Volume 4, Issue 10, October 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Implementation of Spatial FCM for Leaf Image Segmentation in Pest Detection

V. Premalatha
PG Scholar, BIT,
Sathyamangalam, India

Dr. M. G. Sumithra
Professor, BIT,
Sathyamangalam, India

S. Deepak
PG Scholar, BIT,
Sathyamangalam, India

P. Rajeswari
PG scholar, BIT,
Sathyamangalam, India

Abstract: Crop growth is a core element of the field management. Suitable evaluation and diagnosis of crop disease in the field is very critical for the increased production. Agricultural crops in India are under constant threat of pests affecting their roots as well as leaves. Diseased plants can exhibit a variety of symptoms and making diagnosis was extremely difficult. Common symptoms are includes abnormal leaf growth, color distortion, stunted growth, shriveled and damaged pods. Image acquisition devices are used to acquire images of plantations at regular intervals. These images are then subjected to pre-processing, transformation and clustering. The leaf images are segmented using clustering techniques. Clustering is the process of partitioning a group of data points into a small number of clusters. In this paper we present a clustering technique called Spatial FCM (SFCM) to identify the pest & the disease. Also the performance of proposed technique is compared with other clustering techniques such as K-means, Fuzzy C-Means (FCM), Kernel based FCM (KFCM) & Spatial FCM (SFCM). Then the features such as color, texture are extracted from diseased leaf image & then compared with normal leaf image. The neural network method is used to classify the pest & Disease in crops. The evaluation parameters considered for comparison of spatial FCM with other clustering techniques are as follows: Specificity, Sensitivity, Accuracy, Area of affected leaf, Percentage of disease infection, etc. The neural network classification method is used to classify the type of disease and the pathogen or pest that causes that disease.

Keywords: Image Segmentation; Clustering; K-means; FCM; KFCM & Spatial FCM

I. INTRODUCTION

Agriculture is heart of our civilization. But Plant diseases have turned into a dilemma as it can cause significant reduction in both quality and quantity of agricultural products. Nowadays farmers are facing many crucial problems for getting better yield cause of rapid change in climate and unexpected level of insects, in order to get better yield need to reduce the level of pest insect. India is an agricultural country wherein about seventy percentage of the population depends on agriculture. Farmers have wide range of diversity to select suitable fruit and Vegetable crops. However, the cultivation of these crops for optimum yield and quality product is highly technical. It can be improved with the aid of technological support. The management of perennial fruit crops requires close monitoring especially for the management of diseases that can affect production significantly and subsequently the post-harvest life. In case of plant the disease is defined as any impairment of normal physiological function of plants, producing characteristic symptoms. A symptom is a phenomenon accompanying something and is regarded as evidence of its existence.

Disease is caused by pathogen which is any agent causing disease. In most of the cases pests or diseases are seen on the leaves or stems of the plant. Therefore identification of plants, leaves, stems and finding out the pest or diseases, percentage of the pest or disease incidence, symptoms of the pest or disease attack, plays a key role in successful cultivation of crops. In general, there are two types of factors which can bring death and destruction to plants; living (biotic) and nonliving (abiotic) agents. Living agent's including insects, bacteria, fungi and viruses. Nonliving agents include extremes of temperature, excess moisture, poor light, insufficient nutrients, and poor soil pH and air pollutants. Diseased plants can exhibit a variety of symptoms and making diagnosis was extremely difficult. Common symptoms are includes abnormal leaf growth, color distortion, stunted growth, shriveled and damaged pods. Although pests & diseases can cause considerable yield losses or bring death to the plants and it's also was directly affect to human health. However, crop losses can be minimized, and specific treatments can be tailored to combat specific pathogens if plant diseases are correctly diagnosed and identified early. These need-based treatments also translate to economic and environmental gains. Here the segmentation technique used is Spatial FCM & the Neural network (NN) is a mathematical form or computational model based on biological neural networks used to extract the leaf feature as of the database.

II. MATERIALS & METHOD

Image segmentation is the process of partitioning a digital image into multiple segments. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture.

Clustering is the process of partitioning a group of data points into a small number of clusters. Image analysis can be applied for the following purposes:
1. To detect diseased leaf, stem, fruit.
2. To quantify affected area by disease.
3. To find the boundaries of the affected area.
4. To determine the color of the affected area.
5. To determine size & shape of fruits.
6. To identify the object correctly.

