Bark Classification of Trees Using K-Nearest Neighbor & Nearest Neighbor Algorithms

Muhammad Tariq[1]   Muhammad Ibrahim[2]

1.Department of Computer Science Government, College University Faisalabad Layyah Campus, Layyah, 31200, Pakistan

2.Department of Electrical Engineering Government, College University Faisalabad Layyah Campus, Layyah, 31200, Pakistan

Abstract

Pakistan is an agricultural country and less than 4 % of area secured with forests. Tree automatic classification based on computer science and it is the developing trend of classification. In this paper we examine how we can done bark classification of trees using k-nearest neighbor and nearest neighbor algorithms. There we discuss how these algorithms can be used to automatically classify trees from images of bark. We get the images of five kinds of different trees names suppose as **A, B, C, D and E** through using digital camera. We take ten different images of each kind of trees. The capability and information of inspectors are essential to perfectly achieve this process. The all the process will be done in computer vision image processing tool. In this tool we use the Histogram Features, Texture Features, and Pattern Classification. We achieved the final results of five kinds of different trees using nearest neighbor on distance two 82% average and on k-nearest neighbor when k=2 then the average result 82%, when k=3 the average result 82%, when k=4 then the average result 76% and when k=5 the average percentage 72% the result shows the maximum correct result and classifies the trees. These are the best percentage results using these algorithms for classification. In this way we can easily classify the different trees and also these methods provide opportunity to farmer and other people for identify and select the different better different trees for getting more benefit.

Keywords: CVIP Tool, Histogram Features, Texture Features, Pattern Classification, Classification Algorithms

1.   Introduction

Less than 4% of area in Pakistan is secured with forests. Woods assume imperative part in the monetary improvement of our nation. The territory under woodlands in Pakistan is little. As per specialists the region under woodlands ought to be no less than 25 to 30% of the aggregate territory of the nation. Lack of knowledge about trees to identify their varieties for a layman for the purpose of forming is a difficult task to identify tree varieties. This research is to identify the different varieties of trees classification by using machine intelligence identification is used in Pakistan on conversional basis it has some drawbacks lack of human knowledge, so human perception is the difficult task to identify each varieties. So we have introduced machine intelligence to recognize tree varieties by its barks using texture features. Images are very powerful information and communication. Images are not only primary resources of information. They are used for communication among people. An image texture is set of metrics calculation image processing designing to quality the texture of an image. It gives us information about the arrangement of color in an image selected region. The main objective of texture classification is to classify the images into different classes that have the same textural pixels. Song et al. described a texture based bark classification method. It uses two different tissue characteristics one of the co-occurrence matrix and long connection length emphasis (LCLE) the features extracted from the binary image of the skin[1]. Wendel et al. expressed that when a new image classification and calculation features screening histogram of the nearest cluster centers are used as input for SVM. Tests and results have been presented data sets of images of leaves and bark. Automatic identification of leaf with local features to prevent division for automatic recognition of images. This compares to combination skin the wavelet GLCM (gray level co-accurance metrics) and features. Classification Support vector machine is used [2]. Sebastian et al. described that in 1973 Haralick's introduced the co-occurrence matrix and texture features for automated classification of rocks in six categories [3]. Olaode et al. described which unsupervised learning approaches have been successfully applied to the image categorization [4]. Sulc and Matas described that a method to recognize the texture of the skin on a description of the scale of the accession LBP histograms [5]. Ko et al. described a method of hybrid ensemble consisting of sequential and parallel architecture tree genus classification using LiDAR (Light Detection and Ranging) data [6]. Mahajan et al. expressed that identify appropriate wavelet basis function approach using multi resolution based on wavelet transform pyramid for the classification of textures in several classes. For the characterization of texture images at multiple scales, it considered several wavelet basis functions as Haar, Daubechies, Coiflet, Symlet, biorthogonal and reverse biorthogonal wavelets [7]. Othmani et al. described that a new method for individual tree recognition of species based on analysis of 3D Geometric texture of tree bark. The skin to reveal the use of two different data sets D1 and D2 to our credit experimental approach [8]. After reading different researcher methods we conclude that the following methods are slow and

difficult, we introduce new methods for classification of different kinds of trees which is easy and faster. Our work is mainly based on texture and histogram features for classification of bark of trees.

1.1   Problem Statement

Lack of knowledge about trees to identify their varieties for a layman. For the purpose of forming is a difficult task to identify tree varieties. Human perception is the difficult task to identify different varieties of trees. So we have introduced machine intelligence to recognize tree varieties using different algorithms. This research is to identify the different varieties of trees classification by using K-nearest neighbor and nearest neighbor algorithms.

1.2   Objective

The goal of this paper to estimate methods to classify different kinds of trees from images of bark through K-nearest neighbor and nearest neighbor algorithms. The classification will be performed on the image data that was previously gathered from five kinds of trees. The architecture function of tree varieties imaging the texture and histogram features of tree variety classification imaging base on tree features: color, shape, and dead. The tool is developed in CVIP tool to obtain the result. By implementing the purpose method other tools are available like MATLAB, Mazda.

