more effectively and quickly using common distance measures like Euclidean distance or cosine similarity, enabling rapid retrieval of similar images from a database. 6 In the context of image classification, image embeddings can be utilized to predict the class of an input image. The image embedding serves as input to a classifier, which can range from a simple linear classifier to a complex deep neural network. During the training process, the classifier learns to associate the image embeddings with the corresponding class labels using a labeled dataset. In our study, we employed VGG16 and Inception V3 for image embedding on the selected image dataset.

3.1.1 VGG16

With a total of almost 138 million parameters, the VGG16. It uses an input image of size 224x224 and produces an image embedding with a 4096-dimensional feature vector. The final few layers of the network, including the fully connected levels, can be removed in order to use VGG16 as an image embedder, and the output of the final convolutional layer can be used instead. This output is a 512x7x7 tensor, which can be flattened and processed through a few extra layers to produce a compact picture embedding, such as a completely connected layer or a global pooling layer. The extacted feature vectors are now given as input to various machine learning models as input for accurate detection of canker leaf disease.

3.1.2 Inception V3

The 42 layers of the Inception v3 architecture include many convolutional layers with various kernel sizes and pooling techniques, as well as a number of inception modules that enable more effective use of the available computing resources. It uses an input image of size 299x299 and produces an image embedding with a 2048-dimensional feature vector. Remove the final few layers of the network, including the fully connected levels, and use the output of the final convolutional layer as the image embedding to use Inception v3 as an image embedder. This result is an 8x8x2048 tensor, which can be flattened and passed through a few further layers to produce a compact picture embedding, such as a completely connected layer or a global pooling layer. The extacted feature vectors are now given as input to various machine learning models as input for accurate detection of canker leaf disease.

3.2 Machine Learning Classifiers

3.2.1 Neural network

The ability of the neural network model to recognize intricate patterns and relationships in data was a deciding factor. The testing set of the preprocessed data of leaf images with canker disease and other diseases was used to assess the effectiveness of the neural network model. Early pausing during training and regularization strategies like dropout were used to prevent overfitting. The generalizability of the findings, however, might be constrained given the dataset's small size. Additionally, given the complexity of the neural network model, training and evaluating it may take a sizable amount of computational effort. In this study, the Neural Network model was found to be a useful tool for classification problems. The Neural Network model can shed important light on complicated datasets with careful preprocessing and model adjustment. 7

3.2.2 KNN

For the purpose of identifying canker in plant leaves, the classification algorithm KNN (K-Nearest Neighbours) was employed. The algorithm determined which class had the majority of the k nearest neighbours to a given test instance in the feature space, and then it assigned the test instance to that class. The KNN method was trained on a series of labelled images in the context of canker identification, where each image was represented by a set of features extracted using deep learning image embedders. On the basis of the attributes they retrieved, the KNN model was then tested and then the accuracy was attained in identifying plant canker by fine-tuning the KNN model with the proper hyperparameters and feature selection strategies.

3.2.3 SVM

For the classification task of detecting canker spot on leaves in our research study, we employed this model. The SVM model was chosen owing to its capacity for handling complicated tasks and the advantages of having non-linear data, which have been successfully applied in other research disciplines. The dataset utilized in this study was compiled from open-access sources and preprocessed [20] to weed out any inaccurate or missing data. Using the testing set of the preprocessed data, the SVM model's performance was assessed. Since the SVM model that we are utilizing in this instance was trained using a training set of preprocessed data. The kernel function, gamma value, and regularization parameter were among the hyperparameters that were combined in the best way possible. This was done using a grid search method. The final SVM model was trained using the chosen hyperparameters.

3.2.4 Gradient Boosting

Gradient boosting was trained on the features extracted using deep learning image embedders for canker detection, each weak model in the ensemble is trained using the errors of the previous model then the process began with a single model, which is used to make predictions on the training data. The mistakes of this model are then determined, and a second model is trained to predict the errors of the previous model. The predictions of the second model are added to the predictions of the first model, and the errors of the combined model are determined. This procedure is continued with multiple models until the complete ensemble is built, and after careful preprocessing and hyperparameter optimization, the Gradient Boosting with useful insights into datasets is produced.

