CROP DISEASE DETECTION AND CLASSIFICATION BASED ON HYBRID INFORMATION APPROACH

OTKRIVANJE I KLASIFIKACIJA BOLESTI USJEVA NA TEMELJU INFORMACIJSKOG HIBRIDNOG PRISTUPA

S.Vijayalakshmi, D.Murugan
Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Tirunelveli, India

Abstract
The objective of this paper is to identify the diseases in the leaves of all plants. Plant disease diagnosis helps to improve both the quality and quantity of crop productivity. In existing, to detect the diseases they used the spectroscopic techniques. These techniques are very expensive and can only be utilized by trained persons only. This work proposes an approach for the detection of leaf diseases based on the characterization of texture, shape and color properties. The detection of diseases which are detected using ISRC (improved sparse Representation Classifier) technique. First the GENABC clustering approach is applied to the input image to segment the affected area. Then extract the features from the affected area by using feature extraction techniques. In this paper Improved Transform Encoded Local Pattern used to extract the texture feature, Enhanced Gradient Feature (EGF) to extract the shape and Improved Color Histogram Techniques (ICH) are used to extract the color. And then these features are given to the ISRC classifier to get the exact type of disease on affected leaves. To analyze the performance of the proposed method we use four metrics. They are classification accuracy, error rate, precision value and recall value. From the analysis of experimental results ISRC method provides best result than the existing approach.

1. Introduction
Agriculture plays a vital role in the life of an economy. It is the backbone of our economic system. The diseases on leaf can reduce both the quality and quantity of crops and their future growth. Against the background of parasitic diseases, plant diseases, the physiological conditions of changing dirty tissue, for example, photosynthesis changes, neurons, dental contacts, salititsi corrosive and even cellular approval. The absence of sheets can be marked especially from optical sensors today. Use fluorescent light reflected images of mushrooms, cucumber leaves, cucumber and yogurt from isolated wheat. However, infrared can be slightly improved with the region and
evaluate for setting larger areas in control of infection control and their associated diseases as they are susceptible to physiological disorders associated with parasite attacks and diseases. Infrared computer is a useful tool for the effects of disease and asthma on plants.

Spread of plants is the sight of disease in the number of plants. Just as the disease of humans and other animals, plant diseases are caused by parasitic diseases such as bacteria, viruses, mushrooms, oomycetes, neoproto fitoplazmi, protozoa and plant parasites. Scientists try to find out about the outcomes of illness and disease and improve interventions in situations where crop failure may occur. Usually, successful interventions will lead to a degree of acceptable disease based on cultural values. The epidemic of the disease is often examined at plants, with a multi-stage process requiring organic, statistical, geographical and environmental views. Biology is very important to knowledge of the disease, its disease and its lifestyle. It is also important to notify the body's structure of the culture and how it affects it. Agricultural practice always affects the spread of higher or higher diseases. Ecological impact is high. A variety of local flowers can add reservoirs for diseases that are intended to disrupt the vegetation. Statistical models are often performed with a purpose to summarize and describe the complexity of the epidemiology of the disease to understand the process of further illness. For example, the comparison between the style of disease development for special conditions, varieties, management strategies or environmental determinants can help in determining how plant diseases can be managed and of high quality. Coverage can influence the spread of disease through a restricted movement on imports from abuse resources.

2. Related Work

Santanu Phadik ar and Jaya Sil /1/ focused on identifiable infectious evidence and assumed both diseases were specific to leaf blight and brown spots. Restricted discovery and on-site identities are used to include the removal of contaminated parts of plants. The author advised the SOM (Self Organizing Map) system when calculating the scales for preparing unhealthy rice leaves. The technology of creating SOM information vectors is zero hugging and loss of concentration. An approximation calculation provides a reasonable result. Dheeb Al Bashish et.a /2/ suggested five plant infections, especially the end of the hair shaft, the original burning, the form of Ashen is the weakest and lowest of the Jordanian-Ghor region for testing. Using a strategy for grouping the K-Stars, separating the pictures, methods of technology, and delegations (Color co-method methods) are used to examine the surface of the leaves. For the grouping of plant infections, changes in the nervous system are used.

