iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VII **Month of publication:** July 2021

DOI: https://doi.org/10.22214/ijraset.2021.37006

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Leaf Disease Detection using Group Method Data Handling

P. Kumari[1], K. Thulasimani[2]
[1]PG Student, [2]Professor, Dept. of Computer Science, Government College of Engineering, TirunelveliTamil Nadu, India.

Abstract: ***During crop planting, the location of sicknesses in the leaf parts is one of the criticalconnects to the anticipation and control of yield illnesses. This paper takes different leaves as exploratory articles, and uses the profound learningtechnique to remove the illness highlights on leaf surface. After persistent iterative learning, the organization can anticipate the class of each sickness picture. The guided channel is utilized in pre-handling stage and the highlights are extricatedutilizing texture feature extraction, color feature, morphological technique. Then the selected features are fed into Group Method of Data Handling (GMDH) and the comparison experiments are performed. The outcomes illustrate that the method is effective, it can identify whether the plant is the diseased plant or not.***

Keywords: Plant disease classification, Guided filter, GLCM, Morphological, GMDH .

I. INTRODUCTION

The event of plant sicknesses impact affects rural creation, and if the plant sicknesses aren't distinguished on schedule, there'll be Associate in Nursing growth in food weakness.[1] within the application exploration of harvest illness identification, customary computer vision techniques usually have to be compelled to section leaf sores, for instance, pixel-level division, edge division, district division and multi-scale division. Commotion decrease, erosion, improvement and totally different strategies square measure applied to handle image shading area highlights and surface highlights, and afterwards appropriate sore highlights and classifiers square measure picked for recognition. Srivastave Associate in Nursing et al. [1] utilized a measurable limit strategy to accomplish image division of grape sick leaves in a very three-dimensional shading area, and afterwards created a call concerning wool buildup captivated with shading contrasts. Wang S Z et al.[2] planned a touch K-implies grouping calculation for plant leaf infection identifying proof, utilizing vector middle separating to eliminate commotion, extricating leaf infection highlight vectors, and designing input area tests to high- dimensional component area for K-Mean grouping and plant ill health recognizable proof. Wang M L et al. [3] modified over the RGB shading area to the HSV area to separate shading highlights and mathematical highlights, and adequately distinguished four traditional wheat sicknesses. S D Khirade et al. [4] talked concerning the possibility of utilizing leaf footage to acknowledge plant sicknesses, even as some division and highlight extraction calculations used in plant infection recognition. A lot of past works have considered the image acknowledgment, also, a specificclassifier is utilized which classifies the pictures into solid or sick pictures. By and large, the leavesof plants are the first wellspring of plant illness recognizable proof, and the manifestations of most infections may begin to show up on the leaves. In the previous many years, the essential characterization procedures that were prevalently utilized for illness recognizable proofin plants incorporate k-nearest neighbor (KNN), uphold vector machine (SVM) (Deepa and, fisher direct discriminant (FLD), counterfeit neural network (ANN), irregular timber (RF) and then on. As we have a tendency to as a full understand that the ill health acknowledgment paces of the standard approaches rely smartly on the injury division and hand-planned highlights by totally different calculations, like seven invariant minutes, Gabor modification, worldwide neighborhood solitary value, and scanty portrayal and then on…[5]. All the additional as lately, profound learning procedures, particularly convolutional neural organizations (CNNs), square measure quickly turning into the likable techniques to beat a number of difficulties. CNN is that the most well-known classifier for image acknowledgment in each huge and tiny scale problems it's. it's shown exceptional capability in image making ready conjointly, grouping. for example, Mohanty, Hughes and Marcel ready a profound learning model for perceiving fourteen crop species and twenty six yield sicknesses. Their ready model accomplishes Associate in Nursing exactitude of ninety nine.35% on a held-out check set. Mama et al used a profound CNN to direct manifestation shrewd acknowledgment of 4 cucumber sicknesses, i.e., fleece buildup, anthracnose, fine mold, what is more, target leaf spots. They received the acknowledgment exactitude of ninety three.4%. Kawasaki et al.[6] conferred a framework captivated with CNN to understand cucumber leaf illness; it understands Associate in Nursing exactitude of ninety four.9%, and then forth Albeit impressive outcomes are accounted for within the writing, examinations thus far have used image data bases with restricted selection.

©IJRASET: All Rights are Reserved

The foremost photographic materials incorporate footage completely in trial (research center) arrangements, not in real field wild things. [7] while not a doubt, footage caught in development field conditions incorporate a good sort of foundation and a broad assortment of indication qualities. what is more, there square measure Associate in Nursing huge range of boundaries ought to are ready for CNN and its variations, whereas making ready these CNN styles in addition needs totally different marked examples conjointly, vital computer assets with none preparation to guage their exhibition. Gathering a huge named dataset is while not a doubt a tough trip. all the same the restrictions, the past examinations have effectively exhibited the aptitude of profound learning calculations.

