



Research Article

A Comparative Study on Efficacy of CNN VGG-16, DenseNet121, ResNet50V2, And EfficientNetB0 in Toraja Carving Classification

Herman^{1,*}, An'nisa Pratama Putri^{2,*}, Megat Norulazmi Megat Mohamed Noor³, Herdianti Darwis⁴, Lilis Nur Hayati⁵, Irawati⁶, Ihwana As'ad⁷

¹ Universitas Muslim Indonesia, Makassar, Sulawesi Selatan 90234, Indonesia, herman@umi.ac.id

² Universitas Muslim Indonesia, Makassar, Sulawesi Selatan 90234, Indonesia, annisa.iclabs@umi.ac.id

³ MIT University Kuala Lumpur, 50250 Kuala Lumpur, Malaysia, megatnorulazmi@unikl.edu.my

⁴ Universitas Muslim Indonesia, Makassar, Sulawesi Selatan 90234, Indonesia, herdianti.darwis@umi.ac.id

⁵ Universitas Muslim Indonesia, Makassar, Sulawesi Selatan 90234, Indonesia, lilis.nurhayati@umi.ac.id

⁶ Universitas Muslim Indonesia, Makassar, Sulawesi Selatan 90234, Indonesia, irawati@umi.ac.id

⁷ Universitas Muslim Indonesia, Makassar, Sulawesi Selatan 90234, Indonesia, ihwana.asad@umi.ac.id

Correspondence should be addressed to An'nisa Pratama Putri; annisa.iclabs@umi.ac.id

Received 22 December 2024; Accepted 28 March 2025; Published 31 March 2025

© Authors 2025. CC BY-NC 4.0 (non-commercial use with attribution, indicate changes).

License: <https://creativecommons.org/licenses/by-nc/4.0/> — Published by Indonesian Journal of Data and Science.

Abstract:

Introduction: Passura', or Toraja carvings, are an essential element of the cultural heritage of the Toraja people in Indonesia. These carvings feature complex motifs rooted in nature, folklore, and spiritual symbolism. This study aims to evaluate the efficacy of four Convolutional Neural Network (CNN) architectures—VGG-16, DenseNet121, ResNet50V2, and EfficientNetB0—in classifying seven traditional Toraja carving motifs. **Methods:** A dataset of 700 images was collected and categorized into seven motif classes. The dataset was split into 80% for training and 20% for validation. Each CNN model was trained for 25 epochs with standard pre-processing, including resizing to 224×224 and normalization. Performance evaluation was conducted based on validation accuracy and confusion matrix analysis to assess classification precision and model overfitting. **Results:** EfficientNetB0 achieved the highest validation accuracy of 98%, although signs of overfitting were observed. ResNet50V2 followed closely with a validation accuracy of 95.33% and demonstrated the most balanced classification results across all motif categories. VGG-16 and DenseNet121 achieved 94.67% and 81.82%, respectively. Confusion matrix analysis confirmed the robustness of ResNet50V2 in correctly identifying complex patterns. **Conclusions:** The findings indicate that ResNet50V2 provides a reliable balance between accuracy and generalizability for classifying Toraja carvings, making it suitable for digital preservation of cultural heritage. EfficientNetB0, while achieving higher accuracy, may require additional regularization to avoid overfitting. This study contributes to the development of AI-driven cultural documentation and suggests future research with larger and more diverse datasets to improve model robustness.

Keywords: Classification, CNN, DenseNet121, EfficientNetB0, Resnet50V2, Toraja Carving, VGG-16.

1. Introduction

Toraja carvings, known as passura', are a rich part of Toraja's cultural heritage, rich in planar geometry, patterns, shapes, and various motifs. Each motif in the sculpture has not only aesthetic value, but also a deep meaning that reflects the Toraja people's hope for an abundant life [1], [2]. The motifs of Toraja carvings are inspired by natural elements such as animals, plants, and folklore, depicting the harmonious relationship between humans and their gods, fellow humans, and nature [3]. Toraja carvings were originally created on wood to decorate traditional Tongkonan houses and ceremonial tools, but their use has expanded to other mediums, including embroidered fabrics and batiks worn by Torajans [4]. The nearly 70 passura' designs are frequently used in a variety of ceremonies and craft activities [3].

Previous studies have shown that most toraja carvers do not fully grasp the meaning of the motifs they create. Furthermore, although the basic motifs have been preserved, toraja wood carvings have been modified for

aesthetic purposes [1]. These circumstances make it even more difficult to understand the original meaning of toraja carving.