Zulkifli Bin Husin et.al /3/ recommended LABVIEW software tools and Plant Pathology Recognition in MATLAB. Early evaluation of the leaves may be possible due to the combined system of two virtual products. LABVIEW is used to capture sheets, and MATLAB is used as a process of image processing. The location of all sieving operations is terminated morphologically, Fourier uses techniques for the preparation of images and uses techniques for shading groups to break the leaves. Image clarity and characteristics show the strength of the sugar plant. The author advised a technique to look for Malus Domestica /4/. The image in grayscale is obtained by the histogram equation and the texture's analysis in the part of the image is carried out using the algorithm of the matrix method of co-presence - the K-mean algorithm used. In the process of matching thresholds, the comparison between the value of individual pixels and the threshold value. For the discovery of diseased plants, the texture, texture and color are compared to previously obtained leaflets. Murali Krishnan et.al /5/ proposed a strategy for identification of Bacterial leaf singe sullying in plant. In image segmentation by K-means algorithm, right way grouping set of principles is completed for separating closer view and recorded past photographs. Segmentation depends on subtracting the bunched leaf images and power mapping for featuring leaf region. K-means method is capable and simple for discovery of excited region. In the article /6/, the author introduced the technique of detecting leaf orange leaves and the disease: Anthracnose, canker, citrus greening disease etc. Imagery includes the change of color space through the implementation of the YCbCr Color System & L * a * b * and the Color Image Improvement by implementing separate cosine changes. A gray-collar combining matrix is used to derive features to look at different statistics such as power, contrast, likeness, and entropy. Ultimately, SVMRBF and SVMPOLY are used to
look for oranges. The aim of this article to introduced a procedure to find the Orchidaceous burning aqueous leaves and leaf spot disease /7/. Acne treatment is achieved by aligning the histamine, changing the force and improving the image. This section includes three procedural and procedural procedures that have been linked to separation and rescue of small and large protests. Finally, the feature finishes with an estimate of white pixels in the image of the leaves and disease awareness. Usama Mokhtar et.al /8/ describes the procedure of tomatoes, leaving the place of sickness and disease: the accumulation of dust and the original curse. Imagination includes a variety of methods such as smoothness, dismissal, image resizing, image allocation, and laying down images. Gabor wave changes are associated with extraction for the contiguous emphasis in the stack. Cauchy Kernel, Laplacian Kernel and Kernel Invmult Kernel are linked to SVM to select the extraction and preparation for evidence of disease. According to this paper /9/ creators exhibited procedure in which pre-handling included change RGB pictures to dark utilizing the condition $f(x)=0.2989^*R+0.5870^*G+0.114^*B$ and evacuating articles and clamor in leaf images. Limit and spot discovery calculations are designed in division to discover leaf contaminated part. After that H&B segments and shading co-occurrence techniques are utilized to extricate different highlights. Double pictures are made from dark pictures by Otsu limit calculation and ailments are arranged and distinguished utilizing both counterfeit neural system and back proliferation organize alongside K-implies technique. In paper /10/ authors portrayed system to recognize Spot and Scorch illness in which by making shading change structure, shading esteems are changed over to space an incentive in picture pre-preparing like-wise connected K-implies technique for division. Leaf pointless zone is expelled by covering of green pixels and surface highlights are computed for fragmented protest additionally conceal cells are evacuated. Contaminated groups are changed over from RGB to HSI and after that SGDM lattice is produced for H and S. GLCM computations are made for extraction of highlights which are then gone