Pomegranate Fruits Disease Classification with K Means Clustering

Shivanand B Lamani
Faculty in Computer Science Dept, Akkamahadevi Women's University, Vijayapur

***Abstract*: The Identification of pomegranate fruit disease (bacterial blight, scab etc.) and also the remedy for that disease after identification are proposed. Bacterial Blight disease needs to control at initial stages otherwise it makes economic loss to farmers. The captured image of the diseased fruit uploads to the system. The system then makes the image processing and makes the classification of fruit is infected. In Proposed system comparative accuracy analysis is done using K-means segmentation and also with different classifiers like PNN (Probabilistic Neural Network), KNN (K Nearest Neighbours') and SVM (Support Vector machine). To achieve more accuracy closed capturing system, with high resolution camera is used, due to this capturing system 99% accuracy is achieved.**

***Keywords:* K-means segmentation, PNN (Probabilistic Neural Network), KNN (K Nearest Neighbours') and SVM (Support Vector machine).**

I. INTRODUCTION

1.1 Introduction

India is among the country where most of the people depend on agriculture. And the major area which decides economy of the nation is agriculture. The agricultural yield's quality and production quantity is affected by ecological parameters like temperature, rain and other climate related parameters which are out of control of human beings. Another major factor which affects productivity of the yield is the disease; in this factor we can have control to improve the productivity for quality as well as for quantity of yield.

The main threat for pomegranate cultivation is diseases and insect pests. Thereforetimely correct diagnosis and careful treatment essential to defend the yield fromseveredamage and severe loss. Plants diseases may be found in stem, leaves andfruit. Bacterial Blight, Alternaria and Scab are major diseases that affectsthe pomegranate fruits. The diseasesaffects to neighbourhealthy pomegranate plants via wind, sprayed rain and through infected cuttings. In destructive diagnosis methods first the fruit is removed from plant and then measured. In non-destructive methods dimensions of fruit are measured without removing the fruit.The technology makesfarmers to check the possibility of diseases at primary stagesand make possible treatment. A methodology is developed to determine the type of disease the fruit is affected.

For prevention of disease, it is required to be detected at early stage so that treatmentcan be done properly andavoid spreading of the disease. Advances technologies makes it possible to use the images of diseased fruitand detect the type of disease. This achieved by using image processing technology, where features extracted from the images and further used with classification algorithms to make identification.

Controlling the diseases is a challenging and most vital taskwhich can be achievedby proper Disease management. This challenge can be made easy by using image processing for detecting diseases of fruit. With this system it is possible to detect type of disease, theaffected area and severity of the disease.

Pomegranates are among the healthiest fruits. Pomegranateshave a range of beneficial plant compounds, incomparable with other foods. Researchstudies found that they have incredible benefits for human body, and lowers the risk of all sorts of diseases. Human being can enjoy pomegranates; get benefits of pomegranates by consuming in the syrup or eating seeds, juice, paste etc. For cartilage related problems and at baby'sbirth time brain damage problempomegranate is beneficial.

1.2 The most common diseases on Pomegranate are:

ANTHRACNOSE

Symptoms : Small, regular to irregular black spots on fruits which turn

later on as dark brown depressed spots. Figure 1.1(a)

Prevention : Spraying of Difenconazole 25 EC at 1.0 ml/lit or Prochloraz

45 EC at 0.75ml/lit were effective against anthracnose disease.

(a) Anthracnose (b) Bacterial Blight (c) Scab

Figure 1. Pomegranate Deseases and symtoms

BACTERIAL BLIGHT

Symptoms :Fruits develop spots initially looks circular later irregular in shape, color changes from brown to dark brown affecting the fruit. Figure 1.1(b)

Prevention :Use Ethrel spray, Paste on affected area mixture of 0.5g Streptomycin Sulphate + 2.5g Copper oxy chloride + 200g red oxide mixed in one lit of water.