Convolutional Neural Networks (CNNs) are an excellent technique for image classification and pattern recognition [5]–[7]. The core of a CNN, the multilayer perceptron (MLP), is designed to process two-dimensional input data [8], [9]. Various CNN architectures, such as EfficientNet, MobileNet, AlexNet, VGG, ResNet and NasNet, have been shown to be suitable for image classification tasks [10]. CNNs are well suited for image data classification and are particularly effective for the classification of complex and diverse Toraja sculpture patterns [11].

Some previous studies have attempted to use CNN models to classify Toraja carvings. Herman's experiment in this research aims to classify the motifs of Toraja carvings using CNNs, specifically the LeNet-5 architecture. The dataset used consists of 700 images divided into seven motif categories, each consisting of 100 images. This study examined the effect of data separation ratio (60:40, 70:30, 80:20 and 90:10), batch size and target size on classification accuracy. The results show that a data separation ratio of 70:30 leads to a maximum accuracy of 50.67%, while increasing the batch size value has a negative impact on accuracy. However, increasing the target size has a positive effect, with a maximum accuracy of 69.33% at 256×256 . This study shows that the choice of parameters in the LeNet-5 architecture has a significant impact on the classification of truncated torus shapes [11].

Previous research on image classification with CNNs has proven highly effective in analyzing complex patterns, including cultural motifs such as Torajan woodcuts. Wardhani's research shows that machine-learning approaches, particularly CNNs, are essential for cultural preservation. Using the Canny extraction method, the CNN model uses the ResNet50 architecture to identify complex Torajan woodcut patterns. Some classes of Toraja motifs used in the study include Pa' Lulun Pao, Pa' Somba, Pa' Tangke Lumu and Pa' Tumuru. The results show that the Pa' Somba class has an accuracy of 0.9697, a recall of 0.96 and an F1 Score of 0.9648; the Pa' Lulun Pao class has an accuracy and recall of 0.94 and an F1 Score of 0.94; the Pa' Tangke Lumu has an accuracy of 0.9898, a recall of 0.97 and an F1-Score of 0.9798; and the Pa' Tumuru motif has an accuracy of 0.9327, a recall of 0.97 and an F1-Score of 0.9500 [12].

The purpose of this research is to develop and test a better classification method to understand the complex patterns in Torajan carving motifs. By using CNN architectures such as VGG-16, EfisieNetB0, ResNet50V2 and DenseNet121, the classification results are expected to be more accurate and reliable. This study will contribute to the preservation and better understanding of Torajan cultural heritage. This is an important step to protect the rich Torah cultural heritage in the context of modern technological discoveries and to improve digital documentation for future cultural research.

2. Method:

Image dataset collection, data pre-processing methods, classification model building, evaluation, and performance analysis are the research flow in this study, as shown in [Figure 1](#).

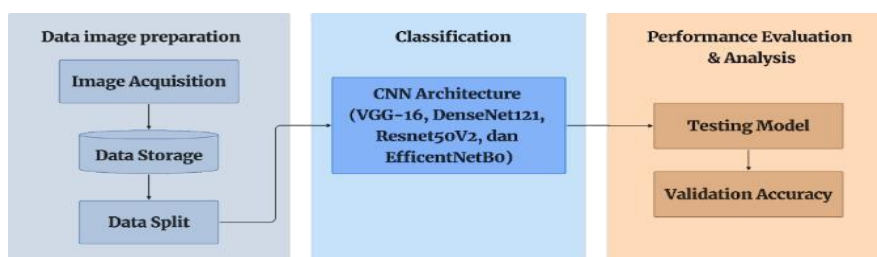


Figure 1. Research Flow

Data Image Preparation

The collection of images for the Passura' dataset was carried out in the field, photographing different types of statues found in traditional houses and cultural monuments in Toraja. Each image was carefully taken to ensure quality resolution so that the details of the subject could be clearly seen in the model created. Once the capture process is complete, all images are saved to the computer and organized for the next step. The amount of image data per category or type of Torajan carvings model is shown in [Table 1](#).

Table 1. Total Data

No	Toraja Carving Class	Amount
1	<i>Paqkapuq Baka</i>	100
2	<i>Paqtanduk Reqpe</i>	100
3	<i>Paqsekong Kandaure</i>	100
4	<i>Patangke Lumuq</i>	100
5	<i>Paqdon Bolu</i>	100
6	<i>Paqtoloqpaku</i>	100
7	<i>Paqtedong</i>	100
	Total	700

Figure 2 shows different Toraja carvings. The photographs were taken and stored on a computer. The data was then divided into two main categories: training data and validation data. These two categories are used to train the model and to evaluate the accuracy of the model and are divided into 80% training data and 20% validation data. Just as Herman has done in his research which uses an 80:20 data split [4].

