Machine Learning-Based Plant Disease Detection for Agricultural Applications: A Review

Himanshu Chanyal[1]; Dr. Rakesh Kumar Yadav[2]
[1]Research Scholar, Department of Computer Science, IFTM University, Moradabad(U.P).
[1]himanshu.chanyal@gmail.com
[2]Associate Professor, Department of Computer Science & Engineering, Maharshi School of Engineering & Technology, Maharshi University of Information Technology, Lucknow(U.P).
[2]rkymuit@gmail.com

2023 Second International Conference On Smart Technologies For Smart Nation (SmartTechCon) | 979-8-3503-0541-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/SmartTechCon57526.2023.10391776

ABSTRACT

The agriculture sector plays a major role in supplying high-quality food and makes the biggest contribution to growing economies and people. Plant ailments have the power to drastically limit food production and eliminate species diversity. Early identification of plant diseases using accurate or automatic detection techniques can increase the quality of food output and decrease financial losses. In recent years, deep learning has considerably improved the identification precision of image classification and object detection systems. The development of agricultural products is crucial for the economy of the country. Farmers have a wide range of alternatives for the production of numerous goods. However, plant diseases can impair any agricultural product. Information technology advancements can be utilised to quickly identify and locate agricultural illnesses. Plant diseases must be categorised in order to evaluate agricultural output, increase market value, and maintain quality standards. Despite the fact that many eminent researchers have contributed their suggestions for developing such systems, there are several shortcomings in the recommended and developed systems. The paper provides in-depth discussions on plant illnesses, disease detection systems, and disease classification using image acquisition, image pre-processing, picture segmentation, feature extraction, and classification.

***Key Words*: *Plant Diseases, Image Processing, Classification, Feature Extraction*.**

1. INTRODUCTION

6.1% of India's GDP is derived from the agricultural industry. Several illnesses pose a danger to plants. The grade of agricultural products will be impacted by this. Early on, manual detection of these illnesses involved eye examination by experienced humans, but this method was inefficient and time-consuming due to its lack of precision. Many farmers want to use contemporary farming techniques, but they are unable to do so for a number of reasons, including a dearth of knowledge about new technology and the high expense of it. However, agricultural losses can be minimized and targeted therapies can be developed to fight particular diseases if plant diseases are correctly recognized and caught early. We can now readily spot diseases at an early period of plant growth thanks to the development of numerous plant disease identification systems and the correct application of these systems. So that we can take the required steps to stop those illnesses. Research on automated plant disease detection systems is essential for keeping track of numerous crops and recognizing disease signs whenever they manifest on plants.

This article surveys various methods for diagnosing diseases. There are four parts in this essay. The literature reviews of scholars from all over the globe are explained in Section 1. Section 2 provides a list of various plant illnesses. The method for identifying plant diseases is described in Section 3, and the result is defined in Section 4.

2. LITERATURE REVIEW

Kishori Patil et al. [1] suggested Deep Learning Algorithm for Leaf Disease Detection. The CNN system has two levels.The feature extraction layer comes first, followed by the feature extraction layer. The precision of the CNN technique for identifying plant leaf disease is up to 86.26%.

Dey et.al [2] Segmenting the rot illness in betel vine plant leaves using the suggested Otsu technique. An picture scanner was used to record 12 sick photos. For precise disease identification, images were cropped during the pre-processing stage and then applied colour conversion. Finally, a severity rating was created after locating the diseased region of the leaf. The use of pesticides can be decreased depending on their intensity, the author found, and this will help to lessen environmental pollution.

Simranjeet kaur et al. [3] suggested Image Processing and Classification Method for Detecting Plant Disease. For feature analysis, the author used the GLCM technique, and for recognition, the KNN classifier. The suggested system's identification accuracy was 95%.

Joshi and Jadhav [4] present a method for image analysis that can identify and categorise four diseases affecting rice: bacterial blight, blast brown spot, sheath rot, and others. Affected rice plant image examples were gathered and saved in jpg format. Color system changes were then carried out. In order to capture colour and form characteristics, RGB pictures were converted to YCbCr. Finally, K closest neighbours and the Minimum Distance Classifier were employed to classify illnesses. The authors came to the conclusion that the same algorithm could even be used to derive various texture characteristics from the samples and identify other rice illnesses.