2.   Related Work

Porebski et al. described that the selection based LBP histogram score color texture classification. Consists in assigning to each Histogram score which measures its effectiveness to characterize the textures appear within the various classes [9]. Kaur and Gupta, expressed that texture classification for do to have used as Brodatz online texture that's database advanced database and three classifiers well known: Support Vector Machine, K-nearest neighbor method Induction and decision tree method [10]. Ougiaroglou et al. explained that the classification based on k-nearest neighbors (KNN classification) is one of the most widely used methods of classification. The number k nearest neighbors used to achieve high accuracy in the classification is given in advance and it is highly dependent on the data used [11]. Todorovic and Ahuja, described a statistical approach to texture 3D classification from a single image obtained under unknown perspective and lighting [12]. Khan et al. explained that the description of the different techniques that classification, such as Decision Tree Induction, Bayesian classification and Neural Networks [13]. Zhao et al. described that there are many ways offered to recognize plants. They expressed the SVM method for classification. Most of the leaves are used as index for the recognition, and some ways to get satisfied results. The detection methods by leaves some have the accuracy of the recognition, the accuracy is still much room to improve. Barks as an additional feature on the leaf characteristics to identify those plants were added. Tree bark texture of any specific feature and the skin is easy achieved [14]. Mzoughi et al. described that the leaves of plants can classify as either simple or connection to it forms. Compound leaves as a collection of simple sheet-like structures called apparent leaflets. However, most computer vision-based approaches describe these two leaf categories similar. A new description and identification process for compound leaves, which takes into account particularities related accepts on the arrangement of its forms. In fact, a new multi- leaflets-based identification approaches [15]. Ramakrishnan and Neelakanta, Described that a parallel study is similar to the effort, instead of coefficients Haralick, a method inspired by bioinformatics. Yet another new song suggested here as an open research question on the interpretation of patterns of DNA the use of artificial neural network (ANN) with a prescribed amount of a teacher of similar Haralick. Finally, as regards the cost of computing, it is understood that the present method as GLCM decision made compatible with the pixel matrix of the image. For the accuracy of the classification, a large number of pixels in a digital image must be made for greater clarity [16]. Vijayalakshmi and Bharathi, Described that feature extraction for tissue analysis is done with the moment Legendre, local Pictures binary pattern, texture spectrum, the spectrum of the moment Legendre, LBP with Legendre moment and proposed. Experimental evaluation shows that local binary Space is able to extract features and patterns of discrimination it have promising performance for various tissues [17]. Taylor et al. described that the bark features can be very useful in tree classification specifically when leaves are vague and branches unreachable however, bark can vary importantly with age, growth rate and habitation. Key identifying features include: texture, color and color patterns, and color and size of variously formed spots on young bark. Branches, twigs, bud, and leaves grow at specific locations on the tree called nodes. Many species grow in an ALTERNATE pattern with one flower or leave per node relatively few classes have leaves or buds that occur in sets at each node. Fewer classes still grow with a WHORLED pattern with three or more structures at a node [18]. Ali et al. Described a tree-based classification detection method for autonomous navigation of vehicles forests in the forest. A mix of color and texture sign to the tree and the background image to be used Objects. Color and texture sign combined to yield better performance than individual symptoms can accomplish. Compilation of LBP (local binary pattern) and color histogram features has shown an excellent performance. Easy Heuristic method is used to measure distance

the estimated distance between the base of the tree and Car tire to tire width in pixels and as a basis reference [19]. Masotti and Campanini, Described that a new method of classification is presented. The study ranklet conversion, image processing method based on the analysis of the relative ranking pixels and their gray-scale value. Some of the features are extracted from the images ranklet Program into different resolutions and orientations of ranklet the image 90 degree rotation invariance is achieved with an average for each resolution correspondent, horizontal, vertical, diagonal and texture characteristics. Finally, a member of the class of tissue using a support vector of dedicated features sport vector machine (SVM) classification. Compared with the three methods and having the same Brodatz and Vistex data sets analyzed, the proposed method work best [20].

3. Methodology

Research will dictate the kinds of research methodologies use to underpin work and methods use in order to collect data. There are different types of research methodology such as quantitative and qualitative research etc. In our research data will be collected physically. It is a quantitative research. We collect quantitative data then probably measuring variables and verifying existing theories or hypotheses or questioning them.

We take pictures of the bark of a trees with digital camera, according to get on the front of the human eye. After taking pictures of different trees bark we use these pictures for different processes and collect which we called the train data set of data. We have five different varieties of trees and taking pictures from the bark of trees of these varieties. Each variety has 10 images and has a total of 10 * 5 = 50 image of bark of trees. All the 5 categories names suppose as **A, B, C, D** and **E**, the one image of each category will be show in the figure 1.

Figure 1: Original Bark Images of Trees

In above figure the all images get through the digital camera. These are all the original images of trees which noised include. Now we can crop the all image one by one and remove the noises in the each images then it will be used in CVIP tools. Crop images will be show in the figure 2, in this figure the one image of each kind of tree will be taken such as

Figure 2: Clear Bark Images of Trees

After removing the noises in the images and clear the bark images we adopt following figure for next process of images in image processing. In which process images will be processed shown in figure 3 and also explained the process.
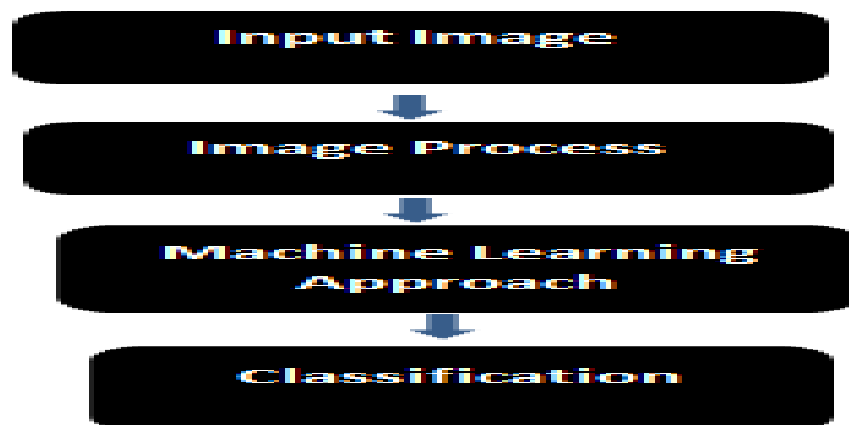
Figure 3: Image Process

It will be explained such as:

- Input Images

In this stage we receipt the pictures through digital cameras and then crop each image one by one. In an input image the all the images will be arranged into a sequence category wise and take one by one for process and getting the result.

