Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

Leaf Pathology Detection in Potato and Pepper Bell Plant using Convolutional Neural Networks

Theyazn H. H Aldhyani
[1]*Al Bilad Bank Scholarly Chair for Food Security in Saudi Arabia, the Deanship of Scientific Research, the Vice Presidency for Graduate Studies and Scientific Research, King Faisal University*
[2]*Applied College in Abqaiq, King Faisal University, P.O.Box 400, Al-Ahsa 31982, Saudi Arabia*
taldhyani@kfu.edu.sa

Hasan Alkahtani
[1]*Al Bilad Bank Scholarly Chair for Food Security in Saudi Arabia, the Deanship of Scientific Research, the Vice Presidency for Graduate Studies and Scientific Research, King Faisal University*
[2]College of Computer Science, King Faisal University
hsalkahtani@kfu.edu.sa

R Jennifer Eunice
Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore, India
jennifer.onesimu@gmail.com

D Jude Hemanth
Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore, India
judehemanth@karunya.edu

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835735

***Abstract*-Agriculture is the backbone of world's economy. This sector faces predominant issues in recognizing crop infection, disease prediction, pest control, weed detection and yield prediction leading to the shortfall in both quality and production of food. To ensure food safety, high resilience and increased crop yields, the precise diagnosis and recognition of underlying plant disease along with classification of crops from weeds is vital. The recent advancements in automatic feature extraction and classification techniques using Artificial Intelligence have gained attraction in the field of agriculture and crop protection. This paper proposes a Novel Convolutional Neural Network model for crop disease classification. The model is trained and tested in publicly available Plant Village Dataset with 38 categories and 15 classes. For the experimental analysis, the model is trained with 5 classes which includes potato and pepper bell categories. Further, the performance of the proposed model is analyzed with machine leaning models such as Support Vector Machine (SVM), K- Nearest Neighborhood (K-NN), Random Forest, Decision Tree and have attained the highest accuracy of 91.28%. In the testing phase, it is observed that this model is superior in terms of accuracy, specificity, precision, recall and F1-Score.**

Keywords—Crop disease Recognition, Support Vector Machine, Early and late blight disease, Convolutional Neural Network

I. INTRODUCTION

Agriculture is a key factor in global economic development since the raw materials for industries are supplied from agricultural farm lands. Besides, it is the foundation for every human existence in terms of supplying food, employment and income to the rural population [1]. Increasing the crop production in agriculture is the prime objective needs to be addressed worldwide [2]. Eventually this could be possible by protecting the crops from environmental pollution, drastic climate change and natural calamities which could be challenging for the farmers as the condition may worsen with sudden downpour of excess rain, drought or absence of snow reducing crop yield causing pests and disease in crops.

On the other hand numerous research effort have be undertaken in assisting the farmers to overcome this issues by making use of the growing technology[3]. Early discovery of underlying disease in plants and crops plays key role in agricultural management and production. Bringing up digital solution for early identification of pests and diseases may prevent the damage of plants and spike the crop yields.

Manual Identification of crop disease in its early stages is strenuous due to lack of knowledge of non -native crop disease, dedicated time and frequent monitoring of the plants is required [4]. Further, the number of Agricultural officers is sparse and may not visit broad-spectrum also equipping these officers is expensive and time consuming[5]. Through manual observation ratio of incorrect diagnosis of crop disease is high. To some extent manual observation are replaced using traditional Image processing techniques. But the recognition accuracy obtained using hand crafted traditional feature engineering is incomparable with machine learning and deep learning models.

To avoid all this issues computer vision can bring substantial outcome in crop production, disease detection and pest control. In recent times, machine learning and deep learning remain hotspot in computer vision which has series of steps and stages for recognition using algorithms [6], [7]. They include pre-processing, feature extraction, classification, feature fusion and selection algorithms to raise recognition systemaccuracy[8].

Feature extraction for recognition purpose focuses on shapes, colours, deep extracted features, and texture. Most of the models are trained on the publicly available dataset such as Plant village, Plant Doc, New Plant disease dataset, plant seedings Dataset, Flowers recognition dataset, Weed Detection in Soya bean crop captured under field data and uniform background [9].

