International Journal of Advance Research in Science and Engineering
Volume No. 11, Issue No. 04, April 2022
www.ijarse.com
IJARSE
ISSN 2319 - 8354

"GRAPES LEAVES DISEASE DETECTION USING K-NN CLASSIFICATION ALGORITHM"

Madan Piske[1], Dipak Kurade[2], Hrushikesh Khaladkar[3], Vishal Kolekar[4], Prof. Sushadevi Adagale[5]

[1,2,3,4,5,]*Computer Engineering, Trinity Academy of Engineering, Pune, India*

Abstract

In this research, we point out a comparative study on Grape leaf disease detection and all the techniques of classification based machine learning by K-NN (K-Nearest Neighbour) algorithm, which had been used for leaf disease detection. We have proposed an automatic system for detecting the diseased leaf using image processing and machine learning technique. The system take the dataset of unhealthy and healthy leaves and then do feature extraction using color and texture and then (do training and testing using k-nn classification supervised machine learning algorithm) and we have obtained a better testing accuracy of 96 Percentage.

I. INTRODUCTION

Plant disease is one of the crucial causes that reduces quantity and degrades quality of the product. The ability of disease diagnosis in earlier stage is then very important task in order to be able to timely cure and control such disease for decreasing satisfactory products. [2]. There are numerous characteristics and behaviors of such plant diseases in which many of them are merely distinguishable. Hence, human expertise is normally required for disease diagnosis. This allows usage of many image processing and computer vision techniques to apply with plant diagnosis problem because disease diagnosis is commonly performed by visual inspection. In present, using such approaches can partially replace plant disease expert, providing usage of a diagnosis system without having any prior knowledge. For examples, techniques of using artificial neural network and image processing are utilized for detection and classification of Grapes Leaves disease using co-occurrence matrix from gray-scale image Classification algorithm.Color, shape, and size information are fed to backpropagation neural network in the stage of disease classification. The proposed work concentrates on detection of unhealthy grape leaves affected by diseases and water deficiency using image processing techniques. It also depicts accuracy of classification algorithms for grape leaves. Leaves are classified as healthy leaf or unhealthy leaf. Unhealthy leaf in this context means that leaf is affected (damaged) by some diseases or due to water deficiency. Classification algorithms like KNearest Neighbor (KNN) is used for classification purpose. Proposed work also includes the comparison of KNN and other and Shirdhonkar [5] presented a review on plant infection discovery in image processing discipline. The features such as color, texture, and morphological are utilized to recognize plant illnesses. Most regularly utilized classifiers for ordering plant infections are Artificial Neural Network (ANN) and SVM. Assessment of recurrent neural network would help in identification of patterns in many applications and reduces the error rate [4]. Xu et al. [8] also presented a study on influence of non independent sampling methods on incremental support vector machines learning performance. Their work includes comparison of Markov re-sampling and random non dependent sampling and it is observed that Markov re-sampling provides

better clarification rates and reduced training time. Identifying the plant diseases automatically would tackle the issue of costly specialists. Vijai Singh and A.K. Misra [3] proposed a method to detect disease in leaf based on segmentation. The experiment is done on plant species such as beans, jack fruits, lime, banana, potato and tomato. Their work includes extraction of texture features including contrast, energy, homogeneity and entropy. They also suggest that the economy of nation profoundly relies upon the agricultural efficiency. Hence it is essential to recognize the infection in plant at extremely starting stage with the goal that the profitability of the yield won't be influenced. In the proposed system, features such as texture and color are used to classify the grape leaves using KNN and SVM classifiers. Texture features include properties contrast, correlation, energy and homogeneity. These properties are explained in the next section.

