Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online at: **http://jurnal.iaii.or.id**

JURNAL RESTI

(Rekayasa Sistem dan Teknologi Informasi)

Vol. 8 No. 4 (2024) 486 - 495 e-ISSN: 2580-0760

Classification of Toraja Wood Carving Motif Images Using Convolutional Neural Network (CNN)

Nurilmiyanti Wardhani[1], Billy Eden William Asrul[2]*, Antonius Riman Tampang[3], Sitti Zuhriyah[4], Abdul Latief Arda[5]

[1,2,3]Faculty of Computer Science, Department of Informatic Engineering, Universitas Handayani Makassar, Indonesia
[4,5]Faculty of Computer Science, Department of Computer Systems, Universitas Handayani Makassar, Indonesia
[1]nurilmi@handayani.ac.id, [2]billy@handayani.ac.id, [3]rimantampang141@gmail.com, [4]zuhriyah@handayani.ac.id, [5]abdullatief@handayani.ac.id

Abstract

Wood carving is a cultural heritage with deep meaning and significance for the Toraja ethnic group's culture. By understanding the meaning of each Toraja carving, both tourists and the local community can gain knowledge about Toraja culture, thereby preserving and maintaining the culture amidst modern developments. Image processing approaches, particularly the development of Convolutional Neural Networks (CNN), offer a solution for extracting information from the diverse and intricate patterns of Toraja wood carvings. This study is highly significant as it implements a deep learning model using the CNN algorithm optimized with the ResNet50 architecture. The methodology in this study involves adjusting the batch size during the model training phase and applying weak-to-strong pixel transformation during the double threshold hysteresis phase in the Canny Feature Extraction process on the edges of Toraja carving images, resulting in ResNet50 architecture accurately recognizing the patterns of Toraja wood carvings. The results demonstrate significant improvements in the performance of the ResNet50 architecture with the preprocessed dataset. average precision, recall, precision, and F1-Score improvements in each Toraja carving class. For the Pa' Lulun Pao class, it was found that the precision and recall values were 0.94, and the F1-Score was 0.94. The Pa' Somba class also showed good results, with a precision value of 0.9697, a recall of 0.96, and an F1-Score of 0.9648. The Pa' Tangke Lumu class showed even better results, with a precision value of 0.9898, a recall of 0.97, and an F1-Score of 0.9798. The Pa' Tumuru class also demonstrated good performance, with a precision value of 0.9327, a recall of 0.97, and an F1-Score of 0.9500. This study not only underscores the effectiveness of processing in enhancing CNN capabilities but also opens opportunities for further research in applying these methods to various image types and exploring different CNN architectures for a more comprehensive understanding.

Keywords: CNN; image classification; Toraja wood carvings

How to Cite: Nurilmiyanti Wardhani, B. E. W. Asrul, Antonius Riman Tampang, Sitti Zuhriyah, and Abdul Latief Arda, "Classification of Toraja Wood Carving Motif Images Using Convolutional Neural Network (CNN)", J. RESTI (Rekayasa Sist. Teknol. Inf.), vol. 8, no. 4, pp. 486 - 495, Aug. 2024.
DOI: https://doi.org/10.29207/resti.v8i4.5897

1. Introduction

Feature extraction mechanisms are used to generate more information about the images which helps us to have an idea of what a particular image is talking about or referring to. Convolutional Neural Networks (CNN) development opens up a huge area in the image classification and object detection [1].

The tongkonan houses of Ke'te Kesu are based on the work by contemporary Torajan woodcarvers, in a tradition passed down from their ancestors [2]. The carvings enhance the character of this vernacular house which is visited by Torajan art and culture-loving guests Each motif has a name and its meaning, and all the details of wooden carving reach every corner inside the house [3]. Nevertheless, the significance of Toraja wood carvings on traditional Tongkonan homes is laced with deep meaning which few visitors can understand what they represent as part of Torajanese culture. The striking aesthetic carvings produce another effect when the common knowledge part is removed; visitors look admiringly upon them yet are not familiar with their names or roots. However, this ends up providing an experience with tourism limited exclusively to the visual aspect and without any sense of cultural meaning [4], [5].

Received: 01-07-2024 | Accepted: 04-08-2024 | Published Online: 07-08-2024

To address this issue, a solution is required that can facilitate a deeper understanding of Toraja wood carvings by providing information about their names and meanings. This would enhance tourists' appreciation of the rich art and culture of Toraja, particularly its wood carving tradition. Implementing computer vision techniques could be an effective method for extracting information from images of Toraja carvings at tourist sites [6] - [8].

Additionally, although Toraja carvings are distinct, their intricate patterns often give them a visually similar appearance, making them difficult to differentiate [4]. Therefore, it is crucial to implement a comprehensive identification process to accurately categorize the types of carvings [9], [10]. This research combines computer vision and artificial intelligence methodologies to achieve optimal image analysis of the carvings.

By leveraging advances in deep learning technology, particularly CNN algorithms and the ResNet50 architecture [6], [11] - [13] the proposed solution is expected to enhance the tourism experience at Ke'te Kesu. Visitors will gain a deeper understanding of Toraja wood carving art, allowing them to appreciate the intricate details and cultural significance embedded in each motif of the Tongkonan traditional houses.

Another review of the literature, on image classification has been conducted, including studies on distinguishing various types of carvings using the ResNet50 Convolutional Neural Network architecture. Through a series of weight testing, it was revealed that Xception achieved the highest accuracy in classifying Jepara carving motifs, with percentages of 0.95%, 0.95%, and 0.94% on datasets in different colour spaces, namely LUV, RGB, and YCrCb. However, when the ResNet50 model was applied to the Jepara motif recognition system, the results showed that ResNet50 outperformed all other architectures, achieving motif identification rates of 84%, 79%, and 80% in the LUV, RGB, and YCrCb colour spaces, respectively. [2]. These findings demonstrate the system's capability to determine whether a carving belongs to the Jepara carving type by recognizing the distinctive motifs characteristic of Jepara embedded in the carvings [14].

