Using Convolutional Neural Network and Saliency Maps for Cirebon Batik Recognition

Yoga Aditiya [1], Marlinda Vasty Overbeek[2], Suwito Pomalingo[3]
[1,2,3]Informatics Study Program, Universitas Multimedia Nusantara, Tangerang, Indonesia
[1]yoga.aditiya@student.umn.ac.id, [2]marlinda.vasty@umn.ac.id, [3]suwito.pomalingo@umn.ac.id

Accepted 20 May 2025
Approved 28 May 2025

***Abstract*— Cirebon Batik is one of Indonesia's cultural heritages that has its own unique patterns and motifs, reflecting the cultural richness and history of its region of origin. This study aims to address the challenges in classifying the complex motifs of Cirebon Batik by implementing Convolutional Neural Network (CNN) and Saliency Map methods. The three main motifs used are Mega Mendung, Singa Barong, and Keratonan. The dataset was obtained from various online sources and processed using image augmentation techniques. CNN is used to recognize complex visual patterns, while Saliency Map highlights important areas in the image that influence the model's decision. The results show that the developed CNN model achieved an accuracy of 82%, precision of 83%, recall of 82%, and F1-score of 82%. The use of Saliency Map provides better interpretability and enhances the understanding of the classification process.**

Index Terms*— Batik Cirebon, classification, CNN, image recognition, Saliency Map.

I. Introduction

Cirebon Batik is a vital part of Indonesia's cultural her- itage, renowned for its intricate motifs like Mega Mendung, Singa Barong, Paksinaga Liman, Keratonan, and Pratan Keris. Preserving these unique designs has become increasingly im- portant amid globalization, which often dilutes local identities [1]. The digitization of Cirebon Batik motifs presents an opportunity for effective preservation and wider international recognition, though the complexity and diversity of these patterns present challenges in classification and recognition efforts [2]. Manual classification is inefficient and requires expertise, making it impractical for large-scale application [3]. However, the diversity and complexity of Cirebon Batik motifs, which are rich in detail and philosophy, pose their own challenges in classification and recognition efforts. There- fore, an innovative approach is needed to ensure that the preservation of these batik motifs can be done accurately and efficiently, which will ultimately support the sustainability of this culture in the modern era.

Artificial intelligence (AI), particularly deep learning, has shown promise in addressing these challenges. Convolutional Neural Networks (CNN) can extract complex visual patterns and are well-suited for batik motif classification due to their ability to handle intricate designs. Prior studies, such as those by Wulandari et al. and Azzalini et al., demonstrate CNN's effectiveness in textile and motif recognition, suggesting its relevance for Cirebon Batik [4][5]. The addition of Saliency Maps, which highlight important image areas, provides further insights into model decisions, improving interpretability and aiding in motif preservation efforts [6][8].

This research explores the use of CNN combined with Saliency Maps to enhance the classification accuracy of Cire- bon Batik motifs while promoting cultural preservation. By increasing transparency and interpretability, this approach can support the development of a more informative database of batik patterns and facilitate their digital documentation. The integration of advanced AI technologies aims to contribute not only to cultural preservation but also to advancing pattern recognition methods in the context of Indonesia's rich artistic heritage [9]–[11].

II. Methodology

Figure 1 presents the entire process of program development from start to finish for the Cirebon batik classification task. It begins with the preparation of the Cirebon batik dataset, which is split into training and validation data. The first stage involves training a Convolutional Neural Network (CNN) with three layers using the training data, followed by validation to evaluate the model's performance. The saliency map is then generated based on the trained model to highlight key features that influence the classification. After obtaining the saliency maps, they are applied to the Cirebon batik dataset, creating a new version of the dataset layered with additional visual insights. This enhanced dataset is again split into training and validation sets and reintroduced into the CNN model for further training and validation. The final step is an evaluation report that summarizes the performance and accuracy of the model after this entire process. This framework ensures that the model not only improves in accuracy but also enhances interpretability by incorporating saliency maps into the training workflow.

