Disease Identification in Plants Using K-means Clustering and Classification with ANN

S.Samyuktha Devi, S.SeethaPriya, S.Shamini & Dr.C.Sujatha, Professor, Department of ECE,

SSM Institute of Engineering and Technology, Dindigul, Tamil Nadu, India

ABSTRACT:

The aim of this study is to design, implement and evaluate an image-processing-based software solution for automatic detection and classification of plant leaf diseases. Studies show that relying on pure naked-eye observation of experts to detect and classify such diseases can be prohibitively expensive, especially in developing countries. Providing fast, automatic, cheap and accurate image-processing-based solutions for that task can be of great realistic significance. The methodology of the proposed solution is image-processing-based and is composed of four main phases; in the first phase we create a color transformation structure for the RGB leaf image and then, we apply device-independent color space transformation for the color transformation structure. Next, in the second phase, the images at hand are segmented using the K-means clustering technique. Finally, in the third phase the extracted features are passed through a pre-trained neural network. Present experimental results indicate that the proposed approach can significantly support an accurate and automatic detection and recognition of leaf diseases. The developed Neural Network classifier that is based on statistical classification perform well in all sampled types of leaf diseases and can successfully detect and classify the examined diseases with a precision of around 93%. In conclusion, the proposed detection models based neural networks are very effective in recognizing leaf diseases, whilst K-means clustering technique provides efficient results in segmentation RGB images.

KEY WORDS: Leaf image segmentation, K-means, Feature extraction, plant disease identification, ANN classification.

INTRODUCTION:

Plant diseases have turned into a nightmare as it can cause significant reduction in both quality and quantity of agricultural products, thus negatively influence the countries that primarily depend on agriculture in its economy. Consequently detection of plant diseases is an essential research topics as it may prove useful in monitoring large field of crops and thus automatically detect the symptom of diseases as soon as they appear on plant leaf. Monitoring crops for to detecting diseases plays a key role in successful cultivation. The naked eye observation of experts is the main approach adapted in practice. However, this requires continues monitoring of experts which might be prohibitively expensive in large forms. Further, in some developing countries, formers may have to go long distance to contact experts, this makes consulting experts to very expensive and time consuming. Therefore, looking for a fast, automatic, less expensive and accurate method to detect plant leaf disease cases is of great realistic significance.

Study shows that image processing can successfully be used as a disease detection mechanism. Since, the late 1970's, computer based image processing technology applied in the agricultural engineering research become a engineering research became a common In this study we propose and experimentally validate the significance of using clustering techniques and neural networks in automatic detection of leaf diseases. The proposed approach is image-processing-based and is composed of four main phases; in the first phase we create a color transformation structure for the RGB leaf image and then, we apply dev ice-independent color space transformation for the color transformation structure. Next, in the second phase, the images at hand are segmented using the K-Means clustering technique In the third phase, we calculate the texture features for the segmented infected objects. Finally, in the fourth phase the extracted features are passed through a pre-trained neural network. As a testbed we use a set of leaf images taken from Al-Ghor area in Jordan. We test our program on five diseases which effect on the plants; they are: Early scorch, Cottony mold, Ashen mold, late scorch and tiny whiteness. Using the proposed framework, we could successfully detect and classify the examined diseases with a precision of around 93% in average. The minimum precision value was 80%.

Present experimental results indicate that the proposed approach can significantly support accurate and automatic detection of leaf diseases.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

A Plant leaf that is infected with the bacterial canker leaf disease

LITERATURE SURVEY

- Lili Wang,Xinda Liu,Weiqing Min Plant disease recognition:A large scale benchmark dataset and a visual region and loss reweighting approach. In this paper we systematically investigate the problem of visual plant disease recognition for plant disease diagnosis,IEEE,2021
- Lili Li, Shujuan Zhang and Bin Wang Plant disease detection and classification by deep learning, The technology used in this paper is Deep learning,IEEE,2021
- Zulkifli Bin Husin,Ali Weon Bin Md Shakaff, Abdul Hallis Bin Abdul Aziz, And Rohani Binti S Mohamed Farrok, Plant chili disease detection using the RGB colour model, This paper describes an image processing technique that identifies the visual symptoms of chili plants,IEEE,2019
- S.W.Zhang,Y.J.Shang and L.Wang, Plant disease recognition based on plant leaf image, This technology method can recognise and classify the plant disease with high recognition rate,IJARCCE,2017
- Arpita patel, Mrs. Barkhs joshi, A survey on the plant leaf disease detection techniques, There is the need of some expert to identify plant disease but manual identification in time consuming ,IJARCCE,2017