**Figure 2.** Example of Toraja Carvings

Classification

A special type of neural network (CNN) is created to analyze the input using Patterns, as shown in the image. The convolution technique allows CNN to extract features by applying filters to the input image. A pooling layer and multiple convolution layers are the basis of the CNN structure [13], [14]. Convolution layers are usually used to see how filters or kernels scan an image to capture certain patterns or features, such as edges and textures. However, pooling layers reduce the spatial size of the information without losing important data. Ultimately, pooling layers help in faster training because they reduce the size of the feature maps after the convolution layers extract them. Finally, a fully connected layer calculates the final prediction using the extracted features [15], [16]. One of the main reasons CNNs are effective in image processing is that they can learn multi-level features, from simple to complex, which makes it easy for CNNs to solve classification and object detection tasks [17].

For image processing, VGG-16 is one of the most popular CNN architectures. VGG-16, which was originally proposed by the Visual Geometry Group at the University of Oxford, consists of 16 invariant layers, with 13 convolutional layers and 3 fully connected layers. The convolution layer uses a 3x3 convolution with padding to preserve the size of the input space. In addition, some networks gradually reduce the size of locally detected maps by using massive Asynchronous training and convergence is accelerated by the ReLU technique. Finally, the full embedding layer combines the learned features to make group decisions. The architecture of VGG-16 as shown in **Figure 3** shows that the network can learn complex features and thus achieve high accuracy for many images classification tasks [18].

DenseNet121 is a CNN architecture consisting of 121 layers as shown in **Figure 4** and is designed to increase the efficiency and accuracy of image recognition by means of dense connections, where each layer is directly connected to all previous layers, enabling better information flow and deeper feature learning. It also reduces the problem of vanishing gradients and the number of required parameters, which increases computational efficiency. The architecture consists of many "dense blocks" with interconnected convolutional layers, as well as transition blocks that reduce the feature dimension

to balance the complexity and accuracy of the model. DenseNet121 has been proven effective in a wide range of applications, including image classification, image segmentation, and object recognition, and is commonly used in medical image analysis, such as tumor detection in radiological images. Its high performance on benchmark datasets such as ImageNet makes it an ideal choice for research and deep learning applications, providing a powerful solution for pattern recognition and image analysis problems [19]–[21].

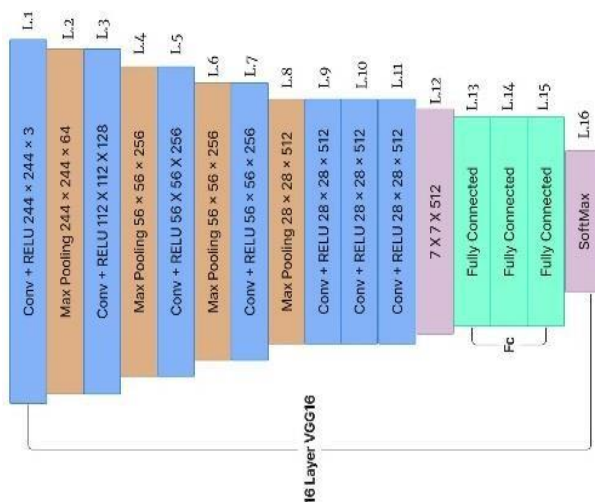


Figure 3. VGG-16 Architecture

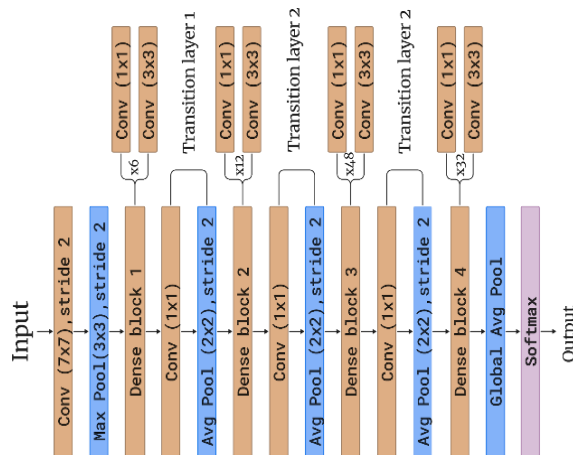


Figure 4. DenseNet121 Architecture

One of the most often used architectures in deep learning is ResNet50V2. The ResNet50 architecture, which is well-known for its capacity to resolve low-dimensional computational issues, is extended by this architecture. Architecture of ResNet50V2 can be seen in Figure 5. Data may readily flow through deep neural networks because to this architecture's usage of residual learning techniques. This architecture makes it easier to train the model by combining 50 layers and connecting non-adjacent layers via shortcuts. Additionally, ResNet50V2 employs extremely effective grouping and communication techniques, which enhance stability and performance throughout training. ResNet50V2 is the suggested option for applications like image classification and data processing because of its great accuracy in a variety of image-related tasks. This architecture is used to analyze medical images and is one of the most widely used tools by researchers and companies [22], [23].