4. Results and Discussion

Our study on detecting canker disease in leaves was performed by collecting images of leaves infected with canker and leaves infected with other diseases. We have preprocessed the images and then performed Image Embedding using two deep 8 learning models namely VGG16 and Inception V3. The features extracted upon the image embedding process are given as input to various machine learning models namely SVM, KNN, Gradient Boosting and Neural Network with a 5-fold cross validation, upon which we have obtained the below accuracies, F1 score, precision and recall represented in Table 1 and Table 2. We can see from Tables 1 and 2 that utilising Inception V3 as an image embedder and a neural network as a machine learning model yields the maximum accuracy of 95.6%. Figure 3 displays the four machine learning models Scores.

Embedders	Model	F1 Score	Precision	Recall
Inception V3	Neural Network	0.956	0.956	0.956
	kNN	0.924	0.924	0.925
	SVM	0.946	0.946	0.946
	GradientBoosting	0.936	0.937	0.937
VGG 16	Neural Network	0.955	0.955	0.955
	kNN	0.908	0.917	0.913
	SVM	0.943	0.944	0.944
	GradientBoosting	0.944	0.945	0.945

Table 1. F1-score, Precision and Recall of the considered ML models with respect to considered Image Embedders

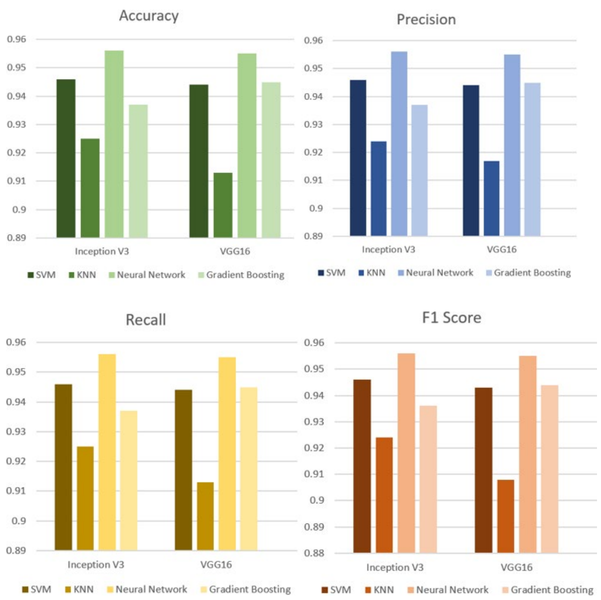

Fig.3. Graphical representation of overall performance of all classifiers

Embedders	Model	Average accuracy
Inception V3	Neural Network	0.956
	kNN	0.925
	SVM	0.946
	GradientBoostig	0.937
VGG 16	Neural Network	0.955
	kNN	0.913
	SVM	0.944
	GradientBoosting	0.945

Table 2. Accuracies of the considered ML models with respect to considered Image Embedders

4. Conclusions

The worldwide annual agricultural loss caused by leaf-affected disease accounts for approximately 14.1%, resulting in an economic loss of $220 billion. To address and minimize the impact of this disease on leaf-affected plants, we have developed an integrated model. This model allows for the early identification of canker leaf disease, enabling farmers to take necessary precautions and prevent substantial crop losses. In our study, we employ Deep Learning image embedders like VGG 16 and Inception V3 to create dense vector representations of the leaf images. Subsequently, we utilize various Machine Learning models. The highest accuracy achieved, at 95.6%, is obtained by employing Inception V3 as the image embedder and neural network as the machine learning model. Our deep learning model utilizes leaf images to detect disease in plants, specifically targeting canker-affected leaves. This integrated approach, combining image embedding and machine learning algorithms, proves to be a valuable tool for agricultural management systems, showcasing significant potential in early identification of canker leaf disease. The importance of early disease detection in crops cannot be overstated, as it leads to substantial savings for farmers and the agricultural industry as a whole. 10 Additionally, this integrated technology outperforms the traditional way of identifying sick leaves and saves farmers a great deal of labour and time. The work emphasises the importance of canker leaf disease early detection and suggests an integrated model that uses deep learning embedding and machine learning techniques for precise and prompt diagnosis. In the future, we also intend to create a mobile application that will enable users to submit real-time photos taken with their smartphones or obtained through picture repositories. With the help of this application, early detection will be even more successful, and farmers will be more aware of the value of taking preventive measures to lessen crop damage and lessen considerable financial losses brought on by canker leaf disease.