Machine learning algorithm using Random Forest technique was proposed by Ramesh [10] for classification of defective crops from healthy crops. The classification accuracy of infected and healthy crops is increased using pre-processing of the dataset using Histogram oriented Gradient technique. However, using random forest, the accuracy achieved in sorting out the crops was 70.14%. Later Anggraeni and Falah [11] in their experiment used K-Nearest Neighbor algorithm to identify disease in tomato plant. He carried out both classification and regression analysis under 4 stages pre-processing, leaf segmentation, feature extraction and classification in Tomato leaf dataset. However, after study the machine learning methods employed to enhance the recognition and classification accuracy of the plant pathology did not meet out satisfactory outcome. It is found that out of all traditional machine learning methods such as SVM, Decision Tree, K-NN, logistic Regression and Random Forest, the Random Forest technique achieves maximum accuracy in recognition and classification of healthy plants from defective plants.

In this paper, we have proposed a novel Convolutional Neural Network (CNN) for the recognition and classification of healthy potato plant and pepper bell plant from infected potato and pepper bell plants. Further, we have trained our model using plant village dataset and achieved outstanding results comparatively with Machine learning models. The rest of the paper is organized as follows section II presents the detail literature works, Section III presents the methodology and methods used in our experiments Section III A gives details on Dataset used, Section III B concentrates on the proposed architecture. Section IV presents the results and discussions and Finally Section V concludes our work with proper guidance towards future directions.

II. RELATED WORKS

Crop disease prediction is subjected to Image recognition, plant pathology, object and pattern recognition, computer vision using machine learning and deep learning techniques to extract the in-depth detail features of the plants to classify and recognize the disease information and categorize them more accurately under exact disease category. Since we are experimenting with machine learning and deep learning models, we have carried our literature review on the previous works done under machine learning. We do provide a comprehensive view on the issues resolved and thereby briefing the advantages of using deep learning over Machine leaning.

A. *Machine Learning based crop disease detection*

Monitoring and identifying the crop disease have attained advancements including X-Ray Imaging, Ultrasound, RGB Imaging, Hyperspectral techniques, image processing. A machine learning model for disease prediction and classification was proposed by [12] which uses Region of Interest for feature extraction in pre-processing with Support vector Machine. However, the proposed model achieved considerable accuracy but they are an absolute fit only for potato plant disease. Further, [13] compared various supervised machine learning algorithms and found promising results using Random forest achieving 79.23%. An ANN based model for artificial plant disease detection was proposed in[14]. Later, [15] analysed his own dataset on tomato disease using machine learning models and found Random forest to be best achieving maximum accuracy of 89%. Disease recognition in citrus fruit and leaves were done using machine learning techniques[16]. On the other side the recent advancement in the field of deep learning many research are carried out in providing best solution for identification of disease in plants and crops. With the advent of deep learning [17] uses Principal Component Analysis for feature extraction and uses SVM CNN and Naives Bayes classification techniques. Later comparison of machine learning techniques along with deep learning models were done and found that deep learning models produced more accurate outcomes when compared with ML [18].

B. *Deep learning based crop disease detection*

To analyze visual imagery a class of deep neural networks namely convolutional neural networks are used. Further the automatic crop disease recognition system is time saving for both the farmers and agriculture officers in early monitoring of crops. To develop most accurate automated system, it is mandatory that the system is trained with large database. Deep learning models outshines the machine learning models in terms of feature engineering and accuracy[19]. Partial least square techniques for feature selection from the deep feature set has been proposed by [20] and achieved accuracy of 90.1 using ensemble classifier. [21] used novel deep learning model for building automated crop disease monitoring system. By using Otsu thresholding technique the noise in the input database is pre-processed and then the preprocessed image is fed to the fully connected CNN[22].

Potato blight disease are the severe disease in potato plants caused by phytophthora Infestans. [23] proposed a deep Convolution neural network using encoder decoder architecture. Later transfer learning approaches were also implemented in recognition of potato blight disease [24] and the outcome seems to be outstanding with 99.91% only for potato blight class from plant village dataset using Efficient Net architecture. Then [25] proposed CNN model for multiclass classification of potato blight disease I potato plants by collecting 900 real time images of potato plants both healthy and with blight disease. Recent works combines different fusion techniques with deep learning for reaching highest accuracy. [26] achieved highest accuracy in identifying tomato disease in tomato plants using CNN. The complexities and optimization issues are addressed using binary solution encoding scheme using Crossover based monarch butterfly optimization algorithm and achieved 99.98 % accuracy using VGG16. Transfer learning approach with deep learning was implemented by [27] for refinement of classification and recognition Tomato plant using RMS prop and SGD optimizers.

