Jurnal Teknik Informatika (JUTIF)
Vol. 5, No. 1, February 2024, pp. 313-320

DOI: https://doi.org/10.52436/1.jutif.2024.5.1.1798
p-ISSN: 2723-3863
e-ISSN: 2723-3871

HYPERPARAMETER OPTIMIZATION OF CONVOLUTIONAL NEURAL NETWORK FOR FLOWER IMAGE CLASSIFICATION USING GRID SEARCH ALGORITHMS

Della Aulia Wibowo*[1], Nanik Suciati*[2], Anny Yuniarti[3]

[1,2,3]Department of Information Technology, Informatics, Engineering Faculty, Institut Teknologi Sepuluh Nopember, Indonesia
Email: [1]6025211027@student.its.ac.id, [2]nanik@if.its.ac.id, [3]anny@if.its.ac.id

(Article received: January 28, 2024; Revision: February 20, 2024; published: February 24, 2024)

Abstract

Indonesia is a country with a tropical climate that greatly affects agriculture. Flowering plants are estimated to account for 25% of species in Indonesia; there are 416 families, 13,164 genera, and 295,383 species of flowering plants. Classification of profit types is a time- and knowledge-intensive job. Convolutional Neural Network (CNN) has revolutionized the field of computer vision by improving the accuracy of image, text, voice, and video recognition. This research is focused on developing a CNN model for Indonesian flower images by optimizing hyperparameters combined with a grid search algorithm and default parameters, as well as comparing two different CNN architectures, namely VGG16 and MobileNetV2. This research aims to improve the classification accuracy of Indonesian flower images by optimizing hyperparameters. The results of CNN research with hyperparameters combined with a grid search algorithm and using data augmentation resulted in MobileNetV2 as the best model. Grid search is designed to get the best value of each parameter. The performance of the grid search algorithm can produce an optimal combination of parameters, with a test accuracy of 89.62%..

Keywords: *Convolutional Neural Network, Grid Search, Hyperparameter, Indonesian flowers image, MobileNetV2, VGG16.*

1. INTRODUCTION

Indonesia is an archipelago country with the largest land area in the world. By 2021, Indonesia's land area will have reached 1.9 million square kilometers [1]. Geographically, Indonesia lies between two main continents, the Asian continent and the Australian continent, as well as between two oceans, the Indian Ocean and the Pacific Ocean. According to its geographical position, it is 6^0 04' 30'' north latitude (LU) with 11^0 00' 36'' south latitude (LS) and between 94^0 58' 21'' and 141^0 01' 10" east latitude (BJ). Indonesia has an evident annual rainfall [2]. Because of its geographical location, the climate in Indonesia is tropical. The tropical climate causes heavy rain but also gets sunlight throughout the year. Therefore, tropical climates strongly influence agriculture because the soil in a country with tropical weather includes fertile soil.

Flowers is one of the most beautiful plants. Flowers are reproductive instruments in plants. Inside the flower, there are reproductive instruments: flower pistil and stamens. The petals are located on the outermost part of the flower. The leaves are green, and the shape of the leaves is like a leaf. The leaf protects the crown and the parts of the flower from external disturbance before the flowers bloom [3]. Botanically, flowers are parts of plants that produce seeds. Fertilization and fertilization occurred inside the flower. After fertilization, the flowers bloom, and the fruit is formed. Botanists have traditionally done flower classification. Flowering plants are estimated to cover 25% of species in Indonesia. This number equals 20,000 species, 40% endemic to Indonesia [4]. There are about 374,000 species of plants. There are 416 families, 13,164 generations, and 295,383 species of flowering plants, making it the most diverse group of terrestrial plants [5].

There's been some research on flower classification using deep learning. Convolutional Neural Network (CNN) is the most common deep learning method to classify flower diversity. In this study, the researchers used CNN techniques to organize flowers. CNN has revolutionized the field of computer vision by improving the accuracy of image, text, voice, and video recognition. With the development of the world of computing and the increase in the power and intelligence of computer processes, the advent of computer science enables computers to acquire information from images and automatically recognize objects.

