Leaf Recognition using Texture Features for Herbal Plant Identification

Zaidah Ibrahim*[1], Nurbaity Sabri[2], Nur Nabilah Abu Mangshor[3]
[1]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor
[2,3]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Campus Jasin, Melaka

Article Info	ABSTRACT
	This research investigates the application of texture features for leaf recognition for herbal plant identification. Malaysia is rich with herbal plants but not many people can identify them and know about their uses. Preservation of the knowledge of these herb plants is important since it enables the general public to gain useful knowledge which they can apply whenever necessary. Leaf image is chosen for plant recognition since it is available and visible all the time. Unlike flowers that are not always available or roots that are not visible and not easy to obtain, leaf is the most abundant type of data available in botanical reference collections. A comparative study has been conducted among three popular texture features that are Histogram of Oriented Gradients (HOG), Local Binary Pattern (LBP) and Speeded-Up Robust Features (SURF) with multiclass Support Vector Machine (SVM) classifier. A new leaf dataset has been constructed from ten different herb plants. Experimental results using the new constructed dataset and Flavia, an existing dataset, indicate that HOG and LBP produce similar leaf recognition performance and they are better than SURF.

Corresponding Author:

Zaidah Ibrahim,
Faculty of Computer and Mathematical Sciences,
Universiti Teknologi MARA,
40450 Shah Alam, Selangor,
Email: zaidah@tmsk.uitm.edu.my

1. INTRODUCTION

Malaysia is rich with herbal plants that are not only useful in cooking but also is beneficial in medical. This useful knowledge should be let known to the public so that they can gain the benefits. Besides that, the access to this knowledge should be easy and fast. Thus, the development of Malaysian medicinal herb recognition system is a necessity. Besides Malaysia [1], researches on herbal plant recognition have been conducted in other countries such as China [2], Indonesia [3] and Vietnam [4].

The most popular part of the plant that has been used for plant recognition is the leaf because the leaf is available throughout the year and it is easily visible. Flowers are only available at certain time while the roots are not easily accessible. Besides that, leaf is the most abundant type of data available in botanical reference collections [4]. Professionals who are working together with botany identify plants through leaves identification [5]. But manual identification in the field is quite troublesome and sometimes time consuming. This is because the knowledge from the botanist is required for the plant recognition. Thus, to overcome these problems, herbal plant identification through leaf recognition application is critical.

A robust feature is needed to cleanly distinguish among the different leaves. Leaves have various characteristics that can be used for recognition such as shape, color and texture. Shape feature has been used in [6] where the leaf's shape contour, the convexity and concavity properties of the arches are measured. Hu invariant moments that represent the shape of the leaf have been utilized in [7]. But the shape can be

affected by the age of the plant where younger leaves may look slightly different compared to the older leaves and this can produce different results for the same plant. Applying color is very challenging since almost all plant leaves have similar colors.

Texture feature for leaf recognition has been investigated in [8] with very good results. Scale Invariant Feature Transform (SIFT) texture features of plant leaves have been utilized in [5] but extracting the SIFT feature is a bit slow. Speeded-Up Robust Features (SURF) texture feature has been applied in [9] but the performance is not as good as SIFT. Fuzzy Local Binary Pattern (LBP) texture features are being used in [3]. Histogram of Oriented Gradients (HOG) has produced good performance in [10]. Even though many of the methods produce satisfactory results, there is still room for improvement. Since different features can produce different plant recognition results, the identification of a good feature or features is critical for plant recognition. Thus, a comparative study is necessary to determine the suitable feature for leaf recognition of the Malaysian medicinal plant. This research investigates three popular texture features namely SURF, LBP and HOG. A multiclass Support Vector Machine (SVM) has been utilized as the classifier of these features for leaf recognition. SVM is chosen because it is widely used for various object recognitions such as watermelon seeds exterior quality recognition [11], remote sensing image classification [12] and brain wave recognition [13].

2. RESEARCH METHOD

Texture features provides measurements for visual patterns in images. The identification of specific textures in an image is achieved by modeling texture as two-dimensional gray-level variations within a segmented region. It can be described based on coarseness, contrast, directions, linelikeness, regularity and roughness [14]. HOG counts the occurrences of gradient orientation in localized regions of an image [15]. Figure 1 illustrates the HOG algorithm. The implementation of HOG algorithm is described as follows:
1. Divide the image into small regions called cells;
2. Compute a histogram of gradient directions for the pixels within the cell;
3. Discretize each cell into angular bins according to the gradient orientation;
4. Group adjacent cells into blocks;
5. Normalize the block histogram.

SURF [16] is achieved by relying on integral images and Hessian matrix-based measure. The implementation of SURF is described as follows:
1. Find image interest points using determinant of Hessian matrix;
2. Find major interest points in scale space;
3. Find feature direction;
4. Generate feature vectors.

LBP [17] labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Figure 2 illustrates a sample computation of the center pixel with its neighboring pixels while Figure 3 demonstrates the general flow of process of LBP. The general implementation of LBP is explained below:
1. Divide the examined window into cells;
2. For each pixel in the cell, compare the pixel to each of its eight neighbors;
3. Produce a vector of binary numbers where if the center pixel's value is greater than the neighbor's value; change it to 0 and 1 otherwise.
4. Compute the histogram of the frequency of each number occurring;
5. Normalize and concatenate the histogram of all cells.