II. LITURATURE SURVEY/BACKGROUND

Diverse research has tended to robotized plant disease identification by utilizing image processing techniques. In 2013, Sannakki et al. [1] proposed a work that intends to distinguish the diseases by applying artificial intelligence and image processing techniques on the grape leaf. The green pixels of the leaves are masked using threshold to eliminate the noise to obtain a processed image. The separation of diseased part from grape leaf image is done by applying K-means clustering so that the affected portion of the plant can be identified clearly. It is observed that feed forward back propagation neural network classification gives the best result. In 2014, Deshpande et al. [6] presented that image processing is not only used for plant disease detection but also for grading plant infections. At the point when SVM is utilized as classifier, accurate results are found by utilizing morphology features compared with color coherence vector (CCV) and color features [3]. Features for example structure of hole on the fruit, morphology, texture and vectors color can be used for classification of fruit diseases in pomegranate, apple and grapes [7]. Grey Level Co-occurrence Matrix (GLCM) is used to find the texture features. Every element (r, c) in matrix indicates that how many times a pixel having value r occurs horizontally adjacent to a pixel having value c [8]. In 2016, P. B. Padol and A. A. Yadav [2] exhibited an examination for leaf contamination recognition fundamentally brought about by infection microscopic organisms, and growths. At first the influenced bit of the leaf is discovered utilizing K-means algorithm and features like texture and color are extracted to train the system. There are several challenges in identifying the plant diseases. Jayme Garcia and Arnal Barbedo [7] presented a paper to clarify the problems related with automatic plant disease identification that makes use of images. The challenges such as occurrence of compound backgrounds which are uneasy to distinguish from the scene of interest, the borders of the scene are also not very much characterized. The changing climatic conditions additionally make the investigation of image more challenging. In 2017, Bharate and Shirdhonkar [5] presented a review on plant infection discovery in image processing discipline. The features such as color, texture, and morphological are utilized to recognize plant illnesses. Most regularly utilized classifiers for ordering plant infections are Artificial Neural Network (ANN) and SVM. Assessment of recurrent neural network would help in identification of patterns in many applications and reduces the error rate [4]. Xu et al. [8] also presented a study on influence of non independent sampling methods on incremental support vector machines learning performance. Their work includes comparison of Markov re-sampling and random non dependent sampling and it is observed that Markov re-sampling provides

better clarification rates and reduced training time. Identifying the plant diseases automatically would tackle the issue of costly specialists. Vijai Singh and A.K. Misra [3] proposed a method to detect disease in leaf based on segmentation. The experiment is done on plant species such as beans, jack fruits, lime, banana, potato and tomato. Their work includes extraction of texture features including contrast, energy, homogeneity and entropy. They also suggest that the economy of nation profoundly relies upon the agricultural efficiency. Hence it is essential to recognize the infection in plant at extremely starting stage with the goal that the profitability of the yield won't be influenced. In the proposed system, features such as texture and color are used to classify the grape leaves using KNN and SVM classifiers. Texture features include properties contrast, correlation, energy and homogeneity. These properties are explained in the next section.

III. PROPOSED WORK/SYSTEM

The proposed work is carried out in different phases like image acquisition, feature extraction and classification as shown in Fig. 1.