In another study, a content-based image retrieval (CBIR) system was developed using pre-trained deep CNN models, namely VGG16 and ResNet-50. This study compared these models in terms of image retrieval efficiency and reliability and tested them on a multi-class digital image dataset. The results showed that the ResNet-50-based system outperformed the VGG16-based system, with an average retrieval precision of 90.18% compared to 81.68%. This demonstrates that the use of deep learning models for feature extraction in CBIR significantly improves performance compared to traditional methods. The improvement in precision was consistent across different image categories and retrieval sizes. Overall, this study successfully implemented an efficient and reliable CBIR system using pre-trained CNN models (VGG16 and ResNet-50) [13].

In another study, Improving Batik Pattern Classification using CNN with Advanced Augmentation and Oversampling on Imbalanced Dataset. This approach enhanced the diversity of the images, encompassing variations in colour, contrast, wrinkles, and warps that may be present in batik garments. This study employed two CNN models, DenseNet169 and VGG-16, along with three different training methods this study. These methods included training without oversampling and advanced augmentation, training with oversampling, and training with both oversampling and advanced augmentation [12].

Another study identified traditional Balinese ornament carvings using the Multilayer Perceptron method. This research aimed to determine the accuracy of identifying traditional Balinese ornament carvings by utilizing digital image processing technology with Gabor filters for feature extraction and Multilayer Perceptron for classification. The study used 18 classes of traditional Balinese ornament carvings with a total dataset of 268 samples. The Multilayer Perceptron method applied a backpropagation learning algorithm with 10,560 input neuron layers, 50 hidden neuron layers, and 18 output neuron layers for classification. The testing achieved an accuracy of 43%. Classification testing using k-fold cross-validation with K = 5 resulted in an average accuracy of 41.14%, with optimal accuracy reaching 56%. The testing results using a Confusion Matrix showed an accuracy of 43.3%, a sensitivity of 42.68%, and a specificity of 96.87% [9], [15].

Subsequent research applied CNN to classify reliefs based on their images. The CNN used had a MobileNetV2 architecture with several adjustments. The results indicated that this method could achieve a classification accuracy of up to 94.5%, demonstrating its potential in automatically classifying types of reliefs [10].

The ResNet architecture employs skip connections in 2-3 layers with ReLU and batch normalization. Residual blocks are used when the input dimensions are the same as the output. Each ResNet block consists of either 2 layers (ResNet 18, 34) or 3 layers (ResNet 50, 101, 152). Initially, similar to GoogleNet, it starts with a 7x7 convolution and 3x3 max pooling (stride 2). ResNet offers variations with 18, 34, 50, 101, or 152 layers [16], [17].

The study presents a novel application of CNN and ResNet50 architecture to classify Toraja wood carving motifs. This contributes to both the field of computer vision and cultural heritage preservation. This research is highly significant as it implements a deep learning model using the optimized CNN algorithm with ResNet50 architecture. By adjusting the batch size and applying weak-to-strong pixel transformation, where pixels with intensities between a low threshold and a high threshold are considered weak, and pixels above the high threshold are considered strong. Subsequently, strong pixels are added to the image edges and iterated to spread strong pixels to weak pixels connected in an image. Furthermore, it calculates the classification accuracy of Toraja wood carving images using deep learning with the CNN algorithm and ResNet50 architecture when analyzing 2000 sample Carving Images. This study is expected to serve as a valuable reference for the research and development of ResNet50 architecture, ensuring it effectively addresses image classification challenges optimally [4], [17].

2. Research Methods

The Canny function takes the image processed in the previous stage and computes the edges of the image using all previously defined functions. The final result is a binary image showing the edges of the original image. The Canny function has two parameters: a low threshold value and a high threshold value, which are used in the double threshold hysteresis stage. The low threshold value is used to filter out pixels with low intensity that are considered not part of the image edges, while the high threshold value is used to filter pixels with high intensity that are considered part of the image edges.

The dataset for this research was collected directly from Tongkonan traditional houses at the Ke'te Kesu Tourism Object in North Toraja Regency, South Sulawesi Province. The dataset used consists of 5 classes: 4 classes of Pa' Lulun Pao carving motif images as shown in Figure 1, Pa' Tumuru carving images in Figure 4, Pa' Somba carving images in Figure 2, and Pa' Tengke Lumu carving dataset in Figure 3. The fifth class is the Undetected Class as depicted in Figure 5. The dataset was divided into training and testing data with an 80% to 20% ratio, respectively, and each image file had a resolution of 224x224 pixels. This resolution was chosen as it is the standard input shape for the ResNet architecture.