Especially, the profound exchange realizing, which eases the issue looked by old style profound learning techniques, for example the arrangements comprising of utilizing a pre- prepared organization where just the boundaries of the last characterization levels should be surmised without any preparation is normally utilized in the pragmatic application. In this work, we study the exchange learning for the profound deep learning in with the point of improving the learning capacity of small injury manifestations alongside diminishing the computational complexity.

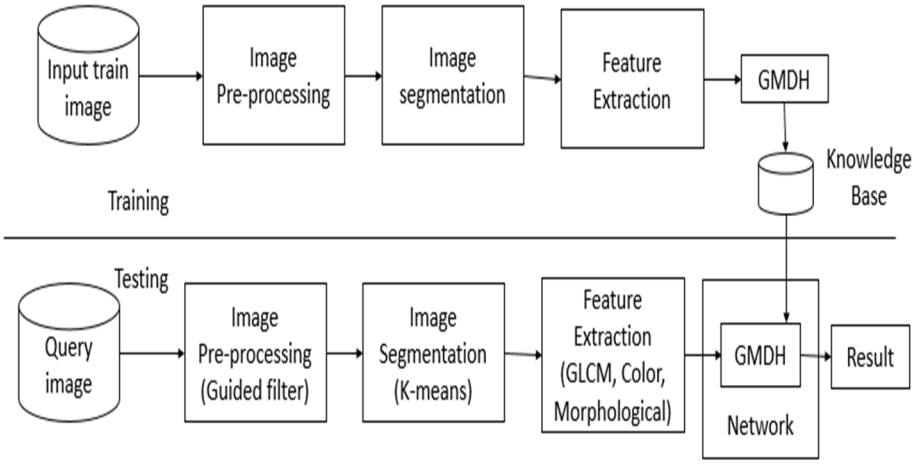


Fig-1 Block Diagram Plant LeafDisease Detection

II. METHODOLOGIES

In the planned technique, leaves were square measure taken as trial article and cluster methodology of information Handling (GMDH) was embraced because the basic model. within the examination, from the leaf options square measure extracted by the feature extracted technique .then it's square measure fed into GMDH. In GMDH, the most improvement points square measure Transfer operate and External criterion. In transfer operate, within the network layer, all doable pairs of the m inputs square measure generated to create the transfer functions of the h=p*(p-1)/2 neurons and the traditional GMDH linear function is replaced by the nonlinear function. In external criterion, traditional GMDH method is often used the accuracy criteria as the external criterion, the main representation of accuracy criteria is the regularity criteria. But as for the plant disease detection, it could be correctly identified whether the plant is diseased or not, which is more than the detection accuracy.

A. Image Acquisition

Images are gained from Gallery. Initially, the images of different leaves are obtained utilizing an advanced camera with required goal for better quality. The information image is then resized to 256x256 pixels. The development of a image data set relies upon the necessary application. The image data set must be deliberately built in that it for the most part chooses the proficiency of the classifier and execution of the proposed strategy. The images of different leaves are procured utilizing an advanced camera with required goal for higher quality. the event of an image information set depends upon the mandatory application. The leaf footage square measure gathered from Kaggle information set. Kaggle is facilitating this opposition for the data science native space to use for the sake of diversion and coaching. The goal of this construction contention is to utilize double leaf footage and removed highlights, together with form, edge and surface, to exactly acknowledge 99 sorts of plants. Leaves, as a result of their volume, generality, and fascinating attributes, square measure a flourishing strategies for separating plant species. They in addition provides a nice introduction to applying ways that embrace image based mostly highlights.

©IJRASET: All Rights are Reserved

B. Image Pre-Processing

The target of the preprocessing stage is to apply conceivable picture upgrade methods to get the necessary visual nature of the pictures. In the proposed technique guided channel is utilized for pre handling. Guided channel is conventional idea for edge protecting smoothing and design moving separating.

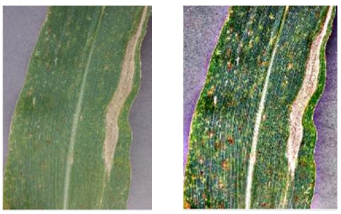


Fig-2 (a) Input image (b) Pre-processed image

C. Image Segmentation

In the exploration and utilization of pictures, individuals are regularly just intrigued by specific pieces of the pictures. These parts are frequently called targets or forefronts, and they by and large relate to explicit, extraordinary zones of the picture. To distinguish and break down the objective, these applicable territories should be isolated and extricated. Picture division alludes to the strategy and cycle of separating a picture into trademark zones and removing objects of revenue. Picture division, which is very significant for PC vision, is brought as parceling a picture into its areas dependent on certain measures where the locales are important and disjoint. Image division is by and large thought to be a middle of the road step of some example acknowledgment applications. In this task, profound leftover organization is utilized for division and arrangement reason.