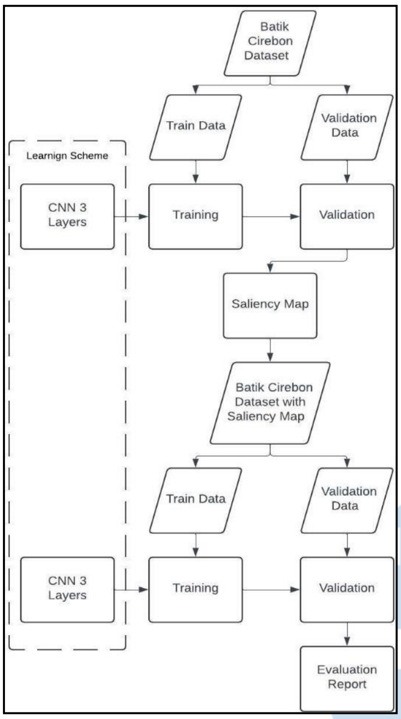

Fig. 1. Framework Full Program

A. Data Gathering

In this research, the dataset used is images of Cirebon batik which consists of three batik motifs. This dataset is obtained from various sources on the internet. The three batik motifs are Mega Mendung motif, Singa Barong motif, Keratonan motif.

B. Pre-processing

At this stage, data will be uploaded to the system that is grouped based on batik motifs. This data will be added based on the augmentation results using "ImageDataGenerator". With "imageDataGenerator" the uploaded data will be rescalled to change the pixel value from the 0-255 range to the 0-1 range. Figure 2 shows the flow of pre-processing.

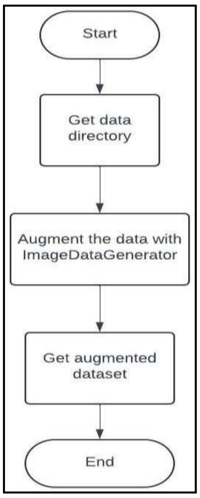

Fig. 2. Flowchart pre-processing

C. Split Dataset

At this stage, we used a split of 75% for training and 25% for validation. A separate test set was not used. We have clarified the revised manuscript under the split dataset detection. Figure 3 shows the flow of the split dataset.

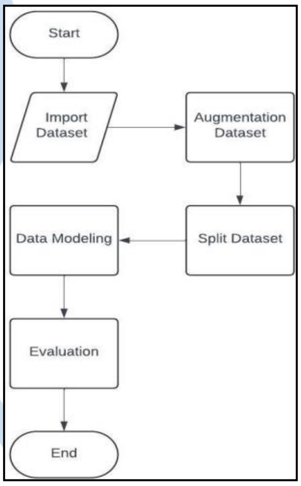

Fig. 3. Flowchart split dataset

D. Model Development

Convolutional Neural Network (CNN) is a type of artificial neural network designed to process grid-like data, such as images. CNN is claimed to be the best model for solving problems in object recognition [12], [13]. CNN is developed based on the Multilayer Perceptron (MLP) which is designed to process two-dimensional data. CNNs consist of several types of main layers, namely Convolutional layer, the Pooling layer, and the Fully Connected layer. [14][18]

At this stage, we will build the model gradually from layer to layer. Three layers are built using ReLu activation with 32, 64, and 128 filters each with a kernel size of 3x3. From each convolution layer, the 'max pooling' feature will be used to reduce the number of

ISSN 2085-4552

parameters generated. After three layer runs, a two-dimensional matrix is converted into a one-dimensional vector that will be input into the dense layer. Before entering the dense layer, the dropout layer is added with a dropout rate of 20% (rate = 0.2). This dropout serves as one of the effective regularization methods to prevent overfitting. In the dropout process, a random number of neurons will be disabled during the training process, so that the model is less dependent on certain neurons and is able to produce a more generalized model. The dropout is applied before the first dense layer, right after the flattening process, to maximize the regularization effect on the dense layers. There are two dense layers used, the first layer uses 64 neurons with ReLu activation. The second layer uses three neurons corresponding to the number of classes to be classified with Softmax activation. Figure 4 shows the flow of model building.

