IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-2, Feb 2019 E-ISSN: 2347-2693

Classification of Healthy and Diseased Arecanuts using SVM Classifier

H. Chandrashekhara[1*], M. Suresha[2]

[1,2]Department of Computer Science, Kuvempu University, Shimoga, India

*Corresponding Author: chandrashekhara5595@ gmail.com, Tel.: +919739993080

DOI: https://doi.org/10.26438/ijcse/v7i2.544548 | Available online at: www.ijcseonline.org

Accepted: 16/Feb/2018, Published: 28/Feb/2019

***Abstract*—** Arecanut is the seed of the areca palm (Areca catechu), Arecanut palm is one of the important commercial crops in India. Majority of arecanut are produced and consumed by Indian populations when compared to other countries. This paper proposes, to Classify Healthy and Diseased Arecanut images. In this paper Healthy and diseased arecanut are have been done. Structured matrix decomposition model (SMD) is used to segment the images and LBP features are extracted using SVM classifier. Experimental results demonstrate proposed method perform well and obtained accuracy of 98%.

***Keywords*—** Arecanut Images, SMD, SVM Classifier.

I. Introduction

Precision Agriculture for the most part characterized as information and innovation-based homestead administration framework to understand examinations and administer spatial and not unusual change inner fields for ideal effectiveness and efficiency, supportability and confirmation of the land useful resource by means of constraining the introduction fees. Expanding ecological cognizance of the overall population is requiring us to regulate horticultural administration for maintainable upkeep of feature property, for instance, water, air and soil best, at the same time as last financially efficient. The essential issue related with the utility of machine vision strategies is that concerning the image segmentation, feature extraction, and classification are the maximum vital responsibilities in this work. In Precision Agriculture Arecanut place significant role in Indian economy. Arecanut is one of the most essential industrial plants in India. India dominates international production of arecanuts generating all maximum half of the production the world. Arecanut plays considerable position inside the livelihood of the human beings. Arecanut is the traditional ayurvedic medicinal properties contains effective against leucoderma, leprosy, cough, fits, worm's anemia, obesity and also used against certain skin diseases. Dry nuts are crafted from fresh of fruit culmination. It is used by the human beings all around the world. For chewing reason in gentle, ripe or processed form. It performs a prominent function- in spiritual, and cultural life regardless of caste and social reputation. India is the principal manufacturer and client of arecanut within the world. The financial produce is the fruit called betel nut, which is used mainly for masticatory purpose and arecanut and betel leaf is a good remedy against bad breath.

In the rest paper, brief introduction of arecanut discussed in section-I. Related work briefly in section-II.Proposed methodology shown in section-III and experimental results illustrated in section-IV. Finally concluded in paper in section-V