SCAB

Symptoms :Wounds on young fruit become brown and corky. Cracks appears in the outer layer and as well as inside, or the fruit may go out of shape. Figure 1.1(c)

Prevention :scab is primarily treated by using fungicide sprays. A variety of fungicide sprays with differing modes of action are available. When and how they are useddepends upon their mode of action. Protectant fungicides prevent the spores from germinating or penetrating leaf tissue. To be effective, they must be applied to the surface of susceptible tissue before infection occurs.

India now one of thefast developing nation. This growth,agricultural field haveforemost impact. Intelligent farming assisting farmers with image processing based smart tools and automatedartificial intelligenttools that effortlessly combine production, information and dealsto improved production, yield quality and income. The traditional approach ofrecognition of fruit infections is using the bare eye analysisfrom the professionalspecialists. Consulting professional experts iscostly and time takingbecause to the unavailability of expert in nearby locations. Classification of fruit diseases and automatically detecting the symptoms as earliest as possible is very important.

II. IMPLEMENTATION

The implementation phase is translating design specification into actual source code, testing and running. The basicaim of software implementation is creationof source code that is easy to read and understand. After designingthe model and executionplan, finalizing specification standards, algorithm to be used the final phase is implementation.

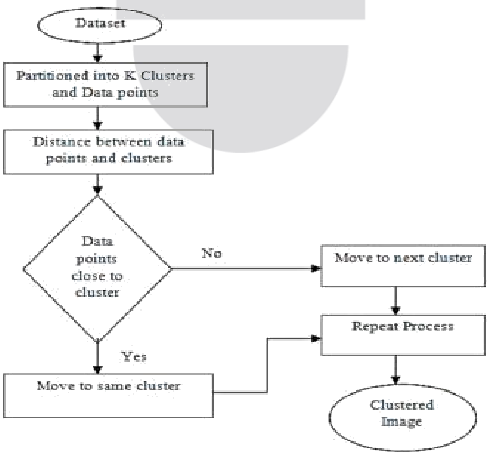

Figure 2: Flow Chart for K means Algorithm

The user requirement specifications and the support from management is very important while implementation. User involvement in the design and implementation is very much beneficial, users can guide in implementation according to business requirements and importance, merging the user ideas and expertise ideas results into superior solutions.

2.1 K-MEANS CLUSTERING

K-Means clustering means the pixels of the image have to be grouped into k distinguishing groups these groups are called clusteres, each cluster having one image pixel in it. How much the clusters non-overlap that much it is favourable to make perfect segmentation.

K-means is unsupervised type algorithms thatmakes effective clustering. Thisalgorithm make use of an easy way to separate given image pixelsby taking specified number of clusters ‘k’. Initially k centroids are defined, belonging for each of ‘k’ clusters. A cluster definespixels “similarity” and “dissimilarity”. K-means important algorithm becauseitgivesbest solutionto the clustering problem.

Theclusterscentroids located in a smart way sincediverse locale causes diverse result. Thereforewhile initialisation centroids are locatedat reasonable distance from one another. From the given image pixels are taken and linked to itsnearestcentroid. When all pixels are linked, then early grouping is over. After early grouping new linking made among the pixel points and the nearest new centroid.
After repeation ofthis loop the k centroids keep on changing their location, then loop stops when no more location changes i.e centroidsgot fixed and notchanginganymore.

2.2 Image Pre-processing

Image pre-processing aimsto removesunwanted areas from image or image features improve which are helpful for processing of remaining steps and to perform analysis task. Image pre-processing do not affect toinformation matter of image. The background is removed using thresholding.Background removed from top, left, bottom and right without affecting fruit area.

