A Novel Approach for Agriculturists on Mobile Platform for Soybean Leaf Disease Detection and Classification Using Proximal Support Vector Machine

Rajhashyamala.M.M
Research Scholar
Dept. Of computer science
Sri Ramakrishna college of arts and science for women
Coimbatore.
shyamli.mohon@gmail.com

Dr.N.Tajunisha
Associate Professor
Dept. Of computer science
Sri Ramakrishna CAS for Women
Coimbatore.
tajkani@gmail.com

Abstract

Agriculture production increases the Indian Economy and vice-versa is also true. An expert system in the agricultural field can be a better option to increase agriculture production. Today, use of smart phones is very common by everyone, including the farmers. Introduction of Information and Communication Technologies (ICT) has seen a keen role in daily life of farmers. For effective diseases identification, an existing system used histogram equalization and Support vector machine for pre-processing and classification respectively. However Support vector machine is does not possible for large data set and the existing system is not capable of removing the noise from the diseased leaf images efficiently. To overcome this problem, the proposed system uses proximal support vector machine (PSVM) for diseases classification and median filter algorithm for removing the noise effectively. In proposed system, the captured leaf image is first pre-processed and then it will be converted to binary values. Finally it is passed to centralized server, in which classification process is achieved by using PSVM. Finally the classification result is sent back to mobile devices. The experiment is done in Soybean data set. The proposed system achieves high performance compared with the existing system.

1. Introduction

Crop diseases are as old as agriculture itself and diseases affect many parts of the plant like leaf, stem, root and fruits. Among these the easily identifiable is the diseases affecting the leaves of the crop. The observation and identification of crop diseases has been done by naked eye as traditional method. But this way of detecting diseases needs constant monitoring by experts which is too costly and time consuming. In addition to this in developing countries the farmers may have to go long distances to get in touch with experts and are also unconscious of non-native diseases [1].

Automated plant leaf detection is an important research topic which can reduce the difficulties in monitoring large fields of crops. This approach can automatically identify the diseases from the symptoms on the plant leaves. Relatively visual identification is labour intensive and less precise.

In some situations the symptoms will not be obvious or it will be more complex, which will be very difficult to distinguish the characteristics. This will delay the control of crop diseases seriously. To overcome this the images of the diseased leaves are analysed by the image processing technology and the features of the disease spot such as colour, texture and other characteristics are extracted from a quantitative point of view. The cause and extend of the disease can be analysed timely and effectively which will prevent and control the entire crop damage.
The rest of the paper follows with brief information about related work. The details of the proposed work and result and discussion are presented and the outcome of the proposed technique and conclusion are discussed.

2. Related work

The proposed work describes about the detection of leaf diseases in the soybean plant. Various approaches have been deployed in detecting leaf diseases in several plant species. P.Revathi et al., [2] has proposed a system for cotton leaf spot diseases which uses enhanced feature selection methods and the selected features were input to SVM, back propagation neural network, Fuzzy with edge CYMK colour feature and GA feature selection. Jiangshen Gui et al., [3] has proposed a method for soybean leaf disease detection based on modified salient regions in which K-means algorithm was used. Shitala Prasad et al., [4] have developed an application called AgroMobile which uses Curvelet Transform and Support Vector Machine (SVM) for crop image analysis. Jiandu Liu et al., [5] proposes a novel method of plant classification from leaf image set based on wavelet transforms and support vector machines (SVMS). N.Valliammal et al., [6] proposes a system which focuses on the pre-processing technique for CAP-LR (Computer aided plant classification through leaf recognition).

Thus many works has been developed for plant disease identification and classification using various techniques. However the classification algorithms have some limitations. And also the algorithms require more classification time and are less accurate. To overcome all these difficulties the proposed work is designed using Proximal Support Vector Machine to improve the overall performance and to provide better results.

3. Proposed work

The proposed framework helps farmers to assist them in detecting and managing the leaf diseases affecting the soybean plant. The leaf image is sent from the smart phone using the android application. Pre-processing is applied using the median filter algorithm to remove the unwanted noise from the images. After the completion of pre-processing the values are given to the server. The server receives these values and gives it to the classifier for classification process. In the proposed system, Proximal-Support Vector Machine (PSVM) algorithm is used for classification. The final results are sent back to the smart phone of the farmer. The results include the name of the diseases and the necessary management steps to be taken as a cure for the disease. The experimental results show that the proposed method has high accuracy with the existing system. The time for classifying the diseases is also less compared to the existing method.