III. MATERIALS & METHODS

This section provides the details on the dataset used for the experimental study, sample data from the dataset and the proposed CNN architecture.

A. *Dataset details*

In this experimental study, we used the dataset collected by Pennsylvania state university called Plant Village. There are around 52k images of 15 classes available in the dataset. The images are RGB with dimension 256 X 256. For our experiment we selected 4,627 images of 5 classes. The classes are Pepper bell Bacterial spot, Pepper bell Healthy, Potato Early Blight, Potato Late Blight, and Potato Healthy. The details of the number of images per class is given in table I and fig.1 shows the sample images from the dataset.

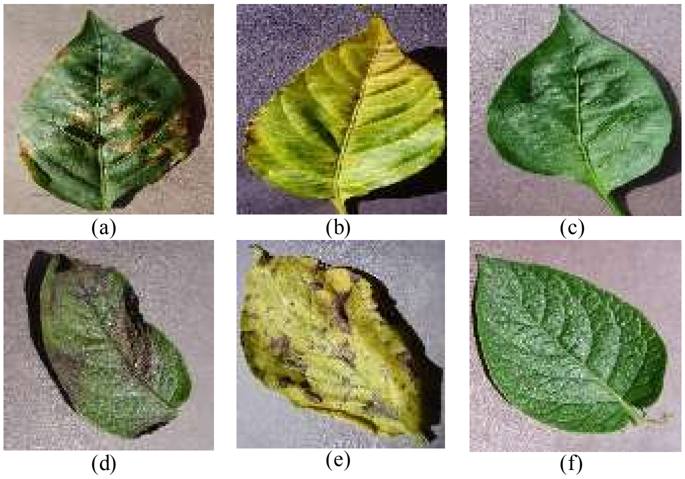

Fig. 1. Sample images from PlantVillage dataset. (a)&(b) Bacterial spots on pepper bell, (c) Healthy pepper bell, (d) Potato early blight, (e) Potato late blight, (f) Healthy Potato

TABLE I. DATASET DETAILS

Classes	No. of Images
Pepper bell Bacterial spot	997
Pepper bell Healthy	1478
Potato Early Blight	1000
Potato Late Blight	1000
Potato Healthy	152
Total	**4627**

B. Proposed Architecture

One prime algorithm that topped and utilized in almost every field with the advancement of Computer vision and deep learning is Convolutional Neural Network. Compared with other classification algorithms, the pre-processing requirement is of less importance as it best suited for both classification and feature extraction. The CNN comprises of three layers convolution layer which is the core building block of CNN, a max pooling layer helps in reducing the spatial size of the representation using several pooling functions L1 norms, L2 Norms weighted average neighborhood, rectangular weighted average neighborhood. A fully connected Layer which prevents overfitting by fine tuning the hyperparameters.

The proposed 2D CNN architecture has 5 convolution layers and 5 pooling layers. The size of the kernel is 3x3 with striding window size as (1,1). The dropout layer is fixed as 0.5 and ReLu is used as an activation function. The Adam optimizer is used in optimization technique. Between every pooling and convolution layers a regularization technique

known as Batch Normalization is applied to prevent internal covariant shift and overfitting.

The input and output of Batch Normalization is four-dimensional tensor represented as $O_{b,\ c,\ x,\ y}$ and $I_{b,\ c,\ x,\ y}$ respectively.

$$O_{b,c,x,y} = \gamma_c \frac{I_{b,c,x,y} - \mu_c}{\sqrt{\sigma_c^2 + \epsilon}} + \beta_c \qquad \forall\, b, c, x, y \tag{1}$$

Where b represents batch, c represents channel, x and y are spatial locations. σ_c denotes Standard deviation, $\in$ denotes numerical Stability and μ_c stands for mean activation.