After Applying Fuzzy-C-Means Clustering

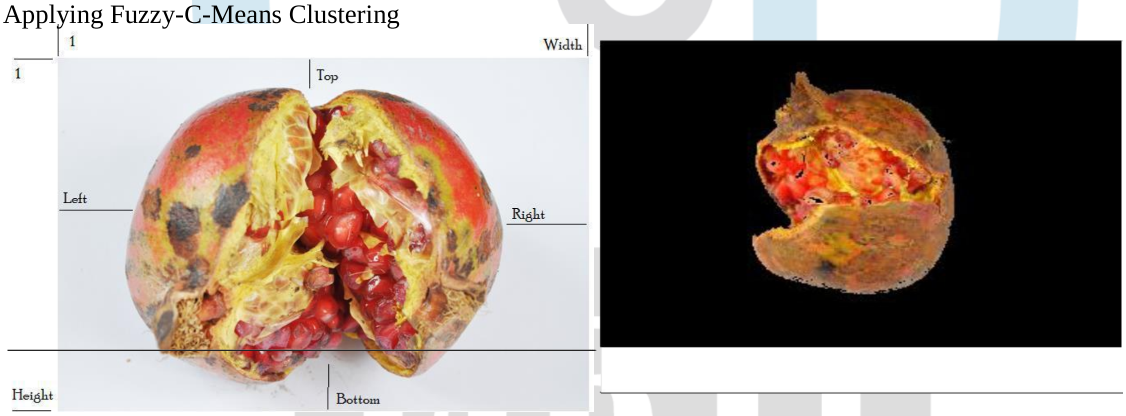

Figure 3. Removal of background

<u>Pseudo Code</u>

```
%Remove from Left
for i=1 to Height
for j=1 to Width
if R,G,B >= 170
set pixel to black i.e make R,G,B as 0
else
break;end
end
end
%Remove from Right
for i=1 to Height
for j=Width to 1
if R,G,B >= 170
set pixel to black i.e make R,G,B as 0
else
break;
end
end
    end
```

```
%Remove from Left
for j=1 to Width
for i=1 to Height
if R,G,B >= 170
set pixel to black i.e make R,G,B as 0
else
break;
end
end
end
%Remove from Bottom
for j=1 to Width
for i=Height to 1
if R,G,B >= 170
set pixel to black i.e make R,G,B as 0
else
break;
end
end
end
```

2.3 Segmentation

Image segmentation makes collections of homogeneous pixels in a regions depending on common similarities. Common similaritiesmay be in terms of pixel colours, texture etc.It's important and must to simplify the image so that, the analysis of image becomes easier and efficient, and this is done by making use of segmentation process.Segmentation makes indirectly separatingobjects and recognising edges of objects in given image.

If only a certain region of an image is important and the rest can be discarded, segmenting the image into different regions is essential. The goal is to segment colour image in an automated fashion using K-means clustering or Fuzzy C Mean clustering. The clustering can be stated as, "the process of organizing objects into groups whose members are similar in some way".

<u>Pseudo Code for K-Means Clustering</u>

```
I=Fruit_I;
mu,mask]=kmeans2(rgb2gray(I),3);
Iseg = label2rgb(mask); %Image After Aplying K-Means Clustering
```

<u>Pseudo Code for Fuzzy C Means Clustering</u>

```
I=Fruit_I;
H=I(:,:,3);
data = im2double(H(:));
[center,U,obj_fcn] = fcm2(data,3,[2.0 NaN NaN 0]); maxU = max(U);
index1 = find(U(1,:) == maxU);
index2 = find(U(2,:) == maxU);
index3 = find(U(3,:) == maxU);
%       Assigning pixel to each class by giving them a specific value fcmImage(1:length(data))=0;
fcmImage(index1)= 1; fcmImage(index2)= 0.8; fcmImage(index3)= 0.6;
%        Reshapeing the array to a image
imagNew = reshape(fcmImage,size(I,1),size(I,2)); imagNew=im2uint8(imagNew);
Iseg = label2rgb(imagNew); %Image after Fuzzy-C-Means Clustering
```

2.4 K-Nearest Neighbour (KNN)

KNN recognises pattern with non-parametric methodby using the given feature space.K-NN is called as lazy learning because the functions guessed close by locally and the actual calculations are done only at the time of classification.KNN is simplestmachinelearning algorithm where classification is done according to majoritychoice from its k neighbours (k selected typically small). If selected k value is 1, then classification done according to its nearest neighbour.The neighbours are taken from feature space.KNN don't have explicit training step. The training step only of stores the features and group labels as training samples.In classification step, k value and the test case vector supplied by the user then the classification made by assigning the grouplabel of samples nearest to test case.The KNN performance depends on value of kand on the distance metric.