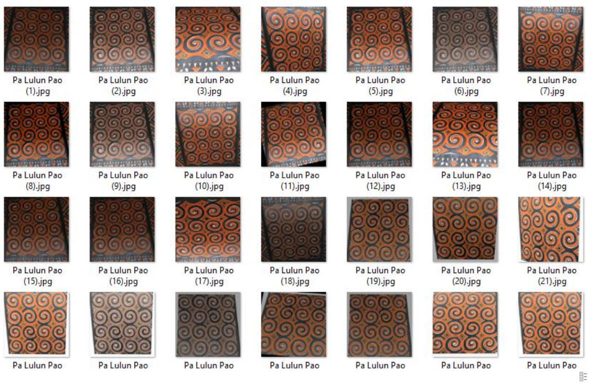


Figure 1. Pa' Lulun Pao Datasets

Figure 2. Pa' Somba Datasets

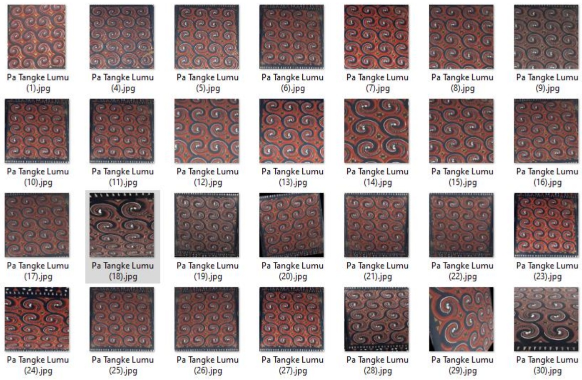


Figure 3. Pa' Tangke Lumu Datasets

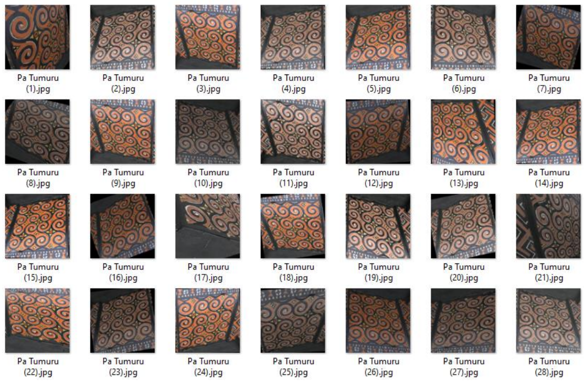


Figure 4. Pa' Tumuru Datasets

Figure 5. Other Object Datasets

In Figure 6, the first preprocessing step performed is image conversion to grayscale, where this process transforms colour images into grayscale images. Converting images to grayscale can facilitate image analysis by reducing them to a single colour channel, which helps save memory and computational time [18] - [20].

Figure 6. Grayscale Process

After converting the image to grayscale, the next step is noise reduction using a Gaussian filter with a 3x3 kernel[21], as shown in Equation 1.

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{1}$$

From the kernel matrix, the total weight value of the matrix is 16. Therefore, the normalization of weights used in the Gaussian filter convolution process is as follows: [22], [23], as shown in Equation 2.

$$G_{(x,y)} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{2}$$

To apply Equation 2, pixel sizes of 5x5 are taken as samples and a 3x3 kernel is applied, starting by placing the kernel at the top-left position of the pixel and inputting the calculated result values into the pixel located at the centre of the kernel, beginning from the first convolution, as shown in Equation 3.

$$G_{(x,y)} = 28.375 = 28 \ (after\ rounded) \tag{3}$$

After performing convolution, the sequence of pixels obtained is as follows:

26	34	26	28	27
29	25	26	27	28
30	31	32	27	35
18	34	25	29	42
66	5	31	39	23

	28	***28***	***28***	
	29	***29***	***30***	
	28	***28***	***31***	

In Gaussian filtering, edge pixels are ignored because they lack neighbours for convolution. In Canny edge detection, the Sobel filter is used to preserve details [20]. The Sobel filter is implemented with kernels G_x and G_y in Canny edge detection, as shown in Equation 4.

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{4}$$

In Figure 7, edge thinning is further performed using the Non-Maximum Suppression (NMS) technique on the gradient magnitude. NMS compares the edge magnitude values of the tested pixel with its neighbors in the gradient direction. If the tested value is higher, the pixel is retained as an edge; otherwise, it is removed as not an edge [24].

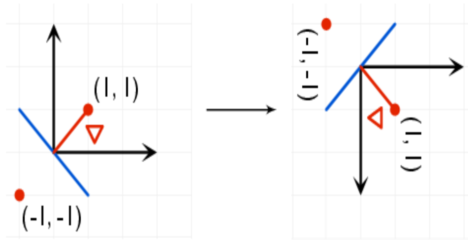


Figure 7. Non-Maximum Suppression technique

In Figure 8, the final step involves converting weak pixels to strong pixels using the Double Threshold Hysteresis technique. In this technique, an image is categorised into two types: strong pixels and weak pixels [25]. Subsequently, weak pixels above a specified threshold value are elevated to strong pixels, while weak pixels below the threshold are removed [26].

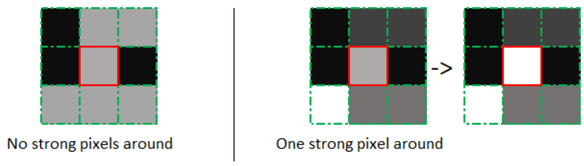


Figure 8. Double Threshold Hysteresis technique

3. Result and Discussions

In this stage, the results and analysis of the research on Toraja wood carving classification using Convolutional Neural Network (CNN) with Canny edge extraction and ResNet-50 architecture will be discussed. The experimental outcomes of implementing this model will be explored, along with a detailed analysis of its performance in classifying wood carvings from Toraja culture. This discussion will include interpreting the results, identifying the strengths and limitations of the proposed model, and discussing the implications of these findings for further developments in the field of art carving image classification.

During the training process, JPG format Carving Image data is used, encompassing various types of carvings, each image having a resolution of 224x224 pixels. This study aims to classify carvings by evaluating Image data. The training process aims to analyze 4 specific pattern classes of carvings: Pa Lulu Pao carvings, Pa Somba carvings, Pa Tengke Lumu carvings, and Pa Tumuru carvings. A total of 2500 images were used in the training data, evenly distributed among the four formed classes. The data was selected based on motifs relevant to the research topic to enhance understanding and prediction of Toraja carving patterns. Training image classification models with objects that differ from the target images is paramount for building robust and reliable models for real-world applications [27],[28]. By utilizing a diverse and representative training dataset, the model can achieve superior generalization, resilience to variations, comprehensive class coverage, and accurate object detection [29],[30]. Table 1 displays the carving motif classes along with their respective data counts.

Table 1. Toraja Carvings and Dataset Counts

No	Class	Total
1	Pa Lulun Pao	500
2	Pa Somba	500
3	Pa Tengke Lumu	500
4	Pa Tumuru	500
5	Other Object	500
Total		2500

3.1. Training Process

The training process was conducted using the Google Colab environment. This platform enables model training using powerful computational resources and web-based accessibility. With Google Colab, Python code can be executed, popular libraries like TensorFlow and PyTorch can be accessed, and GPU or TPU can be utilized to accelerate the machine learning model training process.