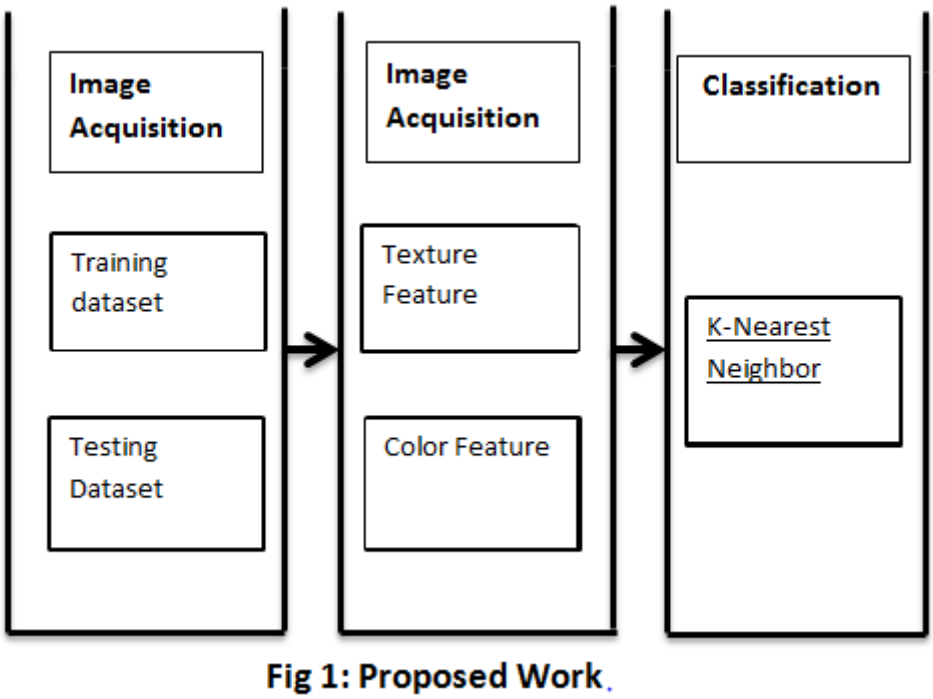

Fig. 1. Propesed Work

A. Image Acquisition First step in our proposed system methodology includes collection and creation of dataset.Images of grape leaves are captured. The dataset is created by capturing images from farms. In the present context, unhealthy leaf is one which is affected by diseases , due to some fungus or by water deficiency.

Fig. 2. Normal Grape leaf

The created dataset includes 50 pics of healthy grape leaves and 50 pics of un- healthy grape leaves. A Sample of captured images of healthy leaf is shown in Fig. 2. In this sample image, leaf is disease free there are no any patches and no spots can be observed .

In this sample image Fig. 3, yellow and dark brown patches on leaf can be ob- served. These patches appear on leaf due to plant disease infection, fungus or some- times due to water deficiency.

B.Feature Extraction In this step, data values related to an image are extracted and these values are called as features. In the proposed work features such as texture and color are extracted and used as input for the classifiers. In general most of the grape leaves have repeated patterns; GLCM is better approach to extract texture features. In addition to that, disease affected part of leaves is usually brown in color. So, in the proposed work both texture and color features are combined to achieve improved results. The texture and color features are described in the following section.

1) Texture Features: GLCM is utilized to calculate texture features [5]. The various steps to find GLCM for the images acquired are as follows -Find number of different intensity levels in the given image I. This identifies the size of a graylevel co-occurrence matrix (GLCM). -Find intermediate matrix A by finding how frequently a pixel p occurs in a particu lar spatial relationship with pixel q. -Calculate GLCM by dividing each element of

matrix A by the sum of elements of matrix A. Once GLCM is computed, properties as such correlation, contrast, homogeneity and energy are computed using GLCM.

Algorithm 1: Texture Feature Extraction

Step 1. Begin Step 2. Read color image (say I). Step 3. Compute GLCM for image I. Step 4. Calculate contrast, correlation, energy and homogeneity for image I using GLCM. Step 5. End

2) Color Features: A fundamental element to watch variation in plant leaves is color shading. In general disease affected part of grape leaves is brown in color. In the proposed work, along with texture features, color features are utilized for identifying the grape leaf influenced by disease. Firstly Red, Green and Blue components are extracted from an image. Extracting red, green and blue components is nothing but identifying the percentage of red, green and blue colors in an image respectively. Then RGB picture is transformed in to Hue, Saturation, Luminance Value (HSV) color spaces. Three components namely hue, saturation and luminance value are extracted. For conversion from RGB(Red Green Blue) to HSV(Hue Saturation Value) color spaces, the values of R, G, and B are divided by 255 for changing the range [0 , 255] to [0 , 1]. Then hue, saturation and luminance values are computed [8]. Totally six color components are extracted and for each of these components, mean, variance and range are computed which results in to eighteen color features. Proposed algorithm to find color features is given below

