

Convolutional Neural Network and Support Vector Machine in Classification of Flower Images

Ari Peryanto ^{1*}, Anton Yudhana ², Rusydi Umar ¹

¹ Informatics Department
Universitas Ahmad Dahlan
Yogyakarta, Indonesia

² Electrical Engineering Department
Universitas Ahmad Dahlan
Yogyakarta, Indonesia

*ari1907048002@webmail.uad.ac.id

Abstract-Flowers are among the raw materials in many industries including the pharmaceuticals and cosmetics. Manual classification of flowers requires expert judgment of a botanist and can be time consuming and inconsistent. The ability to classify flowers using computers and technology is the right solution to solve this problem. There are two algorithms that are popular in image classification, namely Convolutional Neural Network (CNN) and Support Vector Machine (SVM). CNN is one of deep neural network classification algorithms while SVM is one of machine learning algorithm. This research was an effort to determine the best performer of the two methods in flower image classification. Our observation suggests that CNN outperform SVM in flower image classification. CNN gives an accuracy of 91.6%, precision of 91.6%, recall of 91.6% and F1 Score of 91.6%.

Keywords: flower classification, convolutional neural network, support vector machine, machine learning, deep neural network

Article info: submitted August 20, 2021, revised October 6, 2021, accepted October 14, 2021

1. Introduction

Flowers are God's creation that have long been admired to resemble beauty and romance because of their diverse shapes and colors. Flowers are objects of esotericism, witchcraft, medicine, and source of food. Many industries pick flowers as raw materials. Numerous use of flowers and their huge diversity leads to the need of flower classification. Classifying flowers is a strenuous and painstaking task and requires expertise of a botanist. The ability to utilize computer software and technology in flower image classification will greatly help the pharmaceutical, cosmetic and other industries.

Flower image classification has invited many researchers to investigate new methods to obtain accurate results. Among the methods are deep learning and neural networks, which are experiencing rapid progress in artificial intelligence and innovation. The methods have taken a huge leap in recent years and is capable of surpassing humans in several tasks related to detecting and labeling objects.

Convolutional Neural Network (CNN) is one of the most widely used deep learning method, compared

to several image processing methods [1]. CNN has solved some of the categorization problems of real-life object images. This method is relatively easy and flexible so that it is easy to integrate into various platforms. However, the convenience and advantages of this method have the consequence of providing hardware with high capabilities to obtain the desired model [2]. CNN processes images in the same way as the human brain, namely by teaching and learning [3]. Image classification is done by labeling images. Furthermore, with certain equations, the computer begins to classify.

Besides CNN, another method that can be used in image classification is the Support Vector Machine (SVM). SVM is the best and uncomplicated initial classification method [4]. SVM works by defining the boundary between two classes with the maximum distance from the closest data. The maximum distance is obtained by finding the best hyperplane (separation line) in the input space obtained by measuring the hyperplane margin.

When in a high-dimensional input space, SVM still has good generalizability without additional knowledge. This is the advantage of SVM. The weakness of SVM is that its use is limited to linear data. For use with non-

linear data, modifications to the kernel function are required.

Color is one of the most dominant characteristics of flowers. Research related to color in flower images has been carried out by [5] who classifies flowers using the SVM method. The classification results are very good with an accuracy of 85.93%. By applying Artificial Neural Networks (ANN), [6] has carried out pre-processing activities to improve image quality, segmentation and feature extraction. The pre-processing step gives classification results with an accuracy of 81.19%. Research on flower classification using CNN has also been carried out by [7] and the model created produces an overall accuracy of 90% and is able to make predictions in real time of 98%. Meanwhile, [8] in his research conducted flower image classification with CNN, conducted a validation test with k-fold cross validation, and got the highest average accuracy of 76.49%.

So far, there has never been a comparison of the results of flower image classification using Deep Neural Network and classical Machine Learning methods to find out which method shows the best performance. Previous

studies have focused more on the classification process in one method only.

