

Orchid types classification using supervised learning algorithm based on feature and color extraction

Pulung Nurtantio Andono¹, Eko Hari Rachmawanto², Nanna Suryana Herman³, Kunio Kondo⁴

^{1,2}Department of Informatic Engineering, Dian Nuswantoro University, Indonesia

³Department of Software Engineering, Universiti Teknikal Malaysia Melaka, Malaysia

⁴School of Media Science, Tokyo University of Technology, Japan

Article Info

Article history:

Received Oct 15, 2020

Revised Mar 20, 2021

Accepted Jul 1, 2021

Keywords:

K-nearest neighbor

Naïve Bayes

Orchid

Supervised learning

Support vector machine

ABSTRACT

Orchid flower as ornamental plants with a variety of types where one type of orchid has various characteristics in the form of different shapes and colors. Here, we chosen support vector machine (SVM), Naïve Bayes, and k-nearest neighbor algorithm which generates text input. This system aims to assist the community in recognizing orchid plants based on their type. We used more than 2250 and 1500 images for training and testing respectively which consists of 15 types. Testing result shown impact analysis of comparison of three supervised algorithm using extraction or not and several variety distance. Here, we used SVM in Linear, Polynomial, and Gaussian kernel while k-nearest neighbor operated in distance starting from K1 until K11. Based on experimental results provide Linear kernel as best classifier and extraction process had been increase accuracy. Compared with Naïve Bayes in 66%, and a highest KNN in K=1 and d=1 is 98%, SVM had a better accuracy. SVM-GLCM-HSV better than SVM-HSV only that achieved 98.13% and 93.06% respectively both in Linear kernel. On the other side, a combination of SVM-KNN yield highest accuracy better than selected algorithm here.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Pulung Nurtantio Andono
Department of Informatic Engineering
Universitas Dian Nuswantoro
Imam Bonjol 207 Semarang, 50131, Indonesia
Email: pulung@dsn.dinus.ac.id

1. INTRODUCTION

Orchid is the largest monocot plant with an estimated population of 25,000 species in the world. The orchid is a beautiful flower type and is included in the ornamental plants that are widely cultivated in Indonesia [1], besides the orchids are also produced as cut flowers. The uniqueness of orchids is in the shape and color [2] of the lips or labellum [1] that can distinguish with other plants. Dendrobium orchids [3] are able to meet the demands of interest consumers whose tastes change over time. This can be seen from the types of orchids on the market that have varied flower colors and shapes, as well as the presence of new varieties with a more beautiful and attractive appearance [4].

The process of comparing one type of orchid with another orchid can see the color, texture, and flower petals. Knowing these differences can make it easier to classify orchid types, but orchid plants have many similarities in flower petals, texture and color, so people find it difficult to distinguish orchid species, especially lay people who do not yet know the characteristics of orchid species, by therefore a computer system is needed for the automatic classification of flowers that is expected to make it easier to classify orchid types [3]. The classification process [5] is done by digital image processing technology. This information system will produce a level of accuracy that is accurate enough to determine the name and type

of orchid. With this application, people's insight into the names and types of orchids can be increased [6]. Extraction [7], [8] is needed to get a true and accurate information from the object of a smudge contained in the image. Classification is the process of finding a group of models or functions that describe and distinguish each class of data, to predict the class of objects or data whose class labels are unknown.

In this paper, we present the results of a comparison of the accuracy of the supervised learning algorithms, including support vector machine (SVM), k-nearest neighbor (KNN), and Naïve Bayes in the unrealtime classification. This study will provide an overview for novice researchers in developing a classification process in a supervised learning algorithm. For best accuracy we have tried to apply kernel and range variations. We use the SVM, KNN and Naïve Bayes (NB) algorithms which may be classified as quite popular but old school algorithms, in fact we have considered the operating time and parameter setting process using neural networks and their variations that take quite a long time, for example convolutional neural network (CNN), and multilayer perceptron (MLP). In fact, not everything will be good if it is processed with the latest algorithms. We also need to analyze the use of outdated algorithms, because there are still many things that we can gain from the results of these implementations.

