



LBP Feature Based Pest Identification in Rice Crop

Priyadarshi Kanungo^{1*}; Safwan Ghanem²; Shakti Kumari¹; Rubeena Naaz¹; Rashmi Prava Nayak¹

¹ Department of Electronics and Telecommunication Engineering, C. V. Raman Global University, Bhubaneswar, Odisha-752054, India

² Department of Computer Science and Engineering, C. V. Raman Global University, Bhubaneswar, Odisha-752054, India.

*Correspondence E-mail: pkanungo@gmail.com

Abstract

Food is the basic necessity of every living being, unfortunately, vegetables and fruits became harmful because of the overuse of pesticides. Toxins are unsuccessfully spread in farms for increasing the harvest amounts and quantities, needless to say, that causes serious health problems. Therefore, more reliable information regarding pests should be provided to farmers in a user-friendly way. In this work, a novel method for automatic pest identification using local binary pattern (LBP) feature is proposed, implemented, then tested using real-life images. Firstly, the image is resized and the a^* plane of the $L^*a^*b^*$ color space is used to extract the region of interest (ROI). Secondly, the histogram of the LBP of the ROI is calculated. Finally, a correlation matching is performed to measure the similarity between input ROI and predefined trained templates of different pests. The model output includes the pest name and symptoms. Experiments conducted on the variety of pests of rice plant infected leaves showed a 92.4% true identification rate which makes this method reliable comparing with the reported results from other works. The detection time also is 10 ms/ frame which fulfills one of the real-time application requirements.

Keywords: LBP; ROI; CIELAB; Pest Identification.

Introduction

Agriculture contributes to \$991 billion gross domestic product in India and contribution to economy from agriculture in India is much higher than world's average (Phadikar & Sil, 2008). Agriculture in India is one of the most essential source of income for large number of people in rural area as well as urban area. Among crops, rice plays very important role in the agriculture production. Huge amounts of rice are consumed around the world every day. Annual rice intake per individual in Asia is approximately 100 kilograms. Likelihood of 20% global population directly linked with rice

cultivation. The produced rice amounts are the highest in Asia comparing to the other continents. In Asia, over 90% of all rice grown, has more than 200 million rice farms, most of them smaller than one hectare (Macedo-Cruz *et al.* 2011). Hence, focus should be given on rice plant to gain the optimum production without much chemical materials which might cause crop to be destructed. Approximately 15-20 % of crops are destroyed due to pest, diseases or weeds which leads to annual crop loss to approximately Rs. 50,000 crores which is significant in a country where at least 200

million Indians go to bed hungry every night (Huddar *et al.* 2012). Farmers use the inherited experiences to deal with the pests and that can cause the majority of the crops to be polluted by toxins. Agriculture experts can better advise the farmers but farmers cannot stand for the extra charge for that advice. Therefore, the plant pathology has been enriched by technology-based disease detection and identification methods. That motivated us to propose a novel method for automatic identification of the rice plant pests using image processing techniques.

Review of Literature:

Thrips detection in strawberry greenhouses is dealt by Ebrahimi *et al.* (2017). Two indices are passed to the support vector machine (SVM) classifier. Those indices are the color index and the region index. The intensity plane in the (Hue, Saturation, Intensity) color space is considered to be the color index. While the ratio of the major diameter to the minor diameter is used to be the region index. The histogram of the input image is firstly equalized using Gamma operator. Then morphological image operations are applied to fill the image gaps. Finally, an SVM classifier is trained and used to label the test images according to the mean square error measurement. Early disease detection in the greenhouse is introduced (Bhadane *et al.* 2013). In which, the infected leaf image is firstly segmented into a green part and other parts. Thereafter, a customized region growing method is employed to surround the non-green part of the leaf. Finally, an alarm is emitted according to the level of the spread of the diseases.