2. Method

This research was carried out in several stages, namely processing the initial image data (pre-processing), sharing training and testing data, making classification models using CNN and SVM methods, evaluating classification models, results of performance comparisons between classification models, and drawing conclusions. The stages of the research method can be seen in Figure 1.

a. Dataset

The images or objects used as datasets in this study are roses, tulips and aster. The total dataset used is 1200, and was obtained randomly using Google image. The data in the Rose category are 400, Tulip flowers are 400 and Aster flowers are 400. The data in each category is divided into training data and testing data with details of 80% training data and 20% testing data [9]. Table 1 presents the method of dividing the dataset that has been done.

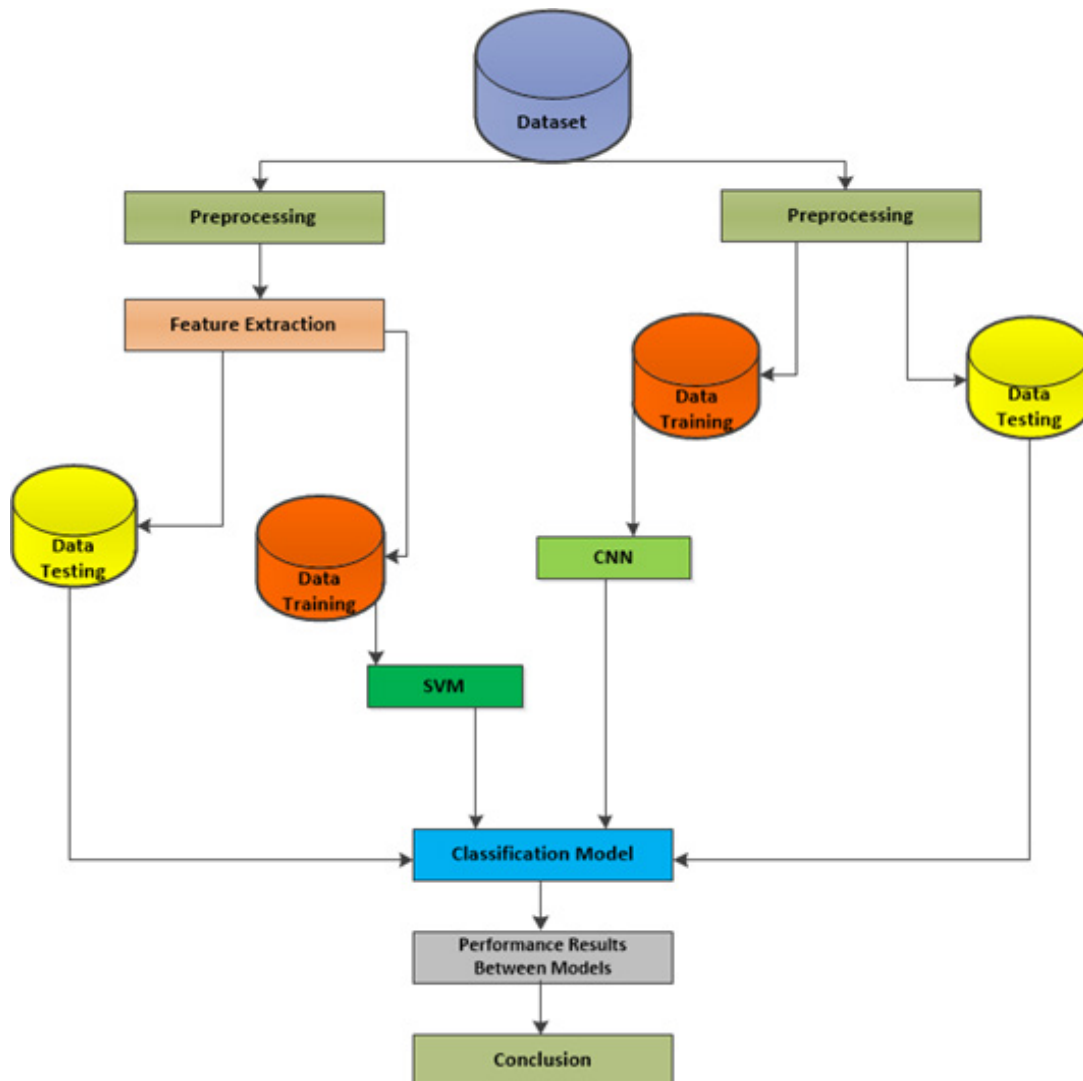


Figure 1. Research Method

Table 1. Dataset Sharing

Category	Total	Train	Test
Rose	400	320	80
Tulips	400	320	80
Aster	400	320	80

b. Preprocessing Data

Image quality can be improved by pre-processing. Noise removal and determination of the part of the image to be used are some of the methods used in pre-processing. Unwanted outer areas can be removed so as to make the image focus better and more uniform, this technique is also known as cropping [10].

As a research test material, good image quality is a must so that the classification results are more optimal, so noise cleaning is one thing that absolutely must be done. Noise is a disturbance in the image that does not reflect the true intensity of the actual scene [11].

All images must have the same size, so the image must be reshaped before modeling. Simple resizing can be done manually by stretching and changing the ratio of each image and then forcing it to a new size. The size of the input image must be chosen which is the most effective for use in classification, because a larger image input will result in a longer classification process, while a smaller image input makes the classification process faster [12]. This study uses an image with a size of 200x200 pixels. An example of an image after preprocessing is shown in Figure 2.