2. RESEARCH METHOD

2.1. Related research

Nilsback [9], has conducted a large-scale investigation of orchid images in 103 image classes using 4 features, namely texture, boundary, petal, and color. In a single feature obtained an accuracy of 55.1% while in combination of all features obtained an accuracy of 72.8%. The challenge in accurately calculating accuracy is influenced by the similarity of large and small classes of each testing data input. SVM has also been used in [10], in this study 13 orchid genotypes were identified by FTIR spectroscopy with 3 models, namely SSAE, SVM, and KNN. SSAE is proven to be more accurate than SVM and KNN. SSAE produces 99.4% accuracy and 97.9% calibration while KNN produces 100% accuracy but only 92.6% calibration. Paphiopedilum orchid species have been recognized using CNN in a study conducted by Arwathananukul *et al.* [11] using 1500 images and 15 classes with an accuracy of 98.6%. Image classification also conducted using Naïve Bayes as simple statistics and probabilities in shallot quality achieved high accuracy [12]. Here, Naïve Bayes combine with hue saturation value (HSV) color model. Using 60 training and testing data, Naïve Bayes produces 91.67% accuracy. Here, the choice of HSV color models and color channels is adjusted to the theory of human vision might change with another preprocessing to get highest accuracy. Jayech and Mabjoud compared with tree augmented Naïve Bayes (TAN) and forest augmented Naïve Bayes (FAN), regular Naïve Bayes (RN) achieved highest mean classification [13]. A comparative study by K. Chandel *et al.* [14], proved that KNN is better than NB with accuracy 93.44% and 22.56% respectively. By H. T. Zaw *et al.* [15] conduct Naïve Bayes to detect brain tumor. Another research by H. M. Zawbaa *et al.* [16] that applying SVM for flower image classification has been done. SVM has been tested and compared with random forest (RF) classifier using 215 flowers. By SIFT and SFTA, the dataset was divided into 70% training and 30% testing result shown that SIFT-SVM yield higher accuracy in 100% than SFTA-SVM. I. Mohamed *et al.* [17], a support vector machine is trained with different strategies according to the organs and species of plants using SIFT and OpponentColor SIFT. This experiment using big data in 5061 leaf uniform background images, 2107 leaf natural background images, and 2167 flower images. Before classify using SVM, images have been clustered using K-Means. Here, optimal parameter were $K=4000$ and $C=100$ has improved score from 0.67 to 0.74. Another research in segmentation in fast and accurate detection of kiwifruit has been done by L. Fu *et al.* [18] and proved their model has small and efficient for real-time kiwifruit detection in the orchard. This experiment need more time to make hardware and high cost using R-CNN with ZFNet, Faster R-CNN with VGG16, YOLOv2 and YOLOv3-tiny, the DY3TNet model has achieved precision of 0.9005 in 27 MB data. Another research by A. Koirala *et al.* [19] using R-CNN in COCO dataset to detect mango fruit. They used of around 400 training tiles. The MangoYOLO(bu) achieved a F1 score of 0.89 on a day-time mango image dataset. Another detection using waxberry image has been done [20]. This research using COCO dataset and performed by MR-CNN and compared with K-Means in verification sample set, while the average detection accuracy and recall rate reaching 97% and 91%, respectively. Based on previous research and urgency of the classification of orchid images, we proposed SVM. It algorithm had been tested and investigated using three kernels are linear, polynomial, and gaussian. We had been investigated the result in accuracy both of using GLCM-HSV or only in HSV and also compare with KNN. We had been investigated the effect of using the extraction feature in achieving the best accuracy.

2.2. Proposed method

The preprocessing steps include cropping, resizing, and extracting the image. Cropping is an image processing process by cutting the image which aims to take an important part of the image, while resizing is the process of resizing an image. Feature extraction is the stage to recognize the characteristics or

information of objects in the image, while the feature is a form that is unique to the image. In this study, the extraction of orchid features using the gray level co-occurrence matrix (GLCM) method. Figure 1 is the stage that we have implemented to find the best accuracy.

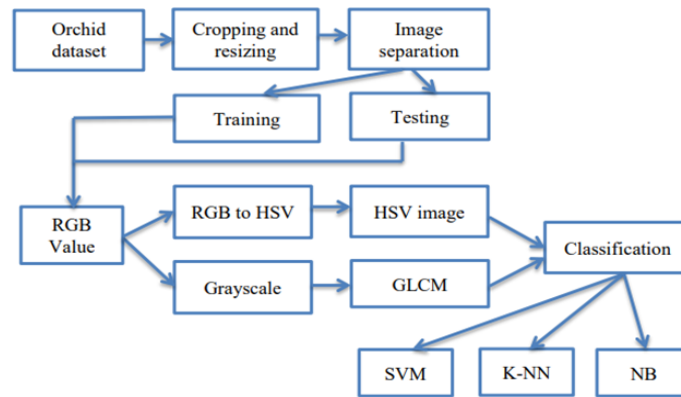


Figure 1. Proposed scheme

Based on Figure 1, orchid data as many as 2250 original images for the entire data. Cropping the 2250 orchid images, after that match all orchid image sizes to 512x512 pixels. From the 2250 data, it is grouped into 2 parts, namely training data and testing data, for training data there are 2250 original images and testing data for 1500 original images. After the data is grouped into 2 parts, each original image will look for the RGB value. After getting the RGB value do the conversion process to grayscale, then create a cohesion matrix with a distance of 1 and an angle of 0°. Finding the value of the four parameters, namely the value of contrast, energy, correlation, homogeneity using the MATLAB R2015a application. Calculate the GLCM using the values obtained from the calculation of the four previous parameters. Finding the HSV value for each orchid image by converting RGB to HSV. Then calculate the average value of hue, saturation and value. After getting the GLCM and HSV values, do the orchid flower classification process using the SVM method. The orchid flower classification process will produce a classification using the SVM method, with the value previously obtained in the extraction of the gray level GLCM and HSV features. When you get the classification results of orchids, calculate the level of accuracy. After that observe and record the results of the classification process of orchids.