One way to improve the performance of the CNN model is to optimize the hyperparameters [6]. Choosing the correct parameters is crucial to producing a good model. With many parameters, it is essential to know the most optimal parameters. One of the primary hyperparameters that need to be adjusted before the training process is the batch size, which is the number of images to be used in the gradient estimation process [7]. On the one hand, a small batch size can converge faster than a large batch, but a large batch can reach an optimal minimum point that a small batch size cannot get [8]. The learning rate can affect the speed of movement toward the gradient. When the movement toward the gradient is too fast, the model training process will be unstable, decreasing the model's performance [9]. The grid search algorithm is a combination method to get a good hyperparameter because it does a one-on-one trial on each combination [10]. The influence of the grid search algorithm on this research aims to obtain the best value of each parameter: optimizer, epoch, learning rate, and batch size [11].

This research conducted a literature study related to relevant theories through previous studies. Flower image research using the CNN model has previously been done [12], producing an accuracy value of 91%. The datasets used in this study are Daisy, Dandelion, Rose, Sunflower, and Tulip Flowers. Research by collecting flower images has been done [13], which consists of 35 flower classes. In his study, the accuracy obtained is 79%.

An approach based on deep learning methods for classifying Indonesian flowers [14]. This research proposes the CNN method using rose flower images. Roses in Malang Flower Park, East Java, Indonesia, are ornamental plants with more than 150 species. This research aims to help the sales system by quickly classifying the rose type. In this study, the accuracy obtained was 96.33%. Research with Indonesian flower datasets has also been carried out. The datasets used are aglonema, sunflowers, bonsai, and calamansi flowers [15].

Previous research used CNN methods with grid search algorithms for image classification [7]. CNN models for classifying brain tumors could be determined. The accuracy of the proposed framework of classifying brain MR images into five classes is 92.98%.

In another study, researchers tried to optimize the model using hyperparameters with the grid search method on the image classification dataset, namely images derived from the face detection process. The research aims to find the optimum value resulting from the hyperparameter, which is 97.56% [16]. The proposed parameters of the Convolutional Neural Network model were optimized using Grid Search. Then, they implemented using Keras and Tensorflow as it yielded better performance and accuracy than the regular classifiers proposed in various models [17].

Therefore, this research will apply the CNN method to classify Indonesian flower images, introduce the types of flora in Indonesia, and find out the best optimization of hyperparameters against grid search algorithms. This method will perform a combination test one by one and then select the combination that produces the best performance, the Confusion matrix used to measure the performance and quality of a model. In addition, the confusion matrix is also used to perform calculations such as finding accuracy, precision, recall, and F1-score values.

2. MATERIAL AND METHOD

This research uses deep learning with the CNN method using Python programming language. The research steps begin with data collection, data preparation, preprocessing, and deep learning model design. In the design of deep learning classification, data training is carried out using the CNN model. Then, the last stage evaluates the CNN model using a confusion matrix to determine the accuracy, precision, recall, and f1-score level. The research diagram can be seen in Figure 1.

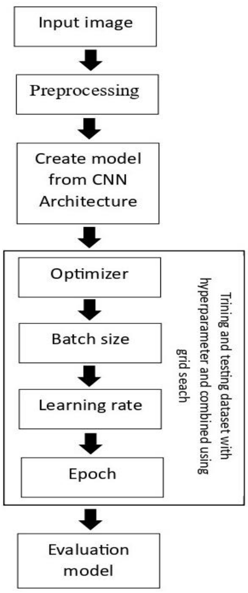

Figure 1. Research flow

2.1. Dataset

In this paper, we collected 1890 flower images. Researchers collect data sets randomly manually via the Google Images search engine with a Creative Commons Zero (CC0) license or images that are considered free to use. The following is a list of flowers in each province [18]. Jasmine gambir is a typical flower of DKI Jakarta province [19]. Corrosion flower was first known in the world of science after it was discovered by Dr Odoardo Beccari in 1878 in the Anai Valley area, West Sumatra [20]. The land of Papua has great potential for orchid wealth. Papua holds almost half of all orchid species found in Indonesia [21]. Manna River is located on the eastern edge of Manna City, South Bengkulu where Dr J Arnoldi discovered Rafflesia Arnoldi [22]. In Indonesia, piper betle is a typical flora of Riau Islands Province [23]. This dataset represents flower images in 33 provinces in Indonesia. Figure 2 Bar chart shows the amount of data used. We randomly selected 80% from each category for training and the remaining 20% for testing. The 80% of training will be split into 80%

training data and 20% validation data. Many images are affected by light, background, and unrelated objects, which will increase the difficulty of image recognition. However, CNN can ignore the influence and recognize the type of flower.