Khaing War Htun et al. [5] suggested a simplified colour conversion technique for classifying paddy diseased leaves. For the categorization and detection of diseased rice leaf, 143 data sets were used. Only four diseases—foli blight, brown spot, leaf blast, and leaf streak—are susceptible to this approach. Statistical, color, and texture characteristics based on SVM can effectively identify and classify paddy illnesses.

G. Saradhambal, et.al [6] suggested a method for early detection of plant diseases. By using the Otsu technique and the K-means clustering algorithm, it forecasts the area of the leaves that are infected. Features of shape and substance were taken. Area, colour axis length, eccentricity, solidity, and periphery are shape-oriented characteristics, whereas contrast, association, energy, uniformity, and mean are texture-oriented features. Finally, a classifier built on a neural network was used to perform categorization in this study.

Ratnasari et.al [7] a model was put forth to assess the intensity of leaf spot in sugarcane plant leaves. On the basis of the concept, a split spot in the L*a*b colour space was produced. The classification was SVM-based. It employs texture characteristics (GLCM) and colour features (L*a*b colour space) to categorise the different types of spot diseases. This model's accuracy was 80% with an average error severity estimate of 5.73. The suggested model has a high accuracy rate and a low error, and in the future, pre-processing can be done to lower this error, according to the authors' conclusion.

Hiteshwari Sabrol, satish kumar[8] suggested a technique for acquiring digital images of both infected and uninfected vegetation. performed image pre-processing, used colour space conversion, segmentation, and feature extraction from segmented images for recognition and classification based on feature analysis, neural networks, support vector machines, fuzzy and rule-based classification, and fuzzy and rule-based classification. Researchers working in the fields of pattern detection and plant pathology should find this model helpful.

A model based on Digital image analysis was proposed by Majumdar et.al [9] to recognize wheat leaves diseases. Fuzzy c means clustering algorithm was used for extracting features. ANN was used for recognition of diseases. Author concluded that this model can identify rust diseases of wheat. In future a web based interface can be developed for efficient detection.

K. R. Aravind, et.al [10] examined maize agricultural diseases in preparation for automating the method for disease identification. Speeded Up Robust Features (SURF) were selected from each picture. The k-means algorithm was used to combine these characteristics. The histogram and the GLCM technique were both used for feature extraction. These techniques were employed to examine different surface characteristics. For categorization, a multi-class SVM based on different kernel functions, such as linear, polynomial, and radial basis function, was used. The obtained precision was 83.7% on average.

A system proposed by Sannakki et.al [11] was built on picture analysis and AI methods for identifying illnesses in grape plants. Using thresholding and masking methods, the background of the original grape leaf picture was eliminated. Segmentation was done using preprocessing, and subsequently texture features were extracted using the GLCM technique to feed into the BPNN classifier. The authors came to the conclusion that other segmentation methods could also be used in this system for the identification of other grape illnesses.

Mokhtar et.al [12] proposed a method for analysing images to find sick tomato leaves. This method was broken down into three stages: pre-processing, feature extraction, and categorization. The condition of tomato leaves was determined by extracting texture characteristics using the GLCM technique. These characteristics were later used as input to an SVM-based classifier to determine the condition of a leaf. According to the author, 99.83% precision was attained using this method.

A survey made by Barbedo [13] was founded on approaches for identifying and categorising agricultural diseases that involved image processing. As a result of some anomalous features in the methods used to handle other plant components such as fruit, roots, seeds, etc., the author gathered pictures of leaves and stalks for disease detection. The three parts of this study were categorization, severity assessment, and detection. A number of methods were outlined based on these parts.

Suresha et.al [14] used KNN classifier to conduct a study on two significant fungi illnesses affecting rice crops: rice blast and rice brown spot. In the segmentation phase, RGB colour pictures were transformed into HSV images. Extractions of features like perimeter, area, major, and axes were given to the KNN algorithm for identification. The precision of 79.59%, according to the authors, was much higher than that of the SVM classification.

Xiao et.al [15] Rice blast disease was classified and found using PCA and BPNN methods. Accuracy and speed are two current issues that are eliminated by the suggested system. F Each image's lesion was examined for color, shape, and material characteristics before performing a stepwise regression analysis. PCA was used to translate 21 features into 6 features as a BPNN input parameter. The suggested model's accuracy was 95.83%.The author suggested using this poll to quickly identify Rice Blast disease in a real-world setting.