2.3. Hue saturation value

HSV is a color extraction feature used for basic color classification and has a tolerance for changes in light intensity [21]. Here, RGB can be converted to HSV using some of calculation steps as in (1) until (10). Some of the advantages of HSV compared to other color spaces are : hue (H), which is a picture of the original color, such as blue, yellow, green, etc. that can be seen clearly by human vision. The angular values in HSV range from 0° to 360°; saturation (S) is the relative purity of the colors represented as the distance from the axis of black and white light with a value of 0 to 100; value (V) is represented as high on the black and white axis or the darkness of a color. Where R is red value before normalized, r is normalized red value, G is green value before normalized, g is normalized green value, B is blue value before normalized, b is normalized blue value, V is value, S is saturation, H is hue in (8) until (10) used to change the image to 8 bit image. The value range from value is 0 to 100, value 0 is black. Based on saturation, 100 as white color or more or less saturation level.

$$r = \frac{R}{(R+G+B)} \quad (1)$$

$$g = \frac{G}{(R+G+B)} \quad (2)$$

$$b = \frac{B}{(R+G+B)} \quad (3)$$

$$V = \max(R, G, B) \quad (4)$$

$$S = \begin{cases} 0, & \text{if } V = 0 \\ 1, & V - \frac{\min(R, G, B)}{V}, V > 0 \end{cases} \quad (5)$$

$$E = \begin{cases} \frac{60(g^0 - b)}{S \cdot V} \\ 60 * \left[2 + \frac{(b-r)}{S \cdot V} \right] \\ 60 * \left[4 + \frac{(r-g)}{S \cdot V} \right] \end{cases} \text{ if } \begin{cases} S = 0 \\ V = r \\ V = g \\ V = b \end{cases} \quad (6)$$

$$= H + 360 \text{ if } H < 0$$

$$V = V \times 225 \quad (7)$$

$$S = S \times 255 \quad (8)$$

$$H = \frac{H}{2} \quad (9)$$

2.4. Gray level co-occurrence matrix

GLCM is feature extraction with texture calculations in the second order, whereas the second order text calculation is the relationship between pairs of two original image pixels [22]. The pixel neighbor can be selected eastward (right). The way to represent this relationship is (1,0), stating that the relationship of two pixels in a row forms a horizontal value of 1 and followed by a pixel of value 0. Based on this composition, the number of groups of pixels that meet the relationship is calculated. The following are the steps to calculate the GLCM, which are: First do the initial matrix formation of the GLCM from a pair of two pixels in the direction parallel to the direction 0° , 45° , 90° , or 135° . Second, forming a symmetric matrix by adding up the initial GLCM matrix with the value of the transpose matrix. And the next steps are normalize the GLCM matrix to eliminate dependence on image size by dividing each matrix element by the number of pixel pairs. Calculating the value of feature extraction in the GLCM method. As in (10) where i is the row value of the i^{th} matrix, j is the value of the j^{th} matrix column and $p(i, j)$ is the value of the co-occurrence matrix element of rows (i) and column (j). As in (11) until (16), i is the matrix row value, j is the matrix column value, $p(i, j)$ is the row (i) and column (j) co-occurrence matrix element value, μ_i , μ_j is the average value of the elements in the row and column matrix, σ_i , σ_j , is the standard deviation value for the rows and columns of the matrix. i jp

$$CON = \sum_{i,j} i - j^2 P_{i,j} \quad (10)$$

$$\mu_i = \sum_i \sum_j i p_{(i,j)} \quad (11)$$

$$\mu_j = \sum_i \sum_j j p_{(i,j)} \quad (12)$$

$$\sigma_i = \sqrt{\sum_i \sum_j (i - \mu_i)^2 p_{(i,j)}} \quad (13)$$

$$\sigma_j = \sqrt{\sum_i \sum_j (j - \mu_j)^2 p_{(i,j)}} \quad (14)$$

$$E = \sum_{i,j} (P_{i,j})^2 \quad (15)$$

$$H = \sum_{i,j} \frac{p(i,j)}{1 + |i-j|} \quad (16)$$

2.5. Support vector machine

SVM works very well on high-dimensional data sets [23]. This method uses a kernel technique that maps original data from the originating dimension to another relatively higher dimension [24]. In the NN method, the training process studies all training data, whereas SVM only studies selected data used in classification [25]. Unlike the k-nearest neighbor method, at the time of prediction it stores all the training data that will be used [26], but for SVM it stores a small portion of the training data to be used at the time of prediction as in (17). Where b is bias value, m =ampunt of support vector, and (x_i) . $\phi(z)$ is kernel function.

