Application of Convolutional Neural Network Method with MobileNet V1 Architecture in Batik Motif Classification

Aulia Chusnyriani Sani Zulkarnaen1, I Gusti Ngurah Rejski Ariantara Putra1, Nada Fauzia Reviana1, Rahmawati Hidayah1, Nur Ibrahim1,2, Nor Kumalasari Caecar Pratiwi1, Yunendah Nur Fuadah1

1 Faculty of Electrical Engineering, Telkom University

auliachusnyriani@student.telkomuniversity.ac.id, [gustiari@student.telkomuniversity.ac.id](mailto:gustiari@student.telkomuniversity.ac.id), [nadafauzia@student.telkomuniversity.ac.id](mailto:nadafauzia@student.telkomuniversity.ac.id), [rahmawatihidayah@student.telkomuniversity.ac.id](mailto:rahmawatihidayah@student.telkomuniversity.ac.id), [nuribrahim@telkomuniversity.ac.id](mailto:nuribrahim@telkomuniversity.ac.id), [caecarnkcp@telkomuniversity.ac.id](mailto:caecarnkcp@telkomuniversity.ac.id), [yunendah@telkomuniversity.ac.id](mailto:yunendah@telkomuniversity.ac.id)

2 Department of Electrical Engineering, Universitas Indonesia

[nur.ibrahim@ui.ac.id](mailto:nur.ibrahim@ui.ac.id)

**Abstract.** Indonesia is a country with diverse natural resources, cultures, and languages. One of the cultural diversities in Indonesia is Batik, which is an Indonesian cultural heritage consisting of a cloth that is hand-drawn using traditional techniques. To assist the community in recognizing various batik motifs, a classification method is developed to identify the types of batik through input images. The classification method utilizes a Convolutional Neural Network (CNN) based on the MobileNet V1 architecture. This research employs a dataset comprising 660 images of batik from six different batik motifs, namely Ceplok, Parang, Nitik, Megamendung, Kawung, and Tambal. The optimal classification model is obtained using the Adamax optimizer, without pre-processing width and zoom range, with an input size of (300, 300), a batch size of 32, and trained for 100 epochs. The model achieves an accuracy of 93.33% on training and 66.66% on validation and testing.

1 Introduction

Indonesia is a country with rich natural resources, diverse ethnicities, cultures, and languages [1]. Indonesia has one culture known as Batik. Batik is a craft and one of the unique cultural heritages of the archipelago that has become a high art and cultural heritage since ancient times [2]. Batik has several definitions, as one of the definitions of Batik is a visual art form from Indonesia that is produced using traditional drawing techniques on materials [3]. Each region in Indonesia has a distinctive batik that characterizes the diversity of types and motifs. In every motif that exists in this type of batik has its own philosophical meaning where each motif has a long historical value [4]. The types and motifs of batik cannot be separated from the elements inherent in each region where it is made. At that time, batik was worn by Indonesians, especially the nobility with certain motifs that were designated as prohibited for use by the general public. However, in order to preserve culture, batik has been allowed to be worn widely by all groups. Even batik has received recognition by the world which is an award from the United Nations Educational, Scientific and Cultural Organization (UNESCO) on October 2, 2009 by stating that batik is the intellectual cultural right of the Indonesian people [5].

The increasing number of batik motifs among the community necessitates the classification of batik motifs in order to identify them based on their batik names. One way to perform classification is by using Machine Learning methods which is implemented in the Mobile Application. This application aims to demonstrate the classification results using Machine Learning. Machine Learning is a subset of Artificial Intelligence. The goal of machine learning is to develop algorithms and computational models that are used to enable a system to learn autonomously and improve its performance based on the provided data. Additionally, machine learning is used for data analysis, image pattern recognition, and making predictions based on the given data. According to the findings, this method is widely applied in the classification of batik motifs.

Based on previous research, batik classification using the Naïve Bayes method based on texture feature extraction can produce accuracy of up to 97.22% using 420 training data and 180 test data covering three different types of batik [3]. Then, in another study, conducted in 2020, it was found that the classification of batik using a combination of Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), Decision Tree, and Support Vector Machine (SVM) methods achieved an accuracy rate of up to 96.43% in classifying two types of batik motifs, with a dataset consisting of 76 batik images [4]. However, both studies still have shortcomings in terms of separate feature extraction processes, simple models, and the limited number of classified batik motifs. Therefore, a new classification method was developed with the CNN (Convolutional Neural Network) model which is a development of ANN. As in research in 2020, the classification of 6 types of batik from a total dataset of 944 images with the CNN method based on the densenet 201 model was able to produce an accuracy rate of up to 99% [1]. This research also compares the accuracy value with other models such as Alexnet which is able to produce up to 98% accuracy, Resnet152 with 100% accuracy, and Squeezenet 1 which produces an accuracy rate of 99%.