**Figure 2. Image after preprocessing**

c. Classification Modeling

Classification modeling is carried out before data processing to determine what parameters are used in the algorithm. The goal is to determine the effect of the parameters on the resulting performance and get the best parameters.

In this research, CNN uses several parameters, namely Epochs, layer depth, and activation layer. The function used for activation is ReLU (Rectified Linear Unit). The number of selected Epochs is 20. The layers used are 4 layers and each layer uses a 3 x 3 kernel. Convolution filters for each layer are 16 in the first layer, 16, 32 in the second layer, 64 in the third layer and 128 in the fourth layer.

On CNN there is a pooling process that is used to reduce the dimensions of the image to the convoluted model which is a type of Non-Linear Down Sampling. The max pooling kernel size used in this research is 2 x 2. In order to avoid overfitting, it is necessary to add the Combined L2 + Dropout parameter, the Dropout value is 0.5 then the L2 value is 0.0001. Dense is the last parameter

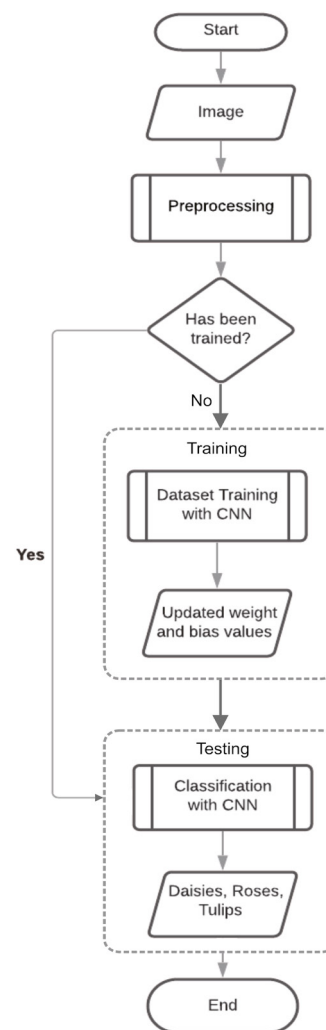
in the classifier network, totaling 3 layers with the number of neurons in each layer is 256, 256 and 128 neurons.

SVM uses several parameters, including the toll, which is the tolerance value that is used as a determinant of when the classification will stop. The toll value used in this study is 0.001.

Feature extraction and preprocessing are performed to transform data from image matrix to vector. The results of the transformation are used as input for the SVM classification which consists of 16 texture features.

d. Convolutional Neural Network

Image classification using the CNN method is carried out in two stages. The first stage is the training process and the second stage is the testing process. Feature extraction in the CNN method is carried out by performing the convolution, ReLU, maxpooling and flatten processes using the image as input. The features generated from this process are different from those generated by GLCM (Gray Level Co-occurrence Matrix) in feature extraction. The CNN method that will be used can be described with a flowchart as shown in Figure 3.