For non-linear data you can use the kernel method in the initial data set feature [27]. The concept of kernel substitution can also be used in other methods in data analysis, but SVM is one of the well-known methods that uses the kernel to represent data [28], [29].

$$f(\phi(x)) = \text{sign}(w \cdot \phi(z) = b) = \text{sign}(\sum_{i=1}^m a_i y_i \phi(x_i) \cdot \phi(z) + b) \quad (17)$$

2.6. K-nearest neighbor

The KNN algorithm is a classification method that works based on the proximity of data to other data, where the results of new query instances are classified based on the majority of the proximity of the existing categories in the KNN [10]. Euclidean distance is a distance search that is widely used in numeric data in the k-nearest neighbor algorithm technique by drawing a straight line from the training data point to the testing data point. According to Partiningsih *et al.* [30], Euclidean distance can adjust the order of the level of image similarity with good results, where $d(x, y)$ is distance between data training and data testing, x_i is data training, y_i is data testing, I is data variable, and p data dimension.

$$d(x, y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (18)$$

2.7. Naïve Bayes

NB algorithm is supervised learning, which means it takes early to make decisions or predictions. The advantage of using the Naïve Bayes algorithm is that it does not use numerical optimization, so it is cheaper matrix [31]. This algorithm is efficient in training and can use binary or polynomial data [13], [32]. Where $P(X | Y)$ is the probability of the data with vector X in class Y , $P(Y)$ is the initial probability of class X , $P(X_i | Y)$ is the independent probability of class Y of the features in vector X , and $P(X)$ is probability of X . For categorical data, it only requires all the possibilities that occur while continuous data can use the following methods [33] are: (1) calculate the probability (prior) of each class; (2) calculate the average (mean) of each feature as in (19), where k is the amount of data and n is data value; (3) compute the standard deviation of these features as in (20); (4) calculate the probability density as in (21); calculate the probability of each class as shown in (22).

$$P(X|Y) = \frac{P(Y) \prod_{i=1}^q P(X_i|Y)}{P(X)} \quad (19)$$

$$sd = \sqrt{\frac{n \sum_{i=1}^n (X_i - \bar{x})^2}{(n-1)}} \quad (20)$$

$$\varphi_{\mu, \sigma}(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} \quad (21)$$

$$P = P(X|Ci) \times P(Ci) \quad (22)$$

3. RESULTS AND DISCUSSION

Here, we used 15 type of orchid taken from ImageNet namely Dendrobium, Brassavola, Cattleya, Cymbidium, Epidendrum, Vanda, Pleurothallis, Oncidium, Calanthe, Coelogyne, Odontoglossum, Masdevallia, Laelia, Caladenia, Helleborine were selected where 100 of them as testing data in 512x512 pixels. Sample of datasets for each type of orchid and sample of preprocessing shown in Figure 2 and Figure 3 respectively. Figure 2 can be used as knowledge about the value of the GLCM and HSV. Table 1 shows the accuracy of the classification results with the SVM algorithm in linear, polynomial and gaussian kernels using and without GLCM HSV. We also compared our SVM with Naïve Bayes. Based on [3], accuracy (23) we has been tested all data.

$$Accuracy = \left| \frac{\text{true value} - \text{analysis value}}{\text{true value}} \right| \times 100\% \quad (23)$$