Fig-3 Segmented image

D. Group Method Data Handling

GMDH algorithm is a system of layers in which there exist neurons. In each layer, there exist a number of neurons. The number of neurons in a layer is defined by the number of input variables. To illustrate, assume that the number of input variables equals to p, since we include two input variables, the number of neurons is to be equal to $h = p \times (p-1)/2$. The architecture of GMDH algorithm is illustrated when there exist three layers and four inputs. In this architecture, the number of inputs is equal to four; therefore, the number of nodes in a layer is determined to be six. In input layer, there exist four input variables. There is no any process at this layer. This is just a starting layer to the algorithm. All plausible combination of four input variables enters to each neuron at first layer. The coefficient is estimated in each neuron. By using estimated coefficients and input variables in each neuron, the desired output is predicted. According to external criteria, p neurons are selected and h-p neurons are eliminated from the network. In this architecture, four neurons are selected while two neurons are eliminated from the network. The outputs obtained from selected neurons become the inputs for the next layer. This process continues until the last layer. At the last layer, only one neuron is selected. The obtained output from last layer is the predicted values for the time series at hand.

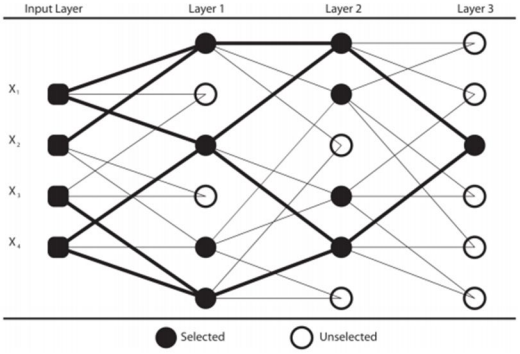


Fig-4 Architecture of GMDH

©IJRASET: All Rights Reserved

Transfer functions (also known as activation functions, utilized throughout interchangeably) are used to capture the better model which explains the relation between inputs and desired outputs. Polynomial function was used to explain the relation between inputs and output in GMDH-type neural network. Also, we include tangent transfer function to consider the sinusoidal relation between covariates and response variable. Transfer functions used in study are as follows, Polynomial Function: $z = y$
In GMDH algorithm:

$$y = \beta 0 + \beta 1xi + \beta 2xj + \beta 3xi\,2 + \beta 4xj\,2 + \beta 5xixj \quad \text{----------(1)}$$

where y is a response variable, xi and xj are the covariate variables defined. β's are the weights where $l = 0, 1, \ldots, 5$.

E. Feature Extraction

1) GLCM Extraction: GLCM highlight extraction is used. Gray Level Co-Occurrence Matrix (GLCM) is that the factual technique for examining surface which worries about the spatial relationship of pixels. The GLCM capacitiesportray the surface of images by processing the spatial relationship among the pixels within the pictures. The factual measures are aloof from this lattice. The Grey Level Co-event Matrix (GLCM) gives an idea regarding the progress of powers between pixels a selected way and distance. Co- event implies the occasions the dim degree of pixel j follows the dim degree of pixel I with a particular goal in mind. During a cerebrum picture of size M * N thedark levels L are often signified by

G = 0, 1, . . ., L - 1.The co-event network of an image may be a L * L square grid and is supposed as
P=[t_ij]_(L*L).

The components of the network indicated by the quantities of advances between all sets of dark levels in G = 0, 1, . . ., L _ 1 and registered for variousestimations of u are often addressed as

$$P(i,j,d,\theta)=((k,l),(m,n))\in(L_y*L_x)*(L_y*L_x)|k-m=0|,|l-n| \quad (2)$$

2) Color Features: The features selected here are considered consistent with the leaf diseases which can be more discriminative. In proposed approach total 10 features are taken, i.e. 8 color features and size of disease, distances of diseased spots from one another. Color features from images are taken by sum and average technique. We've 768x768 size image database, therefore from these images we are getting 8 distinct values. In extraction process, first we take sum of all columns so average of 100 pixels gives 8 values. If the centroid is defined, it's a hard and fast point of all isometries in its symmetry group. Specifically, the geometric centroid of an object within the intersection of all its hyperplanes of symmetry. The centroid of the many figures will be determined by this principle alone.

