INKUBIS: Jurnal Ekonomi dan Bisnis
Volume 8, Issue 4, 3507-3518
e_ISSN: 2775-3913
https://inkubis.politeksci.ac.id/index.php/ink/index
DOI: doi.org/10.59261/inkubis.v8i4.533

Introduction of Songket Motifs from Four Regions Using Inception V3 Image Embedding and Machine Learning in Orange Data Mining

Maya Fitriana Primasari*
Universitas Airlangga, Indonesia

Imam Yuadi
Universitas Airlangga, Indonesia

***Corresponding author:**
Maya Fitriana Primasari,
Universitas Airlangga, Indonesia.
✉maya.fitriana.primasari-2025@pasca.unair.ac.id

Article Info:

Article history:
Received: September 09, 2026
Revised: October 16, 2026
Accepted: October 27, 2026
Available Online: November 12, 2026

Keywords:
Image Embedding;
Inception V3;
Machine Learning;
Motif Classification;
Songket

This is an open access article under the CC BY-SA license.

Copyright © 2026 by Author. Published by Politeknik Siber Cerdika International

Abstract

Background: Manual identification of *songket* motifs is time-consuming and requires specialized expertise, limiting the cataloging and documentation of regional textiles.
Objective: To develop a *songket* motif classification system using image embeddings extracted with Inception v3 and machine learning algorithms implemented in Orange Data Mining.
Methods: The research dataset comprised 200 images of *songket* fabrics representing four regional motif categories: Balinese *songket*, Lombok *songket*, Palembang *songket*, and Riau *songket*, with 50 images per category. Image embeddings were extracted from all images using Inception v3, and the resulting embeddings were used as inputs for several classification models. The classification algorithms evaluated in this study were Logistic Regression, Support Vector Machine (SVM), Random Forest, and Neural Network.
Results: Logistic Regression achieved the best overall classification performance, with an accuracy of 75%, an area under the receiver operating characteristic curve (AUC) of 0.905, and a Matthews correlation coefficient (MCC) of 0.667. Confusion matrix analysis further revealed that Palembang and Riau *songket* motifs had per-class recall values of 88% and 84%, respectively, whereas Balinese and Lombok *songket* motifs had the lowest per-class recall values, at 62% and 66%, respectively. High visual similarity between Balinese and Lombok motifs may explain this difference in classification performance.
Conclusion: Findings demonstrate potential of combining Inception v3-based image embeddings with machine learning algorithms for classifying traditional textile motifs. They also highlight the potential of implementing this approach in Orange Data Mining to support the digitization and cataloging of traditional textiles, thereby contributing to the preservation of Indonesia's textile cultural heritage.

To cite this article: Primasari, M. F., & Yuadi, I. (2026). Introduction of songket motifs from four regions using Inception V3 image embedding and machine learning in Orange Data Mining. *INKUBIS: Jurnal Ekonomi dan Bisnis, 8*(4), 3507–3518. https://doi.org/10.59261/inkubis.v8i4.533

INTRODUCTION

Indonesian traditional textiles are among the country's cultural assets, possessing high aesthetic and philosophical value. One textile that exemplifies this richness is *songket*, a traditional Nusantara textile produced using a supplementary weft-weaving technique. This technique involves weaving the ground fabric while manually incorporating additional gold, silver, or brightly colored threads to create raised, shimmering motifs. Each region has distinct *songket* motifs characterized by variations in color, geometric patterns, floral ornaments, and the distinctive use of metallic threads. These differences reflect regional cultural identities

and constitute important elements of the creative industries and local tourism. Empirical evidence concerning *songket* weaving highlights threats to its preservation. A community service study of Balinese *songket* reports high material costs, a buying and selling system that leaves weavers with low incomes, declining interest among younger generations, and reduced demand for handmade textiles as inexpensive machine-made alternatives become more widespread (Bongon et al., 2024). The same study identifies the documentation and dissemination of best practices as essential to preservation and notes that digital museums and online platforms can help showcase *songket* to local and global audiences (Bongon et al., 2024).

This study is also relevant to the economics and business of the creative industries. The commercial value of *songket* depends partly on its authenticity and regional provenance, while weavers' incomes are constrained by a monopolized buying and selling system and limited distribution channels (Bongon et al., 2024). Reliable, automated motif identification can support provenance labeling, product cataloging, and digital marketing for regional textile products. In this way, the management of cultural heritage can be linked to the competitiveness of local creative enterprises.

Manual identification of *songket* motifs is time-consuming and requires specialized expertise. However, digital and automated approaches to *songket* motif identification remain limited (Anggoro et al., 2024; Ariessaputra et al., 2024). Advances in machine learning and computer vision, particularly image embedding and transfer learning, offer opportunities to address this limitation (Szegedy et al., 2016; Tan et al., 2018; Yu et al., 2022). These developments enable the automated recognition and classification of *songket* motifs, potentially improving the efficiency and consistency of the identification process.