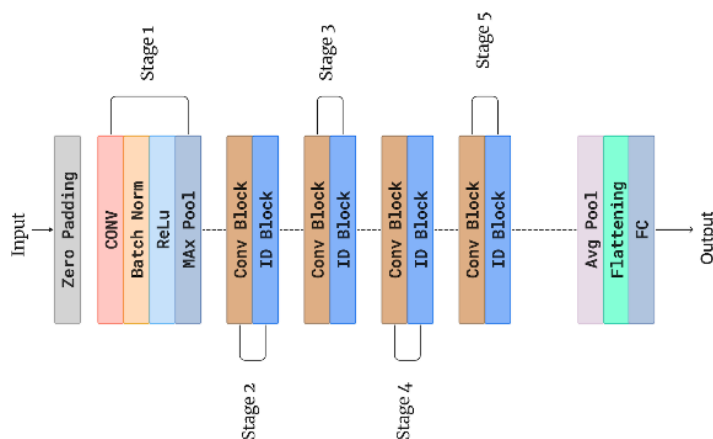


Figure 5. ResNet50V2 Architecture

EfficientNetB0 is a convolutional neural network architecture designed to achieve an optimal balance between accuracy and computational efficiency in image recognition. Architecture of EfficientNetB0 can be seen in Figure 6. EfficientNetB0 launched by Google, is the first model in the EfficientNet family to use customized scaling where network width, depth, and resolution are simultaneously optimized to improve performance. Using techniques such as hybrid scaling and efficient use of convolutional blocks, EfficientNetB0 can achieve higher accuracy with a much smaller number of parameters compared to previous models. This format also combines techniques such as swish activation and group training, which add to training stability and speed. EfficientNetB0 has been proven to be highly efficient for various image recognition tasks and is widely used in industry and research, making it one of the most popular models in deep learning today.

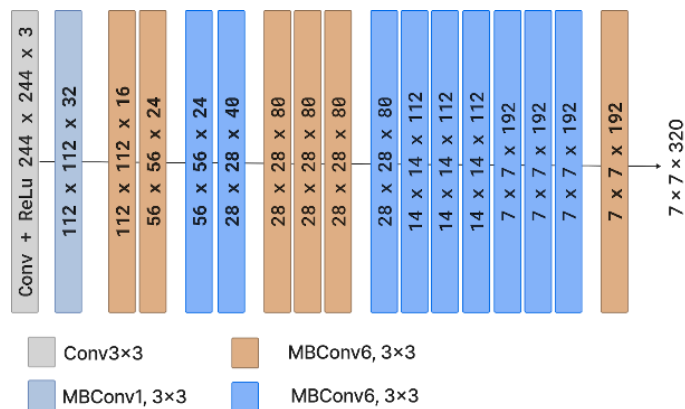


Figure 6. EfficientNetB0 Architecture

Performance Evaluation and Analysis

At this stage, the performance of the models in classifying Toraja carving motifs was measured using VGG-16, DenseNet121, ResNet50V2, and EfficientNetB0 architectures. The training process is performed by splitting the dataset into two parts, namely 80% as training data and 20% as validation data. This division aims to ensure that the model can learn from the most available data while retaining enough data to test its ability to generalize to data it has not seen before. Once the training process is complete, the validation accuracy of each tested CNN architecture is calculated and compared. The accuracy results are analyzed to determine which model provides the best performance in classification. Comparing the validation accuracy of different architectures allows us to identify the most effective and efficient model in completing the task of classifying Toraja carving motifs.

Validation accuracy is calculated by comparing the correct predictions made by the model with the total number of predictions. The formula for calculating accuracy is given in (1).

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

Where FP stands for False Positive, FN for False Negative, TP for True Positive, and TN is for True Negative. This formula provides explicit statistics about model performance by calculating the percentage of images correctly captured among all images. Diagnostic accuracy is determined by comparing the correct predictions made by the model with the total number of predictions [25], [26] as in (1).

