Identification and Classification of Tomato Leaf Diseases Using Machine Learning Techniques

D. Femi
Department of Computer Science and Engineering, Vel Tech Rangarajan Dr Sagunthala R & D Institute of Science and Technology, Avadi, Chennai, India.

R. Murugasami
Department of Electronics and Communication Engineering, Nandha Engineering College, Erode, India.

N. Manikandaprabu
Department of Electronics and Communication Engineering, Nandha Engineering College, Erode, India.

J. Raja Paulsingh
Department of Electronics and Communication Engineering, Sri Nallalaghu Nadar Polytechnic College, Puzhal, Chennai, India.

P. Vanaja
Department of Pharmaceutical Chemistry, Nandha College of Pharmacy, Erode, India.

Received June 02, 2021; Accepted September 04, 2021
ISSN: 1735-188X
DOI: 10.14704/WEB/V18SI05/WEB18297

Abstract

Tomato is cultivated in all countries of the world in fields, glasshouses etc. China, India, USA, Turkey, Egypt, Iran, Italy, Spain and Brazil are the important countries which are cultivating tomatoes. It is most commonly and widely cultivated in India. India is one of the countries in harvesting tomatoes. Tomato is a vital vegetable yield with respect to both income and food. Tomatoes are for the most part summer crops, yet it tends to improve steadily. Naturally, it contains A and C of vitamins which also acts as an antioxidant to prevent cancerous cells. Since the organic product contains novel features, the demand remains the same. A significant and unique feature with high nutrients gains the importance in tomatoes cultivation. Challenges towards the cultivation of tomato made us to plan for an automated machine to detect infection and to increase the productivity. This system automatically detects the infected parts and classify the types of disease which occur on the leaf like early blight, bacterial wilt, Leaf Spot, tomato mosaic virus, septoria leaf spot, leaf curl virus, and tomato spotted wilt disease using gradient anisotropic diffusion filter for pre-processing and then features are extracted using GLCM from the pre-processed leaf
image and the system is trained with the feature values using extreme learning machine which then identifies and diagnoses the type of disease.

**Keywords**
Gradient Anisotropic Filter, GLCM, ELM.

**Introduction**

Agriculture has an important impact on the economy of the world in which our country is no different. India is a country with Agriculture as the major occupation. More than half the population of India is engaged in primary activities like agriculture. In India, agriculture is the main occupation 66.7 percent of population is engaged in agriculture.

In vegetable crop, cultivation of tomato is huge because of its protective nature which transforms into major production. In fleshy fruits division tomatoes are marked their presence and proves their major production. Commercialization of tomatoes cultivation is huge. Lycopersicon esculentum belongs to Lycopersicae family termed the botanical name of tomato. Cultivation and demand increase day by day for the unique nutrient features of tomato. In profitable agribusiness criteria tomatoes show their mark. Vitamin A and C with antioxidants are available in tomato whether it is eaten raw or cooked.

10-25°C is the perfect temperature range for cultivating tomatoes. 10-15°C is the ideal sowing temperature with 400-600 mm rain. 21-24°C is the perfect temperature for cultivating best red coloured tomatoes.

Tomatoes nurtures in variety of soils but grown healthy in drained soils. Medium black, sandy loam, red soils are considered suitable for the cultivation of tomato. 7-8.5 pH value is used for good harvesting (Kumar, V., 2013) (Kumar, T.S., 2013).

Image processing has been demonstrated to be effective means for analysis in various fields and applications (Sadhasivam, J., 2020). In agriculture field image processing plays a vital role in identification and classifying the plant diseases at earlier stage, quality of grains and fruits, yield of product, moisture of the soil monitoring etc from the farmers' point of view. Image processing along with accessibility of communication network farmers can get updates automatically about their fields. Image processing is used along with machine learning techniques which is applied for prediction and classification and for various analysis.
Advancement of technology in the field of agriculture which is used to increase the yield of production. In our work we have taken tomato plant to identify and classify leaf diseases at earlier stage. With the development in technology tomato plant disease detection have become easier and more precise. Anisotropic diffusion method is applied to captured image for removing noise and for applying smoothing to the captured image. Feature values are extracted from the pre-processed image using GLCM four features are obtained like Correlation, contrast, energy, Homogeneity. The obtained values are trained using extreme learning machine which is used for learning and then it identifies and classify the diseases at earlier stage.

**Literature Review**

In (Cruz, A., 2019) author proposed a methodology to recognize and classify diseases which occurs in grapes i.e grapevine yellow in grapes disease. The proposed technique infers Otsu Threshold for segmenting and convolution neural network is used for extracting feature values. After extracting feature values, it is implemented using six neural networks. The system obtained good results in terms of sensitivity and specificity using deep neural networks.