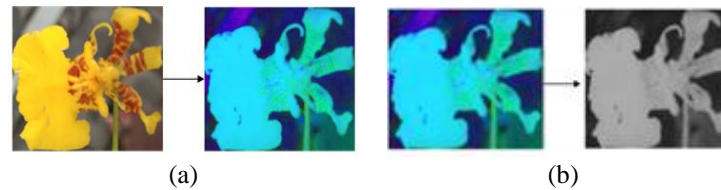


Figure 2. Preprocessing stages, (a) RGB to HSV, (b) HSV to grayscale

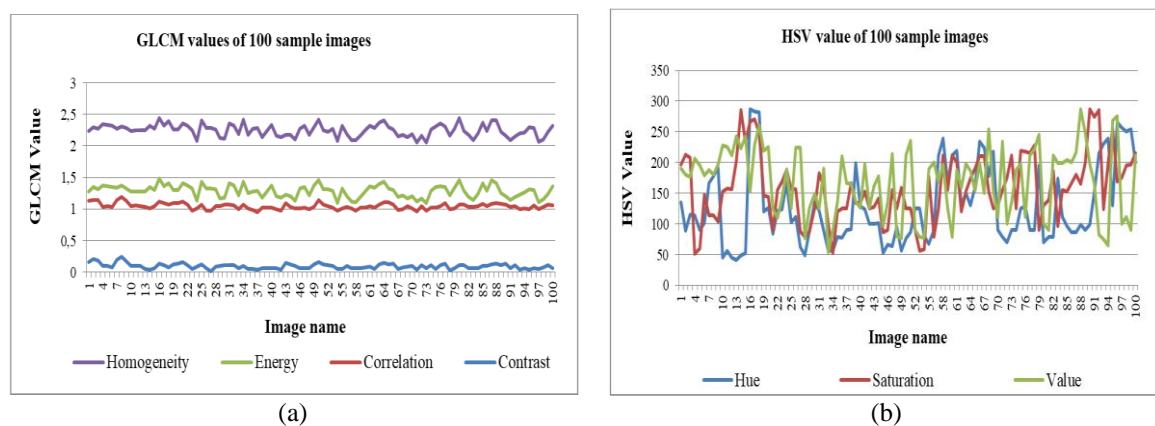


Figure 3. These figures are; (a) representation of GLCM value, (b) representation of HSV value

Here, we known that SVM better than Naïve Bayes that only yield 66% with accuracy far below that of the SVM up to 85%. Based on the classification process that has been done it is known that the classification using the GLCM and HSV feature extraction on the SVM algorithm is the best classification when used to classify orchids. On the other side, due to provide another overview of the research we have made, we conducted another trial. We have tested the classification model by implementing K-Nearest Neighbor. Here we used KNN with HSV and GLCM as explain in Table 2. We intend to compare several supervised learning algorithms as an illustration to prove superiority to the SVM algorithm. In addition to making comparisons, in this paper we also present a combination of SVM and KNN with feature optimization or not. It can be seen that SVM KNN turns out to produce the highest accuracy, so it can be concluded that in fact SVM can classify well, but the combination of SVM KNN produces higher accuracy than SVM alone regardless of whether it uses features or not. According to Table 2 each K represent achieve different accuracy. It caused identification class by pixels distance. Here, we known that $d=1$ in $K=1$ produce a highest accuracy. We also tested our dataset using Naïve Bayes. After this, we compared a best KNN result with SVM and Naïve Bayes as spelled out in Table 1.

Table 1. Comparison result between SVM and Naïve Bayes

Kernels	SVM GLCM HSV			SVM GLCM			Naïve Bayes
	Linear	Polynomial	Gaussian	Linear	Polynomial	Gaussian	
Accuracy	98.13%	97.6%	94%	93.06%	87.86%	86.93%	66%

Table 2. Comparison result between KNN in various values of distance (d) and class (K)

Distance	K Values					
	K1	K3	K5	K7	K9	K11
1	98%	94%	93%	86%	76%	70%
4	93%	92%	86%	79%	74%	65%
8	89%	88%	80%	72%	62%	61%

SVM is a more reliable more of classifiers, however KNN is less computationally intensive than SVM [34]. SVM has better performance than KNN [10], whereas another research by J. Kim *et al.* [35] conclude that KNN better than SVM. Based on thus several research, we had been investigated the accuracy

of KNN and SVM. The results prove that SVM has superior performance compared to KNN which we have tested at values K1 to K11 and pixel distance=1, 4 or 8. For image data per class we use the same as the data used in SVM as shown in Table 1 and Table 3 where the KNN on K1 still produces lower accuracy than SVM.

Table 3. Result Enhancement from a combination of SVM and KNN

	SVM KNN GLCM HSV			SVM KNN GLCM		
Specification	Linear	Polynomial	Gaussian	Linear	Polynomial	Gaussian
Accuracy	99.3%	98.6%	96.8%	94.93%	89%	88%

Based on Table 3 and Figure 4, our proposed method produce highest accuracy compared with Naïve Bayes and KNN. SVM-KNN proved better accuracy than SVM only. Features provide new knowledge which in fact can also improve the accuracy of the machine not significantly. It caused GLCM feature plays a role in identifying image texture, especially in orchid images that have curves and embossed lines such as Dendrobium, Vanda and Laelia with a slightly hairy element in each image. As shown in this Table 3, actually can be concluded that SVM-HSV is better than Naïve Bayes and several K in KNN.