A cluster is usually represented as either grouping of similar data points around a center ie., centroid or a prototype data instance nearest to the centroid. Clusters with well-defined boundaries are called crisp clusters, while those without such feature are called fuzzy clusters.
In this paper Spatial FCM clustering is used for segmentation of leaf images and the results are compared with the existing clustering techniques such as
1. K-means clustering
2. Fuzzy C-Means clustering
3. Kernel based FCM

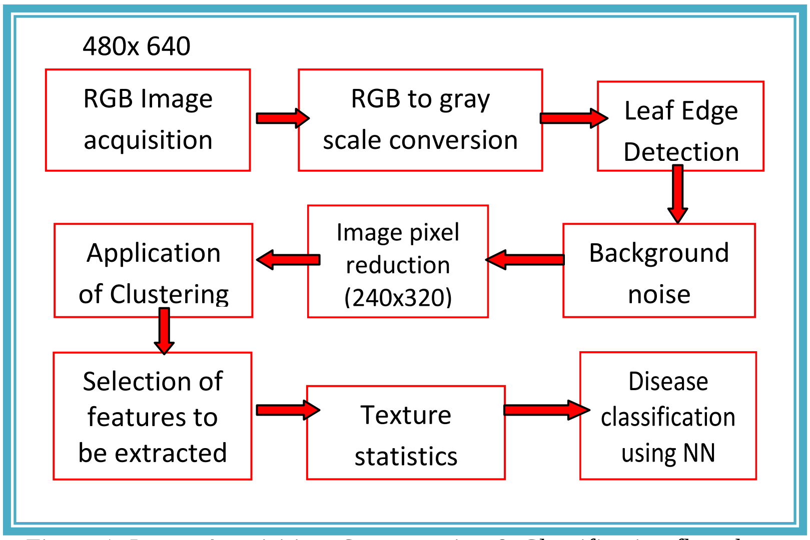


Figure 1: Image Acquisition, Segmentation & Classification flowchart

III. K-MEANS CLUSTERING

K-Means or Hard C-Means clustering is basically a partitioning method applied to analyze data and treats observations of the data as objects based on locations and distance between various input data points. The K-Means clustering algorithm is proposed by Mac Queen in 1967 which is a partition-based cluster analysis method. It is used widely in cluster analysis for that the K-means algorithm has higher efficiency and scalability and computational time is very less. However it also has many inefficiencies: the number of clusters K needs to be initialized, the initial cluster centers are arbitrarily selected, and the algorithm is influenced by the noise points. In view of the shortcomings of the traditional K-Means clustering algorithm, an improved K-means algorithm using noise data filter was developed. It is density based detection methods based on characteristics of noise data where the discovery and processing steps of the noise data are added to the original algorithm. By preprocessing the data to exclude these noise data before clustering data set the cluster cohesion of the clustering results is improved significantly and the impact of noise data on K-means algorithm is decreased.

A. K-means clustering algorithm

K-means clustering algorithm is simply described as follows:
Input: N objects to be cluster ($x_1, x_2 \dots x_n$), the number of clusters k;
Output: k clusters and the sum of dissimilarity between each object and its nearest cluster centre is the smallest;
1. Arbitrarily select k objects as initial cluster centres ($m_1, m_2 \dots m_k$);
2. Calculate the distance between each object X_i and each cluster centre, and then assign each object to the nearest cluster, formula for calculating distance as:

$$d(x_i, m_i) = \sqrt{\sum_{j=1}^{d} (x_{i1} - m_{j1})^2}, \; i=1 \dots N; j=1 \dots k;$$

d (X_i, m_j) is the distance between data i and cluster j;
3. Calculate the mean of objects in each cluster as the new cluster centres,

$$m_i = \frac{1}{N_i}\sum_{j=1}^{N_i} x_{ij} \text{ , i=1, 2... k;}$$

N_i is the number of samples of current cluster i;

4. Repeat 2 & 3 until the criterion function E converged, return $(m_1, m_2 \ldots m_k)$.

B. Merits Of K-Means Clustering

i. Relatively scalable and efficient in processing large data sets
ii. Less computational complexity i.e. O(nkt), where n is the number of objects, k is the number of clusters, and t is the number of iterations
iii. It works well when the clusters are compact clouds that are rather well separated from one another
iv. Very simple method and the results can be easily modeled to deal with streaming data