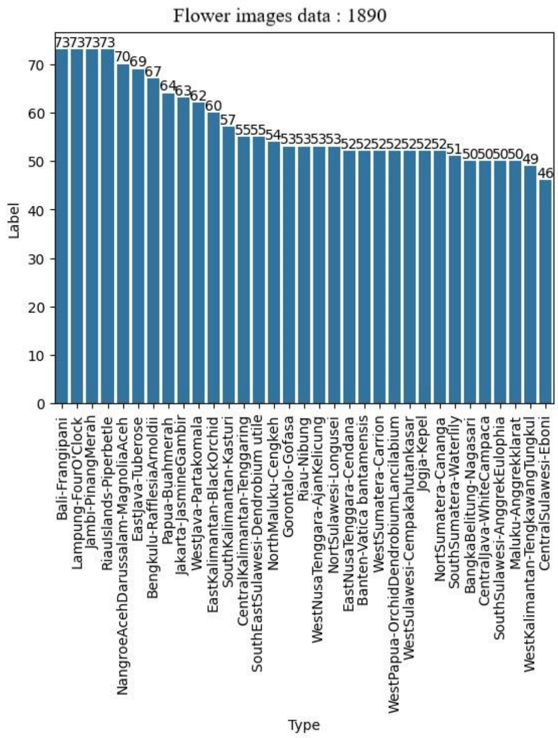

Figure 2. Flower dataset

2.2. Augmentation Data

Data augmentation aims to improve model performance. Data augmentation is the most commonly used method to reduce overfitting problems in image data by artificially enlarging data samples. The augmentation process begins with the initialization of ImageDataGenerator which is a function of the Keras library. Augmentation uses several parameters, namely rotation_range of 20 degrees,width_shift_range of 0.2, height_shift_range of 0.2, Zoom images are randomly enlarged or reduced in the range of 0.2, horizontal_flip and vertical_flip, and fill mode 'Nearest' which is the pixel filling mode used when shifting or rotation occurs. Then, the flow method is used to augment the train data. The result of the flow is used for the model training process.

2.3. Preprocessing Data

The preprocessing stage includes image resizing. The image is resized to align the image size of the dataset so that it can be processed. Then, a preprocessing scenario is put in to determine the optimal performance. In image resizing, the image size varies to become 180x180 pixels. Resizing images in preprocessing makes it easier for the training process and helps the training process get the maximum accuracy.

2.4. Architecture Convolutional Neural Network

A deep hierarchical neural network is called a CNN. It comprises full-connected neural layers, subsampling layers, and convolutional layers. The quantity of each type of layer and how the input data moves through the layers differ between CNNs [24]. Convolutional Neural Network (CNN) was first performed by [25] on the visual cortex of cats. Technically, CNN is an architecture that can be trained and consists of several stages. The architecture of the CNN model proposed in this research consists of an input layer, convolutional layer, pooling layer, dense layer, and output layer.

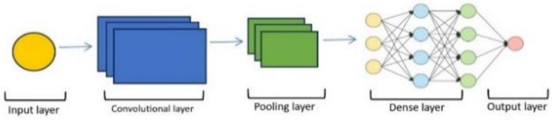

Figure 3. CNN Architecture

CNN works similarly to multilayer perceptron (MLP), but in CNN, each neuron is presented in two dimensions, unlike in MLP, where each neuron has only one size. So CNN can only be used on data with a two-dimensional structure, such as sound and image.

2.5. Visual Geometry Group (VGG)

VGG (Visual Geometry Group) is one of the convolutional neural network (CNN) architectures developed by the Visual Geometry Group at the University of Oxford. This model is known for its good performance in image classification and is used as a reference in CNN model architecture.

The VGG architecture model is characterized by using 3x3 convolutional stack layers. With this convolutional layer size, the depth of the artificial neural network can be modified with more convolutional layers to produce a higher level of accuracy. This characteristic is the reason why we use the VGG16 model in this study. There are 2 types of well-known VGG models, namely the VGG16 model and the VGG19 model.

2.6. MobileNetV2

MobileNetV2 is one of the architectures of Convolutional Neural Network (CNN). Researchers from Google created this architecture for mobile needs. The fundamental difference between MobileNetV2 and other CNN architectures is the use of convolution layers with filter thicknesses that match the input image. MobileNetV2 is also used for depthwise and pointwise convolutions (Wang et al., 2018).