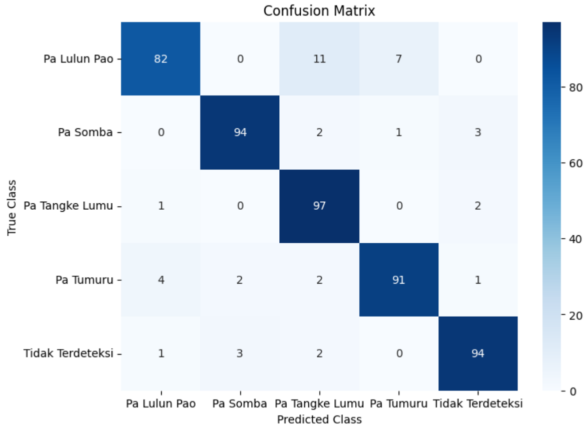
Table 2. Learning Rate Test Scheme

Process	Learning Rate
Testing 1	0.01
Testing 2	0.001
Testing 3	0.0001
Testing 4	0.00001

The selection of appropriate hyperparameters is a critical aspect in training machine learning models, particularly for image classification tasks in computer vision [31],[32]. Hyperparameters, such as batch size, learning rate, and epoch, significantly influence model performance and determine whether the model can learn effectively and achieve optimal performance [33],[34]. In the classification of Toraja wood carvings, the designed model will be implemented into a system for classification using CNN with ResNet50 architecture, thus requiring an optimally functioning model. Therefore, testing of the model to be used is necessary. Parameters used to measure the performance of the model in this study include learning rate and epochs. A batch size of 32 will be used to minimize errors during training, and the Shuffle technique will be employed to randomize the dataset used for testing. [35].

Testing on the learning rate was conducted to find the most optimal weights from the training dataset results with different learning rates, namely 0.01, 0.001, 0.0001, and 0.00001, with a fixed batch size of 32 and an epoch of 100 as seen in Table 2. The following are the results of the learning rate tests. The range of values for this learning rate is from zero (0) to one (1). The higher the learning rate, the faster the training process runs. However, higher learning rates may lead to less accurate networks, and vice versa.

(a)



(b)

Figure 9. Training Results (a) Epoch 100 Testing Graph (b) Confusion Matrix Epoch 100 learning rate 0.01.

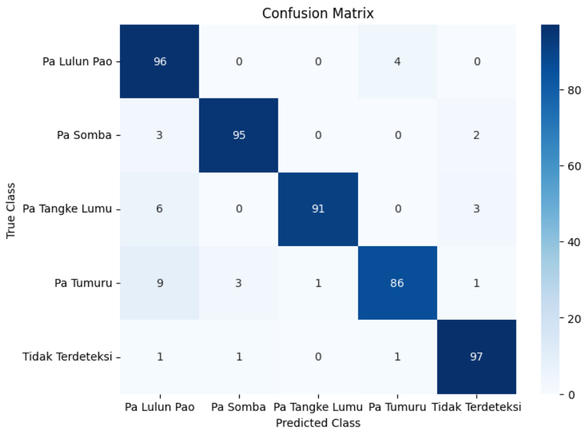
Table 3. Confusion Matrix Calculation Table Learning Rate 0.01

Class	Precision	Recall	Accuracy	F1-Score
Pa Lulun Pao	0,93	0,82	0,95	0,87
Pa Somba	0,95	0,94	0,98	0,94
Pa Tengke Lumu	0,85	0,97	0,96	0,91
Pa Tumuru	0,92	0,91	0,97	0,91
Other Object	0,94	0,94	0,98	0,94

In Figure 9 (a) and (b), the testing results of learning rate 0.01 show improved validation accuracy and training validation at epoch 10. However, the validation accuracy graph shows unstable training improvements and tends to decline. The Classification Report using the Confusion Matrix with a testing dataset of 100 images per class is presented in Figure 9. Table 3 shows that the ability to predict at a learning rate of 0.01 achieved the highest result of 0.98.

In Table 3, the best accuracy is observed in predicting images of Pa' Somba carvings with a value of 0.98, accompanied by a Precision of 0.95 and a Recall of 0.94, resulting in an F1-Score of 0.94.

(a)



(b)

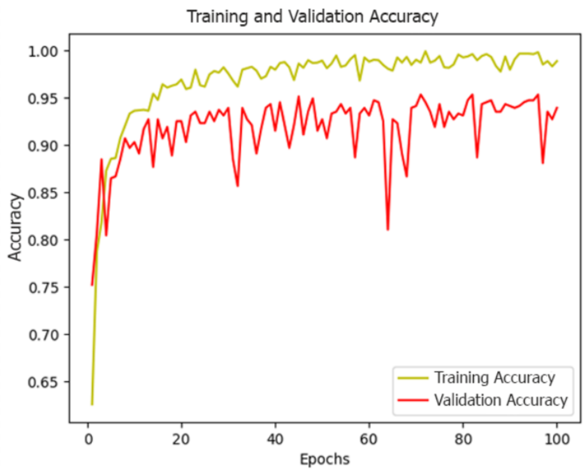
Figure 10. Training Results (a) Epoch 100 Testing Graph (b) Confusion Matrix Epoch 100 learning rate 0.001.