Convolutional Neural Network (CNN) has advantages and disadvantages in the computational process. The advantages of CNN include the ability to select features without modify or develop feature extraction in image data [6]. However, CNN also has a disadvantage, which is the need for data augmentation, requiring data augmentation during the pre-processing stage. Based on the background explanation above, the researcher is interested in conducting a study on "The Application of Convolutional Neural Network Method with MobileNet V1 Architecture in Batik Classification". The aim of this research is to determine the accuracy and loss results of the inputted image data. This research is aimed to fulfill the topic of Multimedia Communication and Applications at the Broadband and Wireless Computing, Communication, and Applications conference, and the hope is that the general public will be able to differentiate or identify batik patterns through the research designed by the researcher.

2 Convolutional Neural Network

Convolutional Neural Network (CNN) is a development of the Multi-Layer Perceptron (MLP) machine learning method designed to process two-dimensional data [7]. As evident from its name, the main process that occurs in CNN is convolution. Convolution is the repeated application of a function to the output of another function. CNN is a popular deep learning method that has been widely applied in previous research and has been proven to achieve good accuracy rates. For example, in a study titled "Wood Species Identification using Convolutional Neural Network with Mobilenet Architecture," a classification system with an accuracy rate of up to 95% was achieved [8]. Currently, various CNN architectures have been developed, such as AlexNet, ResNet, VGG Net, GoogLeNet, NASNet, and MobileNet [9].

2.1 Mobilenet Architecture

MobileNet is one of the CNN architectures specifically developed for mobile computing purposes [10]. The most fundamental difference between the MobileNet architecture and other CNN architectures lies in the use of depthwise separable convolutions, where the thickness of the filters is adjusted according to the thickness of the input image [11]. The operation of the MobileNet architecture is based on the concept of depthwise separable convolutions [12]. Depthwise separable convolutions consist of two main layers: depthwise convolution layers and pointwise convolution layers [13]. The depthwise convolution layer filters the input image without creating new features, while the pointwise convolution layer is responsible for creating new features [14].

3 Methodology

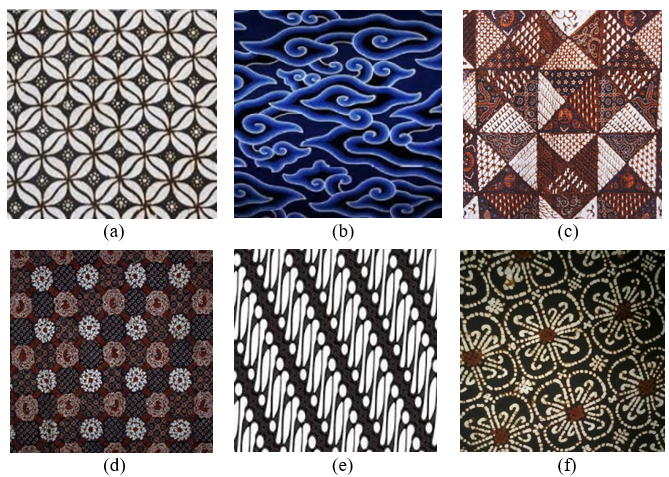
A diagram of a structure

Description automatically generated

The figure above is a flow picture of this research. Starting from the Machine Learning process such as Pre-Processing, Feature Extraction and Classification stages. After the process, the results of Machine Learning are included in the Mobile Application. The pre-processing stage is carried out by performing Rotation, Width, Shear and Brightness using the batik dataset, then from this process the best results will be obtained to proceed to the next stage. Next there is the Feature Extraction stage by performing Weight, Top Layer and Layer Configuration using the best results from the pre-processing process. The last stage of the Machine Learning process is classification. This stage is a continuation of the feature extraction process so that the data used is the data used by the previous stage. At this stage the Optimizer, Learning Rate, Batch, and Epoch are performed. The results that will be included in the Mobile Application are the best results from the entire Machine Learning process. The Mobile Application process is used to produce accuracy equivalent to the Machine Learning process.