C. Demerits of K-Means Clustering

i. The algorithm can only be applied when the mean of a cluster is defined
ii. The numbers of clusters must be specified in advance
iii. This method is not suitable for clusters with non-convex shapes
iv. This method is sensitive to noise and outlier elements

IV. FUZZY C-MEANS CLUSTERING

The Fuzzy C-Means (FCM) clustering algorithm was first introduced by Dunn and later was extended by Bezdek. The algorithm is an iterative clustering method that produces an optimal c partition by minimizing the weighted within group sum of squared error objective function. FCM is an unsupervised clustering algorithm that is applied to wide range of problems connected with feature analysis, clustering and classifier design. FCM is widely applied in agricultural engineering, astronomy, chemistry, geology, image analysis, medical diagnosis, shape analysis and target recognition.

With the development of the fuzzy theory, the FCM clustering algorithm which is actually based on Ruspini Fuzzy clustering theory was proposed in 1980's. This algorithm is used for analysis based on distance between various input data points. The clusters are formed according to the distance between data points and the cluster centers are formed for each cluster.

A. FCM Algorithm

Step 1: Let us consider that M-dimensional N data points represented by x_i (i = 1,2…N) are to be clustered.

Step 2: Assume the number of clusters to be made, that is, C, where $2 \le C \le N$.

Step 3: Choose an appropriate level of cluster fuzziness $f > 1$.

Step 4: Initialize the N × C ×M sized membership matrix U, at random, such that $U_{ijm} \in [0, 1]$ and $\Sigma\, U_{ijm} = 1.0$, for each i, j and a fixed value of m.

Step 5: Determine the cluster centers CC_{jm}, for j^{th} cluster and its m^{th} dimension by using the expression given below:

$$CC_{jm} = \frac{\sum_{i=1}^{N} U_{ijm}^{f} x_{im}}{\sum_{i=1}^{N} U_{ijm}^{f}}.$$

Step 6: Calculate the Euclidean distance between i^{th} data point and j^{th} cluster center with respect to, say m^{th} dimension like the following:

$$D_{ijm} = \|(x_{im} - CC_{jm})\|.$$

Step 7: Update fuzzy membership matrix U according to D_{ijm}. If $D_{ijm} > 0$, then

$$U_{ijm} = \frac{1}{\sum_{c=1}^{C} \left(\frac{D_{ijm}}{D_{icm}}\right)^{\frac{2}{f-1}}}.$$

If $D_{ijm} = 0$, then the data point coincides with the corresponding data point of j^{th} cluster center CC_{jm} & it has the full membership value, that is, $U_{ijm} = 1.0$.

Step 8: Repeat from Step 5 to Step 7 until the changes in $U \le \varrho$, where ϱ is a pre-specified termination criterion.

B. Merits Of FCM

i. Soft clustering technique
ii. Each input pixel or training vector belongs to multiple clusters
iii. Easy tackling of noisy data
iv. Accuracy is high when compared to K-means
v. Fast convergence when compared to other clustering techniques

C. Demerits Of FCM

i. FCM is effectual only in clustering those crisp, spherical, and non-overlapping data.
ii. Some degree of dependence on initial partitioning cluster values in hard C-means
iii. Computational time is very large in fuzzy C-means clustering

V. KERNEL-BASED FUZZY C-MEANS CLUSTERING (KFCM)

KFCM is an unsupervised clustering algorithm that is functional with clustering peceptron design. KFCM is widely practical in agricultural engineering, image analysis, medical diagnosis, and shape analysis and target recognition. The clusters are formed according to the distance between data points and the cluster centers are formed for each cluster.