- Processing images

After input images, the images one by one send for process. In processing the noises will be removed and images make a clear and then process in the CVIP tools for getting the result. In a CVIP tools different tools will be used for classification.

- Machine learning approach

After processing images, the machine learning approach will be adopted. In a machine learning approach we use the K-nearest neighbor and nearest neighbor algorithms for classification. After adopting these two methods we get the final result from the images as categories wise.

- Classification

After adopting the machine learning approaches the different result will be collected from different images. The following results are the final results. In this way we can categorize the trees used these trees bark pictures and categorized the trees.

1. CVIP Tools

CVIP tool means computer vision and image processing tool. It is an image processing software open source. It is an interactive program for image processing and computer vision. This tool is free for windows and older versions of UNIX. CVIP tools are Windows-based software for computer vision and image processing developed at the Computer Vision and Image Processing Laboratory at Southern Illinois University in Edwardsville.

CVIP tools can read many image formats, including GIF, JPEG, BMP, TIFF, PNG, and raw formats. CVIP tools and supports standard image processing functions such as image restoration, image compression, arithmetic and logical operations between images, image sharpening, contrast, and manipulation, and the frequency shift, edge detection, segmentation and geometric transformations.

After open this tool we upload the image for any purpose in this tool using open the image option
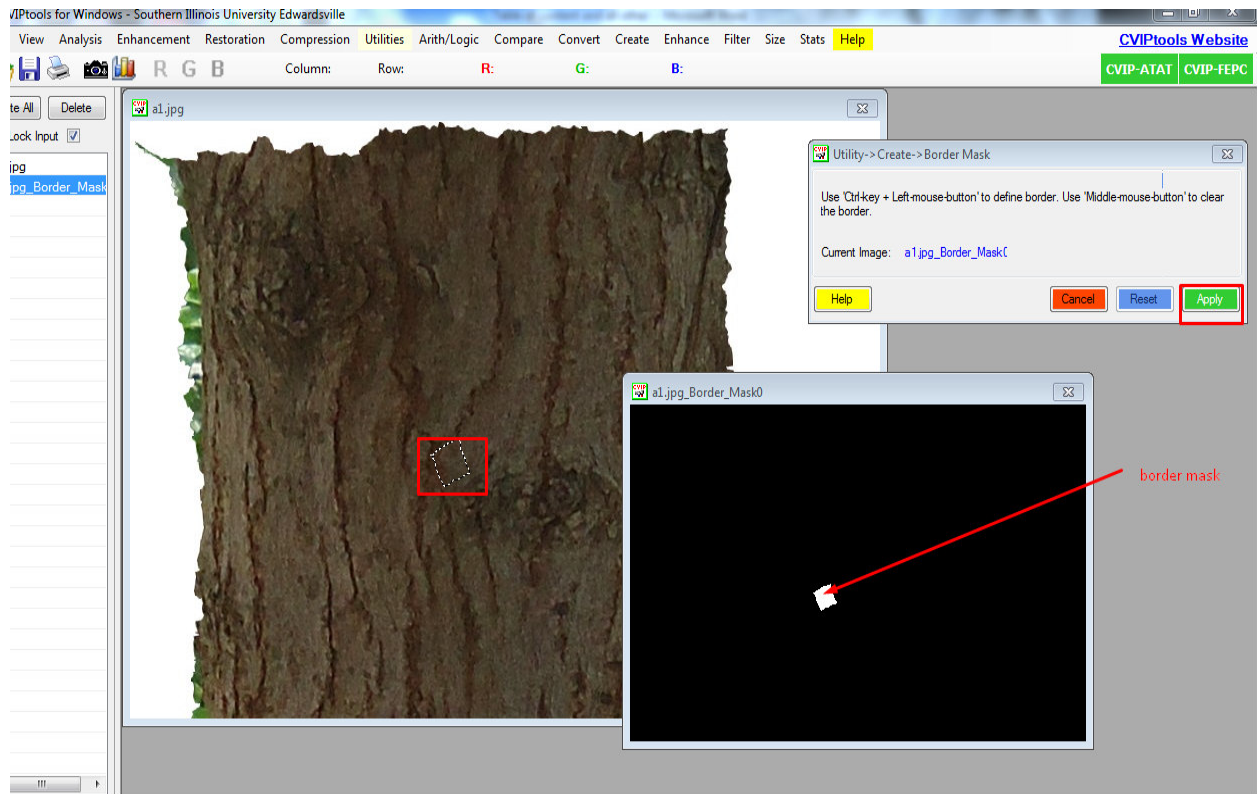

Figure 4: Taking the Border Mask

which is available in the tool. Now this image is used for any result or purpose. Our work depend on two steps, first we can border mask of each image and then take the histogram features and texture features results for collect the train and test data results. Then we can compare train and test data results using nearest neighbor and k-nearest neighbor algorithm for pattern classification of bark of trees.