In this section, only the model with the highest prediction accuracy among the four CNN architectures will also be tested using the confusion matrix. The confusion matrix provides detailed information about the classification performance of each class, allowing for errors in the prediction model, such as motifs being incorrectly classified in the wrong class. By analyzing the confusion matrix, it is possible to critically evaluate how each structure performs in the classification of Toraja sculpture motifs. This competition aims to identify the most suitable and useful designs for the task of classifying Toraja design motifs.

3. Results and Discussion

Each architecture is trained for 25 epochs to ensure that the model has sufficient learning opportunities and that no overfitting occurs, and each architecture uses two equal proportions of data, 80% training data and 20% validation data.

In the VGG-16 architecture, each dataset is processed into 32 batch sizes and the target image size used is (224, 224), which is the standard size of the VGG architecture. In addition, after the source layer, each structure has the same additional layers. A flat layer, a dense layer with 512 units and a ReLU activation function, a rolling layer with a value of 0.5 to avoid overfitting, a dense layer of 256 units with a ReLU activation function and seven classes to distinguish Toraja sculpture patterns, an output layer with 7 units and a Softmax activation function. The model was assembled using the Adam optimizer with a learning rate of 0.0001 after the training process in the VGG-16 model structure achieved the highest evaluation accuracy results of 94.67%. A graph of the accuracy and evaluation results of the VGG-16 model and the structure can be seen in the [Figure 7](#).

In the DenseNet121 architecture, the training parameters of the pair are set to 32 pixels, and the height of the pair image is 224 pixels. Procedural data were acquired by ImageDataGenerator with the parameter rescale=1/255 for normalization without enlarging the interval [0, 1]. The GlobalAveragePooling2D layer is added to reduce the dimensionality of the output, and the dense layer trains the classifier through softmax activation, which increases the number of neurons and the number of classes in the dataset. The model was tuned using the RMSprop optimizer with a learning rate of 0.0001. After the model is trained, an evaluation is performed to calculate the accuracy loss value of the estimation data. The post-training validation results of the DenseNet 121 model architecture achieved a high validation accuracy of 81.82%. A plot of the results of DenseNet 121 model architecture accuracy evaluation is shown in [Figure 8](#).

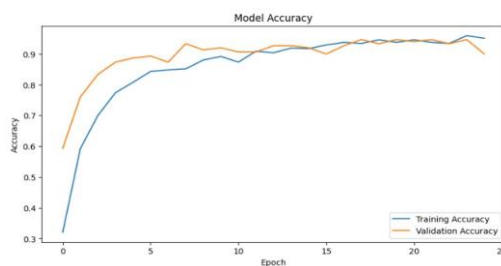


Figure 7. Accuracy of VGG-16

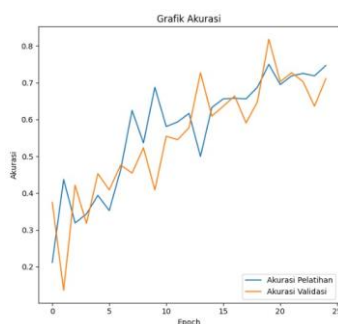


Figure 8. Accuracy of DenseNet121

The training parameters of the EfficientNetB0 architecture include image size set to (224, 224) to ensure consistent input, and batch size set to 32 to batch process the data. In addition, the model also sets the number of classes to 7, depending on the number of classes in the dataset. EfficientNetB0 is loaded without overlapping using pre-trained weights on the ImageNet dataset, followed by a GlobalAveragePooling2D layer, and a Dense layer with softmax activation for classification. The model is then compiled using the Adam Optimizer, which automatically sets the initial learning rate. After the training process, the EfficientNetB0 model shows good performance, with an average validation accuracy of 98%. The graph of the accuracy and validation accuracy results for the EfficientNetB0 model architecture can be seen in the [Figure 10](#).