**Figure 3. Process flowchart with CNN**

e. Support Vector Machine

SVM classification is divided into two stages, namely the training process and the testing process. The training process is carried out after the image undergoes a process of separating background and foreground. The purpose of this stage is that the RGB color is still owned by the foreground and the white color is owned by the background. [13]. The image is then segmented, namely by dividing the image into several areas or objects [14]. In this research, segmentation is done using edge detection. The segmented flower image is used as input for feature extraction data to facilitate the texture feature extraction process. GLCM is used to perform texture feature extraction. The stages of using the SVM method can be modeled with a flowchart as shown in Figure 4.

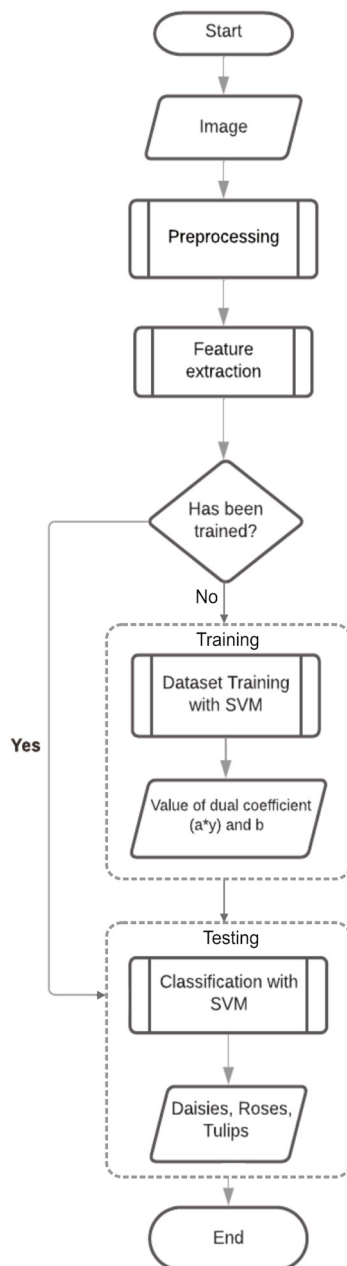


Figure 4. Process flowchart with SVM

f. Classification Model Evaluation

Measuring the performance of the created model is an important step in machine learning. The classification results from a classification model are usually summarized in a confusion matrix. According to [15], test data that are not used in training can be used as an evaluation. Evaluation of the classification model is done by looking at the performance of the classification model. The metrics used are accuracy, precision, recall and F1 score. Accuracy is used to see the percentage of the number of data items that are classified correctly by the system [16]. Equation 1 calculates accuracy.

$$Accuracy = \frac{TP + TN}{Total\ Data} \quad (1)$$

The level of suitability of the information provided by the system and the information required by the user is also called precision, while the success rate of the system in finding information is recall [17]. Equation 2 and equation 3 are used to calculate precision and recall.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F1 Score, is the average formula of the correct recall value and precision value [18]. Equation 4 calculates the F1 score.

$$F1 = 2 \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

3. Result

a. CNN

Modeling with CNN was done in 20 Epochs. The model generated in the 20th epoch provides 91.67% accuracy, 2.59% loss, 91.67% precision, and 92% recall as shown in Figure 5.

Figure 6 shows a graph of changes in Accuracy, Loss, Precision and Recall values during the modeling and classification process with CNN. The loss function is overfitting in the 3rd epoch as indicated by the loss function value of 26.13% while the validation loss value is 74.70%. The loss value indicates the amount of variation in the training data. As the epoch increases, the loss decreases, which means that the error in the model becomes smaller, so that the final loss is 1.54%.

The summary of the classification results shows the final value of precision, recall, and other metrics for each category. The classification summary for the CNN model is shown in Figure 7.