Table 4. Confusion Matrix Calculation Table Learning Rate 0.001

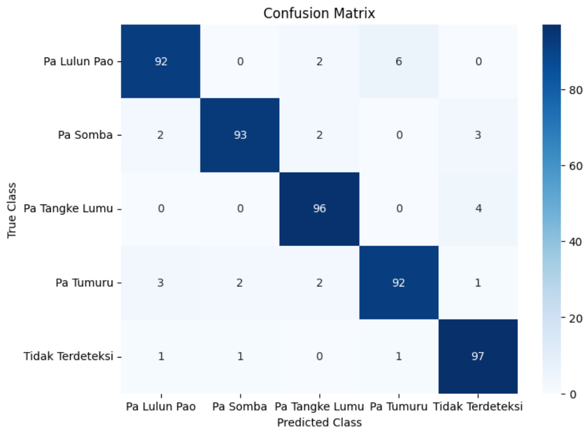
Class	Precision	Recall	Accuracy	F1-Score
Pa Lulun Pao	0,83	0,96	0,95	0,89
Pa Somba	0,96	0,95	0,98	0,95
Pa Tengke Lumu	0,99	0,91	0,98	0,95
Pa Tumuru	0,95	0,86	0,96	0,90
Other Object	0,94	0,97	0,98	0,96

In Figure 10 (a) and (b), the testing results of learning rate 0.001 show a significant increase in validation accuracy and training validation by epoch 15. However, the validation accuracy graph shows unstable training improvements and tends to decline. Table 4 is the Classification Report using the Confusion Matrix with a testing dataset of 100 images per class. Table 4 shows that the predictive ability at a learning rate of 0.001 achieved the highest result of 0.98.

In Table 4, the best accuracy is observed in predicting images of Pa' Somba, and Pa’ Tengke Lumu carvings with a value of 0.98, accompanied by a Precision of 0.95 and a Recall of 0.94, resulting in an F1-Score of 0.94.



(a)



(b)

Figure 11. Training Results (a) Epoch 100 Testing Graph (b) Confusion Matrix Epoch 100 learning rate 0.0001.

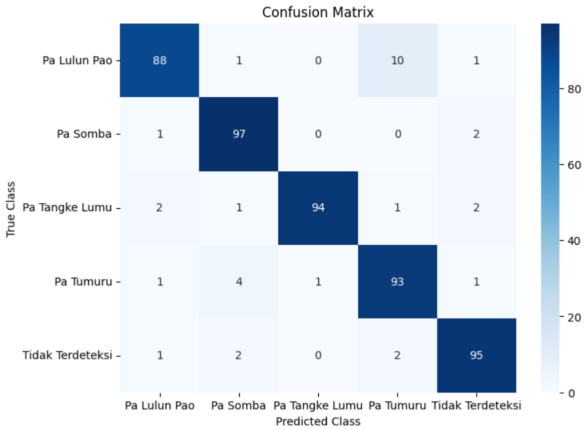
Table 5. Confusion Matrix Calculation Table Learning Rate 0.0001

Class	Precision	Recall	Accuracy	F1-Score
Pa Lulun Pao	0,94	0,92	0,97	0,93
Pa Somba	0,97	0,93	0,98	0,95
Pa Tengke Lumu	0,94	0,96	0,98	0,95
Pa Tumuru	0,93	0,92	0,97	0,92
Other Object	0,92	0,97	0,98	0,95

In Figure 11 (a) and (b), the testing results of learning rate 0.0001 show rapid improvement in validation accuracy and training validation by epoch 10. However, the validation accuracy graph indicates unstable training improvements that tend to decline. Table 4 is the Classification Report using the Confusion Matrix with a testing dataset of 100 images per class. Table 5 shows that the predictive capability at a learning rate of 0.0001 achieved the highest result of 0.98.

In Table 5, the best accuracy is observed in predicting images of Pa' Somba, and Pa’ Tengke Lumu carvings with a value of 0.98, accompanied by a Precision of 0.97 and a Recall of 0.93, resulting in an F1-Score of 0.95.

(a)



(b)

Figure 12. Training Results (a) Epoch 100 Testing Graph (b) Confusion Matrix Epoch 100 learning rate 0.00001.

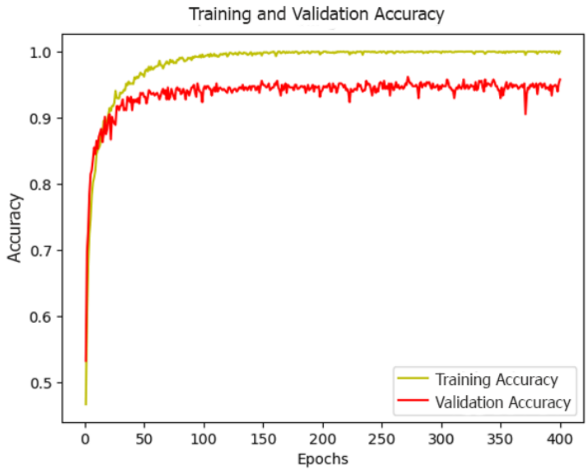
Table 6. Confusion Matrix Calculation Table Learning Rate 0.00001

Class	Precision	Recall	Accuracy	F1-Score
Pa Lulun Pao	0,95	0,88	0,97	0,91
Pa Somba	0,92	0,97	0,98	0,97
Pa Tengke Lumu	0,99	0,94	0,99	0,96
Pa Tumuru	0,88	0,93	0,96	0,90
Other Object	0,94	0,95	0,98	0,95

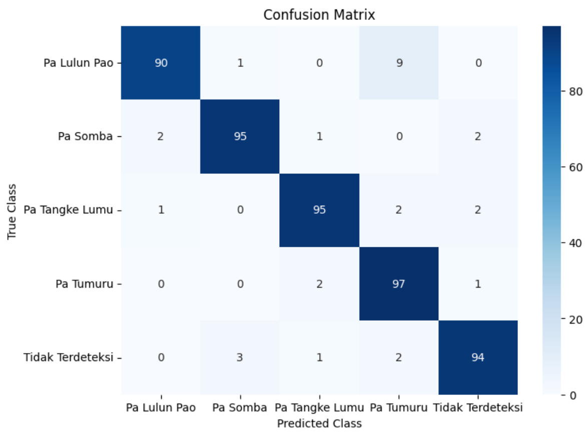
In Figure 12 (a) and (b), the testing results of learning rate 0.00001 show rapid improvement in validation accuracy and training validation by epoch 10. The training and validation graphs appear stable thereafter, but the validation accuracy does not show rapid improvement. Table 6 is the Classification Report using the Confusion Matrix with a testing dataset of 100 images per class. Table 6 shows that the predictive capability at a learning rate of 0.00001 achieved the highest result of 0.99.