References

[1] Syed-Ab-Rahman, Sharifah Farhana, Mohammad Hesam Hesamian, and Mukesh Prasad. "Citrus disease detection and classification using end-to-end anchor-based deep learning model." Applied Intelligence 52.1 (2022): 927-938.

[2] Rauf, Hafiz Tayyab; Saleem, Basharat ALi ; Lali, M. Ikram Ullah ; Khan, Muhammad Attique ; Sharif, Muhammad ; Bukhari, Syed Ahmad Chan (2019), “A Citrus Fruits and Leaves Dataset for Detection and Classification of Citrus Diseases through Machine Learning”, Mendeley Data, V2, doi:10.17632/3f83gxmv57.2.

[3] Gavhale, Kiran R., Ujwalla Gawande, and Kamal O. Hajari. "Unhealthy region of citrus leaf detection using image processing techniques." International Conference for Convergence for Technology-2014. IEEE, 2014.

[4] Pydipati, R., T. F. Burks, and W. S. Lee. "Identification of citrus disease using color texture features and discriminant analysis." Computers and electronics in agriculture 52.1-2 (2006): 49-59.

[5] Barman, Utpal, et al. "Comparison of convolution neural networks for smartphone image based real time classification of citrus leaf disease." Computers and Electronics in Agriculture 177 (2020): 105661.

[6] Shrivastava, Vimal K., et al. "Rice plant disease classification using transfer learning of deep convolution neural network." International archives of the photogrammetry, remote sensing & spatial information sciences 3.6 (2019): 631-635.

[7] Singh, Vijai, and Ak K. Misra. "Detection of plant leaf diseases using image segmentation and soft computing techniques." Information processing in Agriculture 4.1 (2017): 41-49.

[8] Geetharamani, G., and Arun Pandian. "Identification of plant leaf diseases using a nine-layer deep convolutional neural network." Computers & Electrical Engineering 76 (2019): 323-338. W;kmwrmf

[9] Mahum, Rabbia, et al. "A novel framework for potato leaf disease detection using an efficient deep learning model." Human and Ecological Risk Assessment: An International Journal 29.2 (2023): 303-326.

[10] . Mahesh, T. R., et al. "Early Predictive Model for Detection of Plant Leaf Diseases Using MobileNetV2 Architecture." International Journal of Intelligent Systems and Applications in Engineering 11.2 (2023): 46-54.

[11] Fenu, Gianni, and Francesca Maridina Malloci. "Classification of Pear Leaf Diseases Based on Ensemble Convolutional Neural Networks." AgriEngineering 5.1 (2023): 141-152.

[12] Lu, Jinzhu, Lijuan Tan, and Huanyu Jiang. "Review on convolutional neural network (CNN) applied to plant leaf disease classification." Agriculture 11.8 (2021): 707.

[13] Annabel, L. Sherly Puspha, T. Annapoorani, and P. Deepalakshmi. "Machine learning for plant leaf disease detection and classification–a review." 2019 international conference on communication and signal processing (ICCSP). IEEE, 2019.

[14] Sujatha, Radhakrishnan, et al. "Performance of deep learning vs machine learning in plant leaf disease detection." Microprocessors and Microsystems 80 (2021): 10365.

[15] Padol, Pranjali B., and Anjali A. Yadav. "SVM classifier based grape leaf disease detection." 2016 Conference on advances in signal processing (CASP). IEEE, 2016.