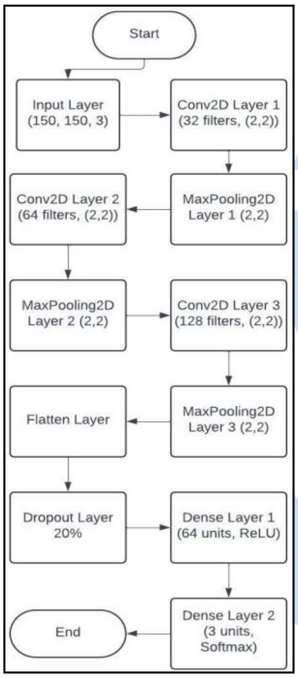

Fig. 4. Flowchart model

E. Model Fit and Callbacks

At this stage, the model that has been built will be tested based on the validation data against the training data. The initial repetition is determined as 30 repetitions with 2 call- backs provided. The 'red_lr' callback is used to reduce the learning rate of the model by three times. my_callback is used to prevent overfit by stopping the repetition when it has reached an accuracy value of 90%. Figure 5 shows the flow of the model and callbacks.

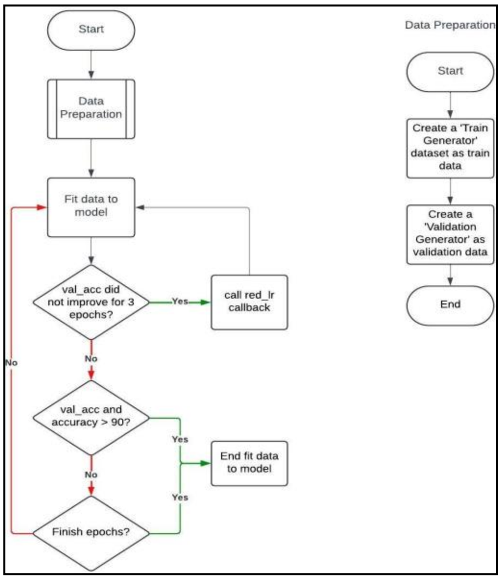

Fig. 5. Flowchart model fit and callbacks

F. Saliency Map

Saliency map is an important concept in pattern recognition and computer vision used to identify the most significant areas in an image that influence the decision of a model. The basic theory of saliency maps is rooted in how the human visual system works, where our brain automatically focuses on parts that stand out or attract attention in a scene. In the context of machine learning and artificial neural networks, saliency maps serve to provide a visual interpretation of the model's decision- making process, by highlighting the pixels in the image that most influence the classification or detection results. Some of the main methods in saliency map generation are gradient, Class Activation Maps (CAM), and Grad-CAM.

At this stage, parts of the batik that most affect the results of the model prediction will be displayed. The part will be displayed with a different color placed above the original image. Figure 6 shows the flow of the saliency map.

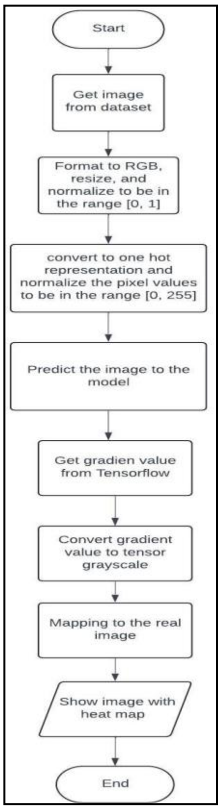

Fig. 6. Flowchart saliency map

G. Modeling with Saliency Map Data

The Cirebon batik dataset, enhanced with saliency maps, will be reused as a new dataset to be fed into the pre- viously constructed model. This dataset, now layered with visual insights from the saliency maps, will undergo the same 75:25 split, where 75% is allocated for training and 25% for validation. The training process will follow the steps outlined in Section D, using the same CNN architecture with convo- lutional, pooling, and fully connected layers. The model will be trained similarly as described in Section E. This approach aims to refine the model's ability to recognize key features of the batik motifs, leveraging the saliency maps to focus on the most relevant patterns, and ensuring both improved accuracy and interpretability in the classification results.