Different classification methods have been introduced to detect diseases from the plant leaf. Color features are extracted from the leaf image using the HIS color model by Park *et al.* (2013). Then, 8-connectivity method was employed to detect the boundary of the image. Finally, self-organizing map (SOM) neural network algorithm for image classification was used. The model can detect and classify two diseases correctly. CIE Lab color space has been used in (Mengistu *et al.* 2016) for damaged crop estimation in the entire field. Image is segmented based on three

thresholding methods, i) Otsu's method, ii) Isodata thresholding, and iii) Fuzzy thresholding. Unsupervised classification is deployed to detect the white flies on the leaves in the greenhouse environment and in the farm environment as well Rani & Amsini 2016. YCbCr color space and relative difference in pixel intensities are used for segmentation of the image. Among Y, Cb and Cr, Cb color plane produces the most appropriate results. It has been used to detect and classify grape diseases like powdery mildew, downy mildew, black rot, and anthracnose. Genetic algorithm-based approach is employed to find the features of the image and extreme learning machine (ELM) classifier is used for the classification purpose (Kale & Sonavane, 2019). In Sharma *et al.* (2020), exponential spider monkey optimizer feature extraction method is used. Extracted features are fed into support vector machine (SVM) to classify the plant leaf as either diseased or healthy. The algorithm shows good performance in bi-classification, but unable to identify the disease name in multilevel classification problem. Novel image segmentation method is proposed to isolate the smitten leaf parts. Researcher applied bacterial foraging optimization based radial basis function neural network on an artificial dataset of leaf images, which have been acquiring under two-sides light sources (Chouhan *et al.* 2018). The algorithm accuracy is high for the pre-build dataset, but it has not been tested for images captured by a farmer in the field under varying illumination conditions. Many patents have been granted to leaf diseases detection researches. Chinese patent (NO. CN102759528B, 2012) used an SVM classification method. Feature extracted from both HSV color space, and Hough of the gradient (HOG) edge features got classified by means of SVM. Another Chinese patent (NO. CN102915446A, 2012) uses dynamic segmentation in the H component of HSI color space which gives better ROI extraction.

Methodology

The proposed method is shown in Fig.1. The input infected leaf image is resized to 256x256 resolution. Then the resized image is converted to L*a*b* color space. ROI is cropped from a* plane. Thereafter, the histogram of LBP is used to build a disease

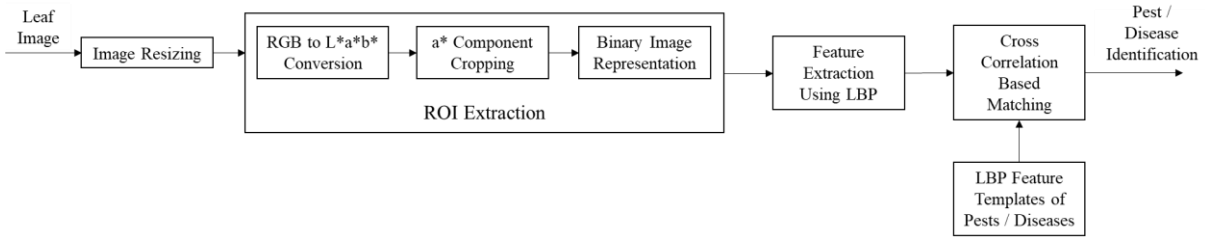


Fig. 1 The proposed method.

identification vector. This vector is yet used as a template and it is matched with the feature vector of the input image using the cross-correlation operator. The proposed method is detailed in the following subsections.