II. Related Work

In [1], proposed classification of diseased and undiseased arecanut have been determined using texture feature of Local Binary Pattern (LBP), Haar Wavelets, GLCM and Gabor. Authors proposed in two stages. In the first stage, LBP have been applied on each color component of HSI and YCbCr color models and histogram of LBP is generated. The statistical method correlation is used to measure the distance between histogram of training set and query sample and obtained a success rate of 92.00%. In [2], proposed work, of neural networks and image processing techniques for detecting and classifying the quality of arecanuts. Defects with diseases or insects of arecanuts were segmented by a detection line (DL) method. Six geometric features i.e., the principle axis length, the secondary axis length, axis number, area, perimeter and compactness of the areca nut image, three color features i.e., mean gray level of an arecanut image on the R, G, and B bands), and defects area were used in the classification procedure. Authors claim 99.05 % of success rate of classification. In [3], propose recognition and classification of White Wholes (WW) grade cashew kernel using artificial neural networks. Classify cashew kernels by using Artificial Neural Network (ANN). They have work on two phases. In first phase, built with a proposed method to extract features, using 16 morphological features. In phase two, a Multilayer Perception ANN used to recognize and classify the given white wholes grades using back propagation learning algorithm and achieves a classification accuracy of 88.93%. In [4], proposed structured matrix decomposition (SMD) model, which formulates the task of salient object detection as a problem of low-rank and structured sparse matrix decomposition. A hierarchical tree-structured sparsity-inducing norm has been proposed to encode the underlying structure of the image in the feature space, while a Laplacian regularization has been introduced to enlarge the distance between the representation of salient objects and that of the background. High-level prior knowledge has also been integrated into the model to enhance the detection. Experiments on five public datasets they have modeled achieves encouraging performance compared to the state-of-the-art methods. In [5], efficient fruit detection using multiple feature-based algorithms is developed and they have proposed. Multiple features like intensity, color, edge, and orientation are analyzed. It computes the feature map for different type of feature points and according to the feature map the fruit regions are extracted. The process is entirely automatic and does not need user intervention. The proposed method is not domain-specific and does not impose limits on the variety of clustered sectional tree image. It can be used for all kind of images provided that there are at least one or more meaningful fruit regions. A simple feature cannot entirely represent the character of the fruit region. Therefore, multiple features analysis is used in the proposed method. that paper, authors consider colors, intensity, edge and orientations as the features of the image. However, it is very likely that there are some other features such as symmetry features also should be considered and accuracy results achieved up to 90%.Effective grading of arecanut using SVM has been discussed in [6], Their model consists of mainly three steps in which first step addresses the segmentation process to segment the arecanut image from the given images using color transformation. In the second step, color features were extracted and SVM classifier is used for classification purpose. Effective Multi-classifier for Arecanut grading based on color has been proposed In [6], and claimed 98% of success rate for their method.

III. PROPOSED METHODOLOGY

(Ajit Danti and Suresh), have conducted field survey to create a dataset of arecanut about thirty agricultural fields fifteen tender markets of arecanut and found different verities of arecanut. As authors found that there are few local varieties of arecanut are considered for this work and proposed work shown in figure [1].

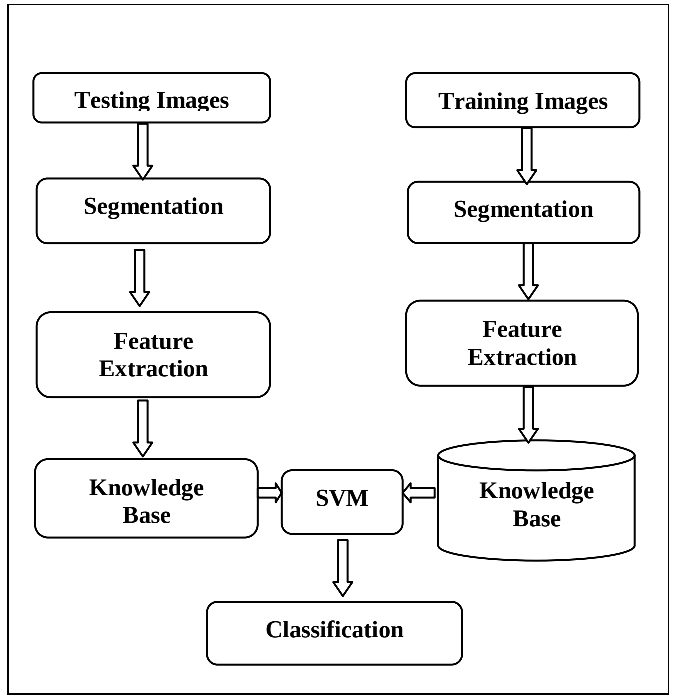


Figure 1: Block Diagram of the proposed work

A. Segmentation

In our work, classified healthy and unhealthy of arecanut. Segmentation achieved by adapting well-known structured matrix decomposition method (SMD) is used and it carried-out in five stages.

Stage-1: Image Abstraction- In this stage, extract the low-level features (F), with RGB color, steerable pyramids and Gabor filter, to construct a 53-dimension feature vector. At that moment, perform the simple linear iterative clustering (SLIC) algorithm. In [4], to over-segment the image into N number of patches (super pixels) $P = f\ P_1; P_2;....P_{Ng.}$

Stage-2: Tree Construction- On top of P, an index tree T is constructed to encode structure information via hierarchical segmentation. To this end, we first compute the affinity of every adjacent patch pair.