Based on the test results, it is found that the model with a learning rate of 0.00001 is the most stable, achieving the highest accuracy of 0.99 as seen from the graph. Hence, it can be used as the parameter for the learning rate in subsequent tests. Once the entire dataset has undergone training in the Neural Network it returns to the beginning for one complete cycle. Four models were generated. In the Epoch testing, four models were created with variations of epochs: 100, 200, 300, and 400.

In Figure 9, from the training results using a learning rate of 0.0001 and 400 epochs, it is observed that this variant model is the most stable based on both the graphical representation and confusion matrix testing, which analyzes and measures the model's performance in classification [36].



(a)



(b)

Figure 13. Training Results (a) Epoch 400 Testing Graph (b) Confusion Matrix Epoch 400 learning rate 0.00001.

In Figure 13, For the class Pa' Lulun Pao, it was found that the Precision and Recall values are 0.94, with an F1-Score of 0.94. Class Pa' Somba also shows promising results, with a Precision value of 0.9697, Recall of 0.96, and F1-Score of 0.9648.

Class Pa' Tangke Lumu demonstrates even better performance, with a Precision of 0.9898, Recall of 0.97, and F1-Score of 0.9798. Class Pa' Tumuru also performs well, with a Precision of 0.9327, Recall of 0.97, and F1-Score of 0.9500. The Not Detected class shows solid performance, with a Precision of 0.9592, Recall of 0.95, and F1-Score of 0.9546.

3.2. Training accuracy testing

In Table 7, the accuracy of the system will be tested by using images manipulated with randomly added lines to the images using 5 sample images from each class. The purpose of this manipulation is to test how well the system can maintain its performance and remain accurate in classifying images that have changed.

Table 7. Results of Training Accuracy Testing

No	Actual	Prediction	Description
1	Pa Lulun Pao	Pa Lulun Pao	TP
2	Pa Lulun Pao	Pa Lulun Pao	TP
3	Pa Lulun Pao	Pa Lulun Pao	TP
4	Pa Lulun Pao	Pa Tumuru	FN
5	Pa Lulun Pao	Pa Lulun Pao	TP
6	Pa Somba	Pa Somba	TP
7	Pa Somba	Pa Somba	TP
8	Pa Somba	Pa Somba	TP
9	Pa Somba	Pa Somba	TP
10	Pa Somba	Pa Somba	TP
11	Pa Tangke Lumu	Pa Tangke Lumu	TP
12	Pa Tangke Lumu	Pa Tangke Lumu	TP
13	Pa Tangke Lumu	Pa Tangke Lumu	TP
14	Pa Tangke Lumu	Pa Tangke Lumu	TP
15	Pa Tangke Lumu	Pa Tangke Lumu	TP
16	Pa Tumuru	Pa Tumuru	TP
17	Pa Tumuru	Pa Tumuru	TP
18	Pa Tumuru	Pa Tumuru	TP
19	Pa Tumuru	Pa Tumuru	TP
20	Pa Tumuru	Pa Somba	FN
21	Other Object	Other Object	TP
22	Other Object	Other Object	TP
23	Other Object	Other Object	TP
24	Other Object	Other Object	TP
25	Other Object	Other Object	TP

To calculate the overall accuracy from all tests in Table 7, the total True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) for each class are summed. The calculation formula is shown in Equation 5.

$$Accuracy = \frac{TP\_total + TN\_total}{TP\_total + TN\_total + FP\_total + FN\_total} = \frac{23 + 98}{23 + 98 + 2 + 2} = \frac{121}{125} = 0.968 \quad (5)$$

Based on the calculations performed, a total accuracy of 0.968 or approximately 96.8% was obtained. Compared to batik motif classification research using the ResNet50 architecture, which achieved an accuracy of 96% [11][12], the current study shows a 0.8% improvement in object recognition performance after optimization. Additionally, when compared to the Content Image ResNet50 research, which was juxtaposed with the VGG16 architecture and achieved an accuracy of 90.18% [13], the ResNet50 architecture in this study outperforms by 6.62%. This demonstrates that ResNet50 improves its performance when weak pixels are optimized to strong pixels using the Double Threshold Hysteresis technique. Accuracy is calculated by dividing the number of correct predictions (True Positive and True Negative) by the total number of evaluated samples (True Positive, True Negative, False Positive, and False Negative). In this test, there were 23 correct positive predictions (True Positive) and 98 correct negative predictions (True Negative). Meanwhile, there were 2 incorrect positive predictions (False Positive) and 2 incorrect negative predictions (False Negative).