<u>Pseudo Code</u>

1. Read 'fet_kmeans.txt'
2. Initialize training set
3. Initialize group

4. Extract test features from Inputted Image
5. Pred_class = knnclassify(test, training, group);

Matlab provides built in KNN classification function knnclassify Class = knnclassify(Sample, Training, Group) Where,

SampleMatrix	:test set features to be classified.
TrainingMatrix	:Feature set i.e. Training set.
GroupVector	:values defining the grouplabeled of the rows in Training.

<u>Pseudo Code</u>

```
%	train for KNN fid=fopen('fet_kmeans.txt','r');
Ftrdata = fscanf(fid, '%g %g %g %g %g %g %g %g %g %g %g %d', [12 inf]); Ftrdata =Ftrdata';
training=Ftrdata(:,1:11);
group=Ftrdata(:,12);
%	prediction using KNN
test=[Cftr,Tftr];
Pred_class = knnclassify(test, training, group);
```

Support Vector Machine

"Support Vector Machine" is a supervised algorithm which is usefulin classification. This algorithmplots n number of features as a point in n-dimensional matrix, feature value represents particular coordinate. Finally classification performed by locating the hyper-plane that distinguishes the classes very well.

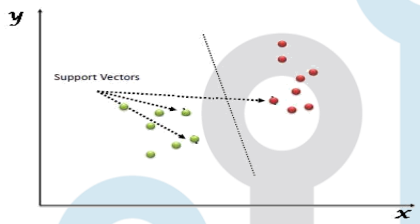

Figure 4. SVM hyper-plane/ line

<u>Pseudo Code</u>

1. Read 'fet_kmeans.txt'
2. Initialize training set
3. Initialize group
4. Train the network using svmtrain function and store in SVMStruct
5. Extract test features from Inputted Image
6. Pred_class = svmclassify(SVMStruct, test)

Matlab provides built in SVM functions:
SVMStruct = svmtrain(Training, Group)
TrainingMatrix of training data, where each row corresponds to an observation or replicate, and each column corresponds to a feature or variable.GroupColumn vector, character array, or cell array of strings for classifying data in Training into two groups. It has the same number of elements as there are rows in Training. Each element specifies the group to which the corresponding row in Training belongs.
Group = svmclassify(SVMStruct, Sample)

III. RESULTS AND ANALYSIS

3.1 K-Mean Results and Anlaysis

Training Single Image Button - One image feature extraction done at a time
Training ImageDatabase Button – Automatically extracts the fetures for images from training set.
Test Single Image Button – Single Image is tested for classification using PNN, KNN or SVM.
Testing-All-PNN Button – Automatically makes testing of all images from testing set using PNN and stored the performance analysis data in fc_pnn_analysis for further comparison.

Testing-All-KNN Button – Automatically makes testing of all images from testing set using KNN and stored the performance analysis data in fc_knn_analysis for further comparison.

Main GUI for K-Means

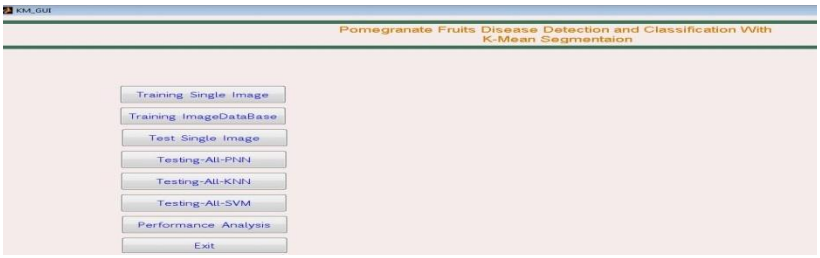

Fig 5 : Main GUI for K-Means

The above figure 5 shows the Main GUI for Classification by using K-Mean Clustering.