Stage-3: Matrix Decomposition- In this stage, decomposition of feature matrix F and the index tree T.

Stage-4: Saliency Assignment-After decomposing F, we transfer the results from the feature domain to the spatial domain for saliency estimation. Based on the structured matrix S.

Stage-5: Finally, Otsu method used for segmentation

B. LBP Features

There exist several methods for extracting the maximum beneficial capabilities from Arecanut images to perform

Diseased Arecanut. One of these function extraction methods is the Local Binary Pattern (LBP) method. Relative new approach was introduced in 1996 by Ojala et al. [7].

The basic local binary pattern operator, introduced by Ojala et al. [7], was based on the assumption that texture has locally two complementary aspects, a pattern, and its strength. In this work, the LBP was proposed as a two-level adaptation of the texture unit to describe local textural patterns.

The original edition of the local binary pattern operator works on a 3 × 3 pixel block of an image. In [8], The pixels in this block are thresholded by its centre pixel value, multiplied by powers of two and then summed to obtain a label for the center pixel. As the neighbourhood consists of 8 pixels, a total of 2^8 = 256 different labels can be obtained depending on the relative gray values of the center and the pixels in the neighbourhood. [7] while processing an image, each processing pixel is compared with its 8 neighbours and the ones whose intensities exceed the processing pixels are marked as 1, otherwise as 0. In this way we get a simple circular point features consisting of only binary bits. Typically the feature ring is considered as a row vector, and then with a binomial weight assigned to each bit, the row vector is transformed into decimal code for further use. In [13],LBP using circular neighbourhoods and linearly interpolating the pixel values allows the choice of any radius, *R*, and number of pixels in the neighbourhood *P*, to form an operator, which can model large scale structure. In basic LBP operation threshold value is the processing pixel, this method is noise sensitive, so we choose average of pixels including processing pixel that encompass by an LBP operator as a threshold. [1], Mathematical model for LBP is shown in equation (1)

$$LBP_{p,r(x,y)} \sum_{p=0}^{p=1} s(g_p - g_c)2^p \tag{1}$$

where g_c is the average gray value of the pixels encompass by LBP operator, g_p is the intensity value of pixels in eight neighbourhood. A descriptor for texture analysis is a histogram, $h(i)$ of the local binary pattern shown in equation (2) and its advantage is that it is invariant to image translation[13].

$$h(i) = \sum\nolimits_{x,y} B(LBPP, R(x,y) = i) \mid i \in [0, 2^{p-1}],$$

$$B(v) = \begin{cases} 1 & v > T \\ 0 & otherwise \end{cases} \tag{2}$$

C. SVM Classifier

Support Vector Machines (SVM) were invented by Vladimir Vapnikin 1979 [9]. Basically the SVMs separate two different classes through hyper planes. If the classes are separable by hyper planes an optimal function can be determined from the empirical data. In the hyperplane is expressed by its normal vector *W*and a bias *b* [10]. The class of hyperplanes can be specified in the scalar product space *H* (feature space) as follows

$$\langle w, x\rangle + b = 0 \quad where \ \ w \in H, b \in R \tag{3}$$

Where $\langle w, x\rangle$ means $x_1 w_1 + ... + x_n w_n$ this yields the corresponding decision function

$$f(x) = \operatorname{sgn}(\langle w, x\rangle + b) \tag{4}$$

Where the sign function extracts the sign of a real number. It is defined as -1 if $f(x) < 0$ and 1 if $f(x) > 0$ which denotes the two different class labels +1 and -1. Usually there exist many hyperplanes which separate the two classes. The basic idea behind SVMs is that the optimal hyperplane maximizes the margin between data sets of opposite classes. In order to construct the optimal hyperplane, the following equation has to be solved.

$$\frac{\min}{w \in H, b \in R} \Im(w) = \frac{1}{2} \quad ||w||^2 \tag{5}$$

Subject to $y_i(\langle w, x_i\rangle + b) \geq 1 \forall i \in \{1,....,n\}$. (6)