Table -1: Plant Disease Identification and Recognition Techniques

Author	Plant	Disorder	Tools Used	Accuracy
Kishori Patil et al.	Cotton	Fungus, foliar leaf, And Alternaria leaf spot.	Convolution Neural network (CNN)	86.26 %
Dey et.al	Betel vine	Rot disease	Otsu Thresholding method	-
Simranjeet kaur et al.	N/A	N/A	GLCM, feature extraction and KNN classifier	95%.
Joshi and Jadhav	Rice	Bacterial blight, Blast brown spot	MDC and KNN	87.02 % and 89.23 %
Khaing War Htun et al.	paddy	Leaf blight, brown spot, leaf blast and leaf streak	SVM, GLCM	90%
G. Saradham bal, et.al	Rice	Leaf blight, brown spot and leaf streak	KNN and Otsu Thresholding method	-
Ratnasari et.al	Sugar cane	Spot disease	SVM and GLCM	80%
Hiteshwari Sabrol,satish kumar	N/A	N/A	PCA and SVM	100%(for SVM)
Majumdar et.al	Wheat	Rust disease	ANN	85%
K. R.Aravind, et.al	Maize		Histogram, GLCM and SVM	83.7%
Sannakki et.al	Grape	Downy Mildew, Powdery	K-Mean clustering and Forward feed	100%(in trainin

		Mildew	BPNN	g phase)
Mokhtar et.al	Tomato	Powdery Mildew, Downy Mildew	K-Mean clustering and SVM	99.5%
Suresha et.al	Paddy	Blast, Brown spot	KNN classifier	76.59 %
Xiao et.al	Rice	Blast	PCA and ANN	95.83 %

Table -2: A Thorough Explanation of the CNN models that were used to identify and categorise plant diseases.

Author	Crop/ Plant	Disorder	Tools Used	Accuracy
Barbedo, J.G.A. et.al	Diverse	Black mould, bacterial blight, citrus canker, etc.	CNN GoogLeNet With tenfold Cross-validation	84%
Shrivastava, V.K.;Pradhan,M.K et.al	Rice plant	Rice blast, bacterial blight, and sheath blight	SVM classifier and pre-trained CNN	91.37 %
Lee, S.H.; Goëau et.al	Diverse	Bacterial spots on pepper bells and early and late tomato blight	Equipped with ImageNet, GoogleLeNet, and VGG-16 models for training	99.09 %
Ferentinos, K.P et.al	Diverse	Downy mildew, cucumber mosaic, powdery mildew, early and late blights, etc.	Pre-trained VGG network on CNN	99.53 %
Fuentes, A.; Yoon, S.; Kim, S.C et.al	Tomato plant	Various pests and diseases affect tomato plants	Faster Region-based Region-based Fully Convolutional Network and SSD with CNN	85.98 %
Militante, S.V.; et.al	Diverse	Early blight, late blight, and black rot	CNN	96.5%
Too, E.C.; Yujian et.al	Diverse	Early blight, late blight, and black rot	ResNet with 50, 101, and 152 layers, VGG-16, Inception V4, and DenseNet with 121 layers	99.75 %

The following graph shows accuracy reported for various classifiers by various researchers:-

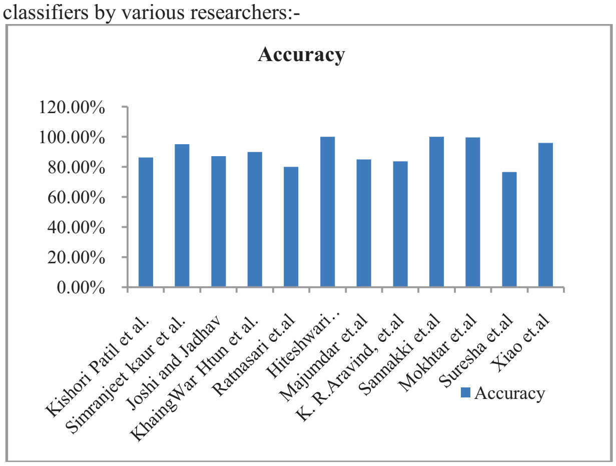

Figure 1: Accuracy of Various Methods Used by Different Researchers

The following graph shows accuracy reported for CNN Based models Used by various researchers:-

