IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sept. 2018 E-ISSN: 2347-2693

A Novel Feature Extraction Method for Identification of Healthy and Diseased Maize and Paddy Leaves Using ECOC Classifier

T. Harisha Naik[1*], M. Suresha[2], Shreekanth K. N.[3]

[1,2,3]Department of P.G. Studies in Computer Science, Kuvempu University, Karnataka, India.

[*]*Corresponding Author: harishtkola@gmail.com, Tel.: 919886897968*

Available online at: www.ijcseonline.org

Accepted: 13/Sept/2018, Published: 30/Sept/2018

***Abstract*—** With the entry of huge databases and the resulting prerequisites for excellent machine learning frameworks, new issues emerge and novel feature extraction methods are in demand.Basic feature reduction methods are feature selection and feature extraction. Feature selection find the subset of those prime features in a given initial set and helps in finding optimal solution. Feature extraction method transform original set of features into new subsets which are smaller number of dimensions. Generally features contain information about the target and more features indicate more information and better discrimination power. In this paper we have proposed a novel feature extraction method for feature extraction of maize and paddy dataset. Global thresholding Otsu method is used for segmentation and Error Correcting Output Codes (ECOC) classifier is used for identification of healthy and diseased maize and paddy leaves and found a success rate of 91.32% for paddy leaves and 92.56% for maize leaves. In this experimentation the similarity difference of Gray with Cb Component has given highest accuracy for both data sets.

***Keywords*—** Disease, ECOC Classifier, Maize, Paddy, Texture Features.

I. INTRODUCTION

Feature selection is the strategy for choosing the subset of features that is important to distinctive application and it enhances the precision of type by method for searching through the good quality function subset, the fixed set of original features according to the feature estimation principles. Feature selection is necessary process for pattern recognition and available all the features may not be helpful. A few features might be repetitive and some other may lead complexity during the learning stage. Entire paper represented with five sections accordingly as follows, section I is about introduction of maize and paddy leaf diseases, section II gives brief literature review and survey on various feature selection and feature extraction methods. Section III explains the methodologies adopted for our work. Section IV is about experimental result obtained in the proposed approach, and section V is conclusion and discussed on possible directions for further research.

Maize is an member of the grass family Poaceae. It is an oat grain which was first invented by the Central America. It is presently the third most vital cereal grain on the world. A paddy field is an overwhelmed product of arable land utilized for developing semiaquatic rice. Paddy fields are the typical features of rice cultivating in east, south and southeast Asia. Fields can be incorporated with soak slopes as porches and contiguous discouraged or steeply slanted highlights, such as rivers or marshes. Leaf diseases are one of the major biotic constraints to reduce crop yield and also deteriorate the quality of product that ultimately reduce the market price.

II. RELATED WORK

Hashim, N et al. developed the transform based image texture analysis using backscattering. The uniqueness of this work is the utilization of a backscattering imaging system with various methodologies of change based image surface examination for the assessment of banana quality at various aging stages was explored with Wavelet Gabor and Tamura changes. The outcomes showed that there were huge changes of elasticity, chlorophyll list and solvent solid substance as the ripening stages improved. Forecast model analysis revealed that the Wavelet transform showed the good results for all of the reference parameters followed by Tamura and the Gabor transform method called image feature selection using genetic programming for figure-ground segmentation. Genetic programing algorithm is utilized to build feature selection methods that purpose to increase the segmentation performance of standard classification techniques. Single and multi-objective GP techniques are examined based on that three novel feature selection methods are suggested. Such as