4. Conclusion

A meticulous Comprehensive Evaluation was conducted on the testing of the ResNet50 architecture model implemented on the CNN model for pattern recognition of Toraja wood carvings. The evaluation employed a transformation from weak pixels to strong pixels in the double threshold hysteresis stage of the Canny Feature Extraction stage on the edges of Toraja carving images to enhance the accuracy of carving image pattern recognition. Additionally, the incorporation of hyperparameter tuning such as batch size, learning rate, and epoch offered a holistic perspective on the CNN training process. This research concludes that the Toraja carving classification model using the Convolutional Neural Network (CNN) technique with Canny feature extraction and ResNet-50 architecture is capable of delivering stable and satisfactory performance. A dataset encompassing carving motifs such as Pa’ Tangke Lumu, Pa’ Lulun Pao, Pa’ Tumuru, Pa’ Somba, and the Undetected class was utilized in the evaluation. From the testing of learning rates 0.01, 0.001, 0.0001, 0.00001, and epochs 100, 200, 300, and 400, the results indicate that the selection of hyperparameters with a learning rate of 0.00001 and epoch 400 yields the highest F1-Score. This is evident from the result graph and Confusion Matrix. All five classes demonstrate promising performance in Precision, Recall, and F1-Score. Specifically, the Precision and Recall values of 0.94 for the Pa' Lulun Pao class, Precision 0.9697 and Recall 0.96 for the Pa' Somba class, and Precision 0.9898 and Recall 0.97 for the Pa' Tangke Lumu class affirm the model's quality. Evaluation using a confusion matrix resulted in an overall accuracy of 96.8%. These evaluation outcomes reinforce the success of applying hyperparameter tuning techniques and ResNet50 architecture optimization to the CNN algorithm, positively impacting the utilization of Toraja wood carving recognition applications. In conclusion, this study opens up significant potential avenues for further research to enhance CNN models in image pattern recognition. However, for future investigations, it is recommended to expand the application of these techniques to various image types beyond wood carving pattern recognition, such as exploring diverse CNN architecture models like AlexNet, VGGNet, MobileNet, EfficientNet, and Xception. Additionally, investigating the Activation function of other hyperparameters could contribute to a more comprehensive understanding of broader applications in diverse image-processing contexts.

Acknowledgements

The author declares that there is no conflict of interest. This research did not receive a specific grant from any funding agency in the public, commercial, or non-profit sectors.

References

[1] S. Liu, L. Wang, and W. Yue, "An efficient medical image classification network based on multi-branch CNN, token grouping Transformer and mixer MLP," *Appl. Soft Comput.*, vol. 153, p. 111323, 2024, doi: https://doi.org/10.1016/j.asoc.2024.111323.

[2] N. S. Ouf, "Leguminous seeds detection based on convolutional neural networks: Comparison of Faster R-CNN and YOLOv4 on a small custom dataset," *Artif. Intell. Agric.*, vol. 8, pp. 30–45, 2023, doi: 10.1016/j.aiia.2023.03.002.

[3] M. Xia, J. Chen, G. Yang, and S. Wang, "Robust detection of seam carving with low ratio via pixel adjacency subtraction and CNN-based transfer learning," *J. Inf. Secur. Appl.*, vol. 75, p. 103522, 2023, doi: https://doi.org/10.1016/j.jisa.2023.103522.

[4] E. Fragoso-Navarro, K. Rangel-Espinoza, M. Nakano-Miyatake, M. Cedillo-Hernandez, and H. Perez-Meana, "Seam Carving based visible watermarking robust to removal attacks," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 7, pp. 4499–4513, 2022, doi: 10.1016/j.jksuci.2021.03.010.

[5] M. Das and M. Panda, "Seam carving, horizontal projection profile and contour tracing for line and word segmentation of language independent handwritten documents," *Results Eng.*, vol. 18, no. September 2022, p. 101110, 2023, doi: 10.1016/j.rineng.2023.101110.

[6] I. Rodriguez-Conde, C. Campos, and F. Fdez-Riverola, *Optimized convolutional neural network architectures for efficient on-device vision-based object detection*, vol. 34, no. 13. Springer London, 2022.

[7] H. Pei, C. Zhang, X. Zhang, X. Liu, and Y. Ma, "Recognizing materials in cultural relic images using computer vision and attention mechanism," *Expert Syst. Appl.*, vol. 239, p. 122399, 2024, doi: https://doi.org/10.1016/j.eswa.2023.122399.

[8] S. Gao, G. Huang, X. Chen, H. Jiang, L. Zhou, and X. Gao, "Two-stage deep learning-based video image recognition of early fires in heritage buildings," *Eng. Appl. Artif. Intell.*, vol. 129, p. 107598, 2024, doi: https://doi.org/10.1016/j.engappai.2023.107598.

[9] I. W. A. S. Darma, N. Suciati, and D. Siahaan, "CARVING-DETC: A network scaling and NMS ensemble for Balinese carving motif detection method," *Vis. Informatics*, vol. 7, no. 3, pp. 1–10, 2023, doi: 10.1016/j.visinf.2023.05.004.

[10] A. P. Wibawa *et al.*, "Decoding and preserving Indonesia's iconic Keris via A CNN-based classification," *Telemat. Informatics Reports*, vol. 13, no. December 2023, p. 100120, 2024, doi: 10.1016/j.teler.2024.100120.

[11] B. S. Negara, E. Satria, S. Sanjaya, and D. R. D. Santoso, "ResNet-50 for Classifying Indonesian Batik with Data Augmentation," in *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, 2021, pp. 1–4, doi: 10.1109/ICOTEN52080.2021.9493488.

[12] B. J. Filia *et al.*, "Improving Batik Pattern Classification using CNN with Advanced Augmentation and Oversampling on Imbalanced Dataset," *Procedia Comput. Sci.*, vol. 227, pp. 508–517, 2023, doi: 10.1016/j.procs.2023.10.552.

[13] G. Gautam and A. Khanna, "Content Based Image Retrieval System Using CNN based Deep Learning Models," *Procedia Comput. Sci.*, vol. 235, no. 2023, pp. 3131–3141, 2024, doi: 10.1016/j.procs.2024.04.296.

[14] L. Z. PCSW and Y. Kristian, "Identification of Jepara Motifs in Carvings Using Convolutional Neural Network," *J. Nas. Tek. Elektro Dan Teknol. Inf.*, vol. 9, no. 4, pp. 403–413, 2020, doi: https://doi.org/10.22146/jnteti.v9i4.541.

[15] M. Irshad, N. F. Law, K. H. Loo, and S. Haider, "IMGCAT: An approach to dismantle the anonymity of a source camera using correlative features and an integrated 1D convolutional neural network," *Array*, vol. 18, no. February, p. 100279, 2023, doi: 10.1016/j.array.2023.100279.