Algorithm 2: Color Feature Extraction Step 1. Begin Step 2. Read color image (say I1).Step 3. Extract three components red, green and blue from image I1. Step 4. Convert color image I1 to HSV image (say I2). Step 5. Extract hue, saturation and intensity of image I2 [13]. Step 6. For each component extracted in step 3 and step 5, compute mean, variance and range. Step 7. End

B. Classification The Classification algorithm is a Supervised Machine Learning Technique. -i.e. Used to identify new observations on the basis on training data. -In Classification, a program

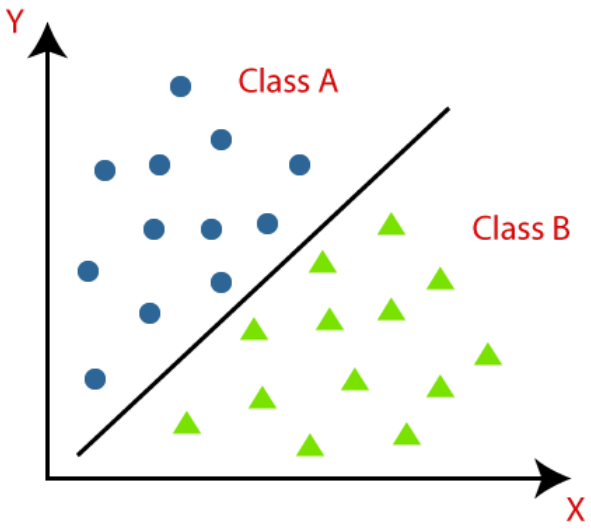

Fig. 4. Classification algorithm

learns from the given observations or and then classifies new observation into a number of groups or classes. Such as,Spam or Not Spam, 0 or 1, cat or dog, Yes or No, etc. Classes can be called as categories or targets labels. -it takes labeled data as input, i.e it contains input with its corresponding output in classification algorithm. Classification Algorithm It will be going to separate on different line or parts or in class A or B etc. etc.

Classification Algorithm

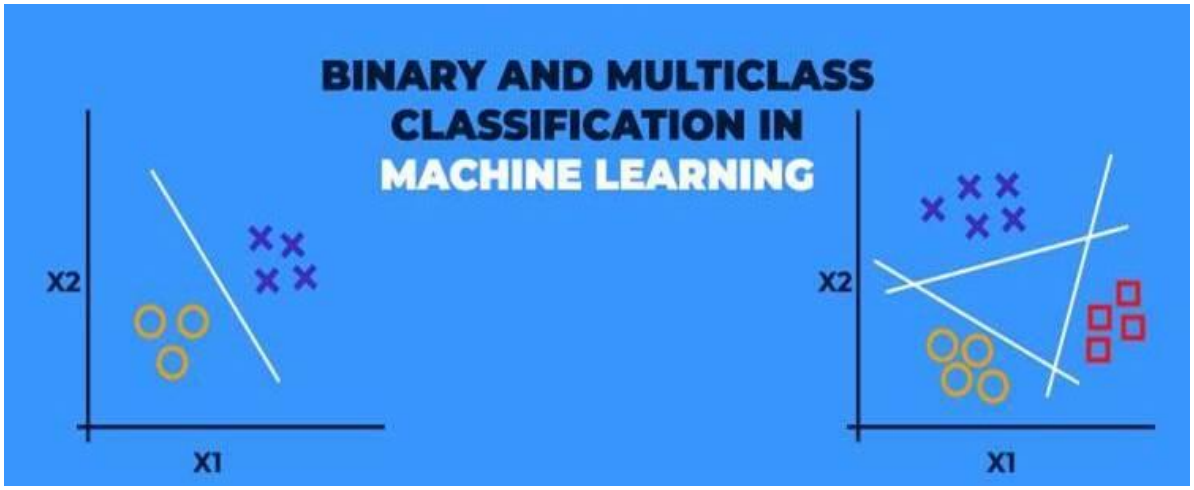

Fig. 5. Types of classification algorithm

Linear Models • Logistic Regression • Support Vector Machines Non-linear Models • K-Nearest Neighbours • Kernel SVM • Na¨ıve Bayes • Decision Tree Classification • Random Forest Classification Types of Classification Algorithm Binary Classifier -If the classification problem's has only 2 possible results then it is called as Binary classifier. -SVM, Logistic. Multi-class Classifier -If the classification problem has more than 2 outcomes then it is called as Multi-Class Classifier. -NB, K-NN, DT