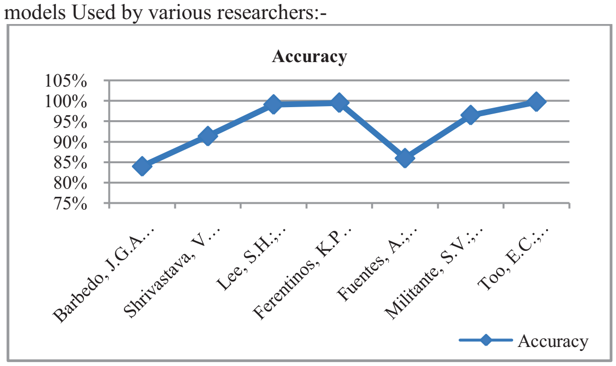

Figure 2: Accuracy of CNN based models used by Various Researchers

3. PLANT DISEASES

Plant disease is an impediment to a plant's essential processes that causes disturbance or change. Disease can affect any type of plant, whether it is untamed or domesticated. Anything that stops a plant from producing at its highest level is known as a plant illness. Plant illness will reduce agriculture's ability to produce. Pathogens are primarily to blame for plant illnesses.

3.1 Pathogens Classification

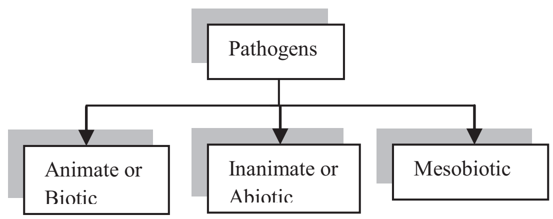

Fig-3: Classification of Pathogens

3.1.1 Animate or Biotic pathogens

The following categories are used to categorise pathogens found in biological things. (i) Phytoplasma (ii) Bacteria (iii) Fungi (iv)Nematodes (v) Algae (vi) Phanerogams (vii) Rickettsia-like organisms(viii) Protozoa

3.1.2 Inanimate or Abiotic Pathogens

Instead of producing disease, these variables disturb the plants.

The major causes are: (i) Deficiencies or excess of nutrients (ii) Light (iii) Moisture (iv) Lack of oxygen (v) Improper cultural practices (vi)Temperature (vii) Toxicity of pesticides (viii) Air pollutants (ix) Abnormality in soil conditions (acidity,alkalinity)

3.1.3 Mesobiotic pathogens

These contagious agents, such as viruses and viroids, are neither living things nor inanimate objects.

3.2 Plant Disease Classification Based on Causal Agents

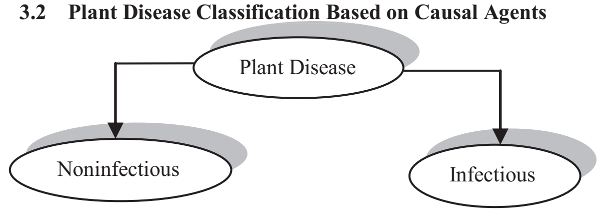

Fig-4: Classification based on casual agents

3.2.1 Non-infectious

Non-infectious pathogens are often thought of as inert or abiotic. Thus, they cannot be transmitted. These are brought on by changes in the structure of the plant brought on by deficiencies in certain innate characteristics, by poor soil and air quality, and by mechanical effects.

Examples:

i. Low/high temperature
ii. Unfavorable oxygen levels
iii. Unfavorable water levels
iv. Hail
v. Wind
vi. Air pollution toxicity etc.

3.2.2 Infectious

Pathogenic organisms or viruses under specific environmental circumstances induce infectious illnesses. Pathogens can be parasitic blooming plants, bacteria, viruses, worms, or fungus. They receive everything they require for reproduction, including food, water, and their host. Bacterial and worm pathogens cause a small number of plant illnesses, while fungal, viral, and other pathogens cause the majority of them. While some viruses can attack different plant species, others require a particular host. The capacity of pathogens, such as fungi and bacteria, to endure, disseminates, and replicate itself varies.

4. PLANT DISEASE DETECTION

There are 5 stages in the Plant Disease Identification System (PDIS). Different techniques for evaluation and detection are used in each stage. The researcher's preference informs the method option. According to the literature analysis, the majority of researchers detect plant diseases using the following stages.

1. Application of image processing techniques

This step comprises of:

1.1 Image Acquisition

It is used to obtain the picture of the ill plant. An picture sensor is utilised for this. The sensor's output data should be able to be digitally processed by the device.

1.2 Image modification

It is used to modify an image so that the finished product is better suited for the intended purpose than the initial image. It is an individualized procedure.

1.3 Image segmentation

Partitioning an image into useful areas for a specific application is the goal of this stage. Based on data obtained from the plant picture, the image segmentation may be based on depth, texture, color, or motion.