In above figure the border mask of any image will be shown. When we take the border mask the
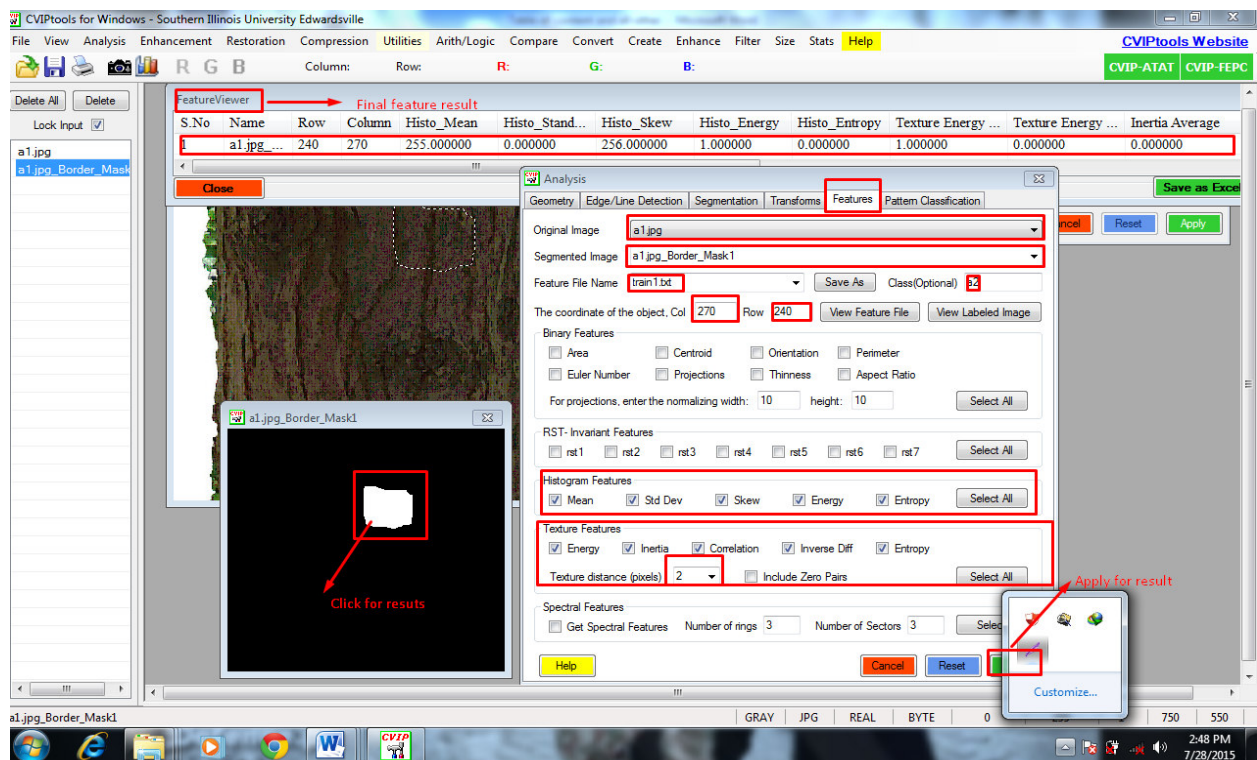
Figure 5: Applying and Taking the train, test Data Features Results

following method is used for taking the border mask of different images. Then these border mask will be used for different results.

- **Histogram features**

In a CVIP tool we choice the histogram feature. In a histogram feature we choice the all histogram feature such as mean, Std Dev, Skew, Energy and Entropy. We develop all the results of each image of bark of different trees.

- **Texture features**

In CVIP tool features we also select the texture features. In a texture features we select all the texture features such as Energy, Inertia, Correlation, Inverse Diff and Entropy. In this way we catch all the results of texture features using all categories of images. In this way we collect train and test data of all images of bark of different trees.

In above figure the test data features results method shown. For test data features results we use following shown options.

2   Pattern Classification

Then these results of train and test data sets used for pattern classification compare train and test data sets and applying the nearest neighbor and k-nearest neighbor for final classification such as shown in figure. When we compare the train and test data results the final pattern classification results achieved which are best for classification of sunflower seeds
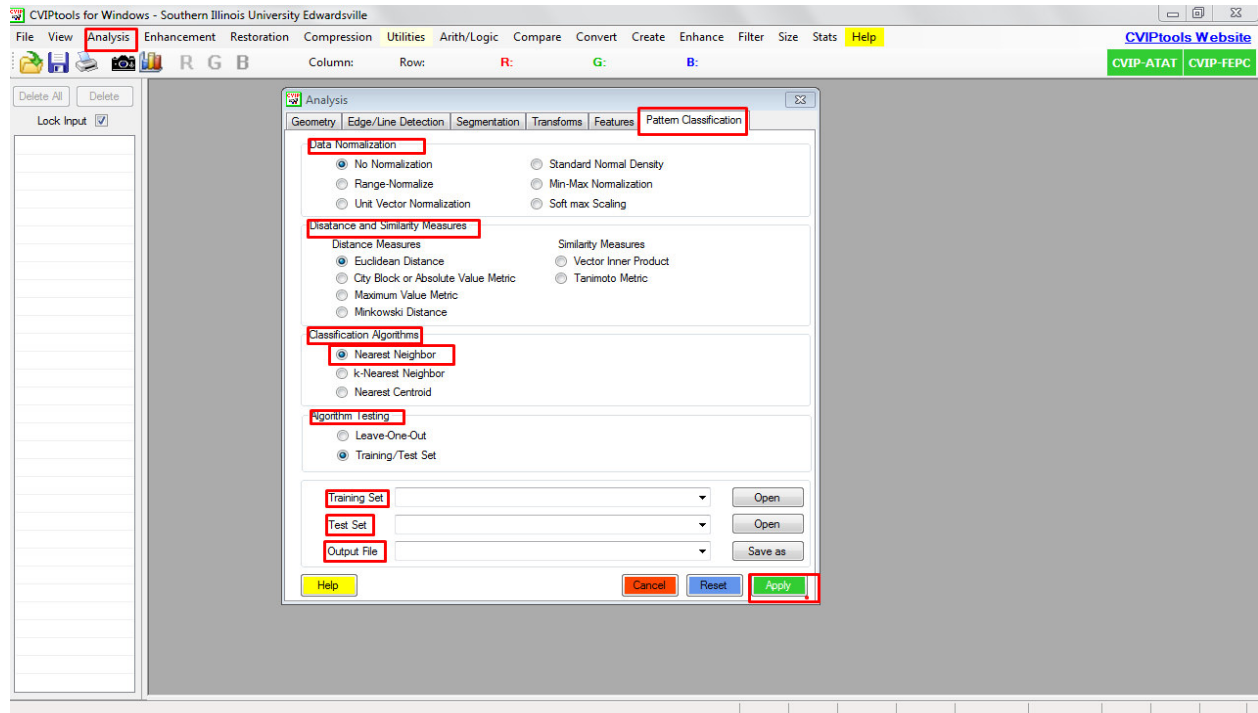
www.iiste.org