K -Nearest Neighbor Algorithm K-Nearest Neighbors is one of the simplest Machine Learning algorithm based on Supervised Learning technique.In KNN, K indicates the quantity of nearest neighbors. K-NN((K-Nearest Neighbours)) algorithm assumes the correspondence between the new data point case and available dataset's cases and put the new case into category that is more nearto the available dataset categories. K-NN(K-Nearest Neighbours) algorithm stores all the collected dataset or data and classifies a new data point based on the correspondence/similarity

International Journal of Advance Research in Science and Engineering
Volume No. 11, Issue No. 04, April 2022
www.ijarse.com
IJARSE
ISSN 2319 - 8354

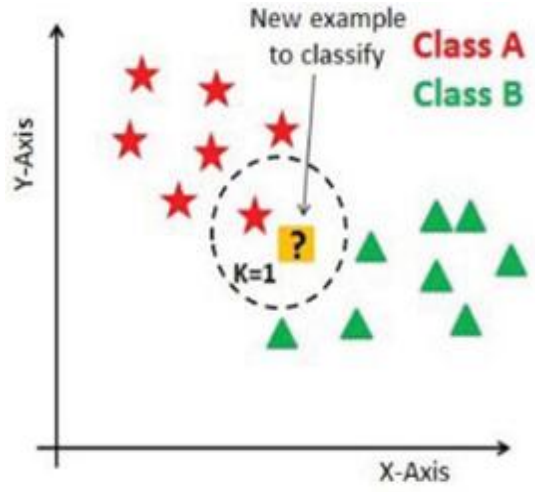

Fig. 6. Illustration of K-NN Classifier

This means if a new data point appears into the model, then it can be classified from its well suitable class/category by using K-NN algorithm. It is a classifier which uses distance between a point and all its neighbors in an image to find nearest neighbors. K-NN algorithm can be used for Classification as well as for Regression but mostly it is used for the Classification problems.

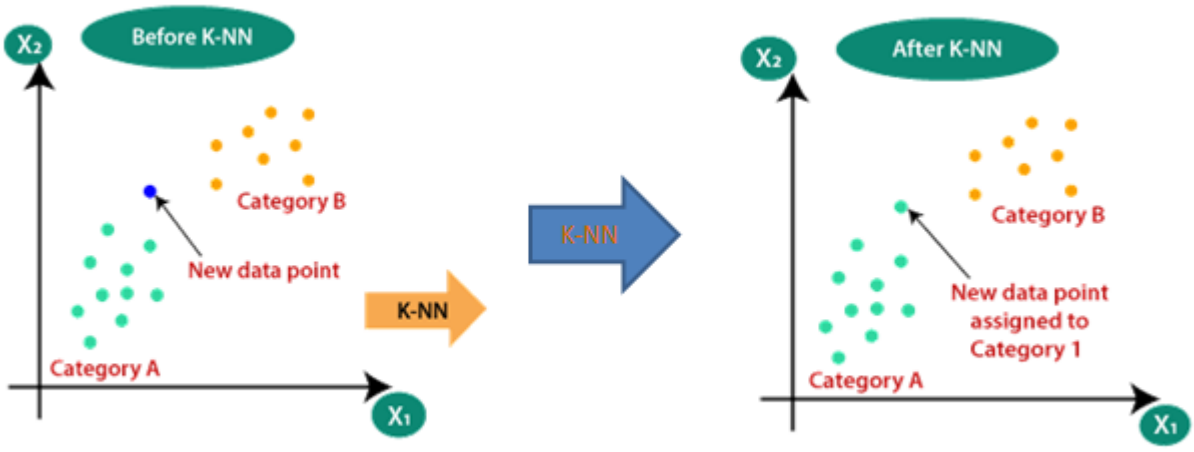

Fig. 7. Working of K-NN Algorithm.