MobileNetV2 can be implemented efficiently using standard operations in modern frameworks and can exceed high-tech products in various performance points using standard benchmarks. In addition, this convolutional module is well-suited for mobile designs and can significantly reduce the memory footprint required for inference. This reduces the need for main memory access in many embedded hardware designs that provide small amounts of software-controlled cache memory (Sandler et al., 2018).

2.7. Parameter Settings

Parameter settings are used to test the model on training data and testing data to produce the most effective parameters with the highest accuracy value of the model. with the highest accuracy value of the proposed model. Parameter settings that compared default parameters and hyperparameter tuning to optimize and compare the accuracy value of each setting. The parameters used to perform hyperparameters in this study are as follows:

Table 1. Parameters

Parameters	
Optmizer	SGD, Adam, RMSprop
Epoch	30, 50
Batch Size	32, 64
Learning rate	0.001, 0.0001

2.8. Grid Search

Grid search is an alternative method to find a model's best parameters. In this research, grid search will look for search for all parameter combinations. A systematic search method to find a combination of parameters that produces the best model performance [26]. The model is trained using training data and evaluated for performance using validation data. After all combinations are evaluated, the parameter that gives the best performance, such as the highest accuracy or something else, is selected as the optimal parameter.

2.9. Evaluation Model

A model evaluation is a step taken to evaluate the performance of a machine learning model. There are many model evaluation methods, including the evaluation metric and the confusion matrix. The confusion matrix measures how well models correctly predict labels and how well they expect the wrong labels [27]. The confusion matrix has four main parts: True Positive (TP), which is the sum of accurate optimistic predictions; False Positive (FP), which represents the number of false positive predictions; and False Negative (FN), which means the total number of wrong pessimistic predictions.

The confusion matrix helps to understand the model's performance in detail and allows for determining appropriate evaluation metrics, such as accuracy, precision, sensitivity, and F1 score. A Confusion matrix is also helpful in identifying problems in the model and helps to determine strategies for improving model performance.

Accuracy is used to evaluate prediction results precisely [28].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Recall is the number of correct results relative to predictions from all actual positives [29].

$$Recall = \frac{TP}{TP+FN} = \frac{TP}{P} \tag{2}$$

Precision is the number of true positives relative to all optimistic predictions [29].

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

F1-score is an evaluation matrix that describes the balance between precision and recall. The F1 score considers the hazardous impact of false positives and false negatives [30].

$$F = \frac{2 \times Precision \times Recall}{Precision+Recall} \tag{4}$$

3. RESULT AND DISCUSSION

3.1. Training process with parameter default

The training process on CNN, VGG16, and MobileNetV2 models without setting parameters. The parameters used are: there are 3 optimizers in this study including Adam optimizer, SGD optimizer, and RMSprop optimizer. 100 for epoch calculation. 0.01 in learning rate, and 64 batch size.

3.1.1. CNN

Table 2. Results of the training process of the CNN model with default parameters

Optimizer	**Trin Loss**	**Trian Accuracy**	**Val Loss**	**Val Accuracy**
Adam	0.2358	0.9264	1.3928	0.6788
SGD	0.8006	0.7628	1.4709	0.5762
RMSprop	0.1779	0.9471	1.3025	0.7053

Based on Table 2, the validation accuracy of the CNN model with default parameters tests a combination of 3 optimizer parameters. In the training process, the best validation accuracy produced is 70.53%. With RMSprop optimizer.

3.1.2. VGG16

Table 3. Results of the training process of the VGG16 model with default parameters

Optimizer	**Trin Loss**	**Trian Accuracy**	**Val Loss**	**Val Accuracy**
Adam	1.3025	0.5405	0.9412	0.6788
SGD	1.8657	0.4810	1.7215	0.6026
RMSprop	1.7678	0.4107	1.1794	0.6354

Based on Table 3, the validation accuracy of the VGG16 model with default parameters tested a combination of 3 optimizer parameters. In the training process, the best validation accuracy produced is 67.88%.

3.1.3. MobileNetV2

Table 4. Results of the training process of the MobileNetV2 model with default parameters

Optimizer	Trin Loss	Trian Accuracy	Val Loss	Val Accuracy
Adam	1.5240	0.5066	0.7967	0.7483
SGD	0.2693	0.9306	0.3994	0.8724
RMSprop	0.8422	0.7125	0.6638	0.8510

Based on Table 4, the validation accuracy of the CNN model with default parameters is testing a combination of 3 optimizer parameters. In the training process, the best validation accuracy produced is 87.24%. With SGD optimizer.