Figure 6: Pattern Classification Using Nearest Neighbor

When we take the pattern classification using training and test data results comparison. We select following shown options in the tool.

For pattern classification of k-nearest neighbor we select k-nearest neighbor in algorithm options. Other options of tool will be remain same as nearest neighbor. In a k-nearest neighbor we can change the k value for taking the different results of comparisons.

4.   Results And Discussions

In this paper it will be discussed that we can get bark images of different five varieties of trees. Five varieties suppose names as A, B, C, D, E which ten images of each kind will be taken. After taking the images it will be used in CVIP tool for border mask and taking the result. In a CVIP tool the each image of each kind of tree will be divided in five border mask and getting different results in training data. In a training data the total 5*10=50*5=250 border mask and results will be taken.

Each border mask provides different result which is used for bark classification of trees. The histogram and texture features will be taken of each border mask. Some of the histogram feature and texture features results will be shown in tables as samples. The histogram and texture have a different more features but there we take some features as a sample and understanding the results of bark images of trees.

When we complete the all results of training data set and test data set according to discussed above. Then we can compare train and test and get the pattern classification results. After comparing these results it will be provide better classification rate for nearest neighbor and k-nearest neighbor. These all work will be performed in CVIP tool such as shown below:

It will be showing the nearest neighbor classification results of bark images. After comparisons of different bark images train and test results. It also shown in the snapshot the all category results are different from each other such as a category, b category, c category, d category and e category. These are the final result of nearest neighbor.
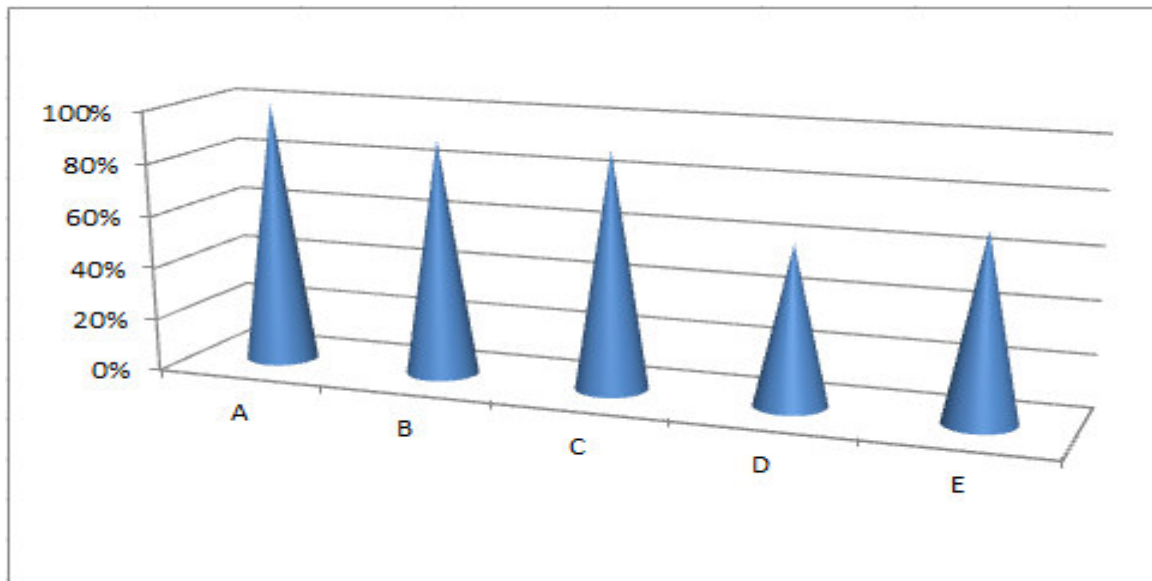
Figure 7: Nearest Neighbor Classification Result

In above diagram the all kind's results will be shown with their final result percentage. These are the all results are best for classification. These are the nearest neighbor final results which A=100%, B=90%, C=90%, D=60% and E=70% which average percentage is 82%. It is the best percentage result for classification of bark image.

Now we go to the k-nearest neighbor results using pattern classification. We will some changes in the analysis. There is we change use the k-nearest neighbor using k=2. Other training set and test set will remain same as use in nearest neighbor. And the other change here is output file name which we use the result 2 of k-nearest neighbor and then apply for final result.