[16] H. Lyu, F. Qiu, L. An, D. Stow, R. Lewison, and E. Bohnett, "Deer survey from drone thermal imagery using enhanced faster R-CNN based on ResNets and FPN," *Ecol. Inform.*, vol. 79, no. July 2023, 2024, doi: 10.1016/j.ecoinf.2023.102383.

[17] F. Ben Nasr Barber and A. Elloumi Oueslati, "Human exons and introns classification using pre-trained Resnet-50 and GoogleNet models and 13-layers CNN model," *J. Genet. Eng. Biotechnol.*, vol. 22, no. 1, p. 100359, 2024, doi: 10.1016/j.jgeb.2024.100359.

[18] D. Chen *et al.*, "Computed tomography reconstruction based on canny edge detection algorithm for acute expansion of epidural hematoma," *J. Radiat. Res. Appl. Sci.*, vol. 15, no. 3, pp. 279–284, 2022, doi: 10.1016/j.jrras.2022.07.011.

[19] H. Hua, M. Liu, Y. Li, S. Deng, and Q. Wang, "An ensemble framework for short-term load forecasting based on parallel CNN and GRU with improved ResNet," *Electr. Power Syst. Res.*, vol. 216, p. 109057, 2023, doi: https://doi.org/10.1016/j.epsr.2022.109057.

[20] M. Huang, Y. Liu, and Y. Yang, "Edge detection of ore and rock on the surface of explosion pile based on improved Canny operator," *Alexandria Eng. J.*, vol. 61, no. 12, pp. 10769–10777, 2022, doi: 10.1016/j.aej.2022.04.019.

[21] D. Z. E. Prastya, D. P. Pamungkas, and R. K. Niswatin, "Implementation of Gaussian Filter and Median Filter for Image Smoothing," in *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 2022, vol. 6, no. 2, pp. 178–187.

[22] S. Sinurat and E. R. Siagian, "Enhancement of Image Quality Using Gaussian Filter on Images Resulting from Robert's Edge Detection," *Pelita Inform. Inf. dan Inform.*, vol. 9, no. 3, pp. 225–231, 2021.

[23] P. Khodaee, A. Esfahanipour, and H. Mehtari Taheri, "Forecasting turning points in stock price by applying a novel hybrid CNN-LSTM-ResNet model fed by 2D segmented images," *Eng. Appl. Artif. Intell.*, vol. 116, p. 105464, 2022, doi: https://doi.org/10.1016/j.engappai.2022.105464.

[24] M. Gong, D. Wang, X. Zhao, H. Guo, D. Luo, and M. Song, "A review of non-maximum suppression algorithms for deep learning target detection," in *Seventh Symposium on Novel Photoelectronic Detection Technology and Applications*, 2021, vol. 11763, pp. 821–828.

[25] I. Kurniasari, Y. B. Utomo, and A. D. Evasari, "Design of a Mobile-Based Clothing Measurement Application," *Inf. (Journal Inform. and Sist. Informasi)*, vol. 14, no. 2, pp. 170–179, 2022.

[26] F. Marpaung, F. Aulia, and R. C. Nabila, "Computer Vision and Digital Image Processing." PUSTAKA AKSARA, 2022.

[27] Z. Li, D. Yu, M. Wu, S. Chan, H. Yu, and Z. Han, "Revisiting single-step adversarial training for robustness and generalization," *Pattern Recognit.*, vol. 151, p. 110356, 2024, doi: https://doi.org/10.1016/j.patcog.2024.110356.

[28] C. Chen, Y. Yang, M. Liu, Z. Rong, and S. Shu, "Regularized joint self-training: A cross-domain generalization method for image classification," *Eng. Appl. Artif. Intell.*, vol. 134, p. 108707, 2024, doi: https://doi.org/10.1016/j.engappai.2024.108707.

[29] J. Guo *et al.*, "A comprehensive evaluation framework for deep model robustness," *Pattern Recognit.*, vol. 137, p. 109308, 2023, doi: https://doi.org/10.1016/j.patcog.2023.109308.

[30] Z. Yin, J. Pu, R. Wan, and X. Xue, "Embrace sustainable AI: Dynamic data subset selection for image classification," *Pattern Recognit.*, vol. 151, p. 110392, 2024, doi: https://doi.org/10.1016/j.patcog.2024.110392.

[31] M. Akhtar, M. Tanveer, and M. Arshad, "Advancing Supervised Learning with the Wave Loss Function: A Robust and Smooth Approach," *Pattern Recognit.*, vol. 155, p. 110637, 2024, doi: https://doi.org/10.1016/j.patcog.2024.110637.

[32] M. Pintor *et al.*, "ImageNet-Patch: A dataset for benchmarking machine learning robustness against adversarial patches," *Pattern Recognit.*, vol. 134, p. 109064, 2023, doi: https://doi.org/10.1016/j.patcog.2022.109064.

[33] Y.-G. Wang, J. Wu, Z.-H. Hu, and G. J. McLachlan, "A new

algorithm for support vector regression with automatic selection of hyperparameters," *Pattern Recognit.*, vol. 133, p. 108989, 2023, doi: https://doi.org/10.1016/j.patcog.2022.108989.

[34] L. Yi, D. Lv, D. Liu, S. Li, and R. Liu, "Self-Training-Transductive-Learning Broad Learning System (STTL-BLS): A model for effective and efficient image classification," *Pattern Recognit.*, vol. 156, p. 110747, 2024, doi: https://doi.org/10.1016/j.patcog.2024.110747.

[35] T. Chen, Y. Tan, Z. Zhang, N. Luo, B. Li, and Y. Li, "Dataflow optimization with layer-wise design variables estimation method for enflame CNN accelerators," *J. Parallel Distrib. Comput.*, vol. 189, p. 104869, 2024, doi: https://doi.org/10.1016/j.jpdc.2024.104869.

[36] B. P. Pratiwi, A. S. Handayani, and S. Sarjana, "Performance Evaluation of Air Quality System Using WSN Technology with Confusion Matrix," *J. Inform. Upgris*, vol. 6, no. 2, 2020.