A. ROI Extraction

The L*a*b* color space has a wider color gamut comparing with other color schemes. Therefore, it is used in pest identification to give more accuracy in differentiation between pests. Commission Internationale d'Eclairage (CEI) came up with 3 hypothetical lights X, Y, and Z. any wavelength λ can be matched perceptually by positive combinations of X,Y,Z. Those X, Y, and Z are calculated using (1).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

X and Z are scaled using (2)

$$\left. \begin{aligned} X &= \frac{X}{X_n} \text{ where } X_n = 0.950456 \\ Z &= \frac{Z}{Z_n} \text{ where } Z_n = 1.088754 \end{aligned} \right\} (2)$$

Then L*, a*, and b* are calculated as in (3)

$$\left. \begin{aligned} L^* &= \begin{cases} 116 * Y^{1/3} - 16 & \text{for } Y > 0.008856 \\ 903.3 * Y & \text{for } Y \leq 0.008856 \end{cases} \\ a^* &= 500(f(X) - f(Y)) + \text{delta} \\ b^* &= 200(f(Y) - f(Z)) + \text{delta} \end{aligned} \right\} (3)$$

where, f(t) is defined in (4) and delta is defined as in (5)

$$f(t) = \begin{cases} t^{1/3} & \text{for } t > 0.008856 \\ 7.787t + 16/116 & \text{for } t \leq 0.008856 \end{cases} \quad (4)$$

$$\text{delta} = \begin{cases} 128 & \text{for 8-bit images} \\ 0 & \text{for floating-point images} \end{cases} \quad (5)$$

The a* color plane is isolated to extract the ROI. Since the infected stains appear in the

center of the resized image, the central part of the central part of the image in a* plane is cropped and considered to be the ROI. The ROI is then converted into binary image using Otsu’s method to be the source of the feature extraction. Fig. 2 shows the ROI extractions results.

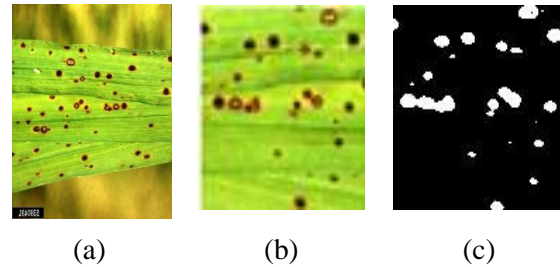


Fig. 2 ROI extraction results. (a) The original image. (b) cropped ROI (c)The image in a* plane in binary format.

B. Feature Extraction

Each pest generally causes a distinguished color on the infected leaf of plant. However, it is possible that for two different pests may have the same color which might mislead the proposed approach. It is observed that textural feature is lonesome for different pests. Therefore, considering the textural feature will increase the efficacy of the classifier. In this work, the Local Binary Pattern (LBP) is used as to generate the feature vector of the ROI. LBP gives a description of the visual feature of the image contents and the texture pattern probabilities can be summarized into a histogram by applying LBP which is carried out as follow:

- i. The gray image of the ROI is divided into 3x3 pixel-size cells.
- ii. Either in clockwise or anti-clockwise direction, each pixel is compared with its 8-neighbouring pixel values.
- iii. If the center pixel value is greater than the neighbor value, then “0” is assigned to the neighbor value otherwise “1” is

- assign to it. These values result 8-digit binary values, which could be converted into one decimal value.
- iv. The frequency of LBP values corresponding to the white area in the a* segmented image is considered for the calculation of LBP histogram and this is considered as an LBP feature vector of that corresponding ROI.
- v. Histogram is normalized by dividing by its maximum value.

Fig. 3 (a) demonstrates the details process for calculating the LBP of a pixel in a 3x3 window. Once the LBP image of the ROI is generated, the corresponding histogram of LBP image is considered as the feature vector of the ROI as shown in Fig. 3 (b).

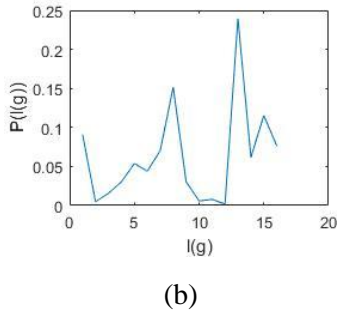
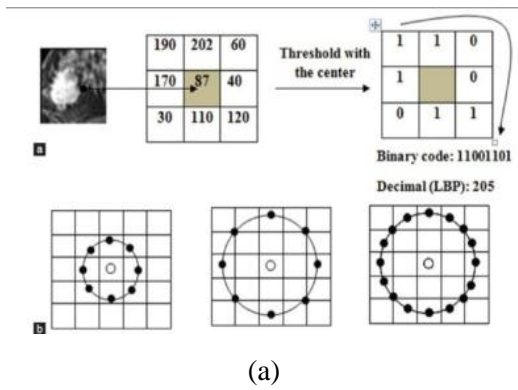


Fig. 3 (a) The LBP evaluation process, (b) the LBP histogram.

