Computer Engineering and Applications Vol. 13, No. 2, June 2024

Image Classification of Traditional Indonesian Cakes Using Convolutional Neural Network (CNN)

Azkiya Nur Azizah[1*], Irwan Budiman[1], Fatma Indriani[1], M.Reza Faisal[1], Rudy Herteno[1]

[1]*Department of Computer Science, Faculty of Mathematics and Natural Science, Lambung Mangkurat University*
**nurazkiyaazizah@gmail.com*

ABSTRACT

Indonesia is one of the countries famous for its traditional culinary. Traditional cakes in Indonesia are traditional snacks typical of the archipelago's culture which have a variety of textures, shapes, colors that vary and some are similar so that there are still many people who do not know the name of the cake from the many types of traditional Indonesian cakes. The problem can be solved by creating a traditional cake image recognition system that can be programmed and trained to classify various types of traditional Indonesian cakes. The Convolutional Neural Network method with the AlexNet architecture model is used in this research to predict various kinds of traditional Indonesian cakes. The dataset used in this research is 1846 datasets with 8 classes of cake images. This study trained the AlexNet model with several optimizers, namely, Adam optimizer, SGD, and RMSprop. The best parameters from the model testing results are at batchsize 16, epoch 50, learning rate 0.01 for SGD optimizer and learning rate 0.001 for Adam and RMSprop optimizers. Each optimizer tested produces different accuracy, precision, recall, and f1_score values. The highest test results that have been carried out on the image dataset of typical Indonesian traditional cakes are obtained by the Adam optimizer with an accuracy value of 79%.

Keywords: Classification, Traditional Cake, AlexNet, Convolutional Neural Network Oprimizer Adam, SGD, RMSprop

1. INTRODUCTION

Indonesia is one of the countries famous for its traditional culinary.. Among the traditional snacks that reflect the cultural richness of the archipelago and are still very popular with the Indonesian people are dadar gulung, kastengel, klepon, layer cake, mud cake, snow princess, risoles, and serabi cake.

Each region has its own characteristics with a variety of traditional cakes that have a unique appearance and taste typical of Indonesian culture with a variety of shapes and textures, there are even cakes that are similar to each other, making it difficult to recognize the various kinds of cakes[1].

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

As time goes by, technological advances are happening rapidly, enabling the development of applications or systems that can recognize traditional cakes through images, thereby preserving their existence even as contemporary food trends threaten to make them forgotten. By collecting information about the texture, shape, and color of the image data of these cakes, accurate recognition is achieved, making it easier to identify a wide variety of traditional cakes, whether for research purposes or documentation. One successful approach to classifying images is using Artificial Neural Networks. Convolutional Neural Networks (CNN) can be employed for object classification problems in the form of images because they utilize a convolution operation that combines several processing layers, employing multiple elements that operate in parallel and taking inspiration from the biological nervous system [2].

CNN has several architectures, one of the popular ones is the AlexNet architecture. AlexNet uses the basic process of CNN, namely, feature extraction and fully-connected layers with a robust architecture consisting of 8 convolution layers and 3 fully connected layers. Alexnet architecture was trained using ImageNet datasets, achieved outstanding accuracy in object image classification and emerged as the winner in the ImageNet LSVRC competition in 2012 under the image classification category[3].

Several researchers have conducted research on the classification of traditional cakes. In a previous study, CNN methods were used with different architectures, such as in traditional cake research using CNN with the mobilenetV2 architecture. The dataset included 845 traditional cake datasets combined from personal collection images and the internet. The research achieved an accuracy value of 0.96 and a loss of 0.14 in its validation. The best classification accuracy value obtained from the model in this study reached 95%[4]. In another research on traditional cakes, the CNN Algorithm technique was applied to 1676 traditional cake datasets from Indonesia, consisting of 80% training data and 20% testing data. The performance evaluation resulted in a conclusion that the classification accuracy reached 65.00%[5].

In another research with a different dataset of 300 signature images using the CNN method with the same architecture, AlexNet with Adam, SGD, and RMSprop optimizers can classify accurately with the highest accuracy obtained by the SGD optimizer compared to the Adam and RMSprop optimizers with an accuracy value of 90%.[6].