The limitations of manual identification extend beyond the time required. Without a consistent automated method, motifs cannot be cataloged uniformly across collections, hindering the verification of regional provenance needed for product authentication in the creative industries and the interpretation of motifs in cultural documentation. Inconsistent labeling also makes it more difficult to transfer knowledge among weavers, curators, and designers and increases the risk that regional motif variations will be lost before they are systematically documented.

The *songket* textiles examined in this study originate from four regions: Bali, Lombok, Palembang, and Riau. The dataset was obtained from the Roboflow Universe platform (https://universe.roboflow.com/) (Roboflow, 2023). *Songket* motif recognition was performed using machine learning in Orange Data Mining, with Inception V3 serving as the convolutional neural network (CNN) backbone for image embedding. This approach is relevant because the model can represent the complex visual structure of textiles and extract discriminative features from intricate motifs, including weaving patterns, thread textures, and traditional ornaments, which may be difficult to distinguish through manual inspection alone (Ariessaputra et al., 2024; Szegedy et al., 2016). Through the Image Embedding widget in Orange Data Mining, each image is transformed into a numerical feature vector that can subsequently be processed by various machine learning algorithms (Kornblith et al., 2019; Nayman et al., 2024). These feature representations may support the development of classification models capable of recognizing visual patterns despite variations in lighting conditions and camera angles, although robustness to such variations must be established empirically (Shorten & Khoshgoftaar, 2019; Zeng, 2024).

The purpose of this study is to classify *songket* motifs from four distinct regions—Bali, Lombok, Palembang, and Riau—using a deep learning-based feature extraction approach in which Inception V3 generates image embeddings that are subsequently processed by machine learning algorithms in Orange Data Mining. In addition, this study evaluates the performance of each classification model using established evaluation metrics and examines prediction errors through confusion matrices. These analyses are intended to characterize the models' abilities to recognize complex visual patterns and distinguish among regional motif classes.

This study addresses three research questions. First, how effectively can machine learning algorithms classify *songket*motifs using feature vectors generated through Inception V3 image embedding, as measured by accuracy, precision, recall, and F1-score? Second, which machine learning model achieves the best classification performance in recognizing the distinctive and complex visual patterns of *songket* motifs? Third, what error patterns emerge in *songket* motif classification, as revealed by confusion matrices Chicco and Jurman (2020) and Szabó et al.

(2024). In particular, which motifs are frequently misclassified as motifs from other regions, and which visual characteristics may contribute to these errors? By addressing these questions, this study aims to clarify the effectiveness of combining image embedding with machine learning in Orange Data Mining to recognize and classify *songket* motifs from Bali, Lombok, Palembang, and Riau.

Overall, this study has several objectives. First, it aims to identify and analyze visual patterns in *songket* textiles using image embeddings generated by the Inception V3 CNN. This approach converts image data into numerical feature vectors, facilitating subsequent computational analysis and classification. Second, it seeks to evaluate several machine learning algorithms for classifying *songket* textiles according to their motifs. Third, it aims to interpret the resulting confusion matrices to determine the sources of misclassification, identify frequently confused motif classes, and examine the visual characteristics that may contribute to classification errors. Finally, the study seeks to compare the strengths and limitations of the evaluated models and assess their suitability for the classification task.

The selection of Inception V3 is motivated by the visual characteristics of *songket* textiles. Their motifs combine fine, thread-level textures with larger geometric and floral elements, requiring a feature extractor capable of representing patterns at multiple spatial scales. Inception V3 incorporates factorized convolutions and multi-branch modules designed to capture features at different scales (Das et al., 2023; Szegedy et al., 2016). In this study, the pretrained network is used as a fixed feature extractor rather than being fine-tuned on the target dataset. This approach is practical for a dataset of 200 images because training a deep neural network from scratch on such a small collection would carry a substantial risk of overfitting (Hill et al., 2025; Raghu et al., 2019; Sun et al., 2017). Alternative architectures offer different trade-offs, including residual networks (He et al., 2016; Xu et al., 2025), densely connected networks (Duggirala & Kollipara, 2025; Huang et al., 2017), lightweight MobileNetV2 models (Miró et al., 2025; Sandler et al., 2018), and Transformer-based models (Bao et al., 2024; Dosovitskiy et al., 2021). Vision Transformers generally benefit from large-scale pretraining, particularly when compared with convolutional neural networks under limited-data conditions. Different backbone architectures also produce feature embeddings with different dimensionalities. Inception V3 produces a 2,048-dimensional feature vector when the output of its standard global average pooling layer is used, and this representation is supported by the Image Embedding workflow in Orange Data Mining. It was therefore selected as a practical balance between representational capacity and implementation feasibility for this dataset. A systematic comparison of these backbone architectures using the same *songket* dataset is left for future research.