PROPOSED APPROACH

Next we describe our proposal with more details. We also compare it to the other solutions already presented in literature. We select from literature the works that have addressed the topic of plant disease detection and classification.We propose an imageprocessing-based solution for the automatic leaf diseases detection and classification. We test our solution on five diseases which effect on the plants. Those diseases are: (1) Early scorch, (2) Cottony mold, (3) ashen mold, (4) late scorch and (5) tiny whiteness.

The concept of automatic plants leaves-disease detection presented in the following sectioris was developed on the plant leaves images acquired from Al-Ghor area in Jordan. Detection and recognition of leaves diseases are likely to give better performance and can provide clues to treat the diseases in its early stages. Visual interpretation of plant diseases manually is both inefficient and difficult; also, it requires the expertise of trained botanist. A closer inspection of the plant diseases images reveals several difficulties for the possible leaves diseases detection.

The authors have worked on the development of methods for the automatic classification of leaf diseases based on high resolution multispectral and stereo images. Leaves of sugar beet are used for evaluating their approach. Suger beet leaves may get infected by several diseases, such as rusts, powdery mildew. The developed system classifies the leafs at hand into infected and notinfected classes. Compared to the work our systems can:

- Identify disease type in addition to disease detection
- Deal with more diseases
- Be directly expanded to cover even more diseases
- Detect diseases that infect plant leaves and stems. Our proposal can identify and classify diseases that infect the stem part of plants as well. Figure I shows an example of such infection cases

A fast and accurate new method is developed based on computer image processing for grading of plant diseases. For that, leaf region was segmented by usirg Otsu method. After that the disease spot regioris were segmented by using Sobel operator to detect the disease spot edges. Finally, plant diseases are graded by calculating the quotient of disease spot and leaf areas. Our proposal is different as it aims at classifying diseased leafs based on disease type.The proposed approach starts first by creating deviceindependent color space transformation structure. Thus, we create the color transformation structure that defines the color space conversion.

The next step in our proposal is that we apply device-independent color space transformation, which converts the color values in the image to the color space specified in the color transformation structure. The color transformation structure specifies various parameters of the transformation. A device dependent color space is the one where the resultant color depends on the equipment used to produce it. For example the color produced using pixel with a given RGB values will be altered as the brightness and contrast on the display device used. Thus the RGB system is a color space that is dependent. To improve the precision of the disease detection and classification process, a device independent color space is required. In a device independent color space, the coordinates used to specify the color will produce the same color regardless of the device used to draw it.

The proposed approach starts first by creating device-independent color space transformation structure. Thus, we create the color transformation structure that defines the color space conversion. Then, we apply device-independent color space transformation, which converts the color values in the image to the color space specified in the color transformation structure. The color transformation structure specifies various parameters of the transformation. Finally, K-means clustering is used to partition the leaf image into four clusters in which one or more clusters contain the disease in case when the leaf is infected by more than one disease. K-means uses squared Euclidean distances.

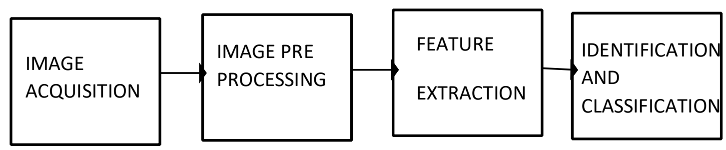
CLUSTERING METHOD

K-means clustering is used to partition the leaf image into four clusters in which one or more clusters contain the disease in case when the leaf is infected by more than one disease. K means clustering algorithm was developed by Macqueen and then by Hartigan and Wong. The k-means clustering algorithms tries to classify objects (pixels in our case) based on a set of features into K number of classes. The classification is done by minimizing the sum of squares of distances between the objects and the corresponding cluster or class centroid .In present experiments, the K-means clustering is set to use squared Euclidean distances.