Based on the background of the Convolutional Neural Network (CNN) method, the best model that can be utilized as a reference in object detection and image classification of the types of Indonesian Traditional Cakes is sought. In this study, the Convolutional Neural Network method with the AlexNet architecture model and three optimizers, namely Adam, SGD, and RMSprop optimizers, is employed. The evaluation is based on the accuracy value to compare and determine the best performance results of each optimizer in recognizing various kinds of traditional Indonesian cakes.

2. MATERIAL AND METHODS

This study uses the Convolutional Neural Network (CNN) method to classify the Indonesian traditional cake dataset with the AlexNet architecture. The AlexNet model is trained using several optimizers, namely Adam optimizer, SGD, and

RMSprop. The stages of this research can be seen in Figure 1. In the first stage, the Indonesian traditional cake dataset will be divided into testing, training, and validation data, as shown in Table 1.

FIGURE 1. Research Stage

2.1 MATERIAL

In designing the classification of typical Indonesian traditional cakes using convolutional neural networks, the dataset used is titled "Indonesian Cakes (Image of various Indonesian traditional cakes)," obtained from the Kaggle.com website. This dataset contains a total of 1846 images of typical Indonesian traditional cakes. The dataset is divided into 8 types of traditional Indonesian cakes, each with diverse and asymmetrical shapes. Most of the image data in this dataset has a size of more than 400 pixels. Before performing classification, the data is first divided into Training Data, Validation Data, and Testing Data, as shown in Table 1.

TABLE 1.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Dataset Division

Deskripsi	Testing	Training	Validation	Jumlah
Kue Dadar Gulung	20	192	20	232
Kue Kastengel	22	180	22	224
Kue Klepon	20	200	20	240
Kue Lapis	20	201	20	241
Kue Lumpur	20	198	20	238
Kue Putri Salju	20	174	20	214
Kue Risoles	20	196	20	236
Kue Serabi	20	181	20	221
TOTAL	162	1.522	162	1.846

2.2 METHODS

2.2.1 CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Network (CNN) is a type of multilayer Perceptron (MLP) specifically designed for processing two-dimensional data, such as images, as part of the Deep Natural Network. The Deep Natural Network is characterized by its high network depth and is well-suited for handling image data. While image classification using MLP is possible, it fails to effectively capture spatial information from the images. This limitation leads to unsatisfactory classification results, as MLP treats each pixel as an independent feature.The discovery of CNN is credited to Hubel and Wiesel, who initially researched the visual cortex in cats to understand their sense of sight. Their work on the visual cortex paved the way for the development of CNNs, which have since become a crucial tool in image recognition and other related tasks[7].

CNNs are mathematical structures, typically comprising three main types of layers: Convolution layer, Pooling layer, and Fully Connected Layer. The initial two layers, convolution and pooling, are responsible for conducting feature extraction. On the other hand, the third layer, which is fully connected, maps these extracted features to the ultimate output, often used for tasks like classification.

The neural network used in CNN enables it to mimic the way a human observes and analyzes an object. While humans may sometimes make mistakes in interpreting objects directly, CNN's presence can significantly enhance the accuracy and speed of object interpretation and recognition processes[8].

2.2.2 ALEXNET ARCHITECTURE

AlexNet Architecture model is made of 8 layers, where the 8 layers are divided into 2 parts as "Convolution" and "Fully Connected". The Convolution part consists of 5 layers while Fully Connected consists of 3 layers[9]. The design of the AlexNet Architecture model can be seen in Figure 2.