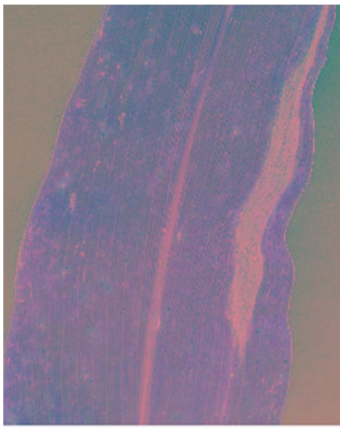


Fig-5 Color Feature Extracted Image

©IJRASET: All Rights are Reserved

3) *Morphological Feature Extraction:* Morphological analysis accustomed collect the form of a picture. Morphological analysis has been verified to exhibit an excellent performance in denoising. This method functions with two basic operators as follows:

Erosion:

$$(f\Theta g)(n) = \min[f(n+m) - g(m)],\ m = 0,1,2,..M-1,\ n = 0,1,2,\dots,N-1$$

----------(3)

Dilation:

$$(f \oplus g)(n) = \max[f(n-m) - g(m)],\ m = 0,1,2,..M-1,\ n = 0,1,2,\dots,N-1$$

----------(4)

Where f(n) is that the original one-dimensional vibration signal, g(m) is that the (Structure Element) SE, Θ and $\oplus$ are the operators of erosion and dilation, respectively. Erosion calculation is used to suppress and smooth the positive and negative impacts, respectively. Against this, dilation calculation is employed to flatten and suppress the positive and negative impacts, respectively. Another two operators are created on the premise of two basic operators as follows:

Opening:

$$(f \circ g)(n) = (f\Theta g \oplus g)(n),\ m = 0,1,2,..M-1,\ n = 0,1,2,\dots,N-1$$

----------(5)

Closing:

$$(f * g)(n) = (f \oplus g\Theta g)(n),\ m = 0,1,2,..M-1,\ n = 0,1,2,\dots,N-1$$

----------(6)

Where $\circ$ and $*$ represent the opening and closing functions, respectively. The opening operator suppresses and preserves the positive and negative impacts, respectively. In contrast, the closing operator suppresses and preserves the negative and positive impacts, respectively. The preceding four operators only calculate the feature information from one aspect and will lose some geometric characteristics of the signal, which is meaningful in fault diagnosis. To detect the impulsive components, the closing and opening operators are combined to determine the difference operator, which might extract the positive and negative fault impacts.

Difference

$$(f * g_n - f \circ g)(n) = (f \oplus g\Theta g - f\Theta g \oplus g)(n)$$

----------(7)

The performance of morphological analysis depends on the operators and SEs; therefore, selecting an appropriate SE is critical. SEs are mainly determined by the length, height, and shape. SEs with a line shape are determined to perform well; hence, considerable attention should be focused to see the length of SEs.

F. Classification

The classical GMDH algorithm uses the linear function as its transfer function, and therefore the result of which is that the detailed numerical data, it's often accustomed to handle the info prediction issue. However, the disease detection within the real business is usually complicated and linear inseparability, during this situation, it is unreasonable to use the linear transfer function to represent the link between response variables and characteristic variables. Therefore, this paper combined the benefits of the Logistic and GMDH algorithms to propose a replacement algorithm GMDH-Logistic, the most improvement points are Transfer function and External criterion.

The detailed steps of GMDH-Logistic are listed as follows.

1) Divide the dataset D into k parts
2) Transfer function formula & external criterion formula are accustomed train GMDH model
3) Every 2 nodes of the previous layer are wont to generated to induce a replacement input nodes. Unknown parameter were estimated by least square
4) Selected model-survived, they're used as input of next layer. Non-selected models – abandoned.
5) Repeat 2-4, until optimal complexity model obtained.
6) Testing set is applied to validate and logistic function is employed for classification
7)

©IJRASET: All Rights are Reserved

III. RESULT AND DISCUSSION

The plant disease is processed and its textural features, color feature and morphological features are extracted is quicker than their features are given to classification to predict the disease in leaf. The plant leaf images are taken from kaggle website. 190 images are used for this experiment. In that, 115 images taken the train the network and 75 images are used for testing purpose.