FEATURE EXTRACTION

The method followed for extracting the feature set is called the color co-occurrence method or ANN method in short. It is a method, in which both the color and texture of an image are taken into account, to arrive at unique features, which represent that image. Next we explain this method in more detailed. Feature extraction is **a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing**. A characteristic of these large data sets is a large number of variables that require a lot of computing resources to process.

BLOCK DIAGRAM



MODULE DISCRIPTION

INPUT – RGB IMAGE:

RGB color model is an additive color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors. The name of the model comes from the initials of the three additive primary colors, red, green, and blue. The main purpose of the RGB color model is for the sensing, representation, and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Before the electronic age, the RGB color model already had a solid theory behind it, based in human perception of colors.

IMAGE PRE-PROCESSING AND SEGMENTATION

The pre-processing involved the procedures to prepare the images for subsequent analysis. The affected lea images were converted from RGB color format t gray scale images. Segmentation refers to the process of clustering the pixels with certain properties into salient regions and these regions correspond to different faces, things or natural parts of the things. We proposed k-means segmentation technique to fragment goal areas. Target regions are those areas in the image that represented visual symptoms of a fungal disease.

FEATURE EXTRACTION:

As the paddy leaf disease consists of several types of disease blast, brown spot and narrow brown spot that had different lesion shape and lesion color.

SHAPE FEATURE EXTRACTION:

Shape is one of the important parameter of the image. Breadth and length of the image are significant characteristic to describe

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

the shape. A simple approach is to measure the breadth and height of the image is to measure the count of the object pixel.

COLOR FEATURE EXTRACTION:

Color plays an important role in image processing. Digital image processing produce quantitative color measurement that are useful for the work of inquiring the lesion for early diagnosis. The pixel in the color images are commonly represented in RGB format, where RGB are RED GREEN BLUE values respectively from the color images capturing device.

K-MEANS SEGMENTATION: k-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.

K MEANS CLUSTERING ALGORITHM

Clustering is a method to divide a set of data into a specific number of groups. It's one of the popular method is k-means clustering. In kmeans clustering, it partitions a collection of data into a k number group of data. It classifies a given set of data into k number of disjoint cluster. K-means algorithm consists of two separate phases.

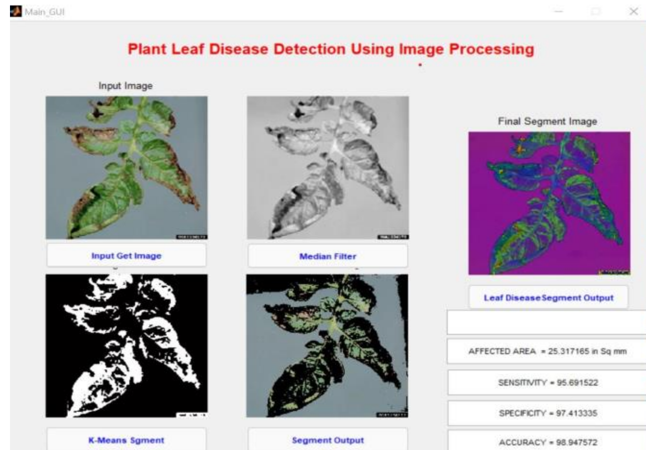
ANN CLASSIFICATION

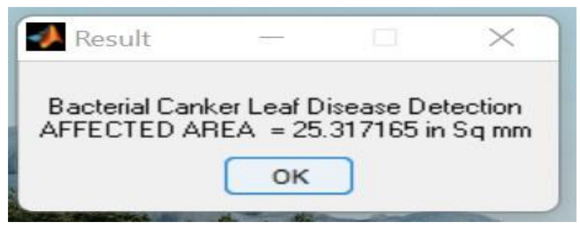
At present ANN is popular classification tool used for pattern recognition and other classification purposes. Artificial neural networks (ANN) are a group of supervised learning methods that can be applied to classification or regression. The normal ANN classifier takes the set of involvement data and calculates to classify them in one of the only two separate classes. ANN classifier is trained by a given set of training data and a model is willing to classify test data established upon this model. Most habitual classification models are established on the empirical risk minimization principle.

ACCURACY:

In pattern recognition and information retrieval with binary classification, precision (also called positive predictive value) is the fraction of retrieved instances that are relevant, while recall (also known as sensitivity) is the fraction of relevant instances that are retrieved.