The learning rate is one of the most vital hyper-parameters used for tuning the deep neural networks. If the learning rate is less (say for e.g., 0.1) the loss will not improve instead it will jeep growing. The loss can be then minimized by exponentially minimizing the learning rate to 0.01, 0.001 and so on. When we start with the lower value, at least at some point the loss functions comes down at few epochs.

$$l_{rate} = Initial\_l_{rate} * \left(1 / (1 + Decay \ * iteration) \right) \tag{2}$$

Where,

l_{rate} – Denotes the learning rate of the current epoch

$Initial\_l_{rate}$ – Denotes the learning rate as argument to SGD

Decay – stands for decay rate which is greater than 0

Iteration – refers to current updater number.

One of the major issues that learning rate faces is co-adaptation. This occurs when all the weights are learned at same time, some connections will have high predictive tendency than the others. The weaker ones are ignored while training and the powerful ones are learning more. This phenomenon is known as co-adaptation. To resolve this issue, many research were carried out and found some regularization technique such as L1 norm, L2 norm, Dropout. Dropout is used to prevent overfitting. The dropout value can be somewhere close to 0.5 and 1. If the dropout value is 0.5 then the probability of retaining output values from hidden layer nodes is possible and if it's close to 1 or 0.8 then it retains the output values from the visible layer nodes.

Figure 2. represents the proposed architecture of 2D CNN which has 1024 neurons in the fully connected layer. And the output classes in our proposed network architecture is 5.

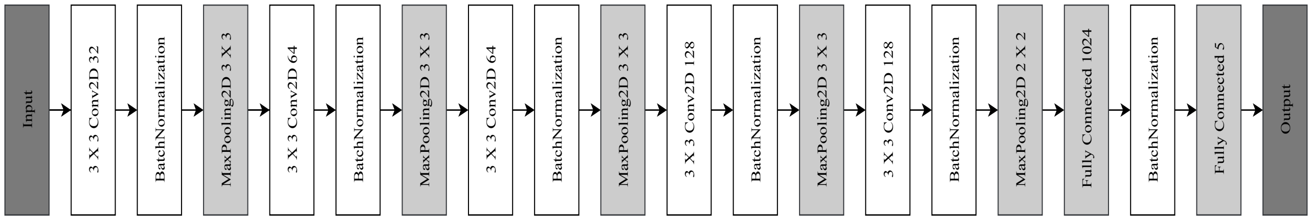

Fig. 2. Proposed Convolutional Neural Networks Architecture

IV. RESULTS & DISCUSSION

The proposed CNN architecture is implemented on publicly available dataset 'Plant Village' comprising 54303 images of healthy as well as unhealthy plants images divided under 38 categories. The dataset holds images of healthy and unhealthy apples, strawberry, corn, cherry, grape, peach, potato, cucumber, tomato, orange, Squash etc. Out of these for our experiments we choose potato and pepper bell plants. The images are taken with uniform background condition and are with the dimension of 256 x 256 x 3.

A. Experimental Analysis

The experiment is carried out by performing pre-processing of the input database. The input dimension of the database is resized to 224 x 224 x 3. After Pre-processing the next step is feature extraction using the proposed 2D CNN model. This model is designed with 5 convolutional layers and 5 max pooling layers. The output dimension of the feature map after max pooling for the feature map having input dimension of N_h x N_w x N_c is given as,

$$\frac{N_h - F + 1}{S} \text{ x } \frac{N_w - F + 1}{S} \text{ x } N_c \quad (3)$$

Where,

N_h – Denotes Height of the feature Map

N_w – Denotes Width of the feature Map

N_c – Represents available Channel Numbers in feature Map

S – represents Striding length

F – Defines Filter Size

For our experimental setup we have trained our model with 5 classes from plant Village dataset. We have chosen Potato and pepper bell plants which includes potato healthy, potato early blight, potato late blight, pepper bell healthy and pepper bell bacterial spot. Overall, our model architecture is trained with 4627 images out of which 997 images under pepper bell bacterial spot, 1478 images under pepper bell healthy, 1000 images each under potato early and late blight, 152 images under potato healthy. Table II briefs the Hyperparameter details of our model.