1.4 Extraction of features

The general objective of feature extraction is to depict the segmented objects in a way that more accurately captures their key characteristics and traits.

1.5 Evaluate the affected region

In this stage, the picture segment's quality is assessed using a variety of tools and approaches to identify the diseased area.

2. Load training data

The database is filled with the diseased pictures' properties for identification.

3. Classifying an image

The primary objective of this stage is to assign land cover classes or themes to every pixel in an image. The key methods used in this phase include ANN, SVM, KNN, K Means clustering, etc.

4. Load the testing data

Here, image characteristics are compared to data from the training process to see if the results match what was anticipated.

5. Evaluate the accuracy

Determines the result's precision.

5. COMPARATIVE ANALYSIS

The evaluation of current classification techniques focuses on a number of issues, such as the features and classifiers in widespread use and their effects on classification accuracy, the testing data sets, and research trends on leaf sorting techniques. Color, form, texture, and vein are just a few of the elements that researchers have included in their techniques. Additionally, we have discovered that some studies integrate many criteria to increase accuracy ratios. In their approaches for classifying objects, most studies concentrate on the properties of forms. The inclusion of numerous characteristics in the classification technique aids in improving the accuracy ratios of Disease Identification, according to an analysis of the accuracy ratios of these approaches. The state of the art includes a number of classifier approaches. Most accuracy ratios indicate that approaches based on CNN perform better than those using alternative classifiers. On the other hand, ANN classifiers are gaining popularity because to their high accuracy rates. The accuracy ratio attained by various classifiers is shown in Fig.5.

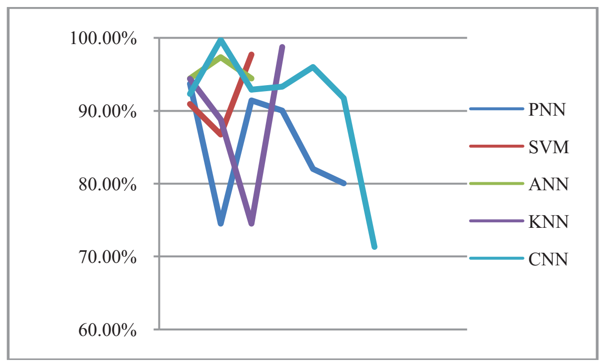

Fig 5. Accuracy ratio attained by various classifiers

6. CONCLUSIONS

This article reviews various image processing methods for plant disease detection that have been employed by numerous scholars globally. The final objective of using image processing methods is to reduce the effect of diseases on farming products. GLCM, KNN Classifier, K Mean Clustering, ANN, CNN, MDC, Otsu Thresholding Method, SVM, Feed Forward BPNN, and PCA are the main methods examined in this article. Various techniques' accuracy rates are described in depth. This would make it easier for academics to choose the best approach for their model.The main conclusions of this study can be used to explore additional issues and potential solutions.

REFERENCES

[1] Kishori Patil, Santosh Chobe,” Leaf Disease Detection using Deep Learning Algorithm,” International Journal of Engineering and Advanced Technology (IJEAT)”, ISSN: 2249 – 8958, Volume-9 Issue-3, February 2020, pp.3172 3175.
[2] Dey, A. K., Sharma, M., & Meshram, M. R. (2016). Image processing based leaf rot disease, detection of betel vine (Piper BetleL.). Procedia Computer Science, 85, 748-754.
[3] Simranjeet kaur, Geetanjali Babbar, Gagandeep,” Image Processing and Classification, A Method for Plant Disease