Confusion Matrix classification results with CNN can be seen in Figure 8. It can be seen that for each class, there are many images that were successfully predicted well, namely 73 out of 80 for Aster class, 71 out of 80 for Mawar class, and 76 out of 80 for Tulip class. Based on equation (1), the accuracy value of classification with CNN is $(73 + 71 + 76) / 240 = 0.916$ or 91.6%.

```

Epoch 11/20
960/960 [=====] - 126s 131ms/step - loss: 0.1505 - accuracy: 0.9678 - precision: 0.9657 - recall: 0.97
02 - val_loss: 0.0259 - val_accuracy: 0.9333 - val_precision: 0.9333 - val_recall: 0.9333
Epoch 12/20
960/960 [=====] - 126s 131ms/step - loss: 0.1522 - accuracy: 0.9676 - precision: 0.9654 - recall: 0.97
03 - val_loss: 0.0241 - val_accuracy: 0.9333 - val_precision: 0.9333 - val_recall: 0.9361
Epoch 13/20
960/960 [=====] - 125s 130ms/step - loss: 0.1550 - accuracy: 0.9678 - precision: 0.9648 - recall: 0.96
95 - val_loss: 0.0229 - val_accuracy: 0.8125 - val_precision: 0.8083 - val_recall: 0.8136
Epoch 14/20
960/960 [=====] - 126s 131ms/step - loss: 0.1526 - accuracy: 0.9671 - precision: 0.9650 - recall: 0.96
96 - val_loss: 0.0216 - val_accuracy: 0.9417 - val_precision: 0.9417 - val_recall: 0.9417
Epoch 15/20
960/960 [=====] - 125s 131ms/step - loss: 0.1468 - accuracy: 0.9699 - precision: 0.9675 - recall: 0.97
17 - val_loss: 0.0206 - val_accuracy: 0.9458 - val_precision: 0.9458 - val_recall: 0.9458
Epoch 16/20
960/960 [=====] - 126s 131ms/step - loss: 0.1412 - accuracy: 0.9678 - precision: 0.9661 - recall: 0.96
95 - val_loss: 0.0195 - val_accuracy: 0.9250 - val_precision: 0.9250 - val_recall: 0.9250
Epoch 17/20
960/960 [=====] - 127s 132ms/step - loss: 0.1416 - accuracy: 0.9662 - precision: 0.9642 - recall: 0.96
83 - val_loss: 0.0184 - val_accuracy: 0.9125 - val_precision: 0.9125 - val_recall: 0.9158
Epoch 18/20
960/960 [=====] - 126s 131ms/step - loss: 0.1538 - accuracy: 0.9654 - precision: 0.9631 - recall: 0.96
72 - val_loss: 0.0184 - val_accuracy: 0.9250 - val_precision: 0.9250 - val_recall: 0.9250
Epoch 19/20
960/960 [=====] - 126s 131ms/step - loss: 0.1540 - accuracy: 0.9667 - precision: 0.9646 - recall: 0.96
87 - val_loss: 0.0173 - val_accuracy: 0.9292 - val_precision: 0.9292 - val_recall: 0.9292
Epoch 20/20
960/960 [=====] - 124s 129ms/step - loss: 0.1730 - accuracy: 0.9622 - precision: 0.9593 - recall: 0.96
51 - val_loss: 0.0259 - val_accuracy: 0.9167 - val_precision: 0.9167 - val_recall: 0.9200
    
```