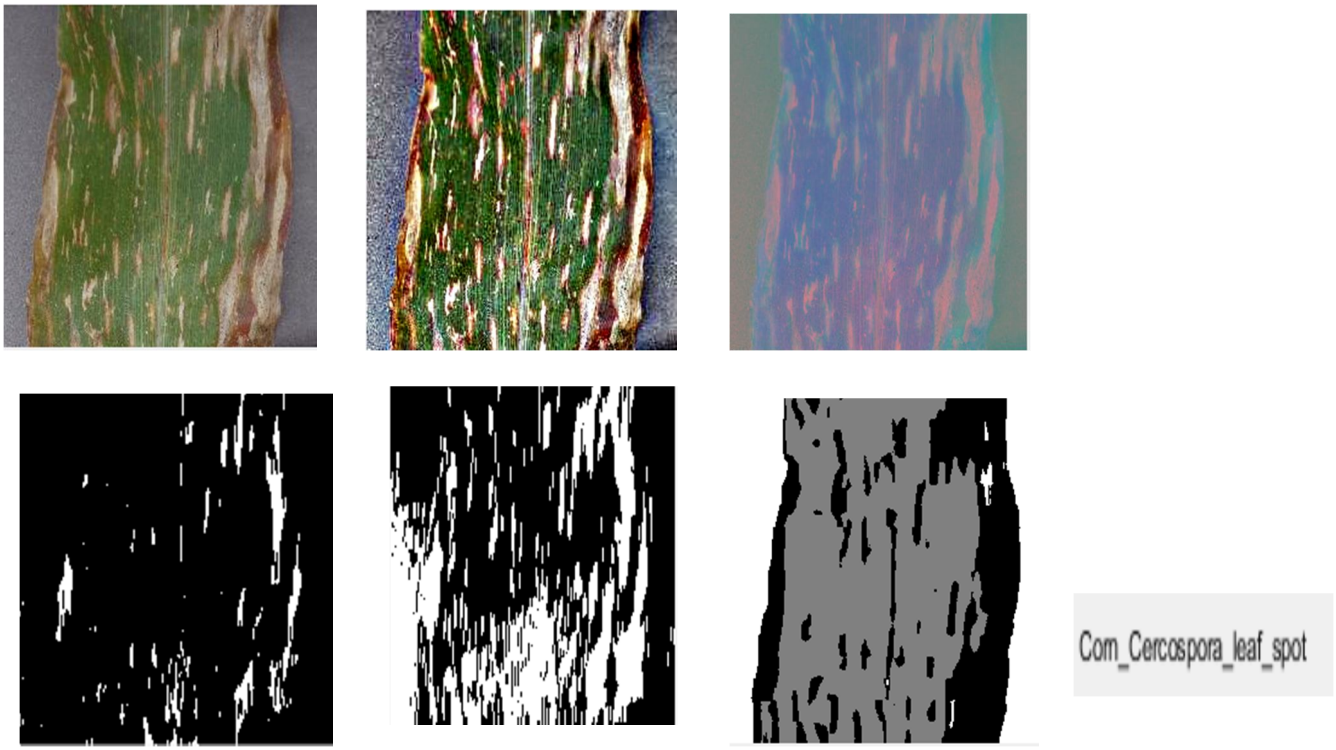


Fig-6 Cercospora Leaf Spot Output Image

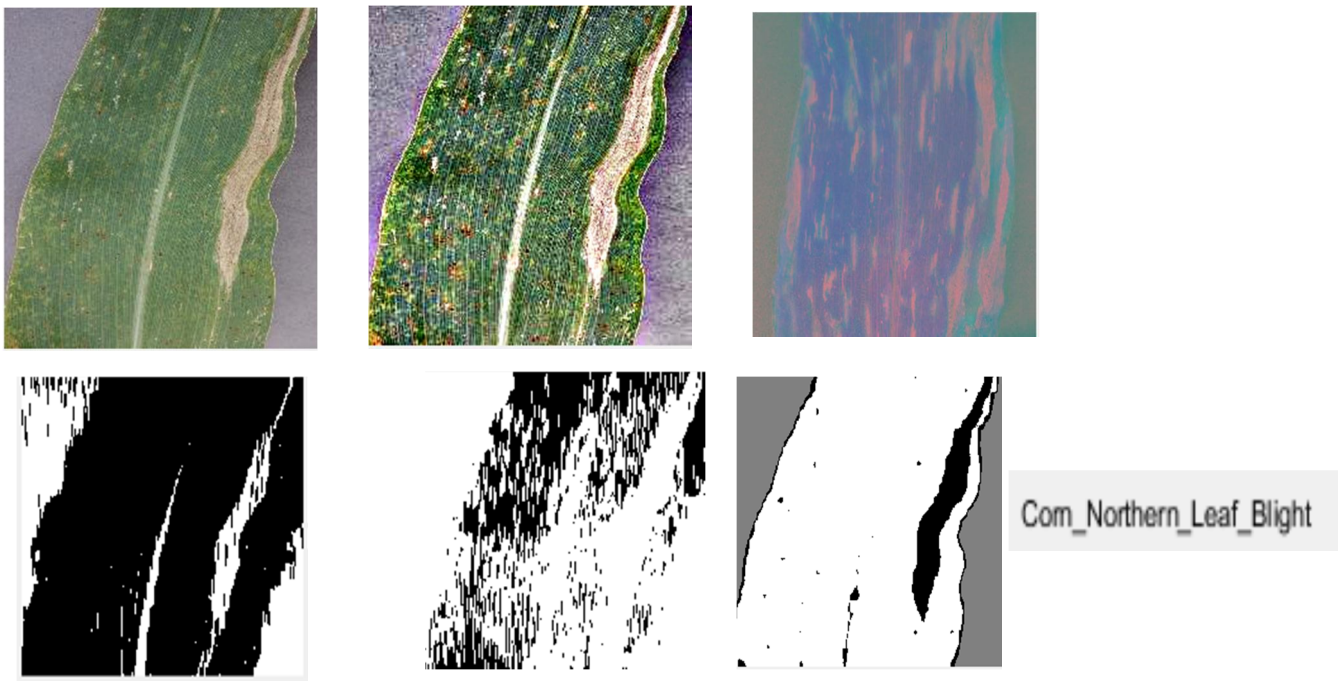


Fig-7 Black_Rotnorthern Leaf Blight Output Image