After taking some changes in the previous process of nearest neighbor, we press apply for final result of k-nearest neighbor using k=2, the following final results appear. It also it will be seen there is different results appears a category, b category and so on.
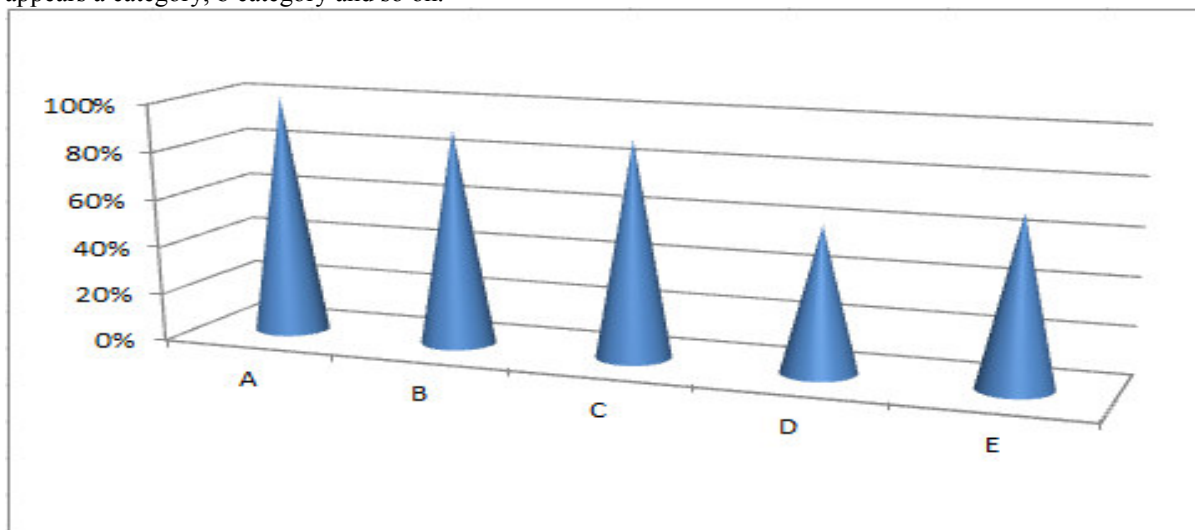

Figure 8: K-Nearest Neighbor Classification Result Using K=2

In the above diagram the final pattern classification results will be shown with its names and percentage. The average percentage here is also is 82%, which is also best for classification.

Now we use the k=3 for final and different results of k-nearest neighbor. It also shown in the above table other things are remaining same and apply for result. When we change the k value it provides different result from other k values. There we also change the output file name. When we press on apply for final result of k-nearest neighbor k=3 the following results appear, which are the different from k-nearest neighbor k=2, because there is k will be change. These percentages are also best for classification which average percentage is also is 82%. It is the best percentage result.
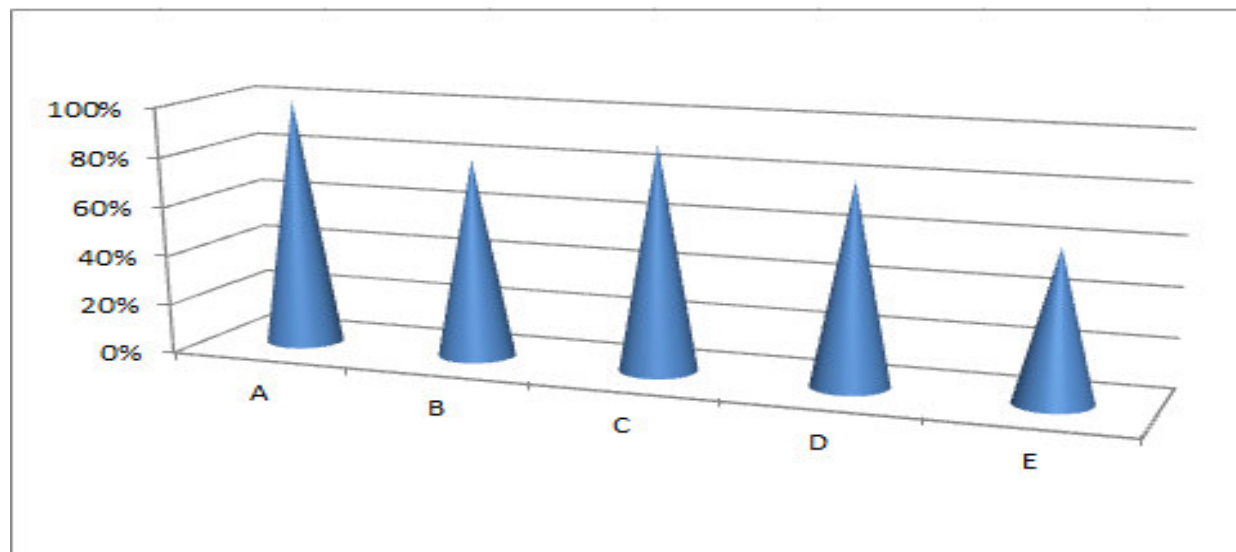
Figure 9: K-Nearest Neighbor Classification Result Using K=3

When we change the k value then the different result will be appeared. In above diagram it will be shown the different result provide different percentage on A, B, C, D, and E. these are the best percentage rate for achieving correct classification.

In above it will be shown there we use k-nearest neighbor k=4 for taking the other results, when we use the k=4 the other results of k-nearest neighbor appears. These all results are different from each other on changing the k values.