Figure 5. Results of the training process with CNN

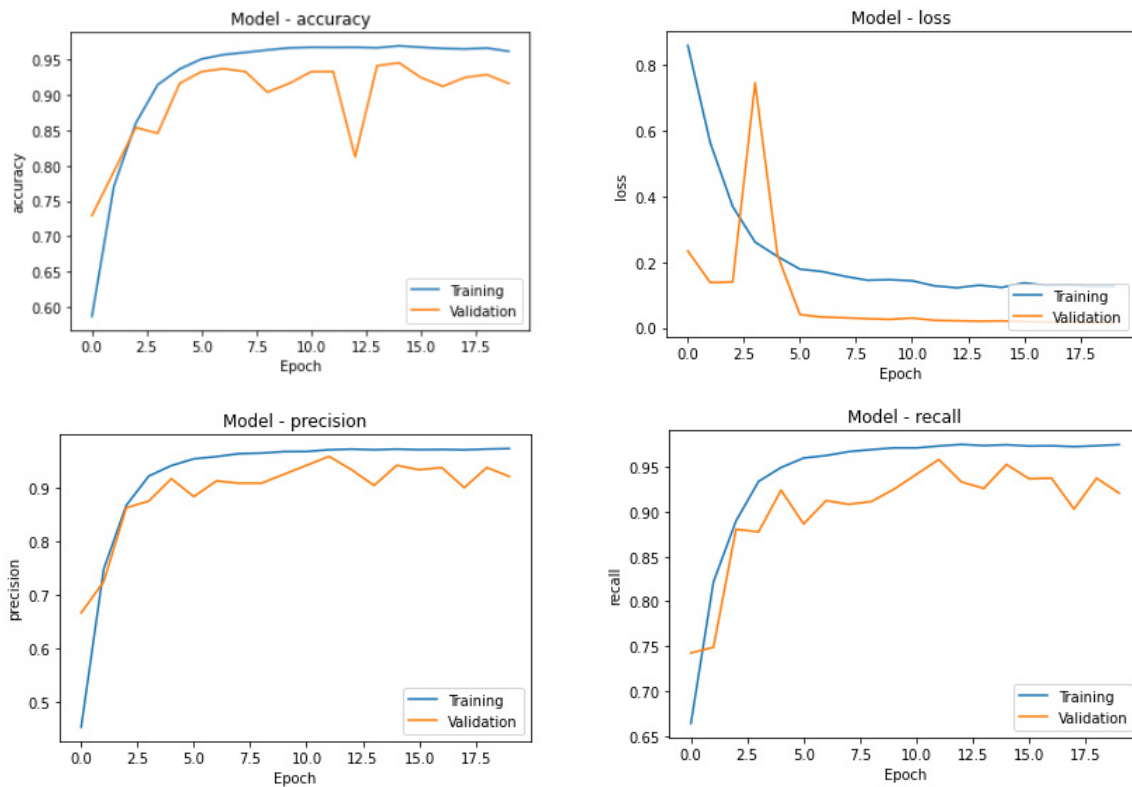


Figure 6. Accuracy, Loss, Precision and Recall Charts with CNN

Classification Report				
	precision	recall	f1-score	support
Daisies	0.92	0.91	0.92	80
Rose	0.91	0.89	0.90	80
Tulip	0.92	0.95	0.93	80
accuracy			0.92	240
macro avg	0.92	0.92	0.92	240
weighted avg	0.92	0.92	0.92	240

Figure 6. CNN Report Classification

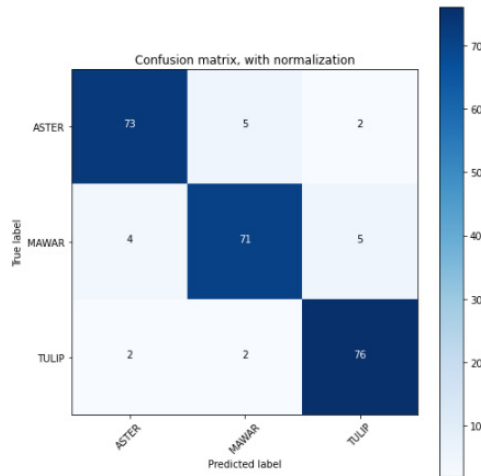


Figure 7. Confusion matrix CNN

b. SVM

Image classification using the SVM method produces accuracy, f1-score, precision and recall values as shown in Figure 9. The classification process with this method provides a Confusion Matrix as shown in Figure 10, with correct predictions for the Aster class of 63, the Mawar class of 64, and Tulip grade of 61.

	precision	recall	f1-score	support
Aster	0.83	0.79	0.81	80
Mawar	0.74	0.80	0.77	80
Tulip	0.79	0.76	0.78	80
accuracy			0.78	240
macro avg	0.79	0.78	0.78	240
weighted avg	0.79	0.78	0.78	240

Figure 8. SVM Report Classification

The accuracy of the classification with SVM can be calculated based on equation (1) by adding up all the data that are True Positive (TP) and True Negative (TN), then dividing the result of the summation by the total number of data. In this way the accuracy is obtained $(63 + 64 + 61) / 240 = 0.783$ or 78.3%.