single-objective method parsimony GP feature selection, multi-objective, named nondominated sorting GP feature selection and strength pareto GP feature selection [1]. Liang, Y et al. a transformative computation procedures, intended to improve the component determination process in face image order known as multi-target wrappers based. In order to guide the search within the space of feasible face classification using multi objective genetic algorithm [2]. Vignolo, L et al. comparative study of feature extraction methods in image classifications evaluate the performance of the data models obtained by the different feature extraction methods in the context of binary and multiclass classifier by using different classifiers [3]. The analysis of performance is based on accuracy rate, recall, precision and f-measure evaluation measures Medjahed, S. A et al. using shape texture feature extraction the shape and directional information are extracted [4]. Three classifiers consisting of Euclidian distance matching criterion, probabilistic neural network (PNN) and SVM are utilized Mohammadi, S. M., et al. review of recent texture Classification methods were discussed and gather the ongoing patterns in surface grouping regarding feature extraction and classification strategies utilized as the texture datasets utilized as a part of the testing and training process [5]. Satyanarayana, C. H et al. work on Dominant local binary patterns for texture classification, it has attempted to a new feature extraction method that is robust to histogram equalization and rotation. First, the conventional LBP approach is extended to the dominant local binary pattern (DLBP) approach in order to effectively capture the dominating patterns in texture images [6]. Liao, S. et al. demonstrated that the advantage of using SVM is that it can easily separate different textures using features after transformed in Wavelet packet [7]. Dong, J. et al.texture classification with a dictionary of basic image features. Focus of this paper has been on image representation using multi-scale texture classification algorithm which, without any tuning of parameters Crosier [8]. M. et al. the role of feature selection on leaf image classification. The work proposed to examine the effect of feature selection algorithm on the predicative classification accuracy of algorithms used for discriminating the different plant leaf images. The Gabor features are invariant to illumination, rotation, scale and translation. The Gabor filters are with its own advantages in feature extraction process compared to Gray Level Co-occurrence Matrix [9]. Kumar, A. et al. Mammography classification using support vector machine with image enhancement and three types of extracted features that proposed [10]. Chaiyakhan, K. et al. optimal feature selection using ranking methods and classification algorithms. The assessment of the practical usefulness of all these as statistical and entropy based, with good performance in various domains Information Gain (IG) attribute evaluation [11]. Novaković, J feature selection for face recognition based on multi-objective evolutionary wrappers. Compared multi-objective wrappers, based on evolutionary computation techniques, designed to optimize the feature selection process in face image classification. The proposed wrappers provide feature sets of different sizes and face class discrimination capabilities, and the choice of the most appropriate wrapper should be guided by the requirements of the problem in hand [12]. Vignolo, L. D. et al. a rotation invariant textural feature extraction for image retrieval using eigen value analysis of intensity gradients and multi-resolution analysis. Novel approach is proposed to estimate the principal texture direction using eigen value analysis of intensity gradients. It gives a successful rotation invariant textural include extraction procedure by concentrate independently the measurable highlights of various recurrence groups along different bearings by multi-resolution investigation as for a reference heading and a definite execution examination of a few multi-determination based texture highlights are performed with regards to texture image retrieval [13]. Gupta, R. D. et al. method for classification of structural textures which combines the characteristics of edge information and second-order neural networks for the classification of structural textures [14]. Lee, B the defect detection and identification in textile fabric. The methods used in this investigation follows the pattern recognition schema image segmentation using Otsu method, feature extraction and classification using SVM [15]. Halimi, A. et al. the evaluation of yarn quality in fabric using image processing techniques, to automatically determine yarn mass parameters using Image Processing (IP) techniques. Feature selection methods are in demand because of improving the performance of classifier for image classification [16]. Shreekanth K N et al. the identification of healthy and diseased paddy leaves used texture based Local Binary Pattern (LBP) features extraction method and Error Correction Output Code (ECOC) classifier for classification and obtained 87.60% of accuracy [17].

III. METHODOLOGY

Plant also suffers from several diseases by disease causing agents like bacterial, Viral and fungal, It reduces the yield loss and quality of grain. In this proposed work identification of healthy and diseased leaves are carried out using Global thresholding Otsu method for segmentation and LBP Features are extracted. ECOC classifier is used to classify healthy and diseased maize and paddy leaves.

A. Image acquisition

Sampled Images are captured using digital camera in day light because clear view of leaf object is required for analysis. Image acquisition is the action of retrieving an image from some source. Sample image of diseased maize and paddy leaves are shown in figure 1.

(a) Maize leaf image (b) Paddy leaf image

Figure 1: Sample images of maize and paddy leaves

B. Image Pre-processing

Pre-processing is the preliminary step before segmentation and feature extraction. In this proposed work images are resized to 300 X 300 to improve the accuracy of an algorithm. Images are converted into RGB to YCbCr Component, individual YCbCr components are extracted for image.

C. Image segmentation

In image analysis and pattern recognition segmentation is the major task, partitioning of an image into groups of pixels based on similarity measurement is known as segmentation. It helps in the feature extraction from the segmented image based on the interested area or based on the objective which is considered for the experiment. Pre- processing of images is significant task. In proposed methodology Cr Component is used for segmentation and global thresholding method of Otsu segmentation.