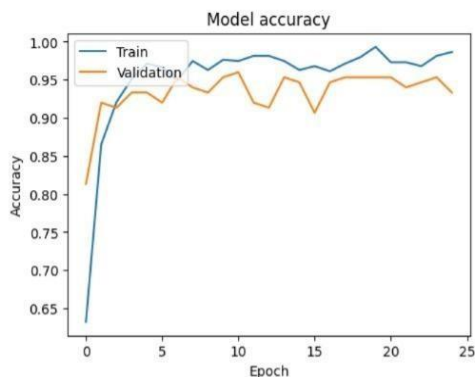


Figure 9. Accuracy of ResNet50V2

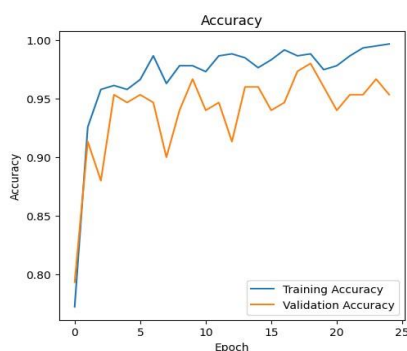


Figure 10. Accuracy of EfficientNetB0

The EfficientNetB0 architecture gets the greatest accuracy of 98%, which is really good, as seen in [Table 2](#). These high findings, nevertheless, might point to overfitting in EfficientNetB0. This implies that there is a chance the model won't be able to maintain the same level of accuracy or might incorrectly classify the toraja carving when a larger or new dataset is included. The results derived from the model's confusion matrix make this very evident depicted in the [Figure 11](#). The ResNet50V2 architecture, the top model in this test with a remarkable validation accuracy of 95.33%, comes next. The data can be efficiently categorized by this model. This illustrates its capacity to classify pertinent patterns and characteristics within the toraja carving.

Table 2. Validation Of The Accuracy Of Each Architecture Model

No	Architecture	Val_Accuracy
1	VGG-16	94.67%
2	DenseNet121	81.82 %
3	ResNet50V2	95.33%
4	EfficientNetB0	98%

Based on the confusion matrix shown in the [Figure 12](#), the ResNet50V2 model has excellent performance in classifying the features of the Toraja carving. Each row of the confusion matrix represents the proportion of correct entries, and the columns represent the prediction model. On the main diagonal, it can be seen that the model is able to accurately classify most of the motifs. For example, the pattern *Paqdon Bolu* had 25 correct guesses and

Paqkapuq Baka had 19 correct guesses. In addition, *Paqsekong Kandaure* and *Paqtanduk Reqpe* had 21 correct predictions, indicating that the model was effective in identifying toraja carving.