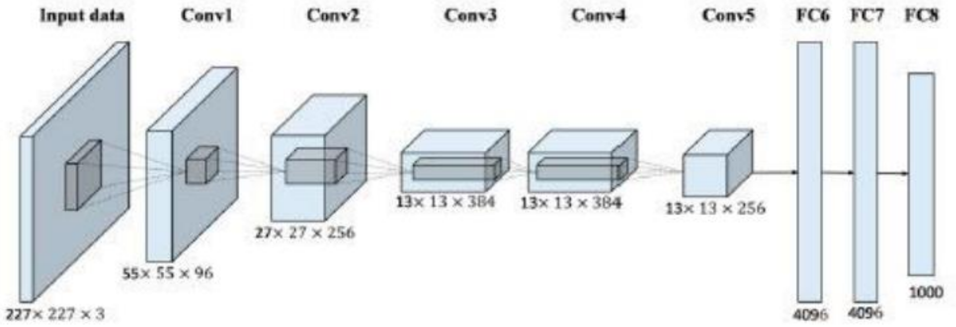

FIGURE 2. Architecture AlexNet

2.2.3 ADAM OPTIMIZER

Adam's optimizer is an effective optimizer and can train data in a fairly short time. This optimizer is a development of the classic Stochastic Gradient Descent (SGD) optimizer which has undergone an update of its network weight.

2.2.4 SGD OPTIMIZER

SGD or Stochastic Gradient Descent is one of the most popular algorithms for optimizing artificial neural networks. This algorithm is used to update parameters, specifically the weights and biases. It is quite simple to understand as it aims to reduce the initial weight by a "part" of the gradient value that has been obtained. Gradient Descent works by minimizing a function $J(\theta)$ that depends on a parameter θ through downward parameter updates. The optimization goal of this algorithm is to find parameters that can minimize the loss function.

2.2.5 RMSPROP OPTIMIZER

Root Mean Square Propagation or RMSprop is an optimizer created by Professor Geoffrey Hinton during his neural network class, and it is a specialized version of Adagrad. Similar to Adaprop, another optimizer aiming to resolve some of Adagrad's unresolved problems, RMSprop takes the latest gradient magnitude to normalize the gradient, maintaining the moving average above the root mean square gradient, which gives it the name "Rms." The RMSprop optimizer utilizes an adaptive approach to adjust the learning rate for each parameter, leading to faster model convergence.

2.2.6 DATA PRE-PROCESSING

Image Resizing is a preprocessing stage where the image data to be used is resized in terms of pixels. The current image dataset has diverse and asymmetrical shapes with a size of more than 400 pixels. In this stage, each previous image data with a size of 400 pixels will be changed to a uniform size of 227 x 227, matching the optimal default image size input for AlexNet architecture modeling. A comparison of the output results of the resizing process can be seen in Figure 3.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

FIGURE 3. Image Output Before and after Resize

2.2.7 CONFUSION MATRIX

For each optimizer, a classification test is conducted to classify each class, and then confusion matrix calculations are performed to obtain classification results for the multiclass data from the traditional cake dataset. The True Positive (TP) represents positive data that is predicted correctly, True Negative (TN) indicates negative data that is predicted correctly, False Positive (FP) represents negative data but predicted as positive data, and False Negative (FN) indicates positive data that is predicted as negative data. These values can be used to calculate accuracy, precision, recall, and F1-score values. The Confusion Matrix can be observed in Figure 4.

Predicted Values	Actual Values: Positive (1)	Actual Values: Negative (0)
Positive (1)	TP	FP
Negative (0)	FN	TN

FIGURE 4. Confusion Matrix

Here are the formulas used to calculate the accuracy, precision, recall, and f1-score values:

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

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{2}$$

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{3}$$

$$F1 - Score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{4}$$

Accuracy is a measure that evaluates the model's performance in classifying all of its classes. It is computed by summing up the correct prediction metrics and then dividing it by the total number of predictions. Precision is employed to calculate the accuracy of predictions for a specific true class, taking into account both correct predictions and misclassifications. Recall, on the other hand, gauges the model's capability to identify instances of the true class. The F1-Score, which is the mean of precision and recall, serves as a measure of the overall balance between these two metrics. A high F1-Score indicates that the classification model has good precision and recall values.