Testing-All-SVM Button – Automatically makes testing of all images from testing set using SVM and stored the performance analysis data in fc_svm_analysis for further comparison.performance Analysis Button – Using the Analysed data from fc_pnn_analysis, fc_knn_analysis and fc_svm_analysis makes tabulations and plots comparative graphs.

GUI for Single Image feature extraction

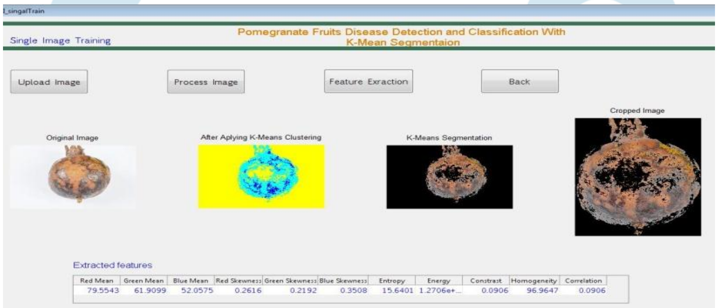

Fig 6: GUI for feature extraction using K-Means

The Figure 6 shows the GUI for single image features extraction by using K-Mean Clustering. Upload Image Button – for Selecting Image.

Process Image – Makes pre-processing, K-Means clustering and cropping. Feature Extraction – Extracted features from processed image.

GUI for Single Image Testing

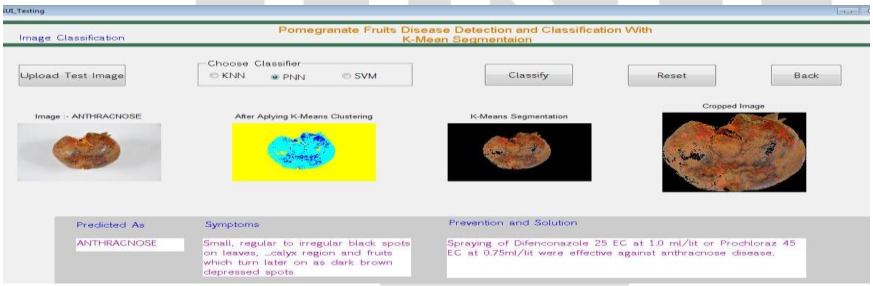

Fig 7. GUI for single image testing.

The Figure 6 shows GUI for single image testing.

Upload Image Button - for Selecting Image. Choose Classifier - Select classifier for testing

Classify Button – Classifies the image according to selected classifier and shows results as in figure 7.

Result Analysis

Perfomance Analysis-With Kmeans Segmentation

	TP	FN	TN	FP	Accuracy	Precision
KNN	60	0	0	0	100	100
PNN	60	0	0	0	100	100
SVM	44	16	0	0	73.3333	100

Fig 8: Accuracy analysis using KNN, PNN or SVM

The Figure 7,8 shows that in K-Mean with KNN has 100% accuracy and SVM has least accuracy for considered test images.