3.2. Training process with hyperparameter and grid search algorithm

In the classification of CNN models using grid search. hyperparameters have an important role in machine learning and deep learning algorithms because the resulting parameters significantly affect the performance of CNN models.

3.2.1. CNN

The CNN model training process uses the grid search algorithm. The following results of the parameters that have been trained on the CNN model are shown in Table 5. The CNN training model using the grid search algorithm performs the training process 24 times. The highest accuracy is 0.7384, with epoch 50, learning rate 0.0001, and batch size 32 with RMSprop optimizer.

Table 5. Results of the training process of the CNN model with hyperparameters

Optimizer	Epoch	Learning rate	Batch size	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
Adam	30	0.001	32	0.2412	0.9190	1.2596	0.7152
SGD	30	0.001	32	3.0481	0.1992	3.0369	0.2682
RMSprop	30	0.001	32	0.3179	0.8992	1.5930	0.6656
Adam	30	0.001	64	0.2701	0.9124	1.4843	0.6656
SGD	30	0.001	64	3.1040	0.1653	3.1046	0.1921
RMSprop	30	0.001	64	0.3062	0.8983	1.4024	0.6987
Adam	50	0.001	32	0.2480	0.9116	1.5300	0.6656
SGD	50	0.001	32	2.3818	0.3603	2.4626	0.3543
RMSprop	50	0.001	32	0.2596	0.9231	1.5796	0.6457
Adam	50	0.001	64	0.2996	0.9017	1.2136	0.7053
SGD	50	0.001	64	2.1457	0.4248	2.3354	0.4001
RMSprop	50	0.001	64	0.2416	0.9281	1.2482	0.7351
Adam	30	0.0001	32	0.2347	0.9207	1.1017	0.6887
SGD	30	0.0001	32	3.4821	0.0463	3.4907	0.0298
RMSprop	30	0.0001	32	0.2421	0.9331	13080	0.6887
Adam	30	0.0001	64	0.2796	0.9149	1.1350	0.6987
SGD	30	0.0001	64	3.4772	0.0388	3.4878	0.0199
RMSprop	30	0.0001	64	0.2405	0.9223	1.1122	0.7086
Adam	50	0.0001	32	0.2507	0.9132	1.0740	0.6887
SGD	50	0.0001	32	3.4669	0.0496	3.4767	0.0532
RMSprop	50	0.0001	32	0.2288	0.9347	1.0789	0.7252
Adam	50	0.0001	64	0.2128	0.9364	1.1934	0.7020
SGD	50	0.0001	64	3.4708	0.0570	3.4683	0.0530
RMSprop	50	0.0001	64	0.9215	0.9215	1.0871	0.7384

3.2.2. VGG16

Table 6. is the VGG16 model training process using hyperparameter and grid search algorithm. The following parameter results have been trained on the VGG16 model. The VGG16 training model using the grid search algorithm performed the training process 24 times. The highest validation accuracy is 0.8964, with epoch 50, and learning rate of 0.0001, and batch size 64 with Adam optimizer.

Table 6. Results of the training process of the VGG16 model with hyperparameters

Optimizer	Epoch	Learning rate	Batch size	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
Adam	30	0.001	32	0.4565	0.8475	0.5601	0.7925
SGD	30	0.001	32	1.5833	0.2644	1.5939	0.2075
RMSprop	30	0.001	32	0.1748	0.8415	0.2345	0.8218
Adam	30	0.001	64	0.6240	0.7932	0.6888	0.7789
SGD	30	0.001	64	1.5854	0.2559	1.5973	0.2075
RMSprop	30	0.001	64	0.4524	0.8534	0.3692	0.8014
Adam	50	0.001	32	0.0055	0.8975	0.1342	0.8728
SGD	50	0.001	32	1.6000	0.2314	1.6016	0.2245
RMSprop	50	0.001	32	0.7840	0.7314	0.8282	0.7177
Adam	50	0.001	64	0.0329	0.8924	0.8666	0.8558
SGD	50	0.001	64	1.5934	0.2517	1.5970	0.2075
RMSprop	50	0.001	64	0.5848	0.8059	0.7315	0.7551