[16] Dhaware, Chaitali G., and K. H. Wanjale. "A modern approach for plant leaf disease classification which depends on leaf image processing." 2017 International Conference on Computer Communication and Informatics (ICCCI). IEEE, 2017.

[17] Ghosh, H., Tusher, M.A., Rahat, I.S., Khasim, S., Mohanty, S.N. (2023). Water Quality Assessment Through Predictive Machine Learning. In: Intelligent Computing and Networking. IC-ICN 2023. Lecture Notes in Networks and Systems, vol 699. Springer, Singapore. https://doi.org/10.1007/978-981-99-3177-4_6

[18] Rahat IS, Ghosh H, Shaik K, Khasim S, Rajaram G. Unraveling the Heterogeneity of Lower-Grade Gliomas: Deep Learning-Assisted Flair Segmentation and Genomic Analysis of Brain MR Images. EAI Endorsed Trans Perv Health Tech [Internet]. 2023 Sep. 29 [cited 2023 Oct. 2];9.https://doi.org/10.4108/eetpht.9.4016

[19] Ghosh H, Rahat IS, Shaik K, Khasim S, Yesubabu M. Potato Leaf Disease Recognition and Prediction using Convolutional Neural Networks. EAI Endorsed Scal Inf Syst [Internet]. 2023 Sep. 21https://doi.org/10.4108/eetsis.3937

[20] Mandava, S. R. Vinta, H. Ghosh, and I. S. Rahat, “An All-Inclusive Machine Learning and Deep Learning Method for Forecasting Cardiovascular Disease in Bangladeshi Population”, EAI Endorsed Trans Perv Health Tech, vol. 9, Oct. 2023.https://doi.org/10.4108/eetpht.9.4052

[21] Mandava, M.; Vinta, S. R.; Ghosh, H.; Rahat, I. S. Identification and Categorization of Yellow Rust Infection in Wheat through Deep Learning Techniques. EAI Endorsed Trans IoT 2023, 10. https://doi.org/10.4108/eetiot.4603

[22] Khasim, I. S. Rahat, H. Ghosh, K. Shaik, and S. K. Panda, “Using Deep Learning and Machine Learning: Real-Time Discernment and Diagnostics of Rice-Leaf Diseases in Bangladesh”, EAI Endorsed Trans IoT, vol. 10, Dec. 2023 https://doi.org/10.4108/eetiot.4579

[23] Khasim, H. Ghosh, I. S. Rahat, K. Shaik, and M. Yesubabu, “Deciphering Microorganisms through Intelligent Image Recognition: Machine Learning and Deep Learning Approaches, Challenges, and Advancements”, EAI Endorsed Trans IoT, vol. 10, Nov. 2023. https://doi.org/10.4108/eetiot.4484

[24] Mohanty, S.N.; Ghosh, H.; Rahat, I.S.; Reddy, C.V.R. Advanced Deep Learning Models for Corn Leaf Disease Classification: A Field Study in Bangladesh. Eng. Proc. 2023, 59, 69.https://doi.org/10.3390/engproc2023059069

[25] Alenezi, F.; Armghan, A.; Mohanty, S.N.; Jhaveri, R.H.; Tiwari, P. Block-Greedy and CNN Based Underwater Image Dehazing for Novel Depth Estimation and Optimal Ambient Light. Water 2021, 13, 3470. https://doi.org/10.3390/w13233470

[26] Rothe, P. R., and R. V. Kshirsagar. "Cotton leaf disease identification using pattern recognition techniques." 2015 International conference on pervasive computing (ICPC). IEEE, 2015.

[27] . Ahmed, Kawcher, et al. "Rice leaf disease detection using machine learning techniques." 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI). IEEE, 2019.

[28] Jaisakthi, S. M., P. Mirunalini, and D. Thenmozhi. "Grape leaf disease identification using machine learning techniques." 2019 International Conference on Computational Intelligence in Data Science (ICCIDS). IEEE, 2019.