Detection", International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-8, Issue-9S, July 2019, pp. 869-871.
[4] Joshi, A. A., & Jadhav, B. D. (2016, December). Monitoring and controlling rice diseases using Image processing techniques. In 2016 International Conference on Computing, Analytics and Security Trends (CAST) (pp. 471-476).IEEE.
[5] Khaing war Htun & Chit Su Htwe"Development of paddy diseased Leaf classification system using modified color conversion".iJournals: International Journal of software & Hardware Research in Engineering, ISSN-2347- 4890, Volume 6 Issue 8 August, 2018, pp 24-32.
[6] Saradhambal, G., Dhivya, R., Latha, S., Rajesh, R., „Plant Disease Detection And Its Solution Using
Image Classification", International Journal of Pure and Applied Mathematics, 2018,Volume 119, No. 14,
pp. 879-884
[7] Ratnasari, E. K., Mentari, M., Dewi, R. K., & Ginardi, R. H. (2014, September). Sugarcane leaf disease
detection and severity estimation based on segmented spots image. In Proceedings of International Conference on Information, Communication Technology and System (ICTS) 2014 (pp. 93-98). IEEE.
[8] Hiteshwari Sabrol, Satish Kumar, "Recent Studies of Image and Soft Computing Techniques for Plant
Disease Recognition and Classification", International Journal of Computer Applications (0975-8887), Volume 126 – No.1, September 2015
[9] Majumdar, D., Kole, D. K., Chakraborty, A., & Majumder, D.D. (2015, August). An integrated digital image analysis system for detection, recognition and diagnosis of disease in wheat leaves. In Proceedings of the Third International Symposium onWomen in Computing and Informatics (pp. 400-405).
[10] Aravind, K.R., Raja, P., Mukesh, K.V., Aniirudh, R., Ashiwin, R., „ Disease Classification in Maize Crop
using Bag of Features and Multiclass Support Vector Machine", Proceedings of the Second International
Conference on Inventive Systems and Control (ICISC), 2018, pp. 1191-1196
[11] Sannakki, S. S., Rajpurohit, V. S., Nargund, V. B., & Kulkarni, P. (2013, July). Diagnosis and classification of grape leaf diseases using neural networks. In 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT) (pp. 1-5). IEEE.
[12] Mokhtar, U., Ali, M. A., Hassenian, A. E., & Hefny, H. (2015, December). Tomato leaves diseases detection approach based on support vector machines. In 2015 11th International Computer Engineering Conference (ICENCO) (pp. 246-250). IEEE.
[13] Barbedo, J. G. A. (2013). Digital image processing techniques for detecting, quantifying and classifying plant diseases. SpringerPlus, 2(1), 660.
[14] Suresha, M., Shreekanth, K. N., & Thirumalesh, B. V. (2017, April). Recognition of diseases in paddy leaves using knn classifier. In 2017 2nd International Conference forConvergence in Technology (I2CT) (pp. 663-666). IEEE.
[15] Xiao, M., Ma, Y., Feng, Z., Deng, Z., Hou, S., Shu, L., & Lu, Z. (2018). Rice blast recognition based on principal component analysis and neural network. Computers and electronics in agriculture, 154, 482- 490.
[16] Dixit, A. K., & Yadav, R. K. (2019). A review on image contrast enhancement in colored images. *Journal of Computer Sciences and Engineering*, *7*(4), 263-273.
[17] Verma, Swati, Rakesh Kumar Yadav, and KuldeepKholiya. "A Scalable Machine Learning-based Ensemble Approach to Enhance the Prediction Accuracy for Identifying Students at-Risk." *International Journal of Advanced Computer Science and Applications* 13, no. 8 (2022).
[18] Barbedo, J.G.A. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Comput. Electron. Agric. **2018**, 153, 46–53.
[19] Shrivastava, V.K.; Pradhan, M.K. Rice plant disease classification using color features: A machine learning paradigm. J. Plant Pathol. **2021**, 103, 17–26.
[20] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
[21] Lee, S.H.; Goëau, H.; Bonnet, P.; Joly, A. New perspectives on plant disease characterization based on deep learning. Comput. Electron. Agric. **2020**, 170, 105220.
[22] Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agric. **2018**, 145, 311–318.
[23] Fuentes, A.; Yoon, S.; Kim, S.C.; Park, D.S. A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition. Sensors **2017**, 17, 2022.
[24] Militante, S.V.; Gerardo, B.D.; Dionisio, N.V. Plant Leaf Detection and Disease Recognition using Deep Learning. In Proceedings of the 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, 3–6 October 2019; pp. 579–582.
[25] Too, E.C.; Yujian, L.; Njuki, S.; Yingchun, L. A comparative study of fine-tuning deep learning models for plant disease identification. Comput. Electron. Agric. **2019**, 161, 272–279.
[26] Kumar, A., Shukla, S. K., Fatima, E., & Yadav, R. K. (2022). A Neoteric Procedure for Spotting and Segregation of Ailments in Mediciative Plants using Image Processing Techniques. *International Journal of Next-Generation Computing*, *13*(4).
[27] Dixit, A. K., & Yadav, R. K. (2019). A review on image contrast enhancement in colored images. *Journal of Computer Sciences and Engineering*, *7*(4), 263-273.