The algorithm is described as follows

**Step 1:** Image of the diseased leaf is sent from the mobile device.
**Step 2:** The leaf image is preprocessed using median filter algorithm.
**Step 3:** The preprocessed image is converted into binary values before sending it to the centralised server.
**Step 4:** The centralized sever receives the binary file and gives it to the Proximal Support Vector Machine (PSVM) classifier.
**Step 5:** The PSVM classifier classifies the disease and send the result back to the mobile device.
**Step 6:** Finally performance comparison between the existing and the proposed system is done.

The flow of the proposed work is represented in the following flow chart.

**Figure 1. Flow chart for proposed system**

In this paper, a method is proposed to detect the leaf disease of the soybean plant using Proximal Support Vector machine for classification and Median filter technique for image preprocessing.
3.1. Sending Image from Mobile Device

The first step is to send the image of the diseased leaf from the mobile device to the centralised server using an android application. The mobile device used is a smart phone with android operating system.

![Image Transfer](image.png)

**Figure 2. Sending image to server using proposed system**

The soybean diseased leaf image is loaded in the android application. Then the mobile device is connected to the server through this android app with the help of Wi-Fi connectivity. Finally the image gets transferred to the server for classification.

3.1.2. Image Preprocessing

Mobile devices act as a client. It sends the leaf images of infected leaves and then preprocessing is applied on them. The median filter per-processing algorithm is used for noise removal from the images. The basic idea behind denoising is the estimation of the uncorrupted image from the distorted or noisy image. There are various methods to repair an image from noisy distortions. To obtain the wanted image, the selection of the correct method plays a very important role. Most of the denoising methods tend to be problem specific.

In our proposed algorithm, noise removal or noise reduction can be done on a given input image by median filtering. In this method the magnitude of all of the vectors within a mask is considered, based on the vectors the magnitudes are sorted. The pixel with the median magnitude is then used to replace the pixel studied. The Simple Median Filter performs better when compared to the Mean filter. It relies on median of the data instead of the mean that is the average of all the image intensities. Finally the pixel with median magnitude is then replaced. The median of a set is more robust with respect to the presence of noise.

\[
\text{Median filter} \ (x_1, \ldots, x_N) = \text{Median} (||x_1||^2, \ldots, ||x_N||^2)
\]

where \(x_1, \ldots, x_N\) are image pixel range.

After completion of image pre-processing it will be converted to byte array values and passed to centralized server.

3.1.3. Conversion of Image into Binary values

The image sent from the mobile device is pre-processed using median filter technique. The pre-processed image must be converted into binary values before giving it to the classifier. This process is done with the help of ByteArray () function.

3.1.4. Classification

The classification process is performed in the centralized server which analyzes the leaf, which is affected by disease. In our proposed system, Proximal Support Vector Machine (PSVM) classifier is used. The PSVM classifier classifies the binary patterns by assigning them to the closer of two parallel planes that are pushed apart as far as possible. The training time for the classifier is found to be faster compared to their previous versions of Support Vector Machines.

On assumption that there are \(N\) training samples, such as \((x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\) among them, the target function of Proximal Support Vector Machine can be denoted by

\[
\text{Min} \ \frac{C}{2} ||y||^2 + \frac{1}{2} (w^T w + r^2) \\
\text{Subject to: } L(Aw - e\mathbf{r}) + y = e
\]

\(C\) is represented as castigation factor, \(y\) denote the sample output, \(w\) figure the normal vector of the classification hyperplane, \(e\) is denoted as units vector, \(r\) is the parameter which can ascertain the position of two dividing-line plane relating to the origin in Proximal Support Vector Machine; A represent the \(n\times m\) dimensional training data set, each sample is corresponding to a list \(A_i\).

**Step 1:** Initialize data points

**Step 2:** Separate hyperplane

**Step 3:** Data points are assigned according to the proximity

**Step 4:** The target function of Proximal Support Vector Machine is denoted by
Min $\frac{c}{2} ||y||^2 + \frac{1}{2} (w^T w + r^2)$

// C is castigation factor, y express sample output, w figure the normal vector of the classification hyperplane, e is units vector, g is parameter.

**Step 5:** The cost function is given as follows:

$$\min_{\omega,b,\xi} f(\omega, b, \xi) = \frac{1}{2} \|[\omega, b]\|^2 + \frac{c}{2} \sum_{i=1}^{m} ||\xi_i||^2$$

**Step 5:** Sample result

### 3.1.5. Sending Result back to the Mobile Device

The classification results are sent back to mobile for diseases identification. The result is displayed in the smart phone with the help of the android app. The name of the disease is displayed along with the steps for managing the disease.