A. Algorithms a) : Steps for K-NN (K -Nearest Neighbor) Algorithm Step-1: Selecting the number K of the neighbors Step-2: now, Calculating the distance by using Euclidean distance of K number of neighbors. Step-3:After calculating Euclidean distance putting the KNearest Neighbor. Step-4: Among these k neighbors,In each category count the number of data points. Step-5: In the data points number of neighbor is maximum assign new data points to that category Step-6: Our model is prepared.

IV. RESULT AND DISCUSSIONS

In the proposed framework, database consisting of 90 images of grape leaves is created manually. Training dataset includes 30 images for healthy and 30 images for unhealthy leaves. Testing dataset includes 30 images including healthy and non-healthy leaves. Matlab tool is used to train the images using 22 combined features (4 texture and 18 color features). An image is randomly selected from the database for testing. Features are extracted from real time images without segmentation is being applied. Testing is done separately using classification techniques i.e. as KNN on 30 test images. When testing is done on randomly selected images, KNN (for K=1) results in 96.66% of true prediction.

V. CONCLUSION

This paper proposes a technique to classify the grape leaf as healthy or unhealthy. Database of 90 images of healthy and unhealthy leaves is created manually. Most relevant features to identify the damaged leaves as such texture and color are used to train and test the system. K-Nearest classifiers is used separately to classify grape leaves. The KNN classifier given better accuracy. Proposed work achieved an accuracy of 96.66% for KNN. As a future work system can be trained to identify the diseases present on the grape leaves and also provide the possible solution.

REFERECNES

[1] S. S. Sannakki, V. S. Rajpurohit, V. B. Nargund and P. Kulkarni, “Diagnosis and classification of grape leaf diseases using neural networks”, In: Fourth International Conference on Computing, Communications and Networking T echnologies (ICCCNT), pp.1-5, Tiruchengode, Tamil Nadu, 2013.

[2] P. B. Padol and A. A. Yadav, “SVM classifier based grape leaf disease detection”, In: Conference on Advances in Signal Processing (CASP), pp.175-179, Pune, 2016.

[3] Manisha Bhange and H. A. Hingoliwala, “Smart farming: Pomegranate disease detection using image processing”, In: proceedings of the Second International Symposium on Computer Vision and the Internet (VisionNet’15), pp.280-286, 2015.

[4] Vijai Singh and A.K. Misra, “Detection of plant leaf diseases using image segmentation and soft computing techniques”, In: Journal of Information Processing in Agriculture, Vol. 4, Issue 1, pp.41-49, 2017.

[5] Anil A. Bharate and M. S. Shirdhonkar, “A review on plant disease detection using image processing”, In: IEEE International Conference on Intelligent Sustainable Systems (ICISS), pp.103- 109, Palladam, Tamil Nadu, 2017.

[6] Tejal Deshpande, Sharmila Sengupta, and K. S. Raghuvanshi, “Grading and identification of disease in pomegranate leaf and fruit”, In: International Journal of Computer Science and Information Technologies, Vol. 5, pp. 4638-4645, 2014.

[7] Ashwini Awate, Damini Deshmankar, Gayatri Amrutkar, Utkarsha Bagul and Samadhan Sonavane, "Fruit disease detection using color, texture analysis and ANN", In: proceedings of International Conference on Green Computing and Internet of Things (ICGCIoT), pp.970-975, 2015.

[8] Jobin Francis, Anto Sahaya Dhas and Annop B. K, "Identification of leaf diseases in pepper plants using soft computing techniques", In proceedings of IEEE Conference on Emerging Devices and Smart Systems, pp.168-173, 2016. [9] https://towardsdatascience.com