A. KFCM Algorithm

The step by step algorithm applied in KFCM is as follows:

Step 1: Update the membership by the following formula. Then we can get the partition matrix $U^{(b)}$ as

$$u_{ik}^{(b)} = \frac{(1/K(x_k,x_k)+K(v_i^{(b)},v_i^{(b)})-2K(x_k,v_i^{(b)}))^{1/(m-1)}}{\sum_{j=1}^{c}(1/K(x_k,x_k)+K(v_i^{(b)},v_i^{(b)})-2K(x_k,v_i^{(b)}))^{1/(m-1)}}$$

Step 2: Update the clustering centers by the following formula. Then we can get the clustering centers matrix

$$V_i^{(b+1)} = \frac{\sum_{k=1}^{n} {u_{ik}^{(b)}}^{m} k(X_k, V_i^{(b)}) X_k}{u_{ik}^{(b)} k(X_k, V_i^{(b)}) X_k}$$

Step 3: If $\| \mathbf{V}^{(b)} - \mathbf{V}^{(b-1)} \| < \boldsymbol{\varepsilon}$ (where $\varepsilon = 0.01$) the algorithm stops, then outputs the partition matrix U and clustering centers matrix V. Otherwise, makes b add 1, then turn back to Step1.

VI. SFCM ALGORITHM

A conventional FCM algorithm does not fully utilize the spatial information in the image. In this paper, we present a fuzzy c-means (FCM) algorithm that incorporates spatial information into the membership function for clustering. The spatial function is the summation of the membership function in the neighborhood of each pixel under consideration. The advantages of the new method are the following: (1) it yields regions more homogeneous than those of other methods, (2) it reduces the spurious blobs, (3) it removes noisy spots, and (4) it is less sensitive to noise than other techniques. This technique is a powerful method for noisy image segmentation and works for both single and multiple-feature data with spatial information.

Let $X = \{x_1, x_2, x_3 ..., x_n\}$ be the set of data points and $C = \{c_1, c_2, c_3 ..., c_n\}$ be the set of centers. The following equations explain the membership and cluster center updation for each iteration respectively.

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c}(d_{ij}/d_{ik})^{(2/m-1)}}$$

$$c_j = \sum_{i=1}^{n}\left(\frac{((\mu_{ij})^m\ x_i)}{(\mu_{ij})^m}\right)$$

where,

$\mathbf{d_{ij}}$ represents the distance between i^{th} data and j^{th} cluster center.
c represents the number of cluster
m is the fuzziness index
$\boldsymbol{\mu_{ij}}$ represents the membership of i^{th} data to j^{th} cluster center.
n is the number of data points.
$\mathbf{c_j}$ represents the j^{th} cluster center

VII. SIMULATION RESULTS & CONCLUSION

The plant leaves considered here for analysis are

i. Cotton plant leaves
ii. Corn plant leaves

A. Results For Cotton Leaves With & Without Disease

It is analyzed that in K-means clustering when the no. of clusters increases, the computational time gets decreased. Thus it shows that no. of clusters is inversely proportional to the computational time. And also computational complexity is also less.

(A). Infected Cotton Leaf (B). Cotton Bollworm (C). Cotton Cercospora

Figure 2: Disease affected Cotton leaves

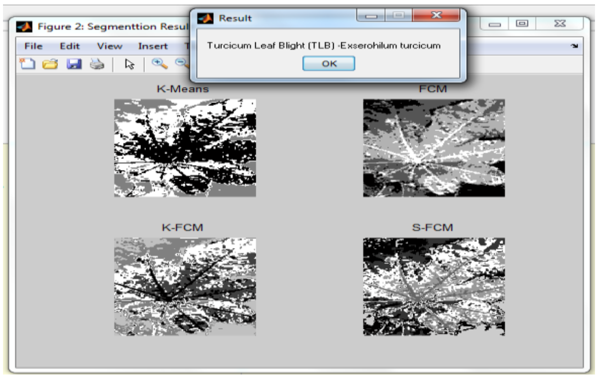


Figure 3: Segmentation results for all four types of clustering & Classification results for cotton leaf blight disease

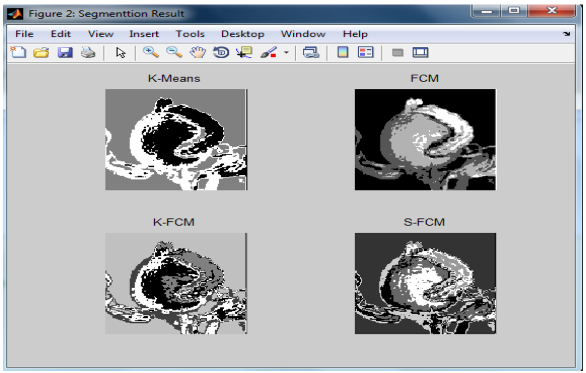


Figure 4: Segmentation results for all four types of clustering for cotton bollworm disease