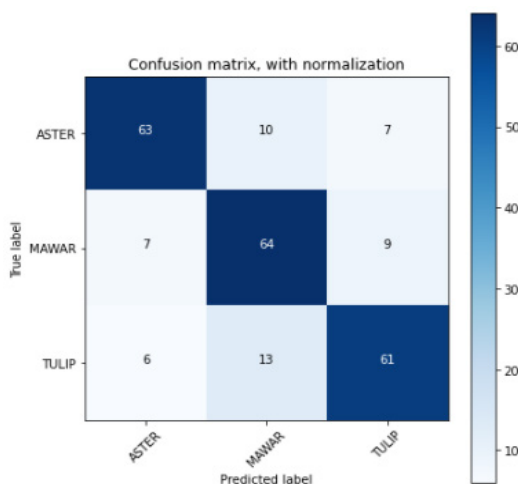


Figure 9. Confusion matrix SVM

c. Comparison

Classification with CNN and with SVM results in different performance. Table 2 presents the performance of both methods for the four metrics measured. The table shows that the classification carried out using the CNN model is in the range of 91.6% for the values of accuracy, precision and recall. Meanwhile, the classification performance using the SVM method provides a lower accuracy of 78.3%, and lower precision and recall values of 78.6% and 78.3%, respectively.

Table 2. Performance data comparison

Method	Metric	Result
CNN	accuracy	0.916
	precision	0.916
	recall	0.916
	f1 score	0.916
SVM	accuracy	0.783
	precision	0.786
	recall	0.783
	f1 score	0.786

Various studies have shown that higher accuracy is obtained when using the CNN model. In the study [7], the CNN classification gave an accuracy value of 90%. Classification using the SVM method in research [5] gives an accuracy value of 85.93, which means it is lower than the results of classification using the CNN model. Some researchers have tried to combine the CNN model for feature extraction and machine learning for the classification process. The use of the hybrid method in research [19] and [20] showed a high classification performance with an accuracy value above 98%.

4. Conclusion

Our investigation shows that the CNN model performance for flower classification gives an accuracy value of 91.6%, precision of 91.6%, recall of 91.6% and F1 Score of 91.6%. The classification with SVM produces lower performance values, where accuracy attains a value at the level of 78.3%, precision at 78.6%, recall at 78.3% and F1 Score at 78.6%. The result suggests that CNN model outperforms SVM methods in flower classification tasks.

References

- [1] R. Waseem and W. Zenghui, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review," *Neural Comput.*, vol. 29, no. 7, pp. 2352–2449, 2017, doi: 10.1162/NECO.
- [2] N. Sharma, V. Jain, and A. Mishra, "An Analysis of Convolutional Neural Networks for Image Classification," *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 377–384, 2018, doi: 10.1016/j.procs.2018.05.198.
- [3] H. Darmanto, "Pengenalan Spesies Ikan