H. Evaluation

Evaluation is done to determine the success rate of the model that has been built. The model will be tested by comparing the training data with the testing data based on several parameters, namely, accuracy, precision, recall, and F1 value calculated based on the confusion matrix. At this stage, an evaluation is carried out using a confusion the actual class. In addition, the accuracy, precision, recall and F1 Score values. Figure 7 also shown flow of the evaluation result matrix to determine the class distribution of the classification.

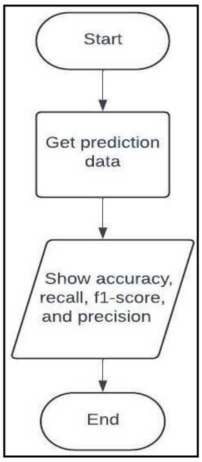

Fig. 7. Flowchart evaluation

III. RESULT AND DISCUSSION

A. Training Model

The CNN model was trained using Cirebon Batik motif data stored in a specific directory. First, the data was generated and resized to 150x150 pixels. The training data is taken from the training data subset, while the validation data is taken from the validation subset. After the data generator was created, the model was trained using the fit function for 100 epochs. Table I shows the last 10 repetitions performed in model in the last repetition is 91.21% and the validation accuracy of the model in the last repetition is 84.51%. training. It can be seen that the training accuracy of the model.

TABLE I
CNN MODEL TRAINING RESULTS FOR EPOCHS 90 TO 100

Epoch	Accuracy	Loss	Val Accuracy	Val Loss
90	0.9680	0.1821	0.8592	0.3755
91	0.9842	0.1397	0.8873	0.3509
92	0.9714	0.1206	0.8732	0.3546
93	0.9757	0.1286	0.8521	0.3903
94	0.9657	0.1695	0.8732	0.3917
95	0.9768	0.1465	0.8310	0.3947
96	0.9479	0.1694	0.7887	0.5655
97	0.9396	0.1933	0.8310	0.3618
98	0.9692	0.1180	0.8521	0.4575
99	0.8961	0.2813	0.8099	0.5629
100	0.9121	0.2514	0.8451	0.4232

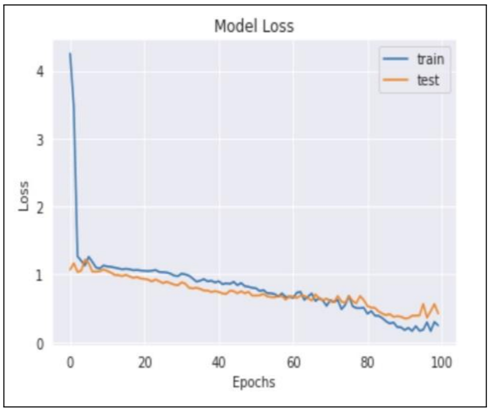

Fig. 8 Loss Graph

In Figure 8, a graph of the loss value of the trained model is shown. The value of loss in the training data has decreased significantly from each epoch, starting from 7 and continuing to decrease until it reaches about 0.02. Meanwhile, the loss value on the validation data also shows a consistent decrease, starting from around 1 until it stabilizes at 0.68. This graph indicates that the model is getting better at minimizing prediction errors in both training and validation data as the number of epochs increases.

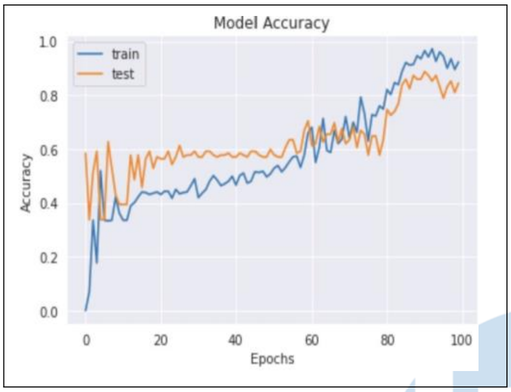

Fig. 9. Accuracy graph

In Figure 9, a graph of the model accuracy value during the training process is shown. The model accuracy on the training data shows a consistent increase as epochs increase, starting from 0.2 until it reaches around 0.98, indicating that the model is able to learn well on the training data. Accuracy on validation data also increased gradually, starting from around 0.4 until it stabilized at around 0.8. However, fluctuations were seen at some points, indicating that the model had some difficulty in maintaining a stable accuracy on the validation data. Overall, however, the model performed well on both training and validation data.