3. RESULT AND DISCUSSION

3.1 TESTINGG RESULT WITH ADAM OPTIMIZER

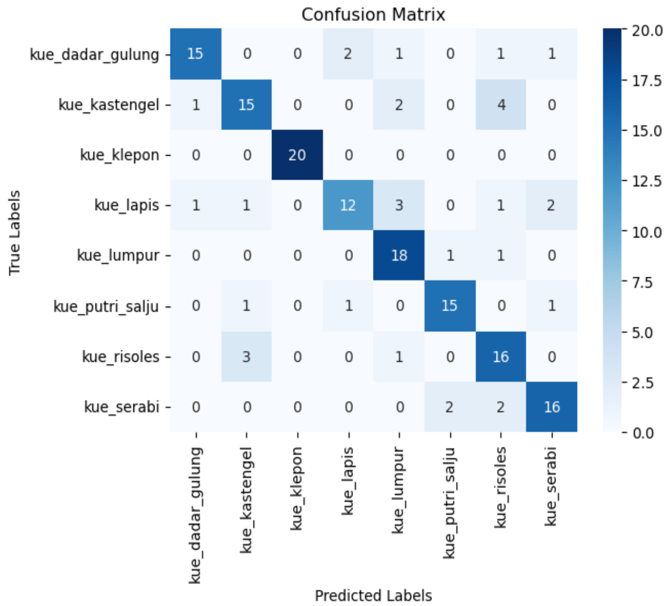

FIGURE 5. Confusion Matrix Results with Adam

Based on the matrix classification results of the AlexNet Architecture using Adam's Optimizer shown in Figure 5, it demonstrates the model's optimality on the kue_klepon image data. Out of 20 image data, all 20 images are classified correctly and accurately without any mispredictions. However, on the kue_lumpur image, 18 images are correctly predicted, while the remaining 2 are mispredicted as kue_putri_salju and kue_risoles, respectively.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

TABLE 2.
Optimizer Adam Classification Results

Class	precision	recall	f1-score
kue_dadar_gulung	0.88	0.75	0.81
kue_kastengel	0.75	0.68	0.71
kue_klepon	1.00	1.00	1.00
kue_lapis	0.80	0.60	0.69
kue_lumpur	0.72	0.90	0.80
kue_putri_salju	0.83	0.83	0.83
kue_risoles	0.64	0.80	0.71
kue_serabi	0.80	0.80	0.80

TABLE 3.
Adam Optimizer Performance Results

Accuracy	79%
Precision	80%
Recall	79%
F1-Score	79%

The results of the Adam optimizer performance classification report in Table 3 show an accuracy value of 79%, and get a precision value of 80%, recall 79%, and f1-score 79%.

3.2 TESTING RESULT WITH SGD OPTIMIZER

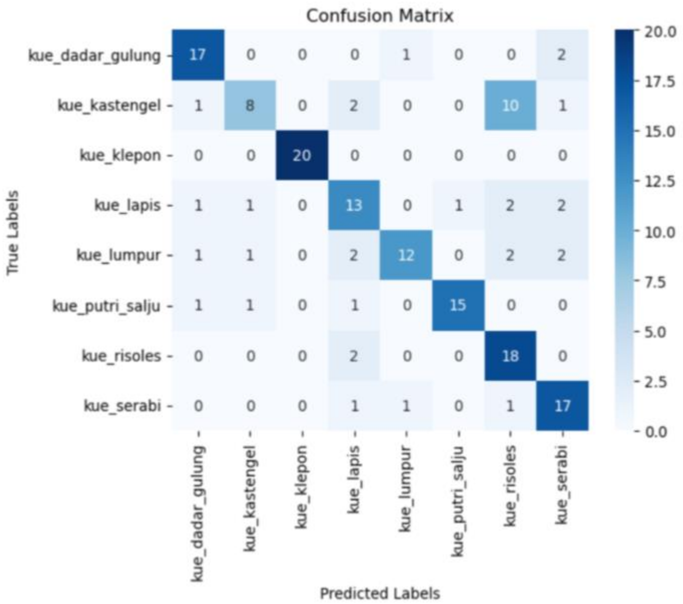

FIGURE 6. Confusion Matrix Results with SGD

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Based on the matrix classification results of the AlexNet Architecture using the SGD Optimizer shown in Figure 6, it demonstrates the model's optimality in the kue_klepon image data. Out of 20 image data, all 20 images are classified correctly and accurately without any mispredictions. However, on the kue_risoles image, 18 images are correctly predicted, while the remaining 2 are mispredicted as kue_lapis.