through the neural system for infection acknowledgment and arrangements. Prajakta Mittal,et.al /13/ depicted strategy of Sugarcane leaf infection discovery and ailments are: Brown Spot, Downy mold, Sugarcane Mosaic, Downy Fungal, Red stripe and Red decay. Pre-handling included change of RGB picture to grayscale and undesirable parts are expelled. Sound territory and possibly contaminated region are situated by division. Straight, Non direct and Multiclass SVM are connected for illness location. Anand Singh Jalal et.al /14/ depicted procedure of Apple organic product malady discovery and ailments are: Apple Scab, Apple Rot and Apple Blotch. Picture pre-handling included RGB to L*a*b* shading space picture change. In highlight extraction Global Color Histogram, Local Binary Pattern, Completed Local Binary Pattern and Color Coherence Vector highlights are utilized for ailment identification. K-implies grouping is connected for ‘a*b*’ space in division to name every pixel and portion picture by shading. CLBP highlights and multiclass SVM are utilized for order of organic product illness. The article introduced strategy of Grapes and Apple organic product illness recognition and ailments are: Black Rot, Powdery Mildew, Rot, and Apple Scab infection. With the assistance of morphology, picture parts are extricated for limits and different visual examples are portrayed by surface element. RGB shading space is changed over to HSI shading space and ANN neural system /15/ and back proliferation calculations are utilized for ailment arrangement and natural product reviewing. Mrunmayee Dhakate and Ingole A.B portrayed strategy of Pomegranate organic product disease identification and ailments are: Bacterial Blight, Leaf Spot, Fruit Rot and Fruit Spot. In picture pre-handling clamor is evacuated utilizing (3*3), (5*5), (7*7) sifting covers.
In division, k-means grouping technique is connected for isolating picture into question and district additionally GLCM equation is arranged for factual surface highlights extraction. Ultimately creators utilized Multilayer Perceptions for preparing the neural systems utilizing backspread calculation. Ridhuna Rajan Nair, et al. introduced improved K-tools strategies. After receiving an RGB image, it will turn gray and HSV. Edge creates a polygonal boundary and histogram to include a key image at the last position to get the desired range. Centrifuging is calculated using the K-Means strategy better than comparing with the resulting database. Ashwini Awate et al. described the strategy for detecting Apple’s disease, grapes and pomegranates, infections of organic grapes, raspberry, manna, fluffy and black coal; Apple’s disease is Apple Blotch, Apple Scab and Apple Rot. The organic substance of this drug is the mass of gray and bacterial Aspergillum. Duplicate images are organized to find objects and bouncy lines of pictures and K-means bunching ways that are used to identify each pixel. In four SURF calculations, the four alterations were removed using the block and neighborhood identification, and false nervous system (ANN) was associated with the coordination and preparation of the infection. Bhavini J. et al. presents the methods of identification of diseases of the body and the body of Apple’s body: apple and apple. The K-Means technique uses the Euclidean separation to search for chaotic neighborhoods and biological products have changed into shades of RGB. Shades, shapes, and skeleton are removed, and the combination of finishing levels to convert more than double. Shadow collection is a universal painting and color histogram. Highlights are the advantages of Gabor, a double-scoring example in the region, the completion of a double-sided model and a local coach model. Finally, random wood distributors have been linked to achievements. Sherlin Varughese et al. presents a system to recognize diseases from Apple’s natural products, and the disease is: infectious diseases, apple scars, and yeast infections. When purchasing an embedded image, this technique is used to identify the area of the fraud and the selection of the infected part. At this point the accent is removed and placed in the database, the vector engine force is arranged for the order of disease and confirmation.