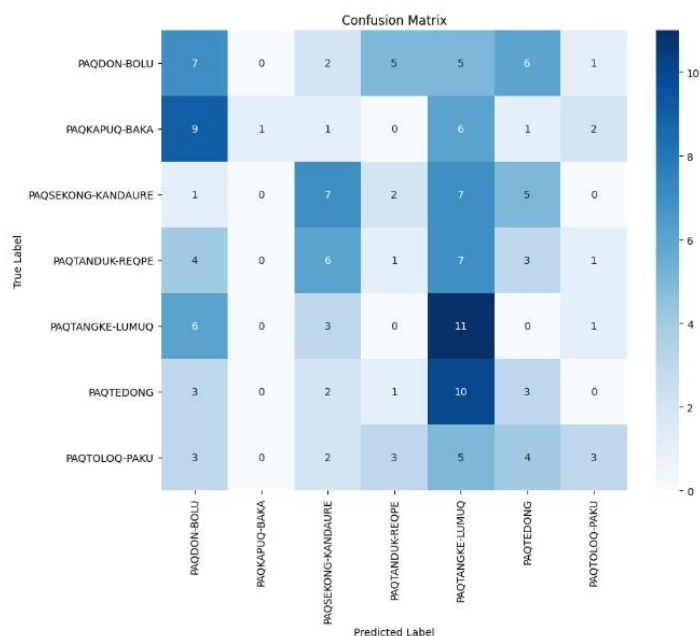


Figure 11. Confusion Matrix Results from EfficientNetB0 Model

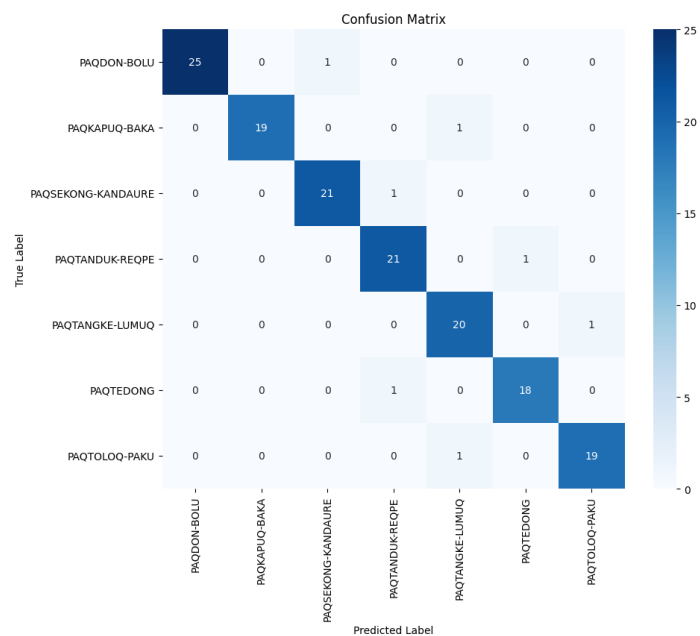


Figure 12. Confusion Matrix Results from ResNet50V2 Model

However, there are some significant classification errors. For example, the *Paqdon Bolu* standard was misclassified as *Paqsekong Kandaure* and the *Paqtanduk Reqpe* standard was misclassified as *Paqtedong*, this error is very small. However, it shows some difficulties in recognizing more accurate patterns. Overall, the ResNet50V2 model shows high performance in classification. The predictions were mostly accurate, but further improvements may be needed to reduce the prediction errors of some models. Allows the model to recognize all fragment classes more accurately.

4. Conclusion

This research measures the performance of VGG-16, DenseNet121, ResNet50V2, and EfficientNetB0 architectures in classifying the motifs of Toraja sculptures trained and validated using an 80:20 split of training

and validation data. DenseNet121 architecture produces the highest classification accuracy of 98%, but the model overfits. This is proven based on the results obtained from the confusion matrix of the EfficientNetB0 model. ResNet50V2 architecture is the most effective for this task as it provides high accuracy and can perform the classification task efficiently and accurately. That is, it achieves a validation accuracy of 95.33%. This study contributes to the preservation and understanding of Toraja cultural heritage through the accuracy of motif classification. This is an important step towards protecting the rich cultural heritage of Toraja against the advancement of modern technology and improving digital documentation for future cultural studies. Future studies should add more datasets to improve the model performance and prevent overfitting in classifying Toraja sculptures.

Acknowledgments

Department of Computer Science and Engineering, Universitas's Muslim Indonesia provided significant support in providing the necessary facilities for this study. The laboratory offers secure and reliable data storage, as well as equipment for testing and data analysis. This support is essential to ensure that the research process runs smoothly and that the data obtained is processed efficiently. With the appropriate facilities from ICLabs, this research can be conducted.

References:

- [1] R. Muslimin, "Toraja glyphs: An ethnocomputation study of passura Indigenous icons," *J. Asian Archit. Build. Eng.*, vol. 16, no. 1, pp. 39–44, 2017, doi: [10.3130/jaabe.16.39](https://doi.org/10.3130/jaabe.16.39).
- [2] Y. S. Nugraha, "Ethnomathematical review of Toraja's typical carving design in geometry transformation learning," 2019. doi: [10.1088/1742-6596/1280/4/042020](https://doi.org/10.1088/1742-6596/1280/4/042020).
- [3] A. Donzelli, "Material words: The aesthetic grammar of Toraja textiles, carvings, and ritual language," *J. Mater. Cult.*, 2020, doi: [10.1177/1359183519858378](https://doi.org/10.1177/1359183519858378).
- [4] L. Sudioanto, "The Encyclopedia of Virtual Art Carving Toraja - Indonesia," 2018. doi: [10.1051/mateconf/201816401025](https://doi.org/10.1051/mateconf/201816401025).
- [5] L. Alzubaidi, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, 2021, doi: [10.1186/s40537-021-00444-8](https://doi.org/10.1186/s40537-021-00444-8).
- [6] S. Minaee, "Image Segmentation Using Deep Learning: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, 2022, doi: [10.1109/TPAMI.2021.3059968](https://doi.org/10.1109/TPAMI.2021.3059968).
- [7] A. Narin, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Anal. Appl.*, vol. 24, no. 3, pp. 1207–1220, 2021, doi: [10.1007/s10044-021-00984-y](https://doi.org/10.1007/s10044-021-00984-y).
- [8] M. Desai, "An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN)," 2021. doi: [10.1016/j.jceh.2020.11.002](https://doi.org/10.1016/j.jceh.2020.11.002).
- [9] J. Liu, "Hourly stepwise forecasting for solar irradiance using integrated hybrid models CNN-LSTM-MLP combined with error correction and VMD," *Energy Convers. Manag.*, vol. 280, 2023, doi: [10.1016/j.enconman.2023.116804](https://doi.org/10.1016/j.enconman.2023.116804).
- [10] Y. Gulzar, "Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique," *Sustain.*, vol. 15, no. 3, 2023, doi: [10.3390/su15031906](https://doi.org/10.3390/su15031906).
- [11] Herman, H. Nasir, M. N. M. M. Noor, T. Hasanuddin, D. Indra, and H. B. Lumentut, "Exploration of CNN Parameters to Measure Performance of LeNet-5 Architecture in Toraja Carving Classification," in *2024 IEEE 8th International Conference on Signal and Image Processing Applications (ICSIPA)*, 2024, pp. 1–6. doi: [10.1109/ICSIPA62061.2024.10686353](https://doi.org/10.1109/ICSIPA62061.2024.10686353).
- [12] N. Wardhani, B. E. W. Asrul, A. R. Tampang, and ..., "Classification of Toraja Wood Carving Motif Images Using Convolutional Neural Network (CNN)," ... (*Rekayasa Sist. dan ...*, 2024, doi: [10.29207/resti.v8i4.5897](https://doi.org/10.29207/resti.v8i4.5897).