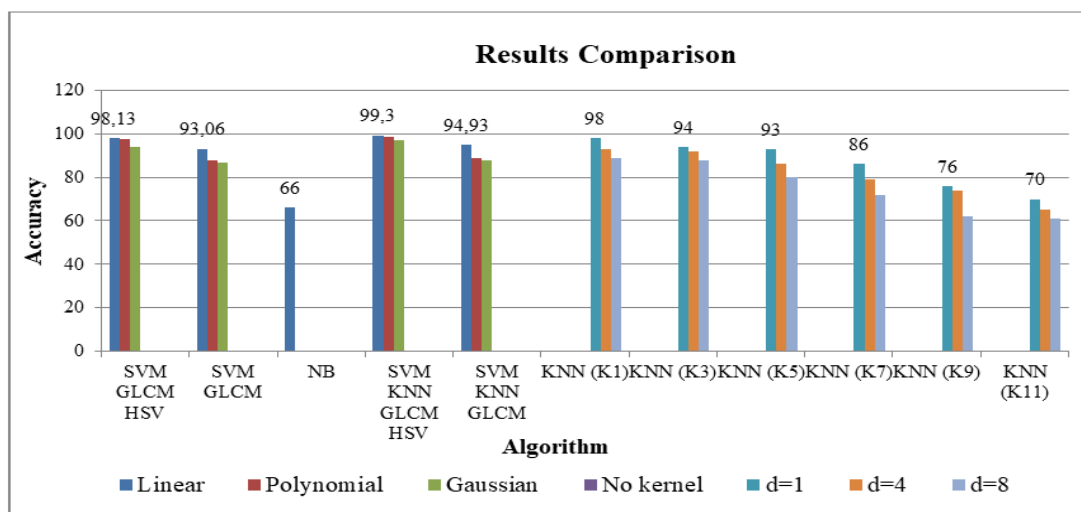


Figure 4. A comparison between SVM, KNN, and NB in various ranges and kernels

4. CONCLUSION

Based on the tests we have done on 1500 orchid image data on 15 types of flowers using 3 supervised learning algorithms, namely SVM, KNN and Naïve Bayes. In the experiment using features, it is known that the KNN and SVM can be obtained higher than the data operated without features. If the algorithm is not combined, then SVM will produce better accuracy than Naïve Bayes both in using features and not. On the other hand, SVM without features is higher than some KNN results with features on certain K, for example at K 1 to K5 values at low d, it has higher accuracy than SVM GLCM on Gaussian kernel. By combining SVM and KNN using either features or not, the accuracy value increases, but in our experiments we haven't been able to produce 100% accuracy. Another finding in our tests is that linear kernels are most suitable for classification processes where the results are better than polynomial or gaussian kernels. This is our challenge to improve accuracy so that it is maximized. SVM may be combined with other algorithms or feature extraction, for example with linear binary processing (LBP) to get a better feature value than GLCM.

REFERENCES

- [1] S. Puttemans, and T. Goedeme, "Visual detection and species classification of orchid flowers," in *Proc. 14th IAPR Int. Conf. Mach. Vis. Appl. MVA 2015*, no. May, pp. 505-509, 2015, doi: 10.1109/MVA.2015.7153241.
- [2] K. W. V Sanjaya, H. M. S. S. Vijesekara, I. M. A. C. Wickramasinghe, and C. R. J. Amalraj, "Orchid classification disease identification and healthiness prediction system," *Int. J. Sci. Technol. Res.*, vol. 4, no. 3, pp. 215-220, 2015.