3.1 Dataset and Pre-processing

In this study, the dataset used consists of images of Indonesian batik with six different motifs: Ceplok, Parang, Nitik, Kawung, Megamendung, and Tambal. The dataset comprises a total of 600 batik images, with 100 images per motif/class. These 600 batik images are divided into 480 images for training data and 120 images for validation data, with 80 training images and 20 validation images per class. Additionally, for testing purposes at the end of the program, the dataset is supplemented with 10 additional images of the same motif per class, resulting in a total of 60 testing images. All images used in this research dataset were sourced from platforms like Kaggle and through independent searches on photo-sharing platforms such as Pinterest, Freepik, and Shutterstock. The following are examples of the images used in this research dataset. Fig 1 shows the images of batik motif and Table 1 shows the details of each class of batik images dataset used in this research dataset.



**Fig. 1.** The batik motif images with details are as follows: (a) Kawung Motif, (b) Megamendung Motif, (c) Tambal Motif, (d) Ceplok Motif, (e) Parang Motif, (f) Nitik Motif.

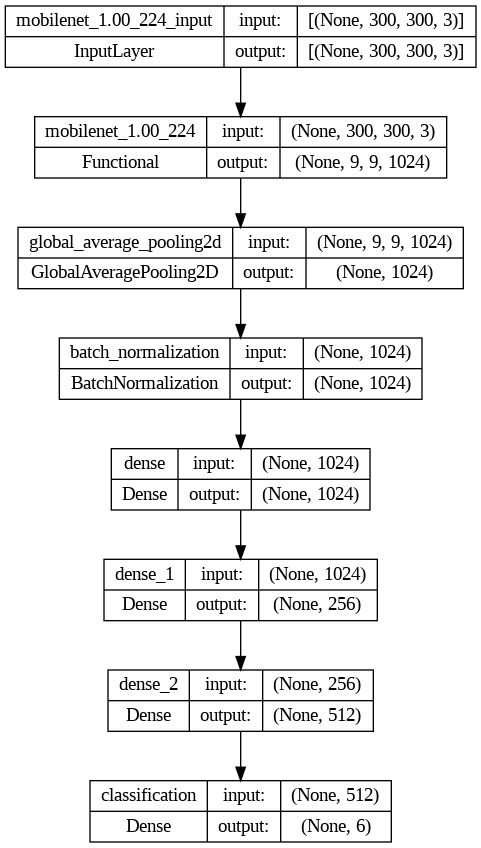
**Table 1.** Font sizes of headings. Table captions should always be positioned *above* the tables.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Motif / Classes of Batik | Data Categories | | | | | Total per Class | |
| Training | Validation | | Testing | | |  | |
| Ceplok | 80 | | 20 | | 10 | 110 | |
| Parang | 80 | | 20 | | 10 | 110 | |
| Nitik | 80 | | 20 | | 10 | 110 | |
| Kawung | 80 | | 20 | | 10 | 110 | |
| Megamendung | 80 | | 20 | | 10 | 110 | |
| Tambal | 80 | | 20 | | 10 | 110 | |
| Total per Categories | 480 | | 120 | | 60 | 660 | |

The data listed in Table 1 then undergoes pre-processing. The pre-processing begins with the use of 'ImageDataGenerator' to perform image augmentation using various methods such as zoom range, width shift range, height shift range, horizontal flip, and vertical flip settings. Next, the batik image data is initialized into three categories: training data, validation data, and testing data.

3.2 MobileNet V1 Architecture Implementation

The MobileNet V1 architecture used in this research serves two purposes: feature extraction from the dataset and conducting batik motif classification. Fig 2 shows an overview of the MobileNet V1 architecture used in this research.