TABLE 4.
Optimizer SGD Classification Results

Class	precision	recall	f1-score
kue_dadar_gulung	0.81	0.85	0.83
kue_kastengel	0.73	0.36	0.48
kue_klepon	1.00	1.00	1.00
kue_lapis	0.62	0.65	0.63
kue_lumpur	0.86	0.60	0.71
kue_putri_salju	0.94	0.83	0.88
kue_risoles	0.55	0.90	0.68
kue_serabi	0.71	0.85	0.77

TABLE 5.
SGD Optimizer Performance Results

Accuracy	75%
Precision	77%
Recall	75%
F1-Score	74%

The results of the SGD optimizer performance classification report in Table 5 are obtained with an accuracy value of 75% and get a precision value of 77%, recall 75%, and f1-score 74%.

3.3 TESTING RESULT WITH RMSPROP OPTIMIZER

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

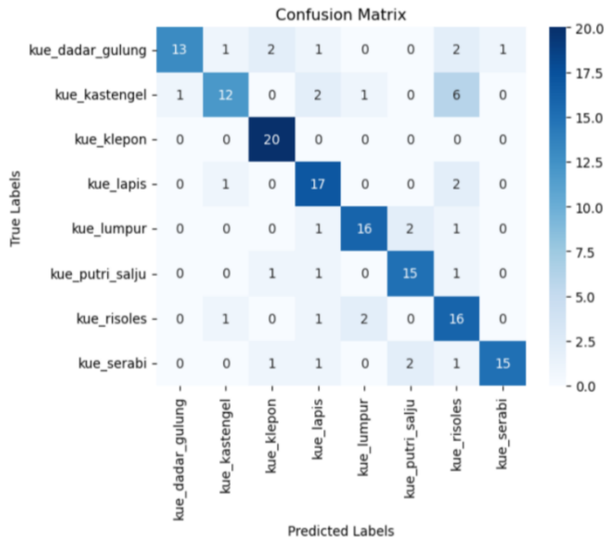

FIGURE 7. Confusion Matrix Results with RMSprop

Based on the matrix classification results of the AlexNet Architecture using the SGD Optimizer shown in Figure 7, it demonstrates the model's optimality on the kue_klepon image data. Out of 20 image data, all 20 images are classified correctly and accurately without any mispredictions. However, on the kue_lapis images, 17 images are correctly predicted, while the remaining 3 are mispredicted as kue_kastengel and kue_risoles, respectively.

TABLE 6.
Optimizer RMSprop Classification Results

Class	precision	recall	f1-score
kue_dadar_gulung	0.93	0.65	0.76
kue_kastengel	0.80	0.55	0.65
kue_klepon	0.83	1.00	0.91
kue_lapis	0.71	0.85	0.77
kue_lumpur	0.84	0.80	0.82
kue_putri_salju	0.79	0.83	0.81
kue_risoles	0.55	0.80	0.65
kue_serabi	0.94	0.75	0.83

TABLE 7.
RMSprop Optimizer Performance Results

Accuracy	77%
Precision	79%
Recall	77%
F1-Score	77%

The results of the RMSprop optimizer performance classification report in Table 7 are obtained with an accuracy value of 77% and get a precision value of 79%, recall 77%, and f1-score 77%.

3.4 ANALYSIS OF MODEL WITH OPTIMIZER

Based on the results of classification testing on the AlexNet Architecture model with the three optimizers: Adam, SGD, and RMSprop, the confusion matrix results on traditional Indonesian cake images show varying outcomes. Particularly, there are interesting accuracy results for the klepon cake image, which can be classified with optimal prediction results. Out of 20 image data, all 20 images are classified correctly and accurately without any prediction errors when testing with the three optimizers: Adam, SGD, and RMSprop.