B. Prediction Result

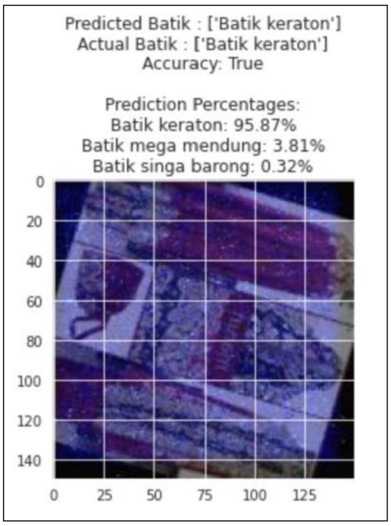

Fig. 10. Prediction result of model

Figure 10 shows an example of predicted data based on the trained model. The model's prediction is "Batik keraton", which also matches the actual data, indicating that the prediction is accurate. The accuracy of the prediction is supported by the model's confidence percentage of the classification, where "Batik keraton" has the highest probability of 85.87%. In addition, the model also gives a small percentage probability for two other motifs, namely "Batik mega mendung" at 3.81% and "Batik singa barong" at 0.32%

C. Saliency Result

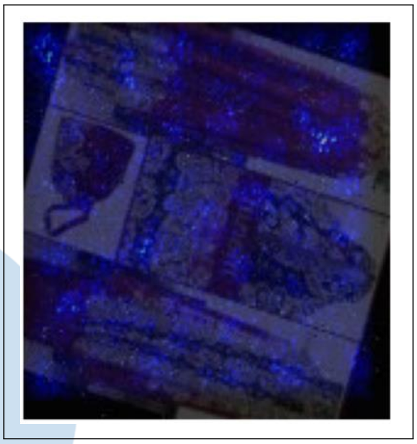

Fig. 11. Saliency result

Figure 11 shows the visualization of batik that has been highlighted by the Saliency Map so that it can be seen which part of the image has the most influence on the model. The part covered in blue indicates the part that is considered important by the model, the brighter the blue color in a part of the image indicates that the part is more important to the model.

D. Evaluation Result

Table II shows an accuracy value of 87%, a weighted precision value of 89%, a weighted recall value of 88%, and a weighted f1-score of 87%.

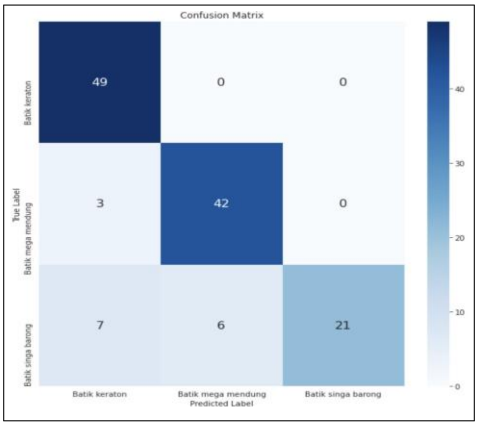

Fig. 12. Confusion matrix

TABLE II

EVALUATION RESULT

Metric	Value
Accuracy	0.87
Precision	0.89
Recall	0.88
F1-Score	0.87

Figure 12 shows the result of the prediction distribution performed by the model. The True Labels (y-axis) represents the original categories of the Batik motifs. The Predicted Labels (x-axis) represents the predicted categories from the model. In True Label batik Keraton, there are 49 samples that correctly classified as "Batik Keraton" and there are no samples that misclassified. In True Label batik Mega Mendung there are 42 samples that correctly classified as "Batik Mega Mendung" and there are 3 samples that misclassified as "Batik Keraton". In True Label batik Singa Barong, there are 21 samples that correctly classified as "Batik Singa Barong" and there are 7 samples that misclassified as "Batik Keraton" with 6 samples that missclassified as "Batik Mega Mendung".