**Fig. 2.** Structure of MobileNet V1 architecture.

From the above picture, it can be seen that the model used has seven layers. The layers used include:

1. One layer of MobileNet is placed at the beginning to invoke the MobileNet architecture. The invoked MobileNet architecture has been previously trained using the ImageNet dataset, to extract features from the input images
2. One Global Average Pooling 2D layer is used to take the average value of the features generated in the form of a 2D matrix
3. One Batch Normalization layer is used to improve the stability of the classification model's performance by normalizing input values, providing regularization, and reducing dependency on weight initialization
4. Three Dense layers with ReLU activation are used to connect the input layer with the output layer
5. One Dense Output layer with Softmax activation is used to generate the output along with the predictions for each class

3.3 Model Testing

In this research, the classification model created will be tested through training, validation, and testing stages to see its performance. Testing will be conducted in several experimental scenarios related to changes in several parameters. The experiments that will be conducted are:

1. The first experiment was conducted by comparing values on the zoom range parameter. The zoom range values to be compared are (0,15), (0,2), and (0,25)
2. The second experiment was conducted by comparing the values of the width shift range parameter. The width shift range values to be compared are (0.15), (0.2), and (0.25)
3. The third experiment was conducted by comparing the type of optimizer used. The optimizer types to be tested are Adam, Adamax, and RMSprop.All experiments will be conducted using epoch value 100 and batch size value 32

4 Experimental Result

In this stage, it contains the experimental results from testing the system designed to classify images of Batik patterns using the MobileNet architecture. The image data of Batik patterns are divided into six classes: Batik Ceplok, Kawung, Megamendung, Nitik, Parang, and the last one is Tambal. The testing conducted involves comparing the image results using different pre-processing techniques and optimizers.

In this testing stage, pre-processing is performed with data augmentation. The zoom range and width range in the designed system are modified with several parameters. The parameter values used for testing are 0.15, 0.2, and 0.25 for the zoom range and width range. This testing will use the MobileNet architecture with an input size 300, batch size of 32, and 100 epochs. Table 2 and Table 3 shows the results of the first test.

**Table 2.** Results of zoom range parameter testing

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Zoom Range | Train Accuracy | Train Loss | Validation Accuracy | Validation Loss | Test Accuracy |
| 0.15 | 92.29% | 0.23 | 68.33% | 1.08 | 64.99% |
| 0.2 | 93.95% | 0.22 | 72.5% | 0.93 | 60% |
| 0.25 | 89.79% | 0.27 | 69.16% | 1.18 | 63.33% |

**Table 3.** Results of width shift range parameter testing

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Zoom Range | Train Accuracy | Train Loss | Validation Accuracy | Validation Loss | Test Accuracy |
| 0.15 | 94.37% | 0.18 | 66.66% | 1.05 | 64.99% |
| 0.2 | 93.54% | 0.19 | 66.66% | 0.99 | 58.33% |
| 0.25 | 91.87% | 0.23 | 69.16% | 0.98 | 64.99% |

Based on table 3, the results of the testing for pre-processing with data augmentation, specifically the zoom range and width range, show that the best test results are obtained with a zoom range value of 0.2 and a width range value of 0.25. The test accuracy is 60% and also 64.99%.

The last testing involves testing the system's optimizer. The parameters used for optimizer testing are Adam, Adamax, and RMSProp. This testing will use the same architecture as the previous testing, with an input size 300, batch size of 32, and 100 epochs. Table 4 shows the results of the last test.

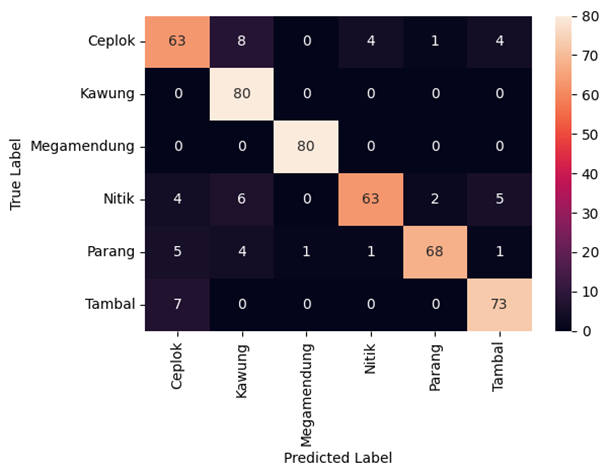
**Table 4.** Results of optimizer testing

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Optimizer | Train Accuracy | Train Loss | Validation Accuracy | Validation Loss | Test Accuracy |
| Adam | 94.37% | 0.18 | 66.66% | 1.05 | 64.99% |
| Adamax | 93.95% | 0.22 | 65.83% | 0.95 | 68.33% |
| RMSProp | 96.04% | 0.14 | 67.5% | 1.96 | 67.5% |

Based on the Table 4, the results of the last optimizer testing show that the best performing optimizer is Adamax. The test accuracy is 68%. As for summary of above test, Table 5 shows the optimized system for batik motif classification and Fig 3 shows the confusion matrix of testing result.