The constraint (6) ensures that $f(x_i)$ yield +1 for y_i = + 1 and -1for y_i = -1 and that the two classes are separated correctly. If ||*w*||=1 the left hand side of (6) is the distance of the training sample xi to the hyperplane. This is the Hessian normal form representation of the hyperplane. The distance of each training sample to the hyperplane can be computed by dividing $y_i(\langle w, x_i\rangle + b)$ by ||*w*||.The overall margin is maximized if the constraint (6) is satisfied for all $i \in \{1,....,n\}$. with *w* of minimal length, as given in (5). The distance of the closest point to the hyperplane is 1/ ||w||This can be illustrated by considering two training samples, one of each class respectively, and by projecting them onto the hyperplane normal vector $\frac{w}{\| w \|}$.

The formulas (5) and (6) specify the constrained optimization problem. It can be transformed to a ‘dual ‘problem; where *w* and *b* are eliminated by introducing Lagrange multipliers α_i.

$$\frac{\max}{\alpha \in R^n} w(\alpha) = \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \langle x_i, x_j\rangle \tag{7}$$

Subject to $\alpha_I \geq 0 \forall i \in \{1,....,n\} \ and \ \sum_{i=1}^{n} \alpha_i y_i = 0$ (8)

This leads to the following decision function

$$f(x) = \operatorname{sgn}\left[\sum_{i=1}^{n} y_i \alpha_i \langle x, x_i\rangle + b\right] \tag{9}$$

Where b can be computed by the Lagrange multipliers which do not equal zero. These are called support vectors. All other samples with α_i =0 are discarded.

Up to now only linearly separable classes were considered, but SVMs are able to classify samples with a non-linear discriminate with a technique called kernel trick. A kernel is a function which takes an input of lower dimensional space and transforms it to a higher dimensional space i.e. it converts not separable problem to separable problem, some of the well-known kernels are polynomial of degree d, Gaussian radial basis function, Laplacian, Sigmoid and etc. The basic idea of SVMs is to map the data into a new feature space and then solve the constrained optimization problem. Obviously it seems to be very expensive to compute the mapping into a high-dimensional space. For this reason a kernel function is introduced to make the computation very simple [11],. This is referred to as the 'kernel trick', which causes an implicit mapping in the feature space without explicitly knowing the mapping function ϕ. Accordingly the scalar product can $\langle x, x_i \rangle$ be substituted by

$$k(x, x_i) := \langle \phi(x_i) \rangle = \langle x, x_i \rangle \qquad (10)$$

So far, we have made the implicit assumption that the datasets are free of noise and may be classified perfectly. In this case SVMs give 'hard margins'. In practice, this assumption does not hold true in most cases. This problem, however, may be handled by 'soft margins' which allow and penalize classification errors. Accordingly, in 1995 modification was introduced where slack-variables ξ_i are used to relax the so-called hard-margin constraints (6) In 12], so that some classification errors depending on ξ_i are allowed. The influence of the classification errors are parameterized with the parameter C. A larger C penalizes a wrong classification more strongly.

IV. EXPERIMENTAL RESULTS

Experiment conducted for identification of healthy and diseased arecanut are performed, here arecanut images are collected [1]. The arecanut images are resized to 255*255 size of pixel resolution to improve the algorithm accuracy and segmentation achieved by SMD method and local binary features (LBP) are extracted features and SVM perform as a classifier for the classification of arecanut. Among the total images 60% of them used for training purpose and 40% of images used for testing and proposed method is obtained success rate 98% of accuracy shown in table.1.

Table. 1:Confution matrix for proposed work.

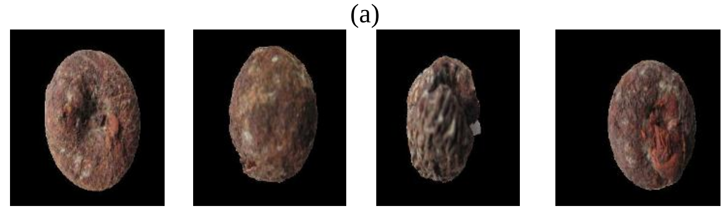
	Healthy	Diseased
Healthy	25	0
Diseased	1	24

(a)

(b)

Figure 2: Sample arecanut images (a) Input images (b) Segmented arecanut images.