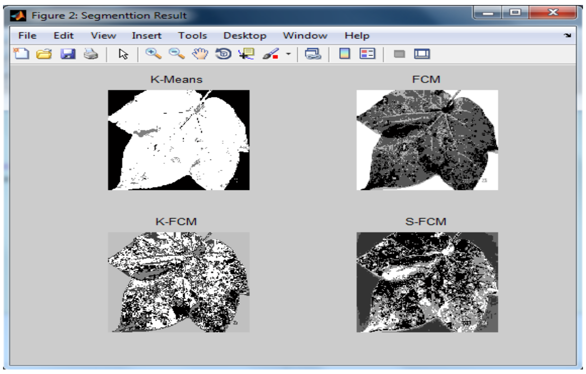
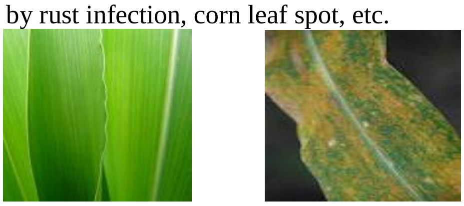


Figure 5: Segmentation results for all four types of clustering for cotton leaf

B. Results For Corn Leaves With & Without Disease

The corn plants are mostly affected by rust infection, corn leaf spot, etc.



(A) Corn Healthy Leaves (B) Corn Rust Infection

Figure 6: Healthy & Disease affected Corn leaves

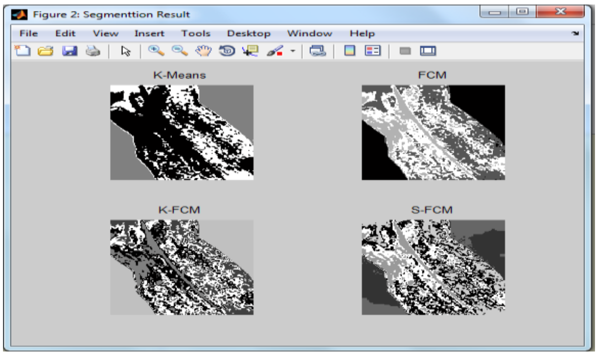


Figure 7: Segmentation results for all four types of clustering rust infected corn leaf

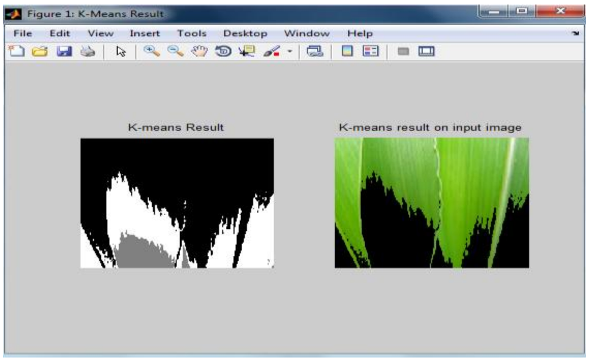


Figure 8: K-means Segmentation results on color & black and white image of healthy corn leaf

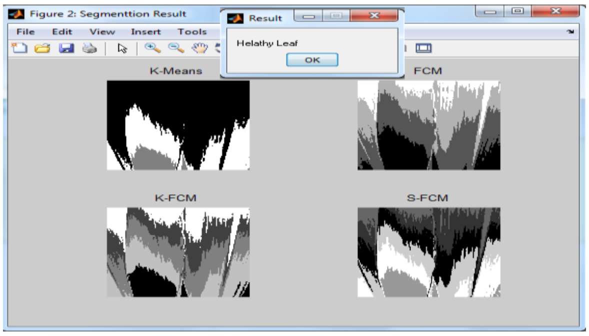


Figure 9: Segmentation results for all four types of clustering & Classification results for Healthy corn leaf

C. Comparison Of Results

The parameters considered for evaluation are as follows:

(i) *Sensitivity:* It specifies the sensitiveness of clustering techniques to noise datas

Se=TP/ (TP+FN)

Sensitivity=se*100

(ii) *Specificity:* It indicates that how specifically the leaves are segmented with respect to infected regions alone.