- [3] N. Sabri, M. F. Kamarudin, R. Hamzah, N. N. A. Manghsor, K. A. F. A. Samah, and N. Hasan, "Combination of color, shape and texture features for orchid classification," in *2019 IEEE 9th Int. Conf. Syst. Eng. Technol. ICSET 2019 - Proceeding*, vol. 6, pp. 315-319, 2019, doi: 10.1109/ICSEngT.2019.8906322.
- [4] M. M. Sani, S. B. Kutty, H. A. Omar, and I. N. M. Isa, "Classification of orchid species using neural network," in *IEEE International Conference on Control System, Computing and Engineering*, 2013, vol. 23, no. 5, pp. 431-435, doi: 10.1109/ICCSCE.2013.6720033.
- [5] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3639-3655, Jul. 2017, doi: 10.1109/TGRS.2016.2636241.
- [6] S. Maji, A. C. Berg, and J. Malik, "Efficient classification for additive kernel SVMs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 66-77, 2013, doi: 10.1109/TPAMI.2012.62.
- [7] C. Junhua, and L. Jing, "Research on color image classification based on HSV color space," in *Proc. 2012 2nd Int. Conf. Instrum. Meas. Comput. Commun. Control. IMCCC 2012*, vol. 255, no. 3, pp. 944-947, 2012, doi: 10.1109/IMCCC.2012.226.
- [8] C.-F. Tsai, K. McGarry, and J. Tait, "Image classification using hybrid neural networks," in *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval.*, 2003, pp. 431-432, doi: 10.1145/860500.860536.
- [9] M. E. Nilsback, and A. Zisserman, "Automated flower classification over a large number of classes," in *Proc. - 6th Indian Conf. Comput. Vision, Graph. Image Process. ICVGIP*, 2008, pp. 722-729, doi: 10.1109/ICVGIP.2008.47.
- [10] Y. Chen, *et al.*, "Variety identification of orchids using Fourier transform infrared spectroscopy combined with stacked sparse auto-encoder," *Molecules*, vol. 24, no. 13, 2019, doi: 10.3390/molecules24132506.
- [11] S. Arwatchananukul, K. Kirimasthong, and N. Aunsri, "A new paphiopedilum orchid database and its recognition using convolutional neural network," *Wirel. Pers. Commun.*, vol. 115, no. 3, no. 0123456789, 2020, doi: 10.1007/s11277-020-07463-3.
- [12] A. Susanto, Z. H. Dewantoro, C. A. Sari, D. R. I. M. Setiadi, E. H. Rachmawanto, and I. U. W. Mulyono, "Shallot quality classification using HSV color models and size identification based on Naive Bayes classifier," in *J. Phys. Conf. Ser.*, 2020, vol. 1577, no. 1, doi: 10.1088/1742-6596/1577/1/012020.
- [13] K. Jayech and M. A. Mahjoub, "New approach using Bayesian Network to improve content based image classification systems," *Int. J. Comput. Sci. Issues*, vol. 7, no. 6, pp. 53-62, 2020.
- [14] K. Chandel, V. Kunwar, S. Sabitha, T. Choudhury, and S. Mukherjee, "A comparative study on thyroid disease detection using K-nearest neighbor and Naive Bayes classification techniques," *CSI Trans. ICT*, vol. 4, no. 2-4, pp. 313-319, 2016, doi: 10.1007/s40012-016-0100-5.
- [15] H. T. Zaw, N. Maneerat, and K. Y. Win, "Brain tumor detection based on Naïve Bayes classification," in *Proceeding - 5th Int. Conf. Eng. Appl. Sci. Technol. ICEAST 2019*, pp. 1-4, doi: 10.1109/ICEAST.2019.8802562.
- [16] H. M. Zawbaa, M. Abbass, S. H. Basha, M. Hazman, and A. E. Hassenian, "An automatic flower classification approach using machine learning algorithms," in *Proc. 2014 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2014*, pp. 895-901, doi: 10.1109/ICACCI.2014.6968612.
- [17] I. Mohamed, L. Diane, and P. Frédéric, "Plant species recognition using Bag-of-Word with SVM classifier in the context of the LifeCLEF challenge," in *CEUR Workshop Proc.*, 2014, vol. 1180, pp. 738-746.
- [18] L. Fu *et al.*, "Fast and accurate detection of kiwifruit in orchard using improved YOLOv3-tiny model," *Precision Agric.*, vol. 22, pp. 754-776, 2021, doi: 10.1007/s11119-020-09754-y.
- [19] A. Koirala, K. B. Walsh, Z. Wang, and C. McCarthy, "Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO,'" *Precision Agric.*, vol. 20, no. 6, pp. 1107-1135, 2019, doi: 10.1007/s11119-019-09642-0.
- [20] Y. Wang, J. Lv, L. Xu, Y. Gu, L. Zou, and Z. Ma, "A segmentation method for waxberry image under orchard environment," *Sci. Hortic. (Amsterdam)*, vol. 266, p. 109309, 2020, doi: 10.1016/j.scienta.2020.109309.
- [21] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, "Tomatoes classification using K-NN based on GLCM and HSV color space," in *Proceedings - 2017 International Conference on Innovative and Creative Information Technology: Computational Intelligence and IoT, ICITech 2017*, 2018, vol. 2018-Janua, doi: 10.1109/INNOCIT.2017.8319133.
- [22] T. Sutojo, P. S. Tirajani, D. R. I. M. Setiadi, C. A. Sari, and E. H. Rachmawanto, "CBIR for classification of cow types using GLCM and color features extraction," in *2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 2017, pp. 182-187, doi: 10.1109/ICITISEE.2017.8285491.
- [23] H. Hiary, H. Saadeh, M. Saadeh, and M. Yaqub, "Flower classification using deep convolutional neural networks," *IET Comput. Vis.*, vol. 12, no. 6, pp. 855-862, 2018, doi: 10.1049/iet-cvi.2017.0155.
- [24] A. Ambarwari, Q. J. Adrian, Y. Herdiyeni, and I. Hermadi, "Plant species identification based on leaf venation features using SVM," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 18, no. 2, pp. 726-732, 2020, doi: 10.12928/telkomnika.v18i2.14062.
- [25] M. T. Ghazal, and K. Abdullah, "Face recognition based on curvelets, invariant moments features and SVM," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 18, no. 2, pp. 733-739, 2020, doi: 10.12928/telkomnika.v18i2.14106.
- [26] H. Nguyen-Quoc, and V. T. Hoang, "Rice seed image classification based on HOG descriptor with missing values imputation," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 18, no. 4, pp. 1897-1903, 2020, doi: 10.12928/telkomnika.v18i4.14069.
- [27] A. Vedaldi, V. Gulshan, M. Varma, and A. Zisserman, "Multiple Kernels for object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, no. Iccv, pp. 606-613, 2009, doi: 10.1109/ICCV.2009.5459183.