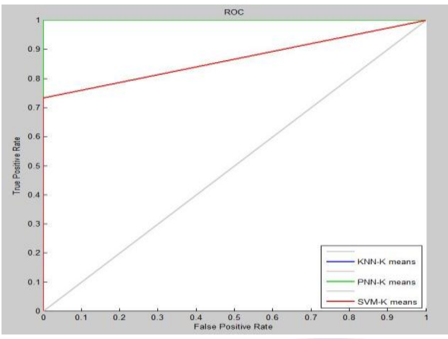

Fig 9: Comparative ROC curve for KNN, PNN or SVM with K-Mean

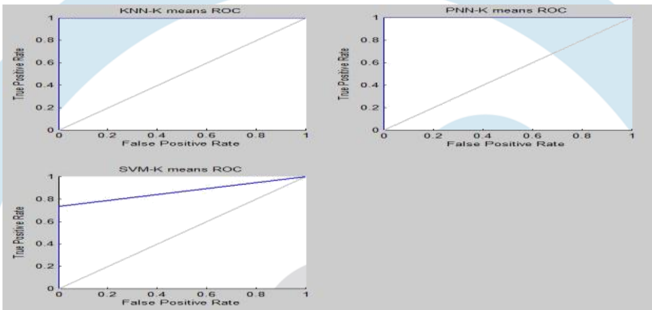

Fig 10: Individual ROC curve for KNN, PNN or SVM with K-Mean

IV. CONCLUSION

Current scenario Suggest to have an approach toautomatically grade the disease on plant.The disadvantages of manualgrading can be overcomeby using this automated system andmayaid the pathologists in terms of making accurate diagnosis.The proposed system implemented by considering fruit features that can be extracted using fuzzy C meanand K means approaches. These approaches have been used for the identification of fruit diseasetypes. The diseases that are affected on pomegranate fruit have been identified using KNN, PNN and SVM classifiers.Theresults analysis showsthat the results found areaccurate and acceptable. Once the disease is identified the symptoms and prevention treatmentsolution provided to prevent further loss. Result analysis shows the KNN and PNN approaches are good.

References

[1] Khot.S.T, Patil Supriya, Mule Gitanjali, Labade Vidya, Pomegranate Disease Detection Using Image Processing Techniques,International Journal of Advanced Research in Electrical,Electronics and Instrumentation EngineeringVol. 5, Issue 4, April 2016.

[2] Swati Dewliya, Ms. Pratibha Singh, Detection and classification for apple fruit diseases using support vector machine and chain code,International Research Journal of Engineering and Technology (IRJET) ,Volume: 02 Issue: 04 | Aug-2015.

[3] Tejal Deshpande1, Sharmila Sengupta2, K.S.Raghuvanshi, Grading & Identification of Disease inPomegranate Leaf and Fruit, International Journal of Computer Science and Information Technologies, Vol. 5 (3), 2014

[4] Monika Jhuria, Ashwani Kumar, Rushikesh Borse,"Image Processing For Smart Farming: Detection of Disease and Fruit Grading", IEEE, Proceedings of the 2013 IEEESecond International Conference on Image Information Processing, 2013

[5] Jagdeesh D. Pujari, Rajesh Yakkundimath , Abdulmunaf S. Byadgi, "Statistical Methods for Quantitatively Detecting Fungal Disease from Fruit's Image", International Journal of Intelligent System and Application in Engineering,vol.1(4),60-67,2013

[6] Shiv Ram Dubey, Anand singh Jalal, "Detection and Classification of Apple Fruit Diseases using Complete Local Binary Patterns" IEEE, Third international conference on Computer and communication Technology, 2012

[7] Ilaria Pertot, Tsvi Kuflik, Igor Gordon, Stanley Freeman, Yigal Elad, Identificator: A web-based tool for visual plant disease identification, aproof of concept with a case study on strawberry, Computers and Electronics in Agriculture, Elsevier, 2012, Vol.88, p.144-154.

[8] Xiaoou Tang, Fang Wen, IntentSearch: Capturing User Intention for One-Click Internet Image Search, IEEE transactions on pattern analysisand machine intelligence, 2012, vol.34, p.1342-1353.

[9] Parag Shinde, Amrita Manjrekar, Efficient Classification of Images using Histogram based Average Distance Computation Algorithm Extendedwith Duplicate Image Detection Elsevier, proc. Of Int. Conf. On advances in Computer Sciences, AETACS, 2013.