After taking some changes in k then the new and different results of k-nearest neighbor will be shown which are different from k=2 and k=3. There is the different category of trees provide different results on comparing the train and test datasets results. There the average percentage is 76% which is also best for classification of bark images of trees
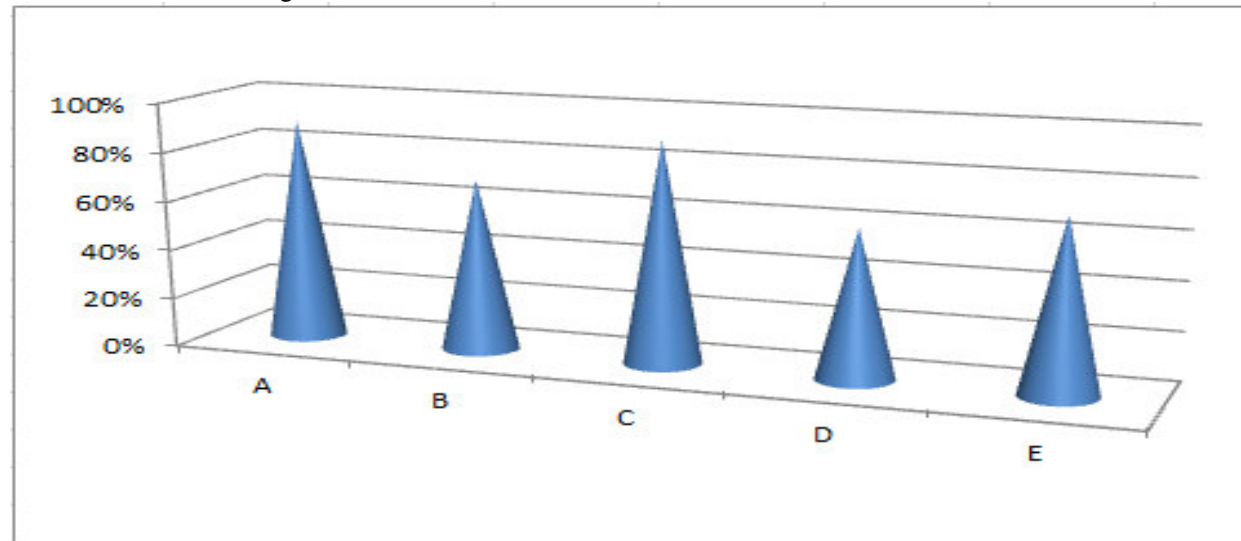

Figure 10: K-Nearest Neighbor Using K=4

The all the kinds of bark images of trees provide different result for classification. Here is average percentage 76% which is best percentage for better classification. For more checking the result we now use the k-nearest k=5 for see other result. After taking the k=5 the different result will be shown and provide best percentage rate for classification. When we use k=5, then the following results will be achieved.

Taking some changes in the k the results will be change appears as above. The different category of different trees provides different result. In above it will be seen the average percentage is 72% of all category which is also best for classification.
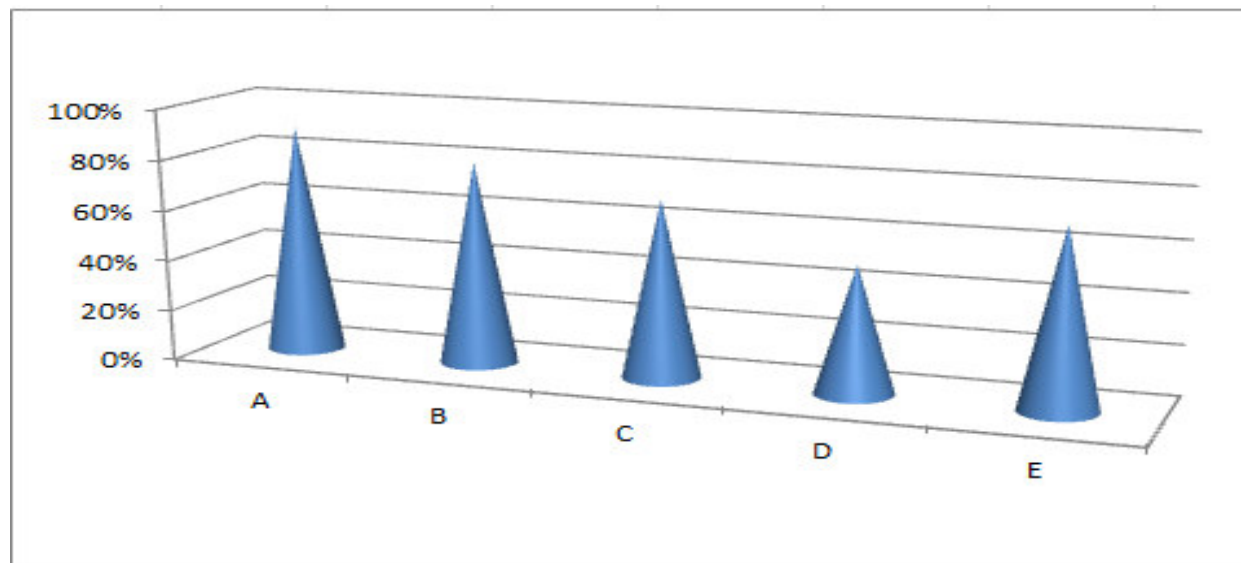
Figure 11: K-Nearest Neighbor Classification Result Using K=5

When we use the k=5, the following results percentage will be appeared on different categories. The average percentage is 72%, which is also best for classification.

When we some changes in above all results such as k values change or distance change, the different results will be achieved. The all results rate are best for classification. In this way we can classify any bark image using these methods.

