Use of Support Vector Machine to Classify Rhizomes Based on Color

M Maimunah*, E R Arumi
Department of Informatics Engineering, Engineering Faculty, Universitas Muhammadiyah Magelang, Jalan Mayjend Bambang Soegeng KM 5 Mertoyudan, Magelang, Indonesia

Corresponding author: maimunah@ummgl.ac.id

Abstract. Rhizome plants are one type of herbal medicinal plants which have various types and benefits. To find out the types of rhizome plants can be done by the classification process. Classification of rhizome plants can be done by identifying the color of the flesh of the rhizome. Identification technology in this study is needed to accelerate and facilitate the process of identifying rhizome species in the form of digital image data. Digital image processing can be used to identify rhizomes based on color using the support vector machine (SVM). In this study, the type of rhizome used as an input image was temulawak, temu hitam and temu mangga. Image preprocessing is done to get the color image of the rhizome flesh which is then used in the feature extraction stage. Image feature extraction process is done to get the color characteristics of the image by calculating the image RGB value. The image color value obtained has been used to identify rhizome species using SVM. The results obtained indicate that the identification of rhizome types using SVM provides an accuracy of 87.5%.

1. Introduction
Very famous medicinal plants are used as raw materials for herbal medicine and traditional medicine. Jamu is the original cultural heritage of the Indonesian nation which has been passed down from generation to generation. Also, it cannot be denied that herbal medicine is a very potential national asset so it should be developed into a superior and useful health commodity. To support herbal medicine as a health commodity, the government through the Minister of Health together with the Indonesian Coordinating Minister for Human Development and Culture inaugurated the Fit and Jamu Movement. The Ministry of Health supports and encourages the Herbal Medicine Movement for health and fitness. Besides, the Ministry of Health has developed a science of herbal medicine which is a service-based research program to obtain scientific evidence to get quality and nutritious herbs. A large number of herbal medicinal plants and the lack of public knowledge about the types and herbal medicinal plants make it difficult for the community in terms of distinguishing the types of herbal medicinal plants so that many people prefer to use chemical drugs. To provide information to the public, an herbal medicinal plant recognition system is needed that is capable of identifying and introducing herbal medicinal plants. The information obtained is in the form of a digital image which is then analyzed and processed by the system. The system identifies leaf images from herbal medicinal plants and performs recognition of a pattern or characteristics of the object [1]. Indonesia is very famous for its abundant natural wealth and very useful in pharmacotherapy. Several variables affect
the decision to purchase herbal medicine in Indonesia. Social system and herbal characteristics showed a significant influence on the decision to purchase medicinal herbs, on the other hand, demographics and communication system shows no positive influence [2]. Identification of herbal plants based on leaf images using texture analysis has been carried out on 10 species. Texture analysis was performed using GLCM and KNN classification. The results showed the accuracy of identification using the 9-fold cross-validation method reached 83.33% using 9 subsets [3]. Classification of medicinal plants is not only based on color characteristics. Another feature that can be used for classification of medicinal plants is the texture feature. Classification of medicinal plant leaves based on texture feature using grey tone spatial dependency matrices method and Local Binary Pattern operators. The Classification that has been done obtained an accuracy of 94.7% [4]. Identification of medicinal plants in India has been carried out using neural networks. The features used for identification are the edge, area, and color of the leaf images of Indian medicinal plants. The classification result gives an accuracy of 75% [5]. Rhizome plants are very diverse and for each type has many varieties as well, an example is a turmeric. Classification of the type of turmeric has been carried out using neural networks. The types of turmeric studied were the samba refinery, samba verali and salmon verali. Classification of the type of turmeric based on histogram and morphology characteristics with an accuracy of 73.33% [6]. The characteristic texture of herbal plants can be classified by using support vector machines on herbal plants. Textural characteristics used include a histogram of oriented gradients (HOG), Local Binary Pattern (LBP) and Speed-Up Robust Features (SURF). The classification produces the best accuracy for HOG and LBP texture characteristics. SURF is more sensitive to rotation and illumination than HOG and LBP [7]. In this study, the classification of rhizome plants which includes temulawak, temu ireng, and temu mangga based on color feature.

2. Methodology

Details of the research stages are presented in Figure 1.

![Figure 1. Detail of classification process](image)

The stages of the research include:

2.1 Image acquisition
The data used in this study were temulawak, temu putih and temu ireng. The image is taken using an 18MP DLSR camera with a distance of 30 cm. Image data is stored in the .jpeg extension format in the size 1296 x 864.

2.2 Feature Extraction
In this stage, the RGB color feature extraction is carried out from the rhizome image. Before feature extraction is carried out, the image of the rhizome is first carried out by cutting the image of the rhizome by taking part in the object which states the color of the rhizome's flesh. Then the RGB color feature extraction from the rhizome image that has been extracted by calculating the average value of the entire pixel. The average RGB value for each component R, G, and B is obtained from the number
of each component divided by the total number of pixels. The average RGB value of the entire pixel is normalized by dividing each value by the number 255 and stored as the RGB parameter which then becomes the input for the classification process.