**Table 5.** Results of pre-processing parameter testing on the Adamax optimizer

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Preprocessing | Train Accuracy | Train Loss | Validation Accuracy | Validation Loss | Test Accuracy |
| None | 93.33% | 0.26 | 66.66% | 0.83 | 66.66% |
| Optimize zoom range and width shift range | 93.12% | 0.24 | 67.5% | 0.88 | 63.33% |



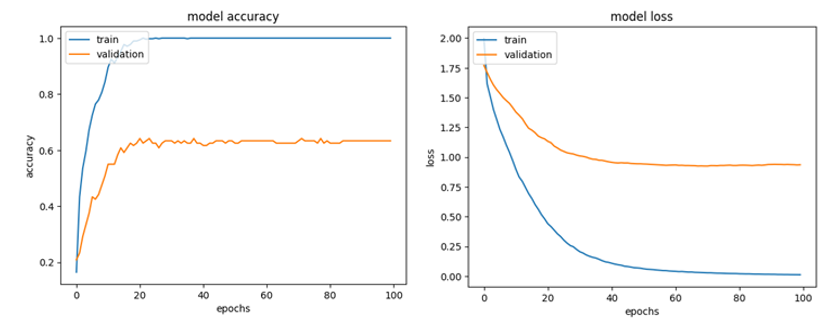
**Fig. 3.** Confusion Matrix of MobileNet V1 classification model.

Based on Table 5, the results of the pre-processing experiment show that the best performance is not using the zoom range and width shift range. And based on Fig 3, the results of the Confusion Matrix from the MobileNet V1 classification model show that there are some misclassifications for four types of batik, namely Batik Ceplok, Batik Nitik, Batik Parang, and Batik Tambal.

**Table 6.** Classification Report of MobileNet V1 classification model

|  |  |  |  |
| --- | --- | --- | --- |
| Batik | Precision | Recall | F1-score |
| Ceplok | 0.80 | 0.79 | 0.79 |
| Kawung | 0.81 | 1.00 | 0.89 |
| Megamendung | 1.00 | 1.00 | 1.00 |
| Nitik | 0.92 | 0.84 | 0.88 |
| Parang | 1.00 | 0.88 | 0.93 |
| Tambal | 0.91 | 0.90 | 0.91 |

In Table 6, the Classification Report of the MobileNet V1 classification model is displayed. This table shows the results of Precision, Recall, and F1-Score for each type of batik. The best precision values are achieved for the batik types "megamendung" and "parang." The best recall values are obtained for the batik types "kawung" and "megamendung." Additionally, the highest F1-Score is achieved for the "megamendung" batik type. The results in this table indicate that the "megamendung" batik type obtains the best results in terms of precision, recall, and F1-score.



**Fig. 4.** MobileNet V1 model accuracy graph (a) and MobileNet V1 model loss graph (b).

Based on Figure 5, the graph shows the results of the MobileNet V1 model for accuracy and loss. From the displayed graph, it can be observed that the training process achieves excellent performance, with accuracy approaching 100% and loss approaching 0. However, when it comes to validation, the performance is not as good and indicates overfitting. Based on previous research, the overfitting experienced in the current study could be attributed to the lack of training dataset [8].

5 Conclusion

Batik is a traditional craft from Indonesia that has been recognized worldwide for its diverse types and motifs. This cultural heritage needs to be preserved. With numerous types and motifs of batik spread across different regions, each has its own historical value, distinctive characteristics, and certain rules to be observed. This calls for innovation in developing a system that can classify batik motifs using Machine Learning. Based on testing the Convolutional Neural Network method with the MobileNet V1 architecture for batik motif classification, the best results were obtained using the Adamax optimizer. It was found that the optimizer Adamax without using the pre-processing methods of zoom range and width shift range achieved higher accuracy. The testing yielded a train accuracy of 93.33% with a loss of 0.26, a validation accuracy of 66.66% with a loss of 0.83, and a test accuracy of 66.66%. The obtained Confusion Matrix shows errors in classifying some types of batik and Batik Megamendung has best score on Precision, Recall and F1-score. As for the graph results of MobileNet V1, it indicates that the training process achieves good performance, but the validation is still inadequate and shows signs of overfitting. The presented results indicate that this research can run well and optimally, ensuring that the system operates according to the designed framework and is aligned with the related topic, which is Multimedia Communication and Applications.