- Berdasarkan Kontur Otolith,” *Joined J. (Journal Informatics Educ.*, vol. 2, no. 1, pp. 41–59, 2019.
- [4] M. Yang and G. Thung, “Classification of Trash for Recyclability Status,” *CS229Project Rep.*, pp. 1–6, 2016.
- [5] A. Lodh and R. Parekh, “Flower recognition system based on color and GIST features,” in *2017 Devices for Integrated Circuit (DevIC)*, 2017, pp. 790–794, doi: 10.1109/DEVIC.2017.8074061.
- [6] H. Almogdady, S. Manaseer, and H. Hiary, “A Flower Recognition System Based On Image Processing And Neural Networks,” *Int. J. Sci. Technol. Res.*, vol. 7, 2018, [Online]. Available: <https://www.ijstr.org/final-print/nov2018/A-Flower-Recognition-System-Based-On-Image-Processing-And-Neural-Networks.pdf>.
- [7] S. N. Parvathy, N. V. Rao, S. B. S, N. Nazer, and P. A. J, “Flower Recognition System Using Cnn,” no. June, pp. 6609–6611, 2020, [Online]. Available: <https://www.irjet.net/archives/V7/i6/IRJET-V7I61229.pdf>.
- [8] A. Peryanto, A. Yudhana, and R. Umar, “Klasifikasi Citra Menggunakan Convolutional Neural Network dan K Fold Cross Validation,” *J. Appl. Informatics Comput.*, vol. 4, no. 1, pp. 45–51, 2020, doi: 10.30871/jaic.v4i1.2017.
- [9] N. S. B. Kusrorong, D. R. Sina, and N. D. Rumlaklak, “Kajian Machine Learning Dengan Komparasi Klasifikasi Prediksi Dataset Tenaga Kerja Non-Aktif,” *J. Komput. Inform.*, vol. 7, no. 1, pp. 37–49, 2019.
- [10] A. Yudhana, Sunardi, and S. Saifullah, “Segmentation comparing eggs watermarking image and original image,” *Bull. Electr. Eng. Informatics*, vol. 6, no. 1, pp. 47–53, 2017, doi: 10.11591/eei.v6i1.595.
- [11] Sunardi, A. Yudhana, and S. Saifullah, “Identity analysis of egg based on digital and thermal imaging: Image processing and counting object concept,” *Int. J. Electr. Comput. Eng.*, vol. 7, no. 1, pp. 200–208, 2017, doi: 10.11591/ijece.v7i1.pp200-208.
- [12] A. Peryanto, A. Yudhana, and R. Umar, “Rancang Bangun Klasifikasi Citra Dengan Teknologi Deep Learning Berbasis Metode Convolutional Neural Network,” *FORMAT J. Ilm. Tek. Inform.*, vol. 8, no. 2, pp. 138–147, 2019.
- [13] K. A. ElDahshan, M. I. Youssef, E. H. Masameer, and M. A. Mustafa, “Segmentation Framework on Digital Microscope Images for Acute Lymphoblastic Leukemia Diagnosis based on HSV Color Space,” *Int. J. Comput. Appl.*, vol. 90, no. 7, pp. 48–51, Mar. 2014, doi: 10.5120/15590-4426.
- [14] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. USA: Prentice-Hall, Inc., 2006.
- [15] Suyanto, *Machine Learning Tingkat Dasar dan Lanjut*. Bandung: Informatika, 2018.
- [16] M. F. Rahman, D. Alamsah, M. I. Darmawidjadja, and I. Nurma, “Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN),” *J. Inform.*, vol. 11, no. 1, pp. 36–45, 2017, doi: 10.26555/jifo.v11i1.a5452.
- [17] I. Saputra and D. Rosiyadi, “Perbandingan Kinerja Algoritma K-Nearest Neighbor , Naïve Bayes Classifier dan Support Vector Machine dalam Klasifikasi Tingkah Laku Bully pada Aplikasi Whatsapp,” *Fakt. Exacta*, vol. 12, no. 2, pp. 101–111, 2019.
- [18] G. Hackeling, *Mastering Machine Learning with scikit-learn*. Birmingham: Packt Publishing, 2014.
- [19] M. Toğaçar, B. Ergen, and Z. Cömert, “Classification of flower species by using features extracted from the intersection of feature selection methods in convolutional neural network models,” *Measurement*, vol. 158, p. 107703, 2020, doi: <https://doi.org/10.1016/j.measurement.2020.107703>.
- [20] B. R. Mete and T. Ensari, “Flower Classification with Deep CNN and Machine Learning Algorithms,” *2019 3rd Int. Symp. Multidiscip. Stud. Innov. Technol.*, pp. 1–5, 2019.