Adam	30	0.0001	32	0.1236	0.9568	0.4242	0.8571
SGD	30	0.0001	32	1.5802	0.2568	1.5915	0.2075
RMSprop	30	0.0001	32	0.1977	0.8271	0.2392	0.8252
Adam	30	0.0001	64	0.3232	0.8898	0.3896	0.8435
SGD	30	0.0001	64	1.5816	0.2602	1.5931	0.2075
RMSprop	30	0.0001	64	0.2882	0.8000	0.2493	0.8354
Adam	50	0.0001	32	0.0067	0.8975	0.1361	0.8626
SGD	50	0.0001	32	1.6020	0.2102	1.6012	0.2109
RMSprop	50	0.0001	32	0.6118	0.7831	0.7235	0.7749
Adam	50	0.0001	64	0.0286	0.8985	0.2788	0.8964
SGD	50	0.0001	64	1.5886	0.2542	1.5933	0.2075
RMSprop	50	0.0001	64	0.7390	0.7373	0.8291	0.7279

3.2.3. MobileNetV2

The MobileNetV2 training model using the grid search algorithm and data augmentation performed the training process 24 times. The highest validation accuracy is 0.9073, with epoch 50, a learning rate of 0.0001, and batch size 64 with Adam optimizer. The following table of MobileNetV2 model training results with hyperparameter and grid search algorithm combinations can be seen in Table 7.

Table 7. Results of the training process of the VGG16 model with hyperparameters

Optimizer	Epoch	Learning rate	Batch size	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
Adam	30	0.001	32	0.1164	0.9587	0.3136	0.8841
SGD	30	0.001	32	1.9190	0.4967	1.7225	0.6854
RMSprop	30	0.001	32	0.1810	0.9331	0.3523	0.8874
Adam	30	0.001	64	0.2403	0.9231	0.3259	0.8775
SGD	30	0.001	64	1.8296	0.5140	1.6208	0.7053
RMSprop	30	0.001	64	0.1779	0.9372	0.3275	0.8808
Adam	50	0.001	32	0.1676	0.9438	0.3015	0.8841
SGD	50	0.001	32	1.1932	0.6942	0.9906	0.8444
RMSprop	50	0.001	32	0.1296	0.9545	0.3361	0.8775
Adam	50	0.001	64	0.2659	0.9083	0.3816	0.8675
SGD	50	0.001	64	1.1410	0.7050	1.0303	0.7848
RMSprop	50	0.001	64	0.1930	0.9190	0.3604	0.8642
Adam	30	0.0001	32	0.3007	0.9215	0.3703	0.8974
SGD	30	0.0001	32	3.4072	0.0785	3.1904	0.1788
RMSprop	30	0.0001	32	0.2833	0.9231	0.3856	0.8709
Adam	30	0.0001	64	0.2924	0.9231	0.3716	0.8808
SGD	30	0.0001	64	3.3840	0.0909	3.2813	0.1060
RMSprop	30	0.0001	64	0.2727	0.9273	0.3542	0.8841
Adam	50	0.0001	32	0.1972	0.9463	0.3313	0.8841
SGD	50	0.0001	32	0.1587	3.1288	3.0669	0.2119
RMSprop	50	0.0001	32	0.1757	0.9455	0.3357	0.8907
Adam	50	0.0001	64	0.1925	0.9455	0.3176	0.9073
SGD	50	0.0001	64	3.2306	0.1215	3.0812	0.2517
RMSprop	50	0.0001	64	0.1493	0.9579	0.3304	0.8742

3.3. Model evaluation with default parameters

The model evaluation uses a confusion matrix, and evaluation metrics (accuracy, precision, recall, and f1-score). The percentage of the evaluation metrics will be averaged by weighted average. Weighted average in the confusion matrix is a method to calculate the average performance of a classification model based on the number of examples in each class. It takes into account how important each class is by considering the number of examples that belong to that class. Confusion Matrix is a representation of the evaluation results of a classification model created using testing data and is used to provide a clearer picture of the accuracy of the model in predicting a class.