References

[1] B. Sartana, E. W. Saptandari, and A. F. Helmi, *Youth Imagination About Being an Indonesian*. Atlantis Press International BV, (2023).

[2] I. Conference, O. N. Art, and F. O. R. Technology, “Proceeding Conference 2,” (2020).

[3] M. A. Rasyidi and T. Bariyah, “Batik pattern recognition using convolutional neural network,” *Bull. Electr. Eng. Informatics*, vol. 9, no. 4 (2020) 1430–1437, doi: 10.11591/eei.v9i4.2385.

[4] I. Qiram, B. -, and G. Rubiono, “Batik Banyuwangi: Aesthetic and Technical Comparison of Coastal Batik,” *Lekesan Interdiscip. J. Asia Pacific Arts*, vol. 1, no. 2 (2018) 79, doi: 10.31091/lekesan.v1i2.407.

[5] N. Dewi Girsang and Muhathir, “Classification Of Batik Images Using Multilayer Perceptron With Histogram Of Oriented Gradient Feature Extraction,” *Proceeding Int. Conf. Sci. Eng.*, vol. 4 (2021) 197–204.

[6] D. Bhatt *et al.*, “Cnn variants for computer vision: History, architecture, application, challenges and future scope,” *Electron.*, vol. 10, no. 20 (2021) 1–28, doi: 10.3390/electronics10202470.

[7] T. Purwaningsih, I. A. Anjani, and P. B. Utami, “Convolutional Neural Networks Implementation for Chili Classification,” *Proceeding - 2018 Int. Symp. Adv. Intell. Informatics Revolutionize Intell. Informatics Spectr. Humanit. SAIN 2018* (2018) 190–194, 2019, doi: 10.1109/SAIN.2018.8673373.

[8] H. Hendriyana and Yazid Hilman Maulana, “Identification of Types of Wood using Convolutional Neural Network with Mobilenet Architecture,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 1 (2020) 70–76, doi: 10.29207/resti.v4i1.1445.

[9] W. Sae-Lim, W. Wettayaprasit, and P. Aiyarak, “Convolutional Neural Networks Using MobileNet for Skin Lesion Classification,” *JCSSE 2019 - 16th Int. Jt. Conf. Comput. Sci. Softw. Eng. Knowl. Evol. Towar. Singul. Man-Machine Intell* (2019) 242–247, doi: 10.1109/JCSSE.2019.8864155.

[10] S. Mohapatra, N. V. S. Abhishek, D. Bardhan, A. A. Ghosh, and S. Mohanty, “Comparison of MobileNet and ResNet CNN Architectures in the CNN-Based Skin Cancer Classifier Model,” *Mach. Learn. Healthc. Appl.*, (2021) 169–186, doi: 10.1002/9781119792611.ch11.

[11] F. H. Tseng, K. H. Yeh, F. Y. Kao, and C. Y. Chen, “MiniNet: Dense squeeze with depthwise separable convolutions for image classification in resource-constrained autonomous systems,” *ISA Trans.*, vol. 132 (2023) 120–130, doi: 10.1016/j.isatra.2022.07.030.

[12] E. Prasetyo, R. Purbaningtyas, R. D. Adityo, N. Suciati, and C. Fatichah, “Combining MobileNetV1 and Depthwise Separable convolution bottleneck with Expansion for classifying the freshness of fish eyes,” *Inf. Process. Agric.*, vol. 9, no. 4 (2022) 485–496, doi: 10.1016/j.inpa.2022.01.002.

[13] N. R. Koluguri, T. Park, and B. Ginsburg, “Titanet: Neural Model for Speaker Representation With 1D Depth-Wise Separable Convolutions and Global Context,” *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2022-May (2022) 8102–8106, doi: 10.1109/ICASSP43922.2022.9746806.

[14] W. Wang, Y. Hu, T. Zou, H. Liu, J. Wang, and X. Wang, “A New Image Classification Approach via Improved MobileNet Models with Local Receptive Field Expansion in Shallow Layers,” *Comput. Intell. Neurosci.*, (2020) doi: 10.1155/2020/8817849.