C. Cross correlation-based matching

The cross-correlation method is employed to measure the similarity between the templates feature vector of different pest and the input image feature vector. In this work, 80% of the database images are used to build up templates for each pest. The LBP histograms for each image is firstly extracted as the aforementioned method. Then the average of feature vectors for each pest class is

considered as template of the respective class. These templates are stored in the model to be matched with the LBP histogram of the input image. The cross-correlation is used as an indicator for the similarity between the template vector and the LBP feature vector for the input image. In this work, a sample estimation of matching, s_k , is done between the ROI image and template vectors using Eq. (6). Considering N dimensional feature vector, the correlation between the input feature vector and the k^{th} template feature vector S_k is evaluated using (6).

$$S_k = \frac{\sum_{i=1}^N (R_k(i) - \bar{R})(T(i) - \bar{T})}{\sqrt{\sum_{i=1}^N (R_k(i) - \bar{R})^2 \sum_{i=1}^N (T(i) - \bar{T})^2}} \quad (6)$$

Where \bar{R}, \bar{T} are the mean of the elements in vector \mathbf{R} and \mathbf{T} respectively, which are evaluated based on (7)

$$\left. \begin{aligned} \bar{R} &= \frac{1}{N} \sum_{i=1}^N R_k \\ \bar{T} &= \frac{1}{N} \sum_{i=1}^N T_k \end{aligned} \right\} \quad (7)$$

If $s_k > s_j \forall k \neq j$ for $j, k=1, 2, 3$; then the input feature vector (\mathbf{T}) is highly correlated with the k^{th} template (\mathbf{R}_k).

Result & Discussion:

The proposed model has been tested for both healthy and infected rice plant leaves. Since no such dataset available for that purpose, a proper dataset is built from collected rice leaf images from multiple resources on the Internet. These images contained leaf photos for multiple diseases and pests, in addition to healthy leaves. Fig. 4 (a) shows three disease related to rice plant leaves, Brown Spot, False Smut, and Plant-Hopper. The extracted ROIs shown in Fig. 4 (b) reflects the high accuracy of the proposed method. LBP histograms displayed in Fig.4 (e) are pretty much different from each other, which makes the identification task much easier. Table (1) shows the percentage of correct detection and identification rate (IR) which is evaluated based on (8).

$$IR = C/N * 100 (\%) \quad (8)$$

Table.1 The detection rate of the proposed method.

	Number of the total leaves	Number of the correct detection	Identification rate
Brown Spot	1000	887	88.7%
False Smut	1000	947	94.7%
Plant-Hopper	1000	940	94%
Average			92.4%

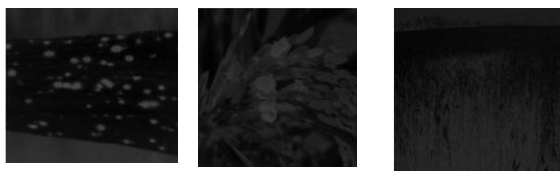
Where C, and N denotes number of correctly identified pests, and the total number of test images in the dataset respectively.



(a) original images



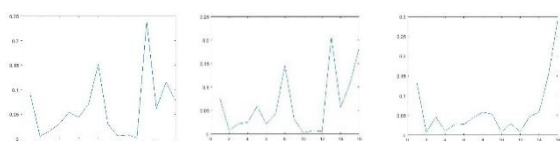
(b) The corresponding L*a*b* image



(c) The cropped image in a* planes



(d) The corresponding binary images for images in (c)



(e) LBP feature vectors

Fig. 4 (a) Rice plant leaves affected by Brown Spot, False Smut, and Plant-Hopper respectively. (b) The corresponding images in L*a*b* color space. (c) The extracted ROIs in a* plane. (d) The binary images for those images in (c). (e) respective LBP feature vectors.