D. Feature Extraction

The local binary pattern texture feature information extraction was proposed by Ojala et al., [17]. Based on theory, texture has locally two corresponding parts like a pattern and its strength with the goal of texture classification and then it is extended to face detection, face recognition, face expression recognition, disease identification [18]. In an input image each processing pixel is coordinated with its eight neighbors and ones whose intensities surpass the processing pixels are set apart as 1, else as 0, in resultant a circular point feature consisting only binary bits were measured. The feature ring is considered as row vector, and binomial weight relegated to each bit. Row vector is changed to decimal for additionally work. Local Binary Pattern utilizing round neighborhoods and straightly joining the pixel values permits the choice of any radius, R, and number of pixel in the area, P, to form an operator, which can model extensive scale structure. Essential LBP is clamor delicate, in light of the fact that the edge esteem is the handling pixel, so normal of pixels including processing pixel that encompass by an LBP operator as a threshold. The mathematical model for LBP is shown in equation (1).

$$LBP_{P,R}(x,) = \sum_{p=0}^{p-1} s\left(g_p - g_c\right)2^p \qquad (1)$$

Where g_c is the average gray value of the pixels and g_p is the intensity value of pixels in eight neighbourhood.

$$h(i) = \sum_{x,y} B\left(LBPP, R(x,y) = i\right) \mid i\ [0, 2^{p-1}], \qquad (2)$$

$$B(v) = \begin{cases} 1 & v > T \\ 0 & Otherwise \end{cases} \qquad (3)$$

For texture analysis histogram is a descriptor h(i), of the LBP shown in equation (2) & (3) and its advantages is, it is invariant to image translation.

In the proposed work the smoothen image obtained by equation (2) are passed to extract LBP features with normalized block size [32 X 32] for normalization of images and histogram is generated, after reconstruction of feature of LBP which is used for classification of the data. Sample segmented result of maize and paddy dataset is as shown in figure 2 and 3 respectively.

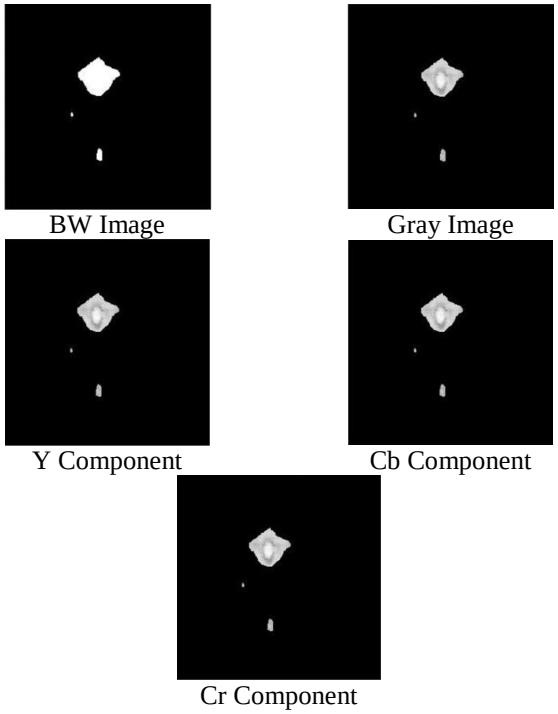
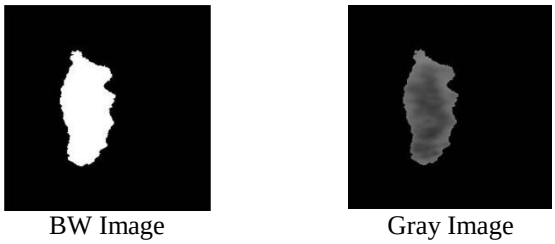