©IJRASET: All Rights are Reserved

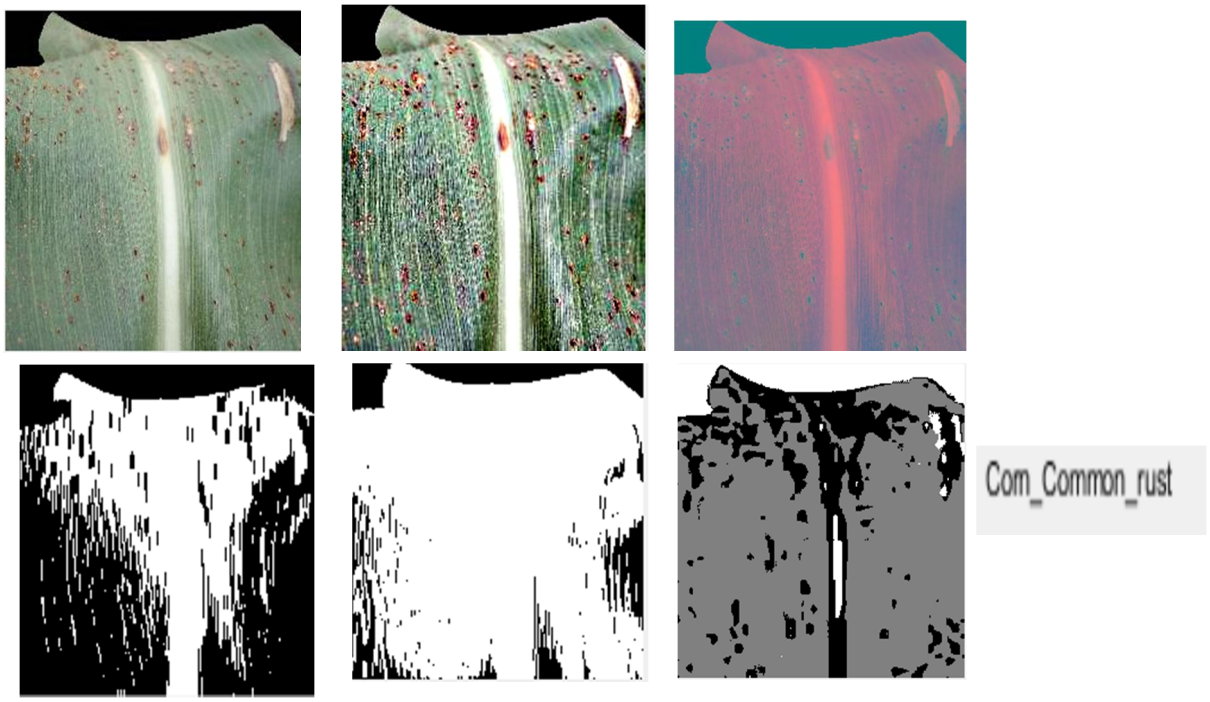


Fig-8 Common Rust Output Image

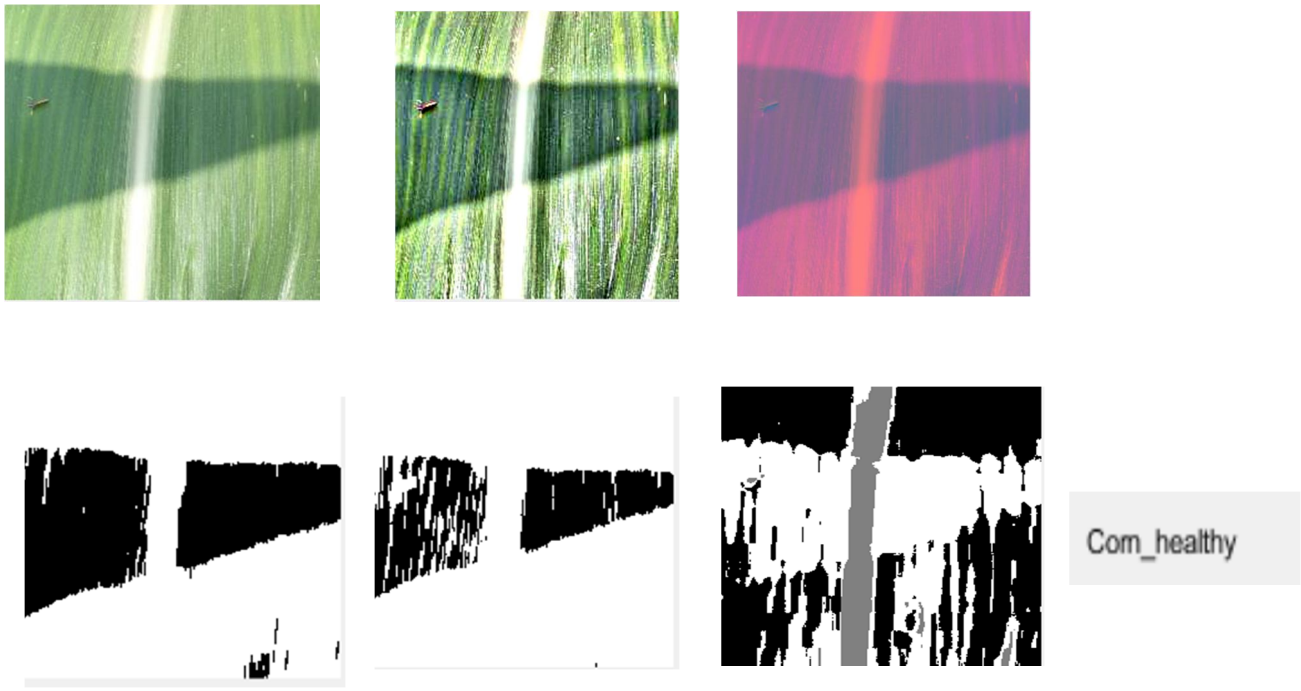


Fig-9 Healthy Leaf Output Image

©IJRASET: All Rights are Reserved

Table-1 Confusion matrix (SVM).

Categorize	Positive	Negative
True	120	5
False	62	3

Table-2 Detection Rate.