The author describes a strategy for the recognition of the disease of organs of pomegranate and disease: Alternaria bacteria, Blight and Anthracnose. Preset includes changes in size, separation, and morphological operations. RGB, La*b, HSV and YCbCr are used for grouping. In the underlined shades underlined on the triangle and the surface is divided and the canals are used as part of the surface and physics to obtain the frames of the image. In this article have some strength and limitations that strength is to achieve a good accuracy and limitations is only four type of diseases are classified in pomegranate plant. Formula formulas were isolated from the voices of organic products, and the smallest particle class (MDC) was linked to the preparation and grouping of the disease or not in the image.
3. Proposed Method

Figure 1. Overall Flow Diagram of Proposed Method

3.1 Image Choosing & Preprocessing

This is the first step of this approach. The leaf images were taken with the camera in the laboratory and saved on the computer in RGB color space. The processing operation includes removing the background, petiole and unwanted noise. To segment the images, they had to be transferred to the device-independent color space. For this, the image is transformed into the lab color space. The result of this step is shown in below figure.

Figure 2. Color conversion from RGB to Lab
3.2 Affected Part for Segmentation

This is the first method of this paper. In this work from the given input image the disease affected part is segmented. Then only find the type of disease and its severity can also be easily measured. To segment the affected part the GENABC Clustering Algorithm is used. It is depicted in Fig.3. GENABC Clustering is an optimized segmentation algorithm. It is based on the intelligent foraging behavior of the honey bee swarm. The result of this step is shown in below figure.

Figure 3. To extracting disease affected part

Algorithm of GENABC:

Input: Attributes and its values, Cross Over Probability CP, Mutation Probability MP

Output: Best Attributes B

1. Select random attribute values from each Attribute and Consider as the Initial Population P.
2. Evaluate the fitness \( f(x) \) of each attribute \( x \) in the population \( P \) by using the below formula
   \[
   f(x) = \frac{1}{N} \sum_{i=1}^{N} x_i
   \]
   where \( x_i \) is ith attribute values in the attribute \( x \) and \( N \) is the total values in the attribute
3. Create a new population \( NP \) by repeating following steps until the new population \( NP \) is complete
4. Select attributes from a population \( P \) according to their fitness. If the fitness value is high then those attributes as selected otherwise rejected.
5. With a crossover probability \( CP \) cross over the selected attributes to form a new attribute. If no crossover was performed, new attribute is an exact copy of selected attributes.
6. With a mutation probability \( MP \) mutate selected attribute.
7. Place selected attribute in a new population \( NP \)
8. Use new generated population \( NP \) for a further run of algorithm
9. If the end condition is satisfied, stop, and return the best solution \( B \) in current attribute.
10. Go to step 2
11. Finally the output of the genetic algorithm is the fitness value \( F \) and the best attribute \( B \).
12. And give the best attribute \( B \) as the weight value of the ABC and execute ABC.
13. From the ABC and return the best attribute \( B \) as the output.

3.3. Feature Extraction

After segmenting process the next step is to calculate the features from the part. These features are used to uniquely identify the disease name and its severity.

3.3.1. Improved Transform Encoded Local Pattern (ITELP)

Input: Image I, Current Pixel value \( S_c \), Current X Index, Current Y Index, Pixel Distance \( P \), Total No of Neighbors \( T \)

Output: IELBP Pattern, IELBP Description

Algorithm:

Apply the below steps for all pixel in an input image
1. Get the neighbouring pixels \( Sp \) of current pixel \( Sc \) Counterlet on the pixel Distance \( P \).
2. Apply Counterlet on \( Sp \) and \( Sc \) to convert pixel into Counterlet Coefficient.
3. Then compare the current Counter let Coeffi-
cient CounterletCc with that the neighboring 
Counterlet Coefficient Counterlet Cp.
4. If the center pixel DFTCc is greater than the 
neighboring pixel CounterletCp put the value 
1
5. Else put the value 0
6. Then convert these binary number into the 
decimal value 
7. This is called local binary pattern description
8. Then store the description into the array. This
array is called bin

3.3.2. Enhanced Gradient Feature (EGF)
The enhanced gradient feature descriptor in this 
paper, that EGF descriptor has two radial bins (a 
centre and a surround) and four angular bins 
(quartering). But there is only one centre cell 
which is similar to the EGF feature. Each circle 
block is in an 8×8 pixels block, and two pixels is 
the best radius of the central bin. That 8×8 block 
is nearly a 32×32 block in an original 64×128 image, 
so our new descriptor has more large-scale inform-
ation than the original histogram of oriented 
gradients (HOG). We use a five-bin histogram to 
represent one circle block, each bin is consists of 
accumulating all gradients within it. Given an n×n 
block in a bin-related 15×31 image I, every pixel 
I(i, j) in the block has a weight wN(i, j)to the bin N 
(N = 1,2,3,4,5) in this block. The weights are relat-
ed to the spatial position of pixels.