Finally we hope that the nearest neighbor and k-nearest neighbor provide all result different from each other and it will be very helpful for bark classification of trees using texture features. These are all the different features and results. If we change any features of any image it provide also different results from other feature

5.   Conclusions

In this paper, we investigated the evolutionary approach for generating texture characteristics. The results of the experiment indicate that both methods makes the prior art for the classification of tree bark using texture features and the best results announced in most standard challenges texture data set. In the classification problem we Bark 100% accuracy in the category, and 90% of the class B, and 90% of category C, 80% of the class D, E, 70% by using the nearest neighbor, which is the best result for the classification of the crust. On the other hand, when we use the nearest neighbor as the title will be achieved the following results in the category of accuracy of 90%, 80% BC 70% 50% D 70% E correction achieved. It is being used in the near K = 5, so it's the best results that can be used as a camera digital, which is simple for the user to take a photo of tree bark. Even experiments using samples of only 10 in the training class do better accuracy, exceeding the best results reported in precision experiments to recognize the bark.

His willingness to common problems of the fabric. Our nearest neighbor average accuracy is 86% and k- nearest neighbor is 72% using K = 5, and the best way to classification. We have shown that the nearest neighbor and nearest road K very fast and suitable neighbor for applications in real-time.

6.   Limitations and Recommendations

In limitations we challenged lack of advance machineries tools and other environmental elements. In this manner if the limitations are resolved we can easily classify the different kinds of trees and also these methods provide opportunity to farmer and other people for identify and select the different better facilities through different kinds of trees. It's not limited to the application of digital cameras, as well as in mobile applications.

7.   Acknowledgement

References:

[1] Song, J., Z. Chi., J. Liu & H. Fu (2004). Bark classification by combining gray scale & binary texture features. *In intelligent Multimedia, Video & Speech Processing. Proceedings ofInternational Symposium.* pp, 450-453.

[2] Wendel, A., S. Sternig & M. Godec (2011). Automated identification of tree species from images of the bark leaves and needles. *In16th Computer Vision Winter Workshop* pp, 67.

[3] Sebastian, V., Unnikrishnan, A & Balakrishnan, K (2012). Gray level co-occurrence matrices: generalisation and some new features. *arXiv preprint arXiv:pp,120-4831.*

[4] Olaode, A., Naghdy, G & Todd, C (2014). Unsupervised Classification of Images: A Review. *International Journal of Image Processing (IJIP)*, vol 5,no *8*,pp, 325.

[5] Sulc, M & J. Matas (2014). Texture-Based Leaf Identification. *In Computer Vision-ECCV 2014 Workshops* pp, 185-200.

[6] Ko, C., Sohn, G., Remmel, T. K & Miller, J (2014). Hybrid Ensemble Classification of Tree Genera Using Airborne LiDAR Data. *Remote Sensing*,vol 11, no *6*, pp, 11225-11243.

[7] Mahajan, P. M., Kolhe, S. R & Patil, P. M (2013). Classification of Texture Images Using Multiresolution Transform. *International Journal of Advanced Research in Computer* and *Communication Engineering*,vol 8,no *2*,pp, 3171-3175.

[8] Othmani, A., A. Piboule & L. F. C. Lew Yan Voon (2013). Hybrid segmentation of depthimages using a watershed & region merging based method for tree species recognition. *In IVMSP Workshop, IEEE 11th* pp., 1-4.

[9] Porebski, A., Vandenbroucke, N & Hamad, D (2013). LBP histogram selection for supervised color texture classification. In *ICIP*, pp, 3239-3243.

[10] Kaur, A & Gupta, S (2012). Texture Classification Based On Gabor Wavelets.

[11] Ougiaroglou, S., Nanopoulos, A., Papadopoulos, A. N., Manolopoulos, Y & Welzer-Druzovec, T.(2007).Adaptive k-nearest-neighbor classification using a dynamic number of nearest neighbors. In *Advances in Databases* and *Information Systems* pp, 66-82.

[12] Todorovic, S & Ahuja, N (2006). 3D texture classification using the belief net of asegmentation tree. In *Pattern Recognition, ICPR. 18th International Conference on* Vol. 4, pp, 33-36.

[13] Khan, M., Ding, Q & Perrizo, W (2002). K-nearest neighbor classification on spatial data streams using p-trees. In *Advances in Knowledge Discovery* and *Data Mining* pp. 517-528.

[14] Zhao, Z., X. Huang. and G. Yang (2015). Plant Recognition Based on Leaf and Bark Images. *Journal of Computational Information Systems*, Vol.11, N0.3, pp, 857-864.

[15] Mzoughi, O., I. Yahiaoui, N. Boujemaa. and E. Zagrouba (2014). Multiple leaflets-based identification approach for compound leaf species.

[16] Ramakrishnan, K. and P. S. Neelakanta (2013). DNA Microarray Data Classification via Haralick Parameters.

[17] Vijayalakshmi, B and V. S. Bharathi (2011). A Novel Approach to Texture classification using statistical feature. *AR Xiv preprint ar Xiv*: 1111.2391.

[18] Taylor, E, R., M. Kirk. and G.Yang (2009). Techniques to identify trees. Damage Prevention and Disaster Recovery Tree Care Kit, Texas Agri Life Extension Service fact sheet 1.2.

[19] Ali, W., F. Georgsson. and T. Hellstrom (2008). Visual tree detection for autonomous navigation in forest environment. *In Intelligent Vehicles Symposium, 2008 IEEE* pp. 560-565.

[20] Masotti, M. and R. Campanini (2008). Texture classification using invariant ranklet features. *Pattern Recognition Letters*, *29*(14), 1980-1986.