SVM was initially developed by for binary classification [18] but multiclass SVM has been produced to cater multiclass problems. A comparison has been conducted among four popular approaches for multiclass SVM namely one-against-all (OAA), one-against-one (OAO), decision directed acyclic graph (DDAG) and adaptive directed acyclic graph (ADAG) and the experimental results indicate that OAO provides the best result [19]. Thus, this research utilizes the OAO approach for multiclass SVM where it consists of multiple binary, linear SVM learners. Classification is performed by a max-wins voting strategy where every classifier assigns the instance to one of the two classes. After that, the vote for the assigned class is increases by one vote and the class with the maximum votes determines the instance classification.
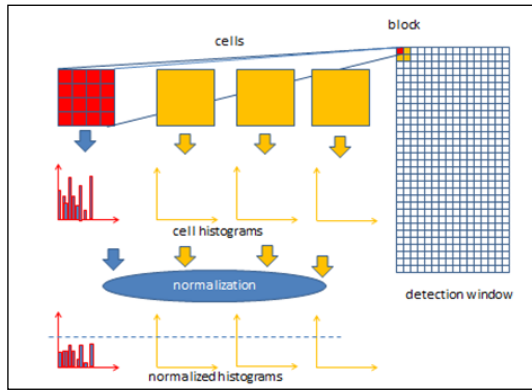
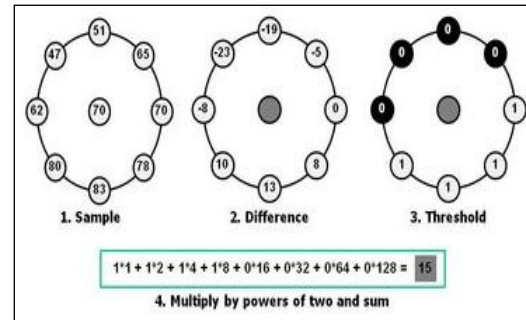Figure 1. Demonstration of HOG algorithm [15]

Figure 2. A sample computation of the interest points for LBP
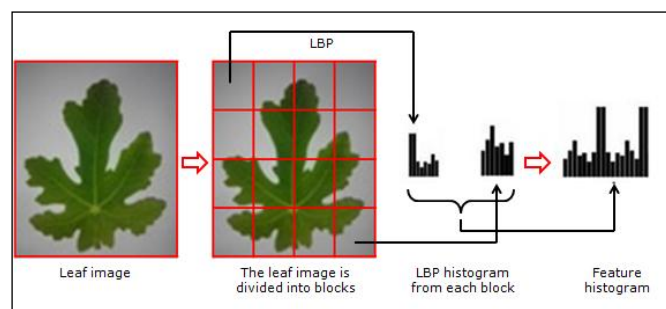

Figure 3. A demonstration of the general flow of LBP

Figure 4 illustrates the flow of process for leaf recognition. Once the image of the leaf has been captured, pre-processing is conducted on the image for resizing and converting the color image into gray-scale image. Then, the texture features are extracted and entered into multiclass SVM for leaf recognition. The three texture features, namely HOG, SURF and LBP are being extracted separately.
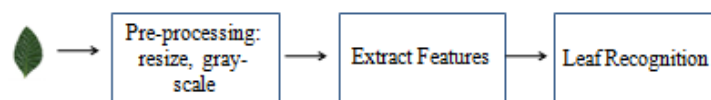

Figure 4. Flow of process for leaf recognition

Flavia dataset [20] consists of various orientations of leaf images from thirty three different plant specifies. Figure 5 shows some sample images of the leaves that have been used in this experiment where 40 samples from each species have been used for training and 10 samples for testing.

Figure 6 shows some sample images from the new dataset that consists of leaves from 10 different Malaysian herb plants where 20 samples from each species have been used for training while 5 samples for testing. For both datasets, the images were in various orientations and sizes.

Figure 5. Sample images of the leaves from Flavia dataset [20]

Figure 6. Sample images of the leaves from new constructed dataset

3.   RESULTS AND DISCUSSION

Table 1 illustrates the recognition accuracy results of the three texture features with multiclass SVM classifier using new constructed dataset while Table 2 shows the results for Flavia dataset. The recognition accuracy is computed by dividing the correctly recognized leaf of the testing images with the total number of tested images. The results of these experiments are constant for both datasets. By looking at these two tables, we can clearly see that HOG and LBP produce similar classification results for both datasets and their results are much better than SURF. Since HOG and LBP descriptors operate on localized cells, the methods uphold invariance to geometric and photometric transformations and robust to monotonic gray-scale changes caused, for instance, by illumination variations. As a result, they perform better than SURF.