3.3.3. Improved Color Histogram (ICH)
The histogram relays on an underlying color 
space and uses its components to determine the 
probability of a certain color being existed in an 
image. YCbCr histograms relay on the YCbCr 
color space components to build a histogram. In 
constructing an YCbCr histogram, each compo-
nent of the YCbCr histogram is quantized into a 
certain number of regions. An implementation of 
YCbCr histogram divides the luminance (Y) into 
eight regions, whereas each of chromic compo-
nents (Cb, Cr) is divided into four regions. The 
three color components then are linked, thus cre-
ating a (8x4x4) histogram of 128 bins /9/. The 
Sample formats of YCbCr color space have three 
types: YCbCr 4:2:0, YCbCr 4:2:2, YCbCr 4:1:1 
and YCbCr 4:4:4. YCbCr 4:1:1 is very commonly. 
In 4:1:1 color space, every light value is stored 
using 8 bits, these are values of Y, 2×2 region is 
used to store a Cb and a Cb. So, in RGB color 
space which R,G,B is presented by 8 bits un-
signed, every pixel is stored by 8×3=24 bits. How-
ever, in YCbCr color space, every pixel is stored 
by 8+ (8/4) + (8/4) =12 bits, in other words, the 
image data is compressed into half. So, if using 
YCbCr 4:2:2 sample format, YCbCr color space 
should be made in non-uniform quantization, that 
is, being quantified into (8,2,2), and the corre-
sponding features of dimensions are 28 ×22 ×22 
=4096. So, an image is expressed by 4096-
dimensional feature vector. By 4096-dimensional 
quantitative level, the three color components can 
be synthesized into matrix L, then L = Y-QCb+Q Cr 
+ Cb-Q Cr + Cr (7) Where, QCb and QCr are the 
quantitative series of component Cb, Cr. Take 
QCb =22 =4, QCr = 22 =4, so formula (7) can be 
expressed as: L=16Y+4Cb+Cr (0 <=L

The extracting algorithm of dominant dolor of the 
image may be experimented as follows:
- Converted the RGB color space to YCbCr 
color space.
- Y, Cb, Cr€[0,255].
- Quantized YCbCr color space to 4096-
dimensional by non-uniform quantiza-
tion.
- Synthesize the three color components of 
HSV into matrix L, L=16Y+4Cb+Cr (0 <=L <4096).
- Order by the attribute value of L, from 
large to small.
- Sum the attribute value, from large to 
small. If the sum of top d attribute value 
is more than 80%, d is 
regard as Intrinsic dimension of the image
- Computed the intersection of two image 
histogram as the similarity for image re-
trieval.

3.4. Classification
The final process is to classify the disease 
name and its severity. To do this process the Im-
proved Sparse Representation Classifier is used.

Improved Sparse Representation Classifier 
(ISRC)
We propose a method to select the kernel and its 
parameters using just one training set that is also 
able to preserve robustness and avoid over fitting, 
thereby avoiding the requirement for a large 
amount of data to create a separate validation set. 
Using the proposed kernel/parameter selection,
the Kernel Group Based SRC algorithm selects the optimal kernel and its parameters automatically. This helps in focusing only on selecting the appropriate feature sources and training data, which become the parameters of the algorithm while a large set of possible kernels and associated parameters, can be set as the search space for the kernel selection algorithm. To preserve robustness while performing parameter selection using the training set, we propose a metric to quantize the goodness of separation created by the transformation enacted by a kernel function.

4. PERFORMANCE ANALYSIS

4.1 EXPERIMENTAL IMAGES

In this paper the images taken from the real cameras are used. The sample images are shown in the below Fig 4. The image size is 512 x 512 color images. All of these images which is affected by any one of the disease. These images are used for experimental purposes.