The novelty of this study lies in combining fixed, ImageNet-pretrained Inception V3 image embeddings with four classical and neural machine learning classifiers and applying this workflow to a regional *songket* dataset representing four distinct motif traditions. Previous CNN-based studies have classified *batik* patterns from Solo Anggoro et al. (2024) and *songket* textiles from Lombok featuring *batik Sasambo* motifs (Ariessaputra et al., 2024). However, these studies focus on individual regional motif collections rather than comparing *songket* traditions across four regions. The present study makes three contributions. First, it compares four classifiers using the same feature embeddings in a workflow that can be reproduced in Orange Data Mining without programming. Second, it evaluates each model using class-level confusion matrices to identify sources of classification errors among regional motif classes. Third, it relates the classification results to the visual characteristics of the motifs, providing a basis for future dataset expansion and the digital documentation of *songket* heritage.

METHOD

The dataset used in this study consisted of 200 images of *songket* textiles divided into four categories according to the regions of origin of their motifs: Balinese *songket*, Lombok *songket*, Palembang *songket*, and Riau *songket*. Each category contained 50 images. According to Fernández et al. (2018) and Johnson and Khoshgoftaar (2019) and Gao et al. (2026), maintaining an equal number of images in each category can help reduce classification bias and ensure that machine learning models have equal opportunities to learn the visual patterns of each *songket* category. Buda et al. (2018) and Sharma and Gosain (2025) also reported that class imbalance in image classification can affect model performance, particularly in complex visual

pattern recognition tasks. The dataset contained representative motifs from each region and reflected the diversity of patterns, colors, and textures characteristic of Nusantara *songket*. The images were obtained from Roboflow Universe (https://universe.roboflow.com/).

The dataset used in this study comprised 200 *songket* images obtained from Roboflow Universe, representing four classes: Balinese, Lombok, Palembang, and Riau *songket*, with 50 images per class. The dataset was used to evaluate the classification performance of Inception V3 image embeddings combined with four machine learning classifiers in Orange Data Mining.

The dataset was obtained from a public repository. The available documentation did not provide sufficient information to establish the specific dataset version, licensing conditions, image acquisition conditions, or procedures used for label verification and image quality screening. These details should be documented where the relevant information is available. The absence of documented validation procedures limited the extent to which the reliability of the original image labels could be assessed.

During the image embedding stage, the Orange Data Mining application automatically resized the images according to the input requirements of the Inception V3 convolutional neural network (CNN). This preprocessing step ensured consistent input dimensions and facilitated the extraction of visual features, including geometric motifs, thread textures, and woven structures. The resulting feature representations were then used to evaluate the performance of several machine learning algorithms in classifying *songket* motifs.

Within the Image Embedding widget, each image was resized to the Inception V3 input resolution of 299 × 299 pixels, and its pixel values were normalized according to the preprocessing convention of the ImageNet-pretrained model before being passed through the network. The 2,048-dimensional activations from the final pooling layer were used as image embeddings. No additional data augmentation, cropping, or color correction was applied. Consequently, the classifiers were trained using the feature representations extracted from the images in their collected form. Variations in lighting and camera angle were not experimentally controlled, and their potential effects on the embeddings were considered limitations of the study.

The system workflow consisted of four stages: image import, image embedding, machine learning classification, and evaluation. In the first stage, the songket motif images were imported into the Orange Data Mining environment for further processing. Images of Balinese, Lombok, Palembang, and Riau *songket* were loaded into the system, with their corresponding class labels retained for subsequent analysis.

Before image embedding, the input data were displayed in a data table to verify that the images and their associated labels had been imported correctly. Next, each image was processed using the Inception V3 CNN to generate a high-dimensional feature representation. These feature representations served as inputs to the classification algorithms in the subsequent stage (Ariessaputra et al., 2024; Kornblith et al., 2019). Four machine learning classifiers were then applied: Logistic Regression, Support Vector Machine (SVM), Random Forest, and Neural Network, implemented as a multilayer perceptron (Szabó et al., 2024; Tharwat, 2021). These classifiers were compared to determine which achieved the best performance in recognizing complex visual patterns in traditional textiles. In the evaluation stage, model performance was assessed using accuracy, precision, recall, F1-score, and confusion matrices (Szabó et al., 2024; Tharwat, 2021). The results were compared to evaluate the relative performance of the algorithms and identify patterns of misclassification among the *songket* categories (Chicco & Jurman, 2020; Tamura et al., 2026). Figure X presents the workflow used to classify the *songket* motifs.

The four classifiers were selected to represent different inductive biases when applied to the same fixed image embeddings. Logistic Regression, a linear model, was used to assess whether the embedding space was sufficiently linearly separable for effective classification. SVM was used to identify margin-based decision boundaries, Random Forest was used to capture nonlinear relationships through an ensemble of decision trees, and the Neural Network, implemented as a multilayer perceptron, was used to model more flexible nonlinear decision boundaries. The selection was guided by the characteristics of the 2,048-dimensional embeddings, which were high-dimensional relative to the 200 available samples, and by established practices in classification benchmarking (Szabó et al., 2024; Tharwat, 2021). These classifiers were also available as standard supervised learning widgets in Orange Data Mining, making the comparison

relevant to users conducting machine learning analyses without programming.