The learning rate of our model is 0.001 and the experiment has been run for 50 epochs and between every convolutional and pooling layer we have done regularization using batch normalization technique. The optimization is done using Adam Optimizer. The dataset has been partitioned into two phases 80% for training phase and 20% for validation. Figure 3 shows the graphical representation of Training Vs Validation Loss and Fig 4 Shows the graphical representation of Training and Validation accuracy.

TABLE II. HYPERPARAMETER DETAILS

Hyperparameter	**Values**
Dropout	0.5
Epochs	50
Convolution Layers	5
Kernel Size	3 X 3
Strides	(1,1)
No. of Pooling Layers	5
Pooling Type & Size	MaxPool 3X3, 2X2
Activation	ReLu
No. of Neurons in Fully Connected Layer	1024
Regularization	Batch Normalization
Optimizer	Adam
Learning rate	0.001
Output classes	5

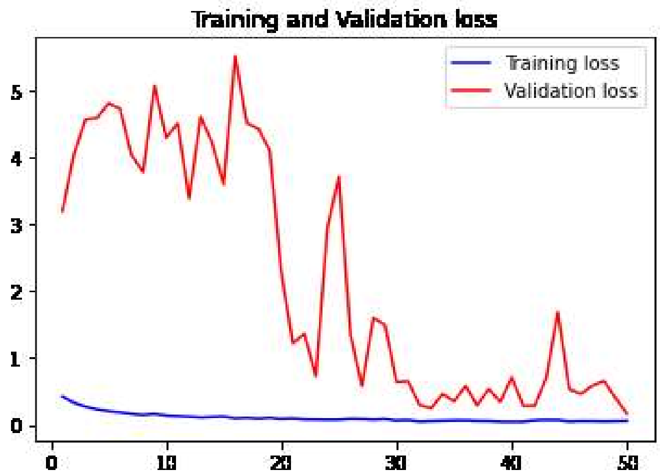

Fig. 3. Training vs Validation loss of the Proposed CNN model

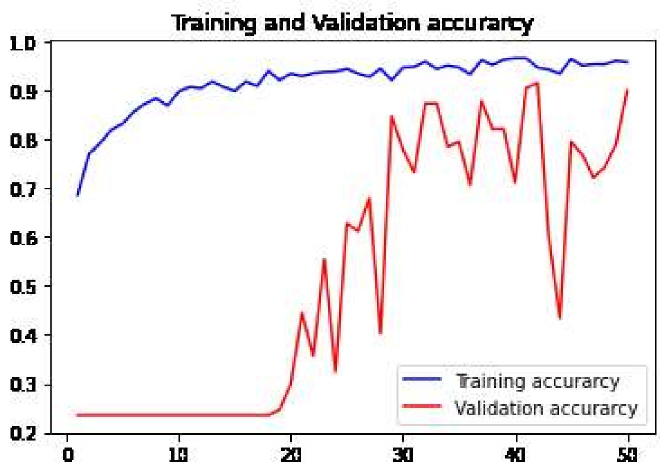

Fig. 4. Training vs Validation Accuracy of the Proposed CNN model

B. Performance Evaluation

The proposed model is evaluated with the existing machine learning algorithm. We have conducted the series of experiments with support vector machine, K-nearest neighbor, decision tree, and Random Forest. The proposed model generated 91.28 % in terms of accuracy, precision, recall and F1 score. Table III gives a detailed comparative analysis of our proposed 2D convolutional neural Network with other deep Learning Models. Figure 5. Shows the performance analysis of our model in terms of recognition accuracy.

TABLE III. COMPARATIVE ANALYSIS OF PROPOSED MODEL WITH MACHINE LEARNING MODELS

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	23.5	0.34	1.00	0.51
K-Nearest Neighbor	37.1	0.79	0.46	0.49
Decision Tree	35.60	0.37	0.36	0.35
Support Vector Machine	55.80	0.85	0.93	0.66
Proposed Model	**91.28**	**0.98**	**0.96**	**0.97**