Figure 4 Experimental Images

(a) Minor Pest  (b) Black Spot  (c) Brown Spot  (d) Black Spot  (f) Black Spot   (g) Alternaria Spot
(h) Alternaria Spot  (i) Alternaria Spot (j) Alternaria Spot (k) Yellow Spot

4.2 PERFORMANCE METRICS

To evaluate the performance of the crop disease detection techniques several performance metrics are available. This paper uses the Classification Accuracy, Error rate, Precision rate value and Recall Rate Value to analyse the performance.

Classification Accuracy
Accuracy is the measurement system, which measure the degree of closeness of measurement between the original disease and the detected disease.

\[
\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}
\]  

(1)

Where, 
- TP – True Positive (equivalent with hit)
- FN – False Negative (equivalent with miss)
- TN – True Negative (equivalent with correct rejection)
- FP – False Positive (equivalent with false alarm)

**Error Rate**

Error Rate is the measurement system, which measure no of falsely identified diseases name from the given input images.

\[
\text{Error Rate} = \frac{\text{No of Images of Falsely Identified Diseases}}{\text{Total No of Images}}
\]  

(2)

<table>
<thead>
<tr>
<th>Crop Disease Detection Methods</th>
<th>Classification Accuracy Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>95%</td>
</tr>
<tr>
<td>ELM</td>
<td>97%</td>
</tr>
<tr>
<td>ISRC</td>
<td>98%</td>
</tr>
</tbody>
</table>

Table 1 Classification Accuracy Value

Figure 5 Classification Accuracy Analysis Graph

<table>
<thead>
<tr>
<th>Crop Disease Detection Methods</th>
<th>Error Rate Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>5%</td>
</tr>
<tr>
<td>ELM</td>
<td>3%</td>
</tr>
<tr>
<td>ISRC</td>
<td>2%</td>
</tr>
</tbody>
</table>

Table 2 Error Rate Value

Figure 6 Error Rate Value Analysis Graph
Table 3  Precision Rate Value

<table>
<thead>
<tr>
<th>Crop Disease Detection Methods</th>
<th>Precision Rate Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>92%</td>
</tr>
<tr>
<td>ELM</td>
<td>94%</td>
</tr>
<tr>
<td>ISRC</td>
<td>96%</td>
</tr>
</tbody>
</table>

Figure 7  Precision rate Analysis Graph

Table 4  Recall Rate Value

<table>
<thead>
<tr>
<th>Crop Disease Detection Methods</th>
<th>Recall Rate Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>93%</td>
</tr>
<tr>
<td>ELM</td>
<td>96%</td>
</tr>
<tr>
<td>ISRC</td>
<td>97%</td>
</tr>
</tbody>
</table>

Figure 8  Recall Rate analysis Graph value
5. Conclusion

In this paper, a new method is proposed for identifying the diseases in the leaves of all plants. Plant pathology is the scientific study of plant diseases caused by pathogens and environmental conditions. In existing, to detect the diseases they used the spectroscopic techniques. These techniques are very expensive and can only be utilized by trained persons only. This work presents the detection of diseases which are detected using ISRC (Improved Sparse representation classifier) technique. First the GENABC clustering Approach is applied to the input image to segment the affected area. Then extract the features from the area by using Improved Transform Encoded Local Pattern, Improved color histogram and Enhanced Gradient Feature Techniques are used. And then these features are given to the ISRC classifier to get the type of disease.

Notes

/2/ D. Al Bashish; M. Braik; S. Bani-Ahmad (2010), A Framework for Detection and Classification of Plant Leaf and Stem Diseases, IEEE International Conference on Signal and Image Processing, pp. 978-1-4244-8594-9/10
/18/ A. Awate; D. Deshmankar; S. Sonavane (2015), Fruit Disease Detection using Color, Texture Analysis and ANN, IEEE International Conference on Green Computing and Internet of Things (ICCGIoT), pp. 978-1-4673-7910.