Table 8. Metric evaluation results with default parameters

Parameters	Accuracy	precision	Recall	F1-Score
CNN	0.6561	0.6932	0.6561	0.6546
VGG16	0.8016	0.8216	0.8016	0.8006
MobileNetV2	0.8386	0.8474	0.8386	0.8280

Table 9. Metric evaluation results with grid search algorithm and hyperparameters

Parameters	Accuracy	precision	Recall	F1-Score
CNN	0.8760	0.8214	0.8760	0.8164
VGG16	0.8959	0.8216	0.8959	0.8006
MobileNetV2	0.8962	0.8916	0.8962	0.8844

Based on Table 8, it can be seen that research by performing several optimation of hyperparameters such as epoch, batch size, learning rate, and optimizer can provide very good results. This is evidenced from aseries of experiments conducted so that MobileNetV2 architecture is higher with an accuracy value of 89.62%, precission 89.16%, recall 89.62% and f1-score 88.44%.

4. CONCLUSIONS

This research aims to get the right parameters to improve the performance of the model by conducting a hyperparameter tuning process using the grid search algorithm on the CNN model. Researchers also conducted research on the VGG16 and MobileNetV2 architectures as a trial on the Indonesian flower image dataset. optimization of the CNN model using several hyperparameters such as epoch, batch size, learning rate, and optimizer for the classification of Indonesian flower images. This research aims to get the optimal hyperparameters to provide good performance on the CNN model. Based on the experiments that have been carried out, the determination of hyperparameters is very influential on the performance of the model. Hyperparameters with the number of epochs 50, batch size 64, learning rate 0.0001, and optimizer Adam provide the most optimal results.

REFERENCES

[1] B. S. Indonesia, “Statistical Yearbook of Indonesia,” Indonesia, 2019.

[2] S. Wirjohamidjojo dan Y. Swarinoto, *Iklim kawasan Indonesia*, Jakarta: Badan Meteorologi Klimatologi dan Geofisika, 2010.

[3] E. Ekawati, *Mata Pelajaran Penyerbukan dan Pemngkasan Tanaman Perkebunan*, Indonesia: Kemendikbud, 2017.

[4] T. Whitmore, K. Sidiyasa dan T. Whitmore, “Tree species enumeration of 0.5 hectare on Halmahera.,” Gardens’ *Bulletin Singapore*, vol. 40(1 & 2), pp. 31-34., 1987.

[5] M. Christenhusz dan J. W. Byng, “The number of known plant species in the world and its annual increase,” *Phytotaxa*, vol. 261(3). pp. 201-217,2016, http://dx.doi.org/10.11646/phytotaxa.261.3.1

[6] J. H. Yoo, H. i. Yoo, H. G. Kim, H. S. Yoon dan S. S. Han, “Optimization of Hyper-parameter for CNN Model using Genetic Algorithm,” dalam *2019 1st International Conference on Electrical, Control and Instrumentation Engineering (ICECIE), Malaysia*, 2019, https://doi.org/10.1109/ICECIE47765.2019.8974762

[7] N. Mukkapati dan D. M. S. Anbarasi, “Multi-Class Classification Framework for Brain Tumor MR Image Classification by Using Deep CNN with Grid-Search Hyper Parameter Optimization Algorithm,” *IJCSNS International Journal of Computer Science and Network Security*, vol. 22, 2022, https://doi.org/10.22937/IJCSNS.2022.22.4.14

[8] I. Kandel dan M. Castelli, “The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset,” *ICT Express*, vol. 6, no. 4, pp. 312-315, 2020, https://doi.org/10.1016/j.icte.2020.04.010

[9] Y. Yoo, “Hyperparameter optimization of deep neural network using univariate dynamic encoding algorithm for searches,” *Knowledge-Based Systems*, vol. 178, pp. 74-83, 2019, https://doi.org/10.1016/j.knosys.2019.04.019

[10] N. A. K. R. Fatmawati, “Klasifikasi Penyakit Diabetes Retinopati Menggunakan Support Vector Machinedengan Algoritma Grid Search Cross-Validation,” *Jurnal Riset Statistik*, vol. 3,pp.79-86,2023, https://doi.org/10.29313/jrs.v3i1.1945

[11] F. Sia dan N. S. Baco, “Hyperparameter Tuning of Convolutional Neural Network for Fresh and Rotten Fruit Recognition,” dalam *2023 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Malaysia*, 2023, https://doi.org/10.1109/IICAIET59451.2023.10291915

[12] G. Yifei, Q. Chuxian, X. Jiexiang, M. Yixuan dan T. T. Toe, “Flower image classification based on improved convolutional neural network,” dalam *2022 12th International Conference on Information Technology in Medicine and Education (ITME)*, China, 2022, https://doi.org/10.1109/ITME56794.2022.00028