Table 1. Results of multiclass SVM for new constructed dataset

Multiclass Classifier	Accuracy (%)
HOG	99
LBP	99
SURF	74

Table 2. Results of multiclass SVM for Flavia dataset

Multiclass Classifier	Accuracy (%)
HOG	97
LBP	97
SURF	63

4.   CONCLUSION

This paper evaluated three texture features for leaf recognition, namely HOG, LBP and SURF. One of the important characteristic for scene images such as herb plants is tolerant to illumination changes. Experimental results on these features extracted from leaf images in a new constructed datasets and Flavia, existing datasets conclude that HOG and LBP are better than SURF. SURF is sensitive to rotation and illumination changes whereas HOG and LBP are not. Besides that, since HOG and LBP descriptors operate on localized cells, the methods uphold invariance to geometric and photometric transformations. Thus, they are very suitable for scene images like leaf recognition for herb plant identification. Future work is to perform a comparative study on the efficiency performance of these features for mobile application.

REFERENCES
[1]   Sainin, M. S., and Alfred, R. (2014). *Feature Selection for Malaysian Medicinal Plant Leaf Shape Identification and Classification*. International Conference on Computational Science and Technology. 1-6.

[2] Xu, H., Li, F., Xue, F., and Shen, L. (2010). *Application of Pattern Recognition to the HNMR Spectra of Chinese Medicinal Herbs for Cold-hot Nature Distinguish*. 6th International Conference on Natural Computation (ICNC). 3278-3281.

[3] Herdiyeni, Y., Ginanjar, A. R., Anggoro, M.R.L., Douady, S. and Zuhud, E.A.M. (2015). *MedLeaf:Mobile Biodiversity Informatics Tool for Mapping and Identifying Indonesian Medicinal Plants*. 7th International Conference of Soft Computing and Pattern Recognition (SoCPaR). 5-9.

[4] Le, T., Tran, D. and Hoang, V. (2014). *Fully automatic leaf-based plant identification, application for Vietnamese medicinal plant search*. Proceedings of the Fifth Symposium on Information and Communication Technology. 146-154.

[5] Priyankara, H. A. C. and Withanage, D. K. (2015). *Computer Assisted Plant Identification System for Android*. IEEE Moratuwa Engineering Research. 148-153.

[6] Wang, B., Brown, D., Gao, Y. and La Salle, J. (2013). *Mobile Plant Leaf Identification Using Smart Phones*. International Conference on Image Processing (ICIP).4417-4421.

[7] Lukic, M., Tuba, E. and Tuba, M. (2017). *Leaf Recognition Algorithm using Support Vector Machine with Hu Moments and Local Binary Patterns*. IEEE 15th International Symposium on Applied Machine Intelligence and Informatics. 485-490.

[8] Jamil, N., Hussin, N.A.C., Nordin, S., Awang, K.: *Automatic plant identification: is shape the key feature?* Procedia Comput. Sci. 76, 436–442 (2015).

[9] Nguyen, Q. K., Le, T. L. and Pham, N. H. (2013). *Leaf based plant identification system for android using surf features in combination with bag of words model and supervised learning*. International Conference on Android Technologies for Communications (ATC). 404-407.

[10] Xiao, X. Y., Hu, R., Zhang, S. and Wang, X. (2010). *HOG-based Approach for Leaf Classification.Advanced Intelligent Computing Theories and Application*. Springer Berlin/Heidelberg. 149-155.

[11] Chen, X., Wang, L., Chen, W. and Gao, Y. (2013). *Detection of Watermelon Seeds Exterior Quality based on Machine Vision*. TELKOMNIKA, Vol. 11, No. 6, 2991 – 2996.

[12] Yang, N., Li, S., Liu, J. and FulingBian. (2014). *Sensitivity of Support Vector Machine Classification to Various Training Features*. TELKOMNIKA Indonesian Journal of Electrical Engineering. Vol. 12, No. 1, 286 – 291.

[13] Zhang, Z., Li, Y. and Peng, X. (2016). *Brain Wave Recognition of Word Imagination based on Support Vector Machines*. TELKOMNIKA. Vol. 14, No. 3A, 277 – 281.

[14] Tamura, H., Mori, S., Yamawaki, T. (1978). *Textural features corresponding to visual perception*. IEEE Trans on Systems, Man and Cybernetics. 8, 460–472.

[15] Dalal, N. and Triggs, B. (2005). *Histograms of oriented gradients for human detection*. IEEE Computer Vision and Pattern Recognition (CVPR). 886-893.

[16] H. Bay, From Wide-baseline Point and Line Correspondences to 3D, Ph.D. Thesis, ETH Zurich, 2006.

[17] He, D. C. and Wang, L. (1990), *Texture Unit, Texture Spectrum and Texture Analysis*. IEEE Transactions on Geoscience and Remote Sensing. vol. 28, 509 - 512.

[18] Cortes, C. and Vapnik, V. (1995). Support-vector network. Machine Learning. vol. 20, 273–297.

[19] Hsu, C. W. and Lin, C. J. (2002). *A comparison of methods for multiclass support vector machines*. IEEE Transactions on Neural Networks, vol.13, no. 2, 415-425.

[20] Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y. X., Chang, Y. F. and Xiang, Q. L. (2007). *A Leaf Recognition Algorithm for Plant classification Using Probabilistic Neural Network*. IEEE 7th International Symposium on Signal Processing and Information Technology. 11-16.