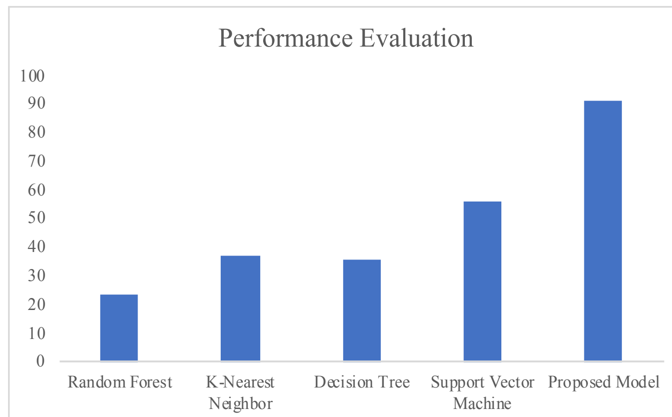

Fig. 5. Performance Measures of Classification in terms of accuracy metrics

V. CONCLUSION

In this paper, a novel 2D Convolutional Neural Network is proposed for the easy recognition and classification of unhealthy potato and pepper bell plants from the healthy ones. The experiments were evaluated on the publicly available Plant Village dataset and the proposed model architecture achieved recognition accuracy of 91.28% which is superior when compared with other machine learning models. The Proposed model in efficient in identifying the plant disease without consuming much time. This model can certainly help the farmers and agriculture officers to diagnose the plant disease with a simple picture of the plant leaves. Further, we are extending this work to improve the classification and recognition accuracy by using different feature fusing techniques and generative models along with pre-trained networks to build a real-time automated leaf pathology classification system.

ACKNOWLEDGMENT

This work is supported by The Al Bilad Bank Scholarly Chair for Food Security in Saudi Arabia, the Deanship of Scientific Research, the Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Al Ahsa, Saudi Arabia [Grant No.79]

REFERENCES

[1] P. S. Neha Khanna, "Agricultural & Horticultural Sciences," in *Agriculture and Horticulture Sciences*, 2014, vol. 2, no. 4, p. 33892.

[2] R. G. Dawod and C. Dobre, "Upper and Lower Leaf Side Detection with Machine Learning Methods," *Sensors*, vol. 22, no. 7, p. 2696,

2022, doi: 10.3390/s22072696.

[3] P. Radoglou-grammatikis, P. Sarigiannidis, and T. Lagkas, "A compilation of UAV applications for precision agriculture," *Comput. Networks*, vol. 172, no. January, p. 107148, 2020, doi: 10.1016/j.comnet.2020.107148.

[4] M. A. Khan, T. Akram, M. Sharif, K. Javed, M. Raza, and T. Saba, "An automated system for cucumber leaf diseased spot detection and classification using improved saliency method and deep features selection," *Multimed. Tools Appl.*, vol. 79, no. July 2020, pp. 18627–18656, 2020.

[5] A. Raju, "An Application of image processing techniques for Detection of Diseases on Brinjal Leaves Using K-Means Clustering Method," in *International Conference On Recent Trends In Information Technology*, 2016, no. October.

[6] T. S. R. Mhathesh, J. Andrew, K. Martin Sagayam, and L. Henesey, "A 3d convolutional neural network for bacterial image classification," in *Advances in Intelligent Systems and Computing*, vol. 1167, Springer, 2021, pp. 419–431.

[7] J. Andrew, R. Fiona, and H. Caleb Andrew, "Comparative study of various deep convolutional neural networks in the early prediction of cancer," in *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, May 2019, pp. 884–890, doi: 10.1109/ICCS45141.2019.9065445.

[8] M. Attique, A. Muhammad, and S. Tanzila, "Fruits diseases classification : exploiting a hierarchical framework for deep features fusion and selection," *Multimed. Tools Appl.*, 2020.

[9] J. Lu and L. Tan, "Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification," *agriculture*, pp. 1–18, 2021.

[10] S. Ramesh, "Plant Disease Detection Using Machine Learning," 2018, doi: 10.1109/ICDI3C.2018.00017.

[11] F. D. Anggraeni and M. A. F. Falah, "Plant Leaf Disease Classification and Detection System Using Machine Learning PLANT LEAF DISEASE CLASSIFICATION AND," 2020, doi: 10.1088/1742-6596/1712/1/012012.