[13] D. Guru, Y. S. Kumar dan S. Manjunath, “Textural features in flower classification,” *Mathematical and Computer Modelling*, vol. 54, pp. 1030-1036, 2011, http://dx.doi.org/10.1016/j.mcm.2010.11.032

[14] I. A. Anjani, Y. R. Pratiwi dan N. B. Nurhuda, “Implementation of Deep Learning Using Convolutional Neural Network Algorithm for Classification Rose Flower,” *International Conference on Science Education and Technology (ICOSETH) 2020*, vol. 1842, 2021, http://dx.doi.org/10.1088/1742-6596/1842/1/012002

[15] M. Fatoni dan Ernastuti, “Ornamental Plants Classification using,” *MIND (Multimedia Artificial Intelligent Networking Database) Journal*, vol. 8, pp. 158-172, 2023, https://doi.org/10.26760/mindjournal.v1i1.49

[16] A. Nurhopipah dan N. A. Larasati, “CNN hyperparameter optimization using random grid coarse-to-fine,” *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, vol. 6, pp. 19-26, 2021, https://doi.org/10.22219/kinetik.v6i1.1185

[17] R. L. Devi dan V. S. V, “Detection and

Automated Classification of Brain Tumor Types in MRI Images using Convolutional Neural Network with Grid Search Optimization," dalam *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), India*, 2021, https://doi.org/10.1109/I-SMAC52330.2021.9640670

[18] Prosea, Daftar flora identitas provinsi di Indonesia, Indonesia, 2011.

[19] J. Dinkes, "Melati Gambir: Ikonik, Puspa Bangsa, Filosofis," Indonesia, 2022.

[20] W. Hetterscheid dan S. Ittenbach, "Everything you always wanted to know about Amorphophallus but were afraid to stick your nose into.," *Aroideana*, vol. 19, pp. 7-131, 1996.

[21] P. B. BBKSDA, "Keanekaragaman Anggrek (Orchidaceae) di region Papuasia," Kementerian Lingkungan Hidup dan Kehutanan, Papua Barat, Indonesia, 2020.

[22] A. Susatya, *Rafflesia Pesona bunga terbesar di dunia, Bengkulu*, Indonesia: Direktorat Kawasan Konservasi dan Bina Hutan Lindung, 2011.

[23] A. Hamsa, T. Aulawi dan B. Solfan, "Perbedaan Waktu Pemanenan Terhadap Mutu Kimia Daun Sirih Merah (Piper Crocatum Ruiz & Pav)," *Jurnal Pertanian Indonesia*, vol. 1, p. 2, Oktober 2020.

[24] F. Hu, F. Yao dan C. Pu, "Learning Salient Features for Flower Classification Using Convolutional Neural Network," dalam *2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIIS),* China, 2020, https://doi.org/10.1109/ICAIIS49377.2020.9194931

[25] D. H. Hubel dan T. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *The Journal of Physiology*, vol. 160, no. 1, pp. 106-154, 1962.

[26] S. Sundhararajan, A. Pahwa dan P. Krishnaswami, "A comparative analysis of genetic algorithms and directed grid search for parametric optimization," *Engineering with computers*, vol. 14, pp. 197-205, 1998.

[27] L.-E. Pommé, R. Bourqui, R. Giot dan D. Auber, "Relative Confusion Matrix: Efficient Comparison of Decision Models," dalam *2022 26th International Conference Information Visualisation (IV),* Austria, 2022, https://doi.org/10.1109/IV56949.2022.00025

[28] A. Jierula, S. Wang, Tae-Min dan P. Wang, "Study on Accuracy Metrics for Evaluating the Predictions of Damage Locations in Deep Piles Using Artificial Neural Networks with Acoustic Emission Data," *Applied Science*, 2021, http://dx.doi.org/10.3390/app11052314

[29] P. Fränti dan R. Mariescu-Istodor, "Soft precision and recall," *Pattern Recognition Letters*, vol. 167, pp. 115-121, 2023, https://doi.org/10.1016/j.patrec.2023.02.005

[30] S. A. Hicks, I. Strümke, V. Thambawita, M. Hammou, M. A. Riegler, P. Halvorsen dan S. Parasa, "On evaluation metrics for medical applications of artificial intelligence," *Scientific Reports*, vol. 12, no. 1, 2022, https://doi.org/10.1038/s41598-022-09954-8.