IV. CONCLUSIONS

Convolutional Neural Network (CNN) and Saliency Map algorithms have been successfully applied in the classification of Cirebon Batik motifs with satisfactory results. The build model shows a good performance with 87% accuracy, 89% precision, 88% recall, and 87% f1-score, reflecting the balance between accuracy and the model's ability to detect the correct motif. In addition, the Saliency Map successfully highlights the parts of the image that have the most influence on the classification decision, providing deeper insight into how the model recognizes and distinguishes batik motifs.

ACKNOWLEDGMENT

The authors would like to thank to Universitas Multimedia Nusantara for funding supporting this research, cultural actors in Batik Pesisir Cirebon in Cirebon Batik UKM, and Universitas Muhammadiyah Cirebon as partner in this research.

REFERENCES

[1]. R. Susanti and D. Purnomo, "Batik sebagai warisan budaya tak benda: Tantangan dan peluang pelestarian di era digital," Jurnal Kebudayaan Indonesia, vol. 12, no. 2, pp. 55–67, 2019.

[2]. Rahman and H. Wicaksono, "Digitalisasi budaya: Strategi pelestarian motif batik melalui teknologi informasi," Jurnal Teknologi dan Infor- masi, vol. 18, no. 1, pp. 43–59, 2020.

[3]. Kompas, "Batik cirebon: Mempertahankan tradisi di tengah modernisasi," 2022, https://www.kompas.com/batik-cirebon- mempertahankan-tradisi.

[4]. R. Wulandari et al., "Batik image classification using convolutional neural network," International Journal of Electrical and Computer Engineering, vol. 10, no. 1, pp. 217–225, 2020.

[5]. M. Azzalini et al., "Textile pattern recognition using convolutional neural networks: A case study on traditional textiles," Journal of Computational Design and Engineering, vol. 8, no. 2, pp. 207–215, 2021.

[6]. K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps," arXiv preprint arXiv:1312.6034, 2013. [Online]. Available: https://arxiv.org/abs/1312.6034

[7]. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolu- tional networks," in European Conference on Computer Vision (ECCV), 2014, retrieved from ECCV.

[8]. Kompas.com, "Digitalisasi batik cirebon dorong pelestarian budaya dan industri kreatif," 2023, https://www.kompas.com.

[9]. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, retrieved from CVPR.

[10]. Sutaryo, "Batikan: Pelestarian batik cirebon dengan teknologi mod- ern," 2023, retrieved from https://www.sutaryo.co.id/batikan-pelestarian- batik-cirebon.

[11]. M. Hadi, "Penerapan teknologi digital dalam pelestarian budaya batik," Jurnal Kebudayaan Indonesia, vol. 17, no. 2, pp. 145–160, 2022.

[12]. Y. Gultom, A. M. Arymurthy, and R. J. Masikome, "Batik classifica- tion using deep convolutional network transfer learning," Jurnal Ilmu Komputer Dan Informasi, vol. 11, no. 2, p. 59, 2018.

[13]. S. F. Tumewu, D. H. Setiabud, and I. Sugiarto, "Klasifikasi motif batik menggunakan metode deep convolutional neural network dengan data augmentation," Jurnal Infra, vol. 8, no. 2, pp. 189–194, 2020.

[14]. R. Refianti, A. B. Mutiara, and R. P. Priyandini, "Classification of melanoma skin cancer using convolutional neural network," Interna- tional Journal of Advanced Computer Science and Applications, vol. 10, no. 3, pp. 409–417, 2019.

[15]. M. Alexander, "Sistem pengenalan wajah untuk pre- sensi dengan algoritma convolutional neural network," https://kc.umn.ac.id/id/eprint/13923, 2020, accessed: 2024-08-09.

[16]. T. A. Hartanto and S. a. Hansun, "Comparative analysis of pre-trained cnn models on retinal diseases classification," international journal of industrial Engineering Production Research, vol. 35, no. 3, 2024. [Online]. Available: http://ijiepr.iust.ac.ir/article-1-1935-en.html

[17]. C. Young, "Applicability of various pre-trained deep convolutional neural networks for pneumonia classification based on x-ray images," International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, 2020.

[18]. Archilles, A., & Wicaksana, A. (2020). Vision: a web service for face recognition using convolutional network. TELKOMNIKA (Telecommunication Computing Electronics and Control), 18(3), 1389-1.