[12] M. Islam, A. Dinh, and K. Wahid, "Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine," in *IEEE 30th Canadian Conference on Electrical and Computer Engineering*, 2017, pp. 8–11.

[13] K. P. Panigrahi, H. Das, and A. K. Sahoo, "Maize Leaf Disease Detection and Classification Using Machine Learning Maize Leaf Disease Detection and Classification Using Machine," *Prog. Comput. Anal. Netw.*, no. January, 2020, doi: 10.1007/978-981-15-2414-1.

[14] S. Ramesh, "Rice Blast Disease Detection and Classification using Machine Learning Algorithm," *2018 2nd Int. Conf. Micro-Electronics Telecommun. Eng.*, pp. 255–259, 2018, doi: 10.1109/ICMETE.2018.00063.

[15] P. Neelakantan, "Materials Today : Proceedings Analyzing the best machine learning algorithm for plant disease classification," *Mater. Today Proc.*, no. xxxx, 2021, doi: 10.1016/j.matpr.2021.07.358.

[16] M. I. Ullah, M. Attique, M. Sharif, S. Ahmad, and C. Bukhari, "Data in brief A citrus fruits and leaves dataset for detection and classi fi cation of citrus diseases through machine learning," *Data Br.*, vol. 26, 2019, doi: 10.1016/j.dib.2019.104340.

[17] H. Pallathadka *et al.*, "Application of machine learning techniques in rice leaf disease detection," *Mater. Today Proc.*, vol. 51, pp. 2277–2280, 2022, doi: https://doi.org/10.1016/j.matpr.2021.11.398.

[18] K. Thaiyalnayaki and C. Joseph, "Classification of plant disease using SVM and deep learning," *Mater. Today Proc.*, vol. 47, pp. 468–470, 2021, doi: https://doi.org/10.1016/j.matpr.2021.05.029.

[19] A. Picon *et al.*, "WHY DEEP LEARNING PERFORMS BETTER THAN CLASSICAL MACHINE," pp. 1–5.

[20] F. Saeed, M. A. Khan, M. Sharif, M. Mittal, L. M. Goyal, and S. Roy, "Deep neural network features fusion and selection based on PLS regression with an application for crops diseases classification," *Appl. Soft Comput.*, vol. 103, p. 107164, 2021, doi: https://doi.org/10.1016/j.asoc.2021.107164.

[21] R. Sangeetha and M. Mary Shanthi Rani, "A Novel Method for Plant Leaf Disease Classification Using Deep Learning Techniques," in *Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication*, 2021, pp. 631–643.

[22] S. K. Upadhyay and A. Kumar, "A novel approach for rice plant diseases classification with deep convolutional neural network A novel approach for rice plant diseases classification with deep convolutional neural network," *Int. J. Inf. Technol.*, no. October 2021, 2022, doi: 10.1007/s41870-021-00817-5.

[23] J. Gao, J. C. Westergaard, E. H. R. Sundmark, M. Bagge, E. Liljeroth, and E. Alexandersson, "Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning," *Knowledge-Based Syst.*, vol. 214, p. 106723, 2021, doi: https://doi.org/10.1016/j.knosys.2020.106723.

[24] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Ecological Informatics Plant leaf disease classification using EfficientNet deep learning model," *Ecol. Inform.*, vol. 61, no. September 2020, p. 101182, 2021, doi: 10.1016/j.ecoinf.2020.101182.

[25] V. Kukreja, A. Baliyan, V. Salonki, and R. K. Kaushal, "Potato Blight: Deep Learning Model for Binary and Multi-Classification," in *2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2021, pp. 672–967, doi: 10.1109/SPIN52536.2021.9566079.

[26] S. Nandhini and K. Ashokkumar, "Improved crossover based monarch butterfly optimization for tomato leaf disease classification using convolutional neural network," *Multimed. Tools Appl.*, vol. 80, no. 12, pp. 18583–18610, 2021, doi: 10.1007/s11042-021-10599-4.

[27] R. Thangaraj, S. Anandamurugan, and V. K. Kaliappan, "Automated tomato leaf disease classification using transfer learning-based deep convolution neural network," *J. Plant Dis. Prot.*, vol. 128, no. 1, pp. 73–86, 2021, doi: 10.1007/s41348-020-00403-0.