In (Pantazi, X.E., 2019), author proposed a segmentation technique used system to detect crop diseases. Black rot downey mildew, powdery mildew is some of the disease identified by the system. The grab cut algorithm is used with accuracy of 95%.

In (Zhang, S., 2019), neural network-based system has been proposed by the author to extract the diseases like target spot, late blight, mosaic virus, early blight etc. 89.29% is the overall accuracy obtained in this system using three channel convolutions neural network.

In (Deepak, A.H., 2019), Deep learning techniques with IOT technology were utilized by the authors for their proposed system to monitor the plants health. and used deep learning and IoT technologies for monitoring the overall health of the plant. Disease in the plant of tomato has been detected with the help of convolutional neural networks. Temperature, Humidity, Moisture sensors used to provide information to the farmers and easy to monitor remotely.

In (Moghadam, P., 2017), Spotted Wilt Virus is the disease is concerned by the author and detected using machine learning and hyperspectral imaging the dimensionality feature increases as that of the performance.
In (Lalli, G., 2014), Author proposed the identification of leaf disease in tomato plant using the proposed LeNet which is a slight variation of CNN. Normal computer resources are enough for obtaining the results. Neural Networks are used to extract the features. 94-95 percent of accuracy is achieved.

**Diseases in Tomato Leaves**

**Bacterial Spot:** Dark brown water-soaked lesion appeared on the leaves; later these lesions become blackish, and ultimately the affected tissue drops out leaving a hole in the leaf.

**Early Blight:** Leaf lesions of early blight are biguneven patches of black, necrotic tissue surrounded by bigger yellow parts.

**Late Blight:** Lesions on leaves appear as large water-soaked areas, that finally turn brown and papery. Green to black uneven lesions are also present on the stems.

**Septoria Leaf Spot:** Round water-soaked lesions occur first on older leaves. These lesions eventually turn brown with graycentres and die, and if infection is severe abundant, the entire leaf will die.

Fig 1 shows the different diseases in tomato leaves.

![Fig. 1 Sample Images of Tomato Leaf Diseases](image-url)
Proposed Work

In our work anisotropic diffusion method is applied to captured image for removing noise and for applying smoothing to the captured image. Feature values are extracted from the pre-processed image using GLCM four features are obtained like colour, contrast, entropy, energy (Lalli, G., 2014). The obtained values are trained using extreme learning machine which is used for learning and then it identifies and classify the diseases at earlier stage. Fig 2 shows the proposed method.

![System Architecture](image)

**Fig. 2 System Architecture**

Dataset

The proposed classical method is trained and tested by means of Plant Village dataset from Kaggle. The dataset is composed of tomato leaves of both healthy and unhealthy. The dataset is separated into two classes one is for training and other one is for testing. The training image includes 80% of total images where the testing image contains 20% images of the total.

Image Pre-Processing

The captured image is made up of RGB with different size. The image resized to 64 X64 pixels, and converted to Gray scale and anisotropic diffusion filter is applied to remove noise in the image without changes in any information from the image. it safeguards the edges of the image. Lastly, a denoised image is got as an output after applying the filter. Fig 3 shows an image after pre-processing.
Feature Extraction

Extraction of feature is one of the important modules in image processing techniques. Feature extraction is the method of transforming raw data into numerical features that can be used in recognition and classification of images. Here we have proposed Gray Level Cooccurrence Matrix for extracting feature values which is based on texture classification it is a histogram of co-occurring grayscale values at a given offset over an image. These features like Correlation, contrast, energy, Homogeneity provide information about the texture of an image. Fig 4 shows the extracted feature values for an image.

```plaintext
Fig. 3 Image After Pre-processing

Fig. 4 Feature Values
```
Classification

The extracted features values are collected from GLCM and it is made as a dataset. Dataset is split into two namely training and test dataset.

Extreme learning machine (ELM) is a novel learning method for a single hidden layer feedforward neural network. Compared with traditional neural network it overcomes the slow training speed and over-fitting problems. Extreme Learning machine is used for classification and identifying the type of disease occurred in tomato leaf at earlier stage when lesions start to appear on the leaf.

Conclusion

In our proposed work for classifying and identifying tomato leaf diseases like early blight, septoria leaf spot, bacterial wilt, bacterial leaf spot, tomato mosaic virus, tomato leaf curl virus, and tomato spotted wilt using ELM by using feature values obtained by GLCM. This technique can also be used for identifying diseases which is occurring in stem, root, fruits and also severity of infection can also be analysed. This technique can also be applied for other plants to diagnose the disease at earlier stage as soon as the lesion starts on the plant.

References