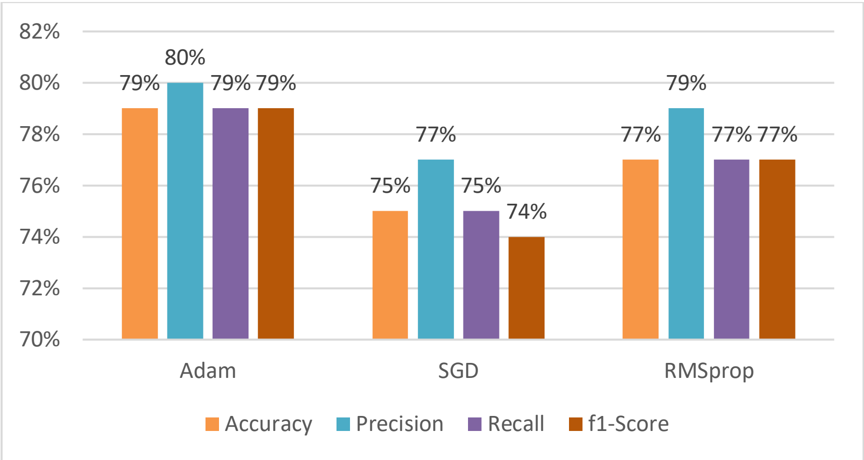

FIGURE 8. Graph of Performance Results of Each Optimizer

Figure 8 illustrates a graphical comparison of the performance results of the three optimizers conducted during the model training stage. Based on these outcomes, it is evident that the Adam optimizer achieved the best performance in AlexNet modeling for image data classification of traditional Indonesian cake images. The Adam optimizer obtained an accuracy value of 79%, precision value of 80%, recall of 79%, and f1-score of 79%.

4. CONCLUSION

Based on the results of the research and the discussions conducted, the following conclusions can be drawn :

a. Image classification on the traditional Indonesian Cake dataset using the AlexNet model demonstrates quite good and accurate performance, with the highest accuracy achieved by the Adam optimizer compared to the SGD and RMSprop optimizers. This is evident from the dataset testing, where the Adam optimizer obtained the highest accuracy value of 79%, along with a precision value of 80%, recall of 79%, and f1-score of 79%.
b. The parameters of the AlexNet model greatly influence the improvement of accuracy performance during classification.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

REFERENCES

[1] D. Sumarlie *et al.*, "PENGENALAN KUE TRADISIONAL INDONESIA," vol. 2, pp. 164–171, 2022.

[2] F. Hu, G. S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery," *Remote Sens.*, vol. 7, no. 11, pp. 14680–14707, 2015, doi: 10.3390/rs71114680.

[3] S. Ilahiyah and A. Nilogiri, "Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network," *JUSTINDO (Jurnal Sist. dan Teknol. Inf. Indones.*, vol. 3, no. 2, pp. 49–56, 2018.

[4] T. Karlita, B. P. Afif, and I. Prasetyaningrum, "Indonesian Traditional Cake Classification Using Convolutional Neural Networks," *Proc. Int. Conf. Appl. Sci. Technol. Soc. Sci. 2021 (iCAST-SS 2021)*, vol. 647, pp. 924–929, 2022, doi: 10.2991/assehr.k.220301.153.

[5] D. A. Kurnia, A. Setiawan, D. R. Amalia, R. W. Arifin, and D. Setiyadi, "Image Processing Identifacation for Indonesian Cake Cuisine using CNN Classification Technique," *J. Phys. Conf. Ser.*, vol. 1783, no. 1, 2021, doi: 10.1088/1742-6596/1783/1/012047.

[6] K. Wijaya and E. P. Widiyanto, "Klasifikasi Kepemilikan Tanda Tangan Menggunakan Convolutional Neural Network dengan Arsitektur AlexNet," *MDP Student Conf.*, vol. 2, no. 1, pp. 133–143, 2023, doi: 10.35957/mdp-sc.v2i1.4328.

[7] I. M. Ihdal, "Klasifikasi Kain Khas Batik Dan Kain Khas Sasirangan Dengan Menggunakan Metode Convolutional Neural Network," *J. Teknol. Inf. Univ. Lambung Mangkurat*, vol. 6, no. 1, pp. 25–30, 2021, doi: 10.20527/jtiulm.v6i1.62.

[8] L. Alzubaidi *et al.*, *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions*, vol. 8, no. 1. Springer International Publishing, 2021. doi: 10.1186/s40537-021-00444-8.

[9] D. Irfansyah, M. Mustikasari, and A. Suroso, "Arsitektur Convolutional Neural Network (CNN) Alexnet Untuk Klasifikasi Hama Pada Citra Daun Tanaman Kopi," *J. Inform. J. Pengemb. IT*, vol. 6, no. 2, pp. 87–92, 2021, [Online].Available:http://ejournal.poltektegal.ac.id/index.php/informatika/article/view/2802