- [13] M. Caldeira, "Comparison study on convolution neural networks (CNNs) vs. Human visual system (HVS)," 2019. doi: [10.1007/978-3030-19093-4_9](https://doi.org/10.1007/978-3030-19093-4_9).
- [14] A. Celeghin, "Convolutional neural networks for vision neuroscience: significance, developments, and outstanding issues," 2023. doi: [10.3389/fncom.2023.1153572](https://doi.org/10.3389/fncom.2023.1153572).
- [15] R. Nirthika, "Pooling in convolutional neural networks for medical image analysis: a survey and an empirical study," 2022. doi: [10.1007/s00521-022-06953-8](https://doi.org/10.1007/s00521-022-06953-8).
- [16] H. Darwis, "Max Feature Map CNN with Support Vector Guided SoftMax for Face Recognition," *Int. J. Informatics Vis.*, vol. 7, no. 3, pp. 959–966, 2023, doi: [10.30630/joiv.7.3.1751](https://doi.org/10.30630/joiv.7.3.1751).
- [17] T. Turay, "Toward Performing Image Classification and Object Detection with Convolutional Neural Networks in Autonomous Driving Systems: A Survey," *IEEE Access*, vol. 10, pp. 14076–14119, 2022, doi: [10.1109/ACCESS.2022.3147495](https://doi.org/10.1109/ACCESS.2022.3147495).
- [18] G. Georgiadis, "Accelerating convolutional neural networks via activation map compression," 2019. doi: [10.1109/CVPR.2019.00725](https://doi.org/10.1109/CVPR.2019.00725).
- [19] G. Huang, "Convolutional Networks with Dense Connectivity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 12, pp. 8704–8716, 2022, doi: [10.1109/TPAMI.2019.2918284](https://doi.org/10.1109/TPAMI.2019.2918284).
- [20] S. A. Albelwi, "Deep Architecture based on DenseNet-121 Model for Weather Image Recognition," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 10, pp. 559–565, 2022, doi: [10.14569/IJACSA.2022.0131065](https://doi.org/10.14569/IJACSA.2022.0131065).
- [21] N. Cinar, "A hybrid DenseNet121-UNet model for brain tumor segmentation from MR Images," *Biomed. Signal Process. Control*, vol. 76, 2022, doi: [10.1016/j.bspc.2022.103647](https://doi.org/10.1016/j.bspc.2022.103647).
- [22] M. Rahimzadeh, "A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2," *Informatics Med. Unlocked*, vol. 19, 2020, doi: [10.1016/j.imu.2020.100360](https://doi.org/10.1016/j.imu.2020.100360).
- [23] A. K. Das, "Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network," *Pattern Anal. Appl.*, vol. 24, no. 3, pp. 1111–1124, 2021, doi: [10.1007/s10044-021-00970-4](https://doi.org/10.1007/s10044-021-00970-4).
- [24] A. Xiao, "Intraoperative Glioma Grading Using Neural Architecture Search and Multi-Modal Imaging," *IEEE Trans. Med. Imaging*, vol. 41, no. 10, pp. 2570–2581, 2022, doi: [10.1109/TMI.2022.3166129](https://doi.org/10.1109/TMI.2022.3166129).
- [25] M. A. Joodi, "Increasing validation accuracy of a face mask detection by new deep learning model-based classification," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 29, no. 1, pp. 304–314, 2023, doi: [10.11591/ijeecs.v29.i1.pp304-314](https://doi.org/10.11591/ijeecs.v29.i1.pp304-314).
- [26] Y. I. Sulistya, E. T. Br Bangun, and D. A. Tyas, "CNN Ensemble Learning Method for Transfer learning: A Review," *Ilk. J. Ilm.*, vol. 15, no. 1, pp. 45–63, 2023, doi: [10.33096/ilkom.v15i1.1541.45-63](https://doi.org/10.33096/ilkom.v15i1.1541.45-63)