(a)



(b)

Figure 3: Sample arecanut diseased images (a) Input images (b) Segmented arecanut images.

V. CONCLUSION

In this proposed work classification of Healthy and Diseases arecanut is carried out. In this method, segmentation of arecanut using existing method such as structured matrix decomposition (SMD) method. LBP features are extracted features from both training and testing samples. Using SVM classifier for the classification and obtained success rate 98% of accuracy.

REFERENCES

[1] M. Suresha, A. Danti and S. K. Narasimhamurthy, "*Classification of Diseased Arecanut based on TextureFeatures*", 2014.

[2] S. H. Tsai, Y. H. Tseng, " *A novel color detection method based on HSL color space for robotic soccer competition*", Computers & Mathematics with Applications, 64(5), 1291-1300, 2012.

[3] N.V , Ganganagowdar, H. K Siddaramappa "*Recognition and 'classification of White Wholes (WW) grade cashew kernel using artificial neural networks*", ActaScientiarum. Agronomy, 38(2), 145-155., 2012, 2016.

[4] H. Peng, B. Li, H. Ling, W. Hu, W. Xiong, & S.J. Maybank, "*Salient object detection via structured matrix decomposition*.", IEEE transactions on pattern analysis and machine intelligence, 39(4), 818-832, 2017.

[5] H. N. Patel, R. K. Jain, & M. V. Joshi, "*Fruit detection using improved multiple features based algorithm*.", International journal of computer applications13(2), 1-5, 2011

[6] A. Danti & M. Suresha, "*Arecanut grading based on three sigma controls and SVM*.", In Advances in Engineering, Science and Management (ICAESM), International Conference on (pp. 372-376). IEEE. 2012.

[7] T. Ojala, M. Pietikainen and D. Harwood. "*A comparative study of texture measures with classification based on feature distributions*.", International Pattern Recognition vol. 29, 1996.

[8] T. Mäenpaa, "*The local binary pattern approach to texture analysis—extensions and applications*.", PhD thesis,Acta Universitatis Ouluensis C 187, University of Oulu 2003.

[9] V. N. Vapnik, "*Estimation of dependencies based on empirical data* ",Springer. New York, 1982.

[10] M. Weis, T. Rumpf, R. Gerhards, & L. Plümer, "*Comparison of different classification algorithms forWeed detection from images based on shape parameters*.", Bornimer Agrartechn. , 69, 53-64. 2009.

[11] B. E. Boser, I. M. Guyon & V. N. Vapnik, "*A training algorithm for optimal margin classifiers.*", In Proceedings of the fifth annual workshop on Computational learning theory (pp. 144-152). ACM.1992 July.

[12] C. Cortes & V. Vapnik, "*Support-vector networks. Machine learning*.", 20(3), 273-297.1995.

[13] M. Suresha K. N. Shreekanth and B. V. Thirumalesh B V. "*Diagnosis and Classification of diseases in Paddy Leaves*'.", Journal of Current Advanced Research , 06(10), pp. 6448-6452.DOI: http://dx.doi.org/10.24327/ijcar.2017.6452.0945, 2017.

Authors Profile

Chandrashekhara H obtained his BSc. (Electronics, Mathematics, and Computer Science) and MCA during 2007 to 2010 from Kuvempu University and currently he is pursuing Ph.D. in Kuvempu University, India. Digital Image Processing and Pattern Recognition is his research area.

Dr. Suresha M received the BSc degree in Physics, Chemistry and Mathematics from Kuvempu University, India during 2000 and MCA and Ph.D., from Kuvempu University during 2003 and 2014 respectively. Since 2007, he has been with the Department of Computer Science and MCA, Kuvempu University, India, where he is currently working as Assistant professor. His research interests include Image Processing, Pattern Recognition, and Machine Learning.