Total Images	Training Images	Testing Images	Detection Rate
190	115	75	98.25%

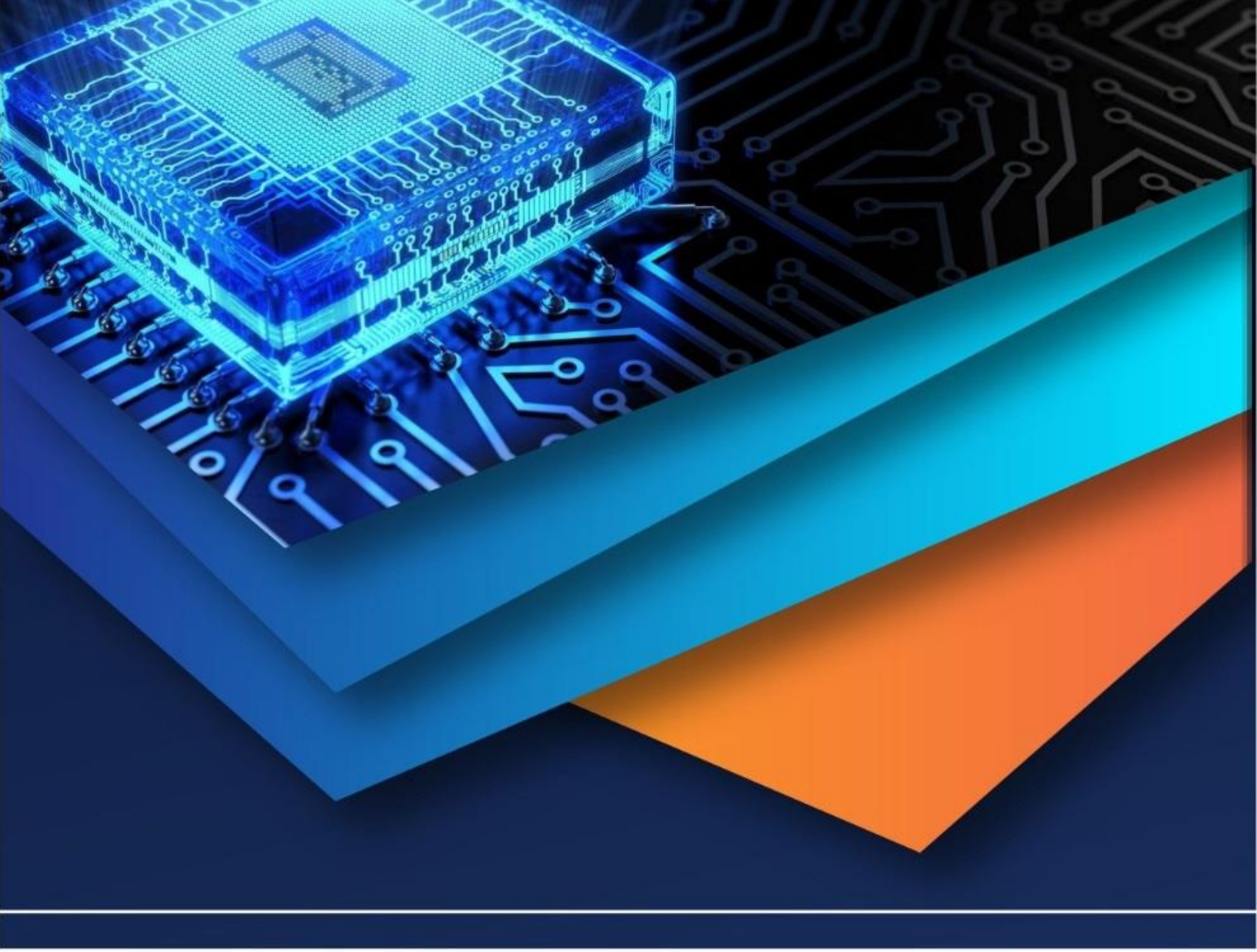
IV. CONCLUSION

In this paper, a classification method of leaf diseases supported on deep learning was introduced to spot and classify leaf diseases. Supported the improved Resnet-50 model, leaf pictures (after data augmentation) of various diseases were selected and ready for batch learning and training. First of all, for the classification of multiple diseases of a one species using GMDH. The experiment has achieved a particular effect.

REFERENCES

[1] Srivastava A, Ma S, Inoue K. Development of a sensor for automatic detection of downey mildew disease [J]. 2004:562 - 567.
[2] WANG S Z, HE D J, LI W, et al. Plant leaf disease identification based on kernel k-means clustering algorithm [J]. Agricultural machinery, 2009, 40(3):152-155.
[3] WANG M L, NIU X J, ZHANG H M, et al. Researchon feature extraction and recognition of common diseases of wheat leaf [J]. Computer Engineering and Applications, 2014, 50(7): 154 – 157.
[4] KHIRADE S D, PATIL A B. Plant diseasedetection using image processing[C]//Proceedings of the 2015 International Conference clustering algorithm and correspondence filters.Int. J. Adv. Biotechnol. Res. 4 (2), 189–199.
[5] Al-Hiary, H., Bani-Ahmad, S., Reyalat, M., Braik, M., Alrahamneh, Z., 2011. Fast and accurate detection and classification of plant diseases. Int. J.Comput. Appl. 17 (1), 31–38.
[6] Bai, X., Cao, Z., Zhao, L., Zhang, J., Lv, C., Li, C., Xie,J., 2018. Rice heading stage automatic observation by multi-classifier cascade based rice spike detection method. Agricul. Forest Meteorol. 259, 260–270.
[7] Al Bashish, D., Braik, M., Bani-Ahmad, S., 2011. Detection and classification of leaf diseases using K-means-based segmentation and Information Technology Journal 10 (2), 267–275.
[8] Pooja, V., Das, R., & Kanchana, V. (2017, April). Identification of plant leaf diseases using image processing techniques. In 2017 IEEE Technological Innovations in ICT for Agriculture and RuralDevelopment (TIAR) (pp. 130-133). IEEE.
[9] Khirade, S. D., & Patil, A. B. (2015, February). Plant disease detection using image processing. In 2015 International conference on computing communication control and automation (pp. 768- 771). IEEE.
[10] Ebrahimi, M.A., Khoshtaghaza, M.H., Minaei, S., Jamshidi, B., 2017. Vision-based pest detection based on SVM classification method. Comput. Electron. Agricul. 137, 52–58.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)