Figure 2: Sample segmented result of maize leaf



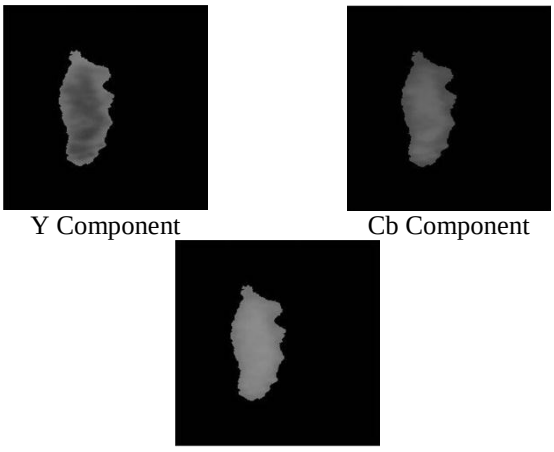


Figure 3: Sample segmented result of paddy leaf

E. Classification

In the proposed work using Estimate Posterior Probabilities Using ECOC Classifiers identification of healthy and diseased maize and paddy leaves are classified. In this proposed work the template of support vector machine (SVM) is formed using Gaussian kernel function to compute the Gram Matrix values shown in equation 3, it holds good for standardized predictors template holds all the properties are empty. ECOC classifier is used by using SVM template with class names (ie Healthy and Diseased), Fit Posterior argument for displaying the message at the time of a training using verbose argument to specify the class order. Binary learner used in these samples is labelled by negative class indices and positive class indices. Binary learners are computing posterior probabilities binary loss by using quadratic function. Predictions from all learners have been shown in figure 4.

$$G(x_1, x_2) = exp(-\|x_1 - x_2\|^2) \text{ ^2} \quad (4)$$

LBP features are extracted using difference of different components of yellow, chromatic blue, and chromatic red with block size 32X32.The equations for these are given in the following equations (5) to (7).

TrueLabel	PredLabel	Posterior	
'Diseased'	'Diseased'	0.78227	0.21773
'Diseased'	'Diseased'	0.79234	0.20766
'Health'	'Diseased'	0.78012	0.21988
'Diseased'	'Diseased'	0.83417	0.16583
'Diseased'	'Diseased'	0.78665	0.21335
'Diseased'	'Diseased'	0.78228	0.21772
'Diseased'	'Diseased'	0.79306	0.20694
'Health'	'Diseased'	0.74958	0.25042
'Diseased'	'Diseased'	0.78243	0.21757
'Health'	'Diseased'	0.77042	0.22958

Figure 4: corresponding posterior values with respect to class labels

$$f(GRAY, Y) = GRAY(x, y) - Y(x, y) \quad (5)$$
$$f(GRAY, Cb) = GRAY(x, y) - Cb(x, y) \quad (6)$$
$$f(GRAY, Cr) = GRAY(x, y) - Cr(x, y) \quad (7)$$

Where GRAY – Gray component,
Y – Yellow component,
Cb – Chromatic blue component,
Cr – Chromatic red component.

IV. EXPERIMENTAL RESULTS

The collected leaf images are resized to 300 X 300 pixel resolution to improve the algorithm accuracy. Global segmentation method is used for maize and paddy leaves segmentation. In this proposed work binary image of segmented image were interpolated with individual components of gray image and yellow(Y), chromium blue(Cb), chromium red (Cr).

The data set difference is calculated by subtracting the Y, Cb and Cr components with gray component as shown in equation 5-7(Here the similarity between LBP Features are computing the squared error between them). For these obtained features classification has been carried out individually to observe the accuracy of the feature for difference data set and also for individual component of Y, Cb, Cr and gray. ECOC classifier is used to obtained result as shown in table 1.

In this proposed work different dataset is considered with ECOC classifier, in this experimentation, when the similarity difference of Gray with Cb Component has given highest accuracy when compared to gray, yellow, Cb and Cr components for both maize and paddy leaves.

Table 1: classification of results for different dataset with different normalized block size.

Normalized Block Size	Sample data set	Accuracy w.r.t paddy leaves	Accuracy w.r.t maize leaves
32 X 32	Gray images	87.60%	88.42%
	Similarity difference of Gray with Y component	90.08%	91.73%
	Similarity difference of Gray with Cb component	91.32%	92.56%
	Similarity difference of Gray with Cr component	89.25%	91.32%
	Y component	84.71%	85.53%
	Cb component	78.92%	80.16%
	Cr component	76.85%	78.51%

V. CONCLUSION

In this paper, identification and classification of heathy and diseased maize and paddy leaves has been done using difference of LBP gray with YCbCr features. Global thresholding Otsu method is used for segmentation and Error Correcting Output Codes (ECOC) classifier is used for better accuracy in classification of the data and found a success rate of 91.32% and 92.56 % respectively. The proposed novel feature extraction method has given the better result comparing with LBP standard features.