- [28] J. C. Coetsier, and R. Jiamthapthaksin, "Parallelized FPA-SVM: Parallelized parameter selection and classification using flower pollination algorithm and support vector machine," in *Proc. 2017 14th Int. Jt. Conf. Comput. Sci. Softw. Eng. JCSSE 2017*, 2017, doi: 10.1109/JCSSE.2017.8025899.
- [29] B. C. Kuo, H. H. Ho, C. H. Li, C. C. Hung, and J. S. Taur, "A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 1, pp. 317-326, 2014, doi: 10.1109/JSTARS.2013.2262926.
- [30] N. D. A. Partiningsih, R. R. Fratama, C. A. Sari, D. R. I. M. Setiadi, and E. H. Rachmawanto, "Handwriting Ownership Recognition using Contrast Enhancement and LBP Feature Extraction based on KNN," in *Proc. - 2018 5th Int. Conf. Inf. Technol. Comput. Electr. Eng. ICITACEE 2018*, 2018, pp. 342-346, doi: 10.1109/ICITACEE.2018.8576945.
- [31] A. Abbasi, and N. Fahlgren, "Naive Bayes pixel-level plant segmentation," in *IEEE western New York image and signal processing workshop (WNYISPW)*, 2016, pp. 1-4, doi: 10.1109/WNYIPW.2016.7904790.
- [32] X. Shi, and R. Manduchi, "A study on bayes feature fusion for image classification," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 8, 2003, doi: 10.1109/CVPRW.2003.10090.
- [33] S. McCann, and D. G. Lowe, "Local Naive Bayes nearest neighbor for image classification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 3650-3656.
- [34] R. Muralidharan, and C. Chandrasekar, "Object Recognition using SVM-KNN based on geometric moment invariant," *International Journal of Emerging Trends & Technology in Computer Science*, vol. 1, no. 3, pp. 215 - 220, 2011.
- [35] J. Kim, B.-S. Kim, and S. Savarese, "Comparing image classification methods: k-nearest-neighbor and support-vector-machines," *Appl. Math. Electr. Comput. Eng.*, pp. 133-138, 2012, doi: 10.5555/2209654.2209684.

BIOGRAPHIES OF AUTHORS



Pulung Nurtantio Andono, is currently a Lecture of Departement Masters in Informatic engineering Universitas Dian Nuswantoro (Udinus). He is currently a lecturer in faculty of computer science and his interest research is in image procesisng and computer vision. Dr. Pulung has obtained his Dr. degree from Institut Teknologi Sepuluh November (ITS) and has been the General Chair of International Seminar on Application for Technology of Information and Communicatic 2018. He has authored several scientific publications including journal papers, conference papers, book chapters and book. Scopus ID 55924269600 and Orchid ID 0000-0001-7408-0558.



Eko Hari Rachmawanto, received his Master's degree in Informatic Engineering from Universitas Dian Nuswantoro and University Teknikal Malaysia Melaka (UTeM) in 2012. Now he is lecture in Universitas Dian Nuswantoro since 2012 and has interest research in image processing and data security. He also achieved awarded from Ristekbrin DIKTI as the top 50 best researchers in 2020, several times awarded as the best author and best paper of international conferences, also authored in scientific scopus indexed publications and also as speaker at several national scientific meetings. Scopus ID 57193850466, Orchid ID 0000-0001-6014-1903.



Nanna Suryana Herman received his Engineer's degree from Padjajaran University, Master's degree from Institute of Geoinformatics and Earth Observation (ITC), and Ph.D from Wageningen Research University Holland respectively in 1980, 1987, and 1996. He is active as lecture in University Teknikal Malaysia Melaka (UTeM) in Department of Software Engineering and currently serves as Director of International Office. He is active as a speaker in international scientific activities. He is interest in System and GIS interoperability and Spatial Modelling. Scopus ID 57208235865, Orchid ID 0000-0003-3695-639X.



Kunio Kondo received his Bachelor of Engineering from Nagoya Institute of Technology (1978) and Ph.D in Engineering from The University of Tokyo (1988). He already served as Vice President of Japan Society of Graphic Science in (2005-2009), Chair and president of several organizations and since 2016 he is President of ADADA International and ADADA Japan. He also as reviewer of many international journals and got many awards as best paper. His research interest in Computer Graphics, Content Production Technology for animation, games, and interactive modeling. SCOPUS ID 35488253300, Orchid ID 0000-0001-9306-9901.