Percentage = (Affected area / Total area)*100





CONCLUSION

In this project an image-processing-based approach is proposed and used for leaf disease detection. We test our program on five diseases which effect on the plants; they are: Early scorch, Cottony mold, ashen mold, late scorch, tiny whiteness. The proposed approach is image-processing-based and is highly based on K-Means clustering technique and Artificial Neural Network (ANN). The approach is composed of four main phases; after the preprocessing phase, the images at hand are segmented using the Kmeans technique, then some texture features are extracted in which they are passed through a pre-trained neural network. As a testbed we use a set of leaf images taken from the Ghor area in Jordan.

Present experimental results indicate that the proposed approach is a valuable approach and can significantly support accurate and automatic detection of leaf diseases. Based on our experiments, the developed Neural Network classifier that is based on statistical classification perform well and can successfully detect and classify the tested diseases with a precision of around 93%

FUTURE WORK

For future research,They have been some directions,such as: Developing better segmentation techniques; selecting better feature extraction using deep learning in plant disease recognition can avoid the disadvantages caused by artificial intelligence selection of disease spot features.

REFERENCE

- Lili Wang,Xinda Liu,Weiqing Min “Plant disease recognition:A large scale benchmark dataset and a visual region and loss reweighting approach”,IEEE Journal, 2021.
- Lili Li, Shujuan Zhang and Bin Wang “Plant disease detection and classification by deep learning- A review” IEEE Journal, 2021.
- K. Chen, Y. Chen, C. Han, N. Sang, and C. Gao, "Hard sample mining makes person re-identification more efficient and accurate," Neurocomputing, vol. 382, pp. 259-267, Mar. 2020.
- S. Jiang, W. Min, L. Liu, and Z. Luo, "Multi-scale multi-view deep feature aggregation for food recognition," IEEE Trans. Image Process ., vol. 29, pp. 265-276, 2020.
- S. Li, J. Li, H. Tang, R. Qian, and W. Lin, "ATRW; A benchmark for Amur tiger re-identification in the wild,"inProc.28thACMInt.Conf. Multimedia, Oct. 2020, pp. 2590-2598.
- Y. Peng, X. He, and J. Zhao, "Object-part attention model for finegrained image classification," IEEE Trans. Image Process ., vol. 27, no. 3, pp. 1487-1500, Mar. 2018.
- Z. Yang, T. Luo, D. Wang, Z. Hu, J. Gao, and L. Wang, "Learning to navigate for fine-grained classification," in Proc. Eur. Conf. Comput. Vis ., Sep. 2018, pp. 438-454.
- W. Min, S. Jiang, L. Liu, Y. Rui, and R. Jain, "A survey on food computing," ACM Comput. Surv., vol. 52, no. 5, pp. 92:1— 92:36, 2019.
- E. Moen, D. Bannon, T. Kudo, W. Graf, M. Covert, and D. Van Valen, "Deep learning for cellular image analysis," Nature Methods, vol. 16, no. 12, pp. 1233-1246, 2019.
- T. Hu and H. Qi, "See better before looking closer: Weakly supervised data augmentation network for fine-grained visual classification," CoRR, vol. abs/1901.09891, 2019.
- Y. Chen, Y. Bai, W. Zhang, and T. Mei, "Destruction and construction learning for fine-grained image recognition," in Proc. IEEE/CVF Conf. Comput. Vis, Pattern Recognit. (CVPR), Jun. 2019, pp. 5157-5166.
- Zulkifli Bin Husin,Ali Weon Bin Md Shakaff, Abdul Hallis Bin Abdul Aziz, And Rohani Binti S Mohamed Farrok “Plant chili disease detection using the RGB colour model”, IEEE Journal, 2019.
- C. Liu et al ., "Progressive neural architecture search," in Proc. Eur. Conf. Comput.Vis .,Sep. 2018,pp. 19-35.
- S.W.Zhang,Y.J.Shang and L.Wang “Plant disease recognition based on plant leaf image” ,IJARCCE Journal,2017.
- Arpita patel, Mrs. Barkhs joshi “A survey on the plant leaf disease detection techniques”, ISSN Journal, 2017.