2.3 Classification
At the classification stage, the method used is multiclass SVM. By using the tree method, first, compare temu ireng and temulawak. The winner between temu ireng and temulawak is compared to temu mangga. To solve this problem, an equation is made between the temu ireng and temulawak, temu ireng and temu mangga, temulawak and temu mangga.

3. Result and Discussion
The data used in this research are the image of temu ireng, temulawak and temu mangga obtained from the market with a total of 40 images each. The image is taken using a mini studio using a Canon EOS 600D camera and the image capture distance is 30 cm. The image capture process is carried out as in Figure 2 with the results of taking the image of the rhizome as in Figure 3.

Rhizomes image used consisting of temu ireng, temulawak and temu mangga which has been obtained is then performed cropping to get the flesh of the findings only without any background objects. In this study, the classification process was carried out using Matlab software. The feature used for the classification of temu ireng, temulawak and temu mangga are color feature using RGB color space. Input images are calculated pixel values and normalized R, G and B values have a range of values between 0 and 1. Normalized values are expressed in variables r, g, and b which will then be used as input in the classification stage. The results of feature extraction are presented in table 1.
Table 1. Color feature extraction

|       | r      | g      | b      |
|-------|--------|--------|--------|
| Temu Ireng | 0.738852 | 0.384302 | 0.009729 |
|        | 0.794535 | 0.436512 | 0.00824  |
|        | 0.757932 | 0.409909 | 0.015808 |
| Temulawak | 0.768538 | 0.638383 | 0.075956 |
|        | 0.713804 | 0.598794 | 0.043371 |
|        | 0.773416 | 0.659666 | 0.059808 |
| Temu Mangga | 0.669368 | 0.670359 | 0.269494 |
|        | 0.59195  | 0.600923 | 0.247807 |
|        | 0.696839 | 0.681288 | 0.241195 |

To be able to classify temu ireng, temulawak and temu mangga, data learning is needed. In this study, training data are used in several of 32 pieces for each class. The multiclass SVM learning process is done using the tree method. After the training process is carried out, the testing process is done to get the classification results. The test data used were 8 pieces for each class. The classification process is presented in Figure 4.

![Figure 4. Rhizomes classification](image)

The test results obtained are presented in the confusion matrix as in table 2.

Table 2. Confusion matrix

|       | Temu Ireng | Temulawak | Temu Mangga |
|-------|------------|-----------|-------------|
| Actual|            |           |             |
| Temu Ireng | 7          | 1         | 0           |
| Temulawak  | 0          | 6         | 2           |
| Temu Mangga | 0         | 0         | 8           |
4. Conclusion
Research has been carried out to classify rhizome plants which include temu ireng, temulawak and temu mangga. The classification of the types of findings is done because given the benefits of the types of findings are very many and on the other hand, have the same feature. One distinguishing characteristic is the color. The Classification has been carried out on types of rhizome plants using SVM based on RGB color. The test results obtained an accuracy of 87.5%. Further research is needed to identify temu ireng, temulawak and temu mangga based on odor feature.

Acknowledgments
The author of this paper would like to acknowledge great support and encouragement by LP3M Universitas Muhammadiyah Magelang for providing all the equipment used for Penelitian Revitalisasi Visi Institusi (PRVI) 2019 scheme.

References
[1] Kementrian Kesehatan RI D J B K dan A K 2015 Informasi kefarmasian dan alat kesehatan 1
[2] P. Kautsar A, Ayunovani F. S. M and Surahman E 2015 The Influence of Demographic, Social System, Communication System, and Herbal Characteristics on Purchase Decisions of Herbal Medicine in Indonesia J. Econ. Bus. Manag. 4 235–8
[3] Ni’mah F S, Sutojo T and Setiadi D R I M 2018 Identifikasi Tumbuhan Obat Herbal Berdasarkan Citra Daun Menggunakan Algoritma Gray Level Co-occurrence Matrix dan K-Nearest Neighbor J. Teknol. dan Sist. Komput. 6 51
[4] Arun C H, Sam Emmanuel W R and Christopher Durairaj D 2013 Texture Feature Extraction for Identification of Medicinal Plants and Comparison of Different Classifiers Int. J. Comput. Appl. 62 975–887
[5] Aitwadkar P P, Deshpande S C and Savant A V 2018 Identification of Indian Medicinal Plant by using Artificial Neural Network Int. Res. J. Eng. Technol. 05 1669–71
[6] Kaur A, Saini N, Kaur R and Das A 2016 Automatic classification of turmeric rhizomes using the external morphological characteristics Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2016 1507–10
[7] Ibrahim Z, Sabri N and Mangshor N N A 2018 Leaf recognition using texture features for herbal plant identification Indones. J. Electr. Eng. Comput. Sci. 9 152–6