REFERENCES

[1] S N.Hashim, S E. Adebayo & K. Abdan, M. Hanafi, "Comparative study of transform-based image texture analysis for the evaluation of banana quality using an optical backscattering system", Postharvest Biology and Technology, Vol.135, pp.38-50, 2018.

[2] Y. Liang, M. Zhang, & W N. Browne, " Image feature selection using genetic programming for figure-ground segmentation", Engineering Applications of Artificial Intelligence, Vol.62, pp.96-108, 2017.

[3] L D. Vignolo, D H. Milone & J Scharcanski, "Feature selection for face recognition based on multi-objective evolutionary wrappers", Expert Systems with Applications, Vol.40,Issue.13,pp. 5077-5084, 2013.

[4] S A. Medjahed, "A Comparative Study of Feature Extraction Methods in Images Classification", International Journal of Image, Graphics and Signal Processing,Vol.7, pp.16-23, 2015.

[5] S M. Mohammadi, M S. Helfroush & K. Kazemi, "Novel shape-texture feature extraction for medical x-ray image classification" , Int J Innov Comput Inf Control,Vol.8, pp.659-76,2012.

[6] C H. Satyanarayana, S. Anuradha, "A Review of Recent Texture Classification Methods", October 2013

[7] S. Liao, M W, A C. Chung, " Dominant local binary patterns for texture classification", IEEE transactions on image processing, Vol.18,Issue.5, pp.1107-1118, 2009.

[8] J. Dong, Y. Duan, Z. Yang, "Three-dimensional surface texture classification based on support vector machines and wavelet packets", In Intelligent Information Technology Application, 2008. IITA'08. Second International Symposium Vol. 3, pp. 124-127. IEEE, 2008.

[9] M. Crosier, L D. Griffin, " Texture classification with a dictionary of basic image features", In Computer Vision and Pattern Recognition, IEEE Conference. pp. 1-7. 2008.

[10] A. Kumar, V. Patidar & D. Khazanchi, P. Saini, " Role of Feature Selection on Leaf Image Classification", Journal of Data Analysis and Information Processing,Vol. 3 Issue.04, pp.175-183, 2015.

[11] K. Chaiyakhan, N. Kerdprasop & K. Kerdprasop, " Feature selection techniques for breast cancer image classification with support vector machine", In Proc Int Multi Conf Eng Comp Sci Hong Kong, 2016

[12] J. Novaković, " Toward optimal feature selection using ranking methods and classification algorithms", Yugoslav Journal of Operations Research,Vol. 21,Issue.1, 2016.

[13] L D. Vignolo, D H. Milone & J. Scharcanski, " Feature selection for face recognition based on multi-objective evolutionary wrappers", Expert Systems with Applications, Vol.40,Issue.13,pp. 5077-5084, 2013.

[14] R D. Gupta, J K. Dash & M. Sudipta, " Rotation invariant textural feature extraction for image retrieval using eigen value analysis of intensity gradients and multi-resolution analysis", Pattern Recognition,Vol. 46,Issue.12,pp. 3256-3267, 2013.

[15] B. Lee, " A new method for classification of structural textures", International Journal of Control, Automation, and Systems,Vol. 2,Issue.1,pp. 125-133, 2004.

[16] A. Halimi, A. Roukhe & Ouhamd, " Defect Detection and Identification in Textile Fabric by SVM Method", IOSR Journal of Engineering (IOSRJEN) ,Vol. 04, PP.69-77, 2014.

[17] Shreekanth K N , Suresha M , " Identification of Healthy and Diseased Paddy Leaves using Texture Features with ECOC Classifier" , IPASJ INTERNATIONAL JOURNAL OF COMPUTER SCIENCE(IIJCS) , Vol. 6, Issue.2, pp. 034-038, 2018.

Authors Profile

Harisha Naik T received MCA degree from Kuvempu University, Karnataka, India and M.Phil degree from Periyar University, Tamil Nadu, India. Presently persuing Ph.D from Kuvempu University Karnataka India. His research interests include Machine Learning and Pattern recognition.

Suresha M received MCA and Ph.D degree from Kuvempu University, Karnataka, India, Presently working as an Assistant Professor, Dept of computer Science, Jnana sahyadri, Kuvempu University, Karnataka, India. His research interests include Machine learning and Pattern Recognition.

Shreekanth K N received MCA from VTU, Belgavi, Karnataka, India in 2008. Currently he is pursuing Ph.D. in Kuvempu University, India. His research interests include Image Processing and Pattern Recognition.