The overall process to make the decision for any input image is approximately 10 milliseconds, which makes the proposed method capable to fulfill real-time application requirements.

Conclusion:

In this paper, image processing techniques are used to identify the infected leaves of the rice plant. The LBP histogram feature is employed as a texture feature to identify the pest. The a* color plane from L*a*b* color space is used because it gives the best vision of the infected spots on the plant leaf. The proposed method shows 92.4% correct detection rate using the ground truth images and cross-correlation based matching. The dataset has been built by collecting images from the Internet. MATLAB implementation, dataset, and the used templates are available for free access on: Google Drive. This model can be extended to be more robust and reliable by employing more features in the templates and machine learning approaches. Further because of its simplicity, it is suitable for real-time applications at the use of farmers in the field by developing an appropriate and user friendly mobile application.

Acknowledgments:

The authors would like to thank the faculty of information technology at Damascus university, and Image Processing and Computer Vision lab in C. V. Raman Global university, India for providing the required facilities to conduct this research.

Conflicts of Interest:

The authors declare that the research review was conducted in the absence of any commercial or economic associations that could be construed as a potential conflict of interest.

References

- Bhadane, G., Sharma, S., & Nerkar, V. B. (2013). Early pest identification in agricultural crops using image processing techniques. *International Journal of Electrical, Electronics and Computer Engineering*, 2(2), 77-82.
- Chouhan, S. S., Kaul, A., Singh, U. P., & Jain, S. (2018). Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: An automatic approach towards plant pathology. *IEEE Access*, 6, 8852-8863.
- Ebrahimi, M. A., Khoshtaghaza, M. H., Minaei, S., & Jamshidi, B. (2017). Vision-based pest detection based on SVM classification method. *Computers and Electronics in Agriculture*, 137, 52-58.
- Huddar, S. R., Gowri, S., Keerthana, K., Vasanthi, S., & Rupanagudi, S. R. (2012, July). Novel algorithm for segmentation and automatic identification of pests on plants using image processing. In 2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12) (pp. 1-5). IEEE.
- Kale, A. P., & Sonavane, S. P. (2019). IoT based Smart Farming: Feature subset selection for optimized high-dimensional data using improved GA based approach for ELM. *Computers and Electronics in Agriculture*, 161, 225-232.
- Macedo-Cruz, A., Pajares, G., Santos, M., & Villegas-Romero, I. (2011). Digital image sensor-based assessment of the status of oat (*Avena sativa* L.) crops after frost damage. *Sensors*, 11(6), 6015-6036.
- Mengistu, A. D., Mengistu, S. G., & Alemayehu, D. M. (2016). Image analysis for Ethiopian coffee plant diseases identification. *International Journal of Biometrics and Bioinformatics (IJBB)*, 10(1), 1.
- Park, Y. S., Chung, Y. J., & Moon, Y. S. (2013). Hazard ratings of pine forests to a pine wilt disease at two spatial scales (individual trees and stands) using self-organizing map and random forest. *Ecological informatics*, 13, 40-46.
- Phadikar, S., & Sil, J. (2008). Rice disease identification using pattern recognition techniques. In 2008 11th International Conference on Computer and Information Technology, IEEE Xplore, 420-423.
- Rani, R. U., & Amsini, P. (2016). Pest identification in leaf images using SVM classifier. *International Journal of Computational Intelligence and Informatics*, 6(1), 248-260.
- Sharma, A., Sharma, N., Sharma, H., & Bansal, J. C. (2020). Exponential Adaptive Strategy in Spider Monkey Optimization Algorithm. In *Soft Computing for Problem Solving 2019* (pp. 1-15). Springer, Singapore.