Sp=TN/(TN+FP)

Specificity=sp*100

(iii) *Accuracy:* It indicates the accurate cluster centres & how accurately the leaf image is clustered.

acc=(TP+TN)/(TP+TN+FN+FP)

acc=acc*100

ACCURACY=acc+20

And also how much area of leaf is infected by pest & how much percentage of leaf is affected by pest.

Table I: Result Analysis for all above discussed clustering techniques

Features	**K-means**	**FCM**	**Spatial FCM (SFCM)**
Sensitivity	71.42857143	80	75
Specificity	75	66.66666667	85.71428571
Accuracy	92.72727273	95	98.57142857
Affected Area	25377	39303	60777
Total Area	196608	196608	196608
Percentage	12.90740967	20.13071159	30.91293799

From table 1, it is evident that all the evaluation parameters are high in our proposed technique SFCM when compared to other clustering techniques**.**

Table II: Comparative analysis of K-means & FCM in terms of time complexity

Algorithm	Time Complexity	Elapsed time(Seconds)
K-means	O(ncdi)	0.448755
FCM	O(ncdi^2)	0.781679

From Table 2 & 3, it was evident that the time consumed for clustering process is less in K-Means but it is very high in FCM techniques. Thus accuracy is very high in FCM techniques but time complexity is very high.

Table III: Time complexity of K-Means & FCM for different no. of iterations

S.No.	Number of Iterations	K-Means Time Complexity	FCM Time Complexity
1	5	3000	6000
2	10	6000	12000
3	15	9000	18000
4	20	12000	24000

REFERENCES

[1] S.Ananthi, S.Vishnu Varthini, "Detection and Classification Of Plant Leaf Diseases", International Journal of Research in Engineering & Applied Sciences, Volume 2, Issue 2,2012. ISSN: 2249-3905.

[2] "Fuzzy c-means clustering with spatial information for image segmentation", Keh-Shih Chuang , Hong-Long Tzeng, Sharon Chen, Jay Wu, Tzong-Jer Chen, Computerized Medical Imaging and Graphics 30, 9–15, 2006.

[3] Fuzzy c-means clustering with spatial information for color image segmentation, Jaffar, M.A, Naveed, N, Ahmed, B, Hussain, A, Mirza, A.M, ICEE '09. Third International Conference on 9-11, 1- 6, 2009.

[4] Tarun Kumar , Karun Verma,"A Theory Based on Conversion of RGB image to Gray image", International Journal of Computer Applications (0975 – 8887) Volume 7– No.2, September 2010.

[5] Rafael C. gonzalez, Richard E.Woods," Digital Image Processing", Second Edition, pearson Education.

[6] Rajeshwar Dass, Priyanka, Swapna Devi, "Image Segmentation Techniques",*IJECT*, Vol. 3, Issue 1, Jan. - March 2012.

[7] Amanpreet Kaur, Bhogal Neeru, Singla Maninder Kaur, "Comparison of Algorithms for Segmentation of Complex Scene Images*", (IJAEST)* Vol No. 8, Issue No. 2, 306 – 310

[8] Krishna Kant Singh1 , Akansha Singh, "A Study Of Image Segmentation Algorithms For Different Types Of Images", *IJCSI International Journal of Computer Science Issues*, Vol. 7, Issue 5,September 2010.

[10] Pratt, W. K., 2007, "Digital Image Processing", Fourth Edition, A John Wiley & Sons Inc. Publication.

[11] S. Thilagamani and N. Shanthi, "A Survey on Image Segmentation Through Clustering", *International Journal of Research and Reviews in Information Sciences*, Vol. 1, No. 1, March 2011.

[12] Prabhjot Kaur, Pallavi Gupta, Poonam Sharma, "Review and Comparison of Kernel Based Fuzzy Image Segmentation Techniques", *I.J. Intelligent Systems and Applications*, 2012, 7, 50- 60.

[13] Mr. V. A. Gulhane & Dr. A. A. Gurjar, "Detection of diseases on Cotton Leaves and Its Possible Diagnosis", *International Journal of Image Processing (IJIP)*, Volume (5), Issue (5), 2011