![Figure 3. Receiving results of disease using proposed system](image)

### 3.1.6. Performance Comparison between Existing and Proposed System

The final step is to compare the performance of the existing and the proposed system. The proposed system is more efficient than the existing system in terms of classification accuracy and classification time.

The proposed system has many advantages which includes high accuracy when compared with the existing system. It removes the noise efficiently and the rate of identifying the disease is high.

Thus the system architecture of the proposed algorithm is explained along with the classification of soybean leaf disease dataset by the existing and the proposed algorithms. The pseudo code and disadvantages of the existing algorithms are also presented. Finally, the proposed algorithm flowchart, pseudo code and advantages are discussed.

### 4. Results and discussion

#### 4.1. Dataset Used For Experimental Study

The soybean dataset have been collected from UCI Repository database. There are totally 307 instances and 35 attributes in this dataset. In this analysis the attributes related to the leaf characteristics are considered for classification. Some the attributes include leaves, leafspot-halo, leafspots-marg, leafspot-size and so on.

#### 4.1.2 Performance Measures

In the simulation the existing Support vector machine classification based leaf diseases detection method and the proposed Proximal Support vector machine classification based leaf diseases detection method is evaluated in terms of Accuracy and time.

**Accuracy:** The term accuracy refers to the correctly classified instances by the total number of instances present in the dataset. The accuracy of the system is calculated with the values of the True Negative, True Positive, False Positive, False negative actual class and predicted class outcome it is defined as follows,

![Figure 4. Accuracy comparison of existing and proposed system](image)
support vector machine is better than existing one. From this graph, we can say that the accuracy of proposed system approach is increased, which will be the best one.

**Time:** Time is another measure by which the performance of both the existing and proposed systems has been compared. The following is a graph which represents the time comparison between the two systems.

From the Figure 5 it can be proved that the proposed methodology provides better result than the existing approach by reducing classification time. In this graph, the proposed and existing methods are plotted in the x axis and the execution time is plotted in the y axis.

![Figure 5. Time comparison](image)

Figure 5. Time comparison

The classification time is high while using the method of support vector machine classification. The classification is decreased significantly by using the proposed proximal support vector machine method. From the experimental result we can conclude that proposed method is superior and efficient to existing system.

4.1.3. Classification of Soybean Dataset

Consecutively, in this thesis a new classification algorithm is proposed to classify the soybean dataset. We compare the proposed algorithm with the existing algorithm which is the Support Vector Machine classifier.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Accuracy (%)</th>
<th>Time(ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support Vector Machine(Existing)</td>
<td>89.11</td>
<td>2812.01</td>
</tr>
<tr>
<td>Proximal Support Vector Machine(Proposed)</td>
<td>92.08</td>
<td>2398.62</td>
</tr>
</tbody>
</table>

It has been found that the proposed algorithm which is the Proximal Support Vector Machine classifier is better than the existing algorithm for classifying the soybean dataset. The above table represent the comparison of performance measures between both the existing and proposed system.

The results and discussion of the existing and proposed algorithms have been discussed. The performance measures such as accuracy and time are used to compare the efficiency of these algorithms. Finally the comparison of existing and proposed algorithm is presented. Thus from the experimental results, it has been found that the proposed algorithm performs better than the existing method.

5. Conclusion and future enhancement

The proposed new framework is used for farmers especially the Indian farmers to assist them in agricultural needs. We presented a new way of interaction for soybean leaf disease detection using a simple smart phone. The proposed system is aimed to reduce the workload on the farmer’s devices by running the classification on the server instead of the smart phone itself. The proposed system helps farmers in crop disease analysis and understanding it more clearly. It provides the information about the diseases and the necessary measures on the figure tips at any time anywhere to the farmer having an android smart phone.

As a future work, we can improve the proposed system architecture by introducing many more algorithms for crop analysis. Another future work is to increase the dataset for plant species and leaf diseased image. In future, the system is aimed to utilize all the advantages of cloud computing technologies along with the mobile computing techniques to develop a complete real-time system using Mobile Cloud Computing (MCC).

6. References


[12] Jon Atli Benediktsson, Martino Pesaresi, Kolbeinn Arnason ,"Classification and feature extraction for remote sensing images from urban areas based on morphological transformations “, IEEE transactions on geoscience and remote sensing.


[16] Sean M. White, Dominic M. Marino” LeafView: A User Interface for Automated Botanical Species Identification and Data Collection” ACM UIST 2006 Conference Companion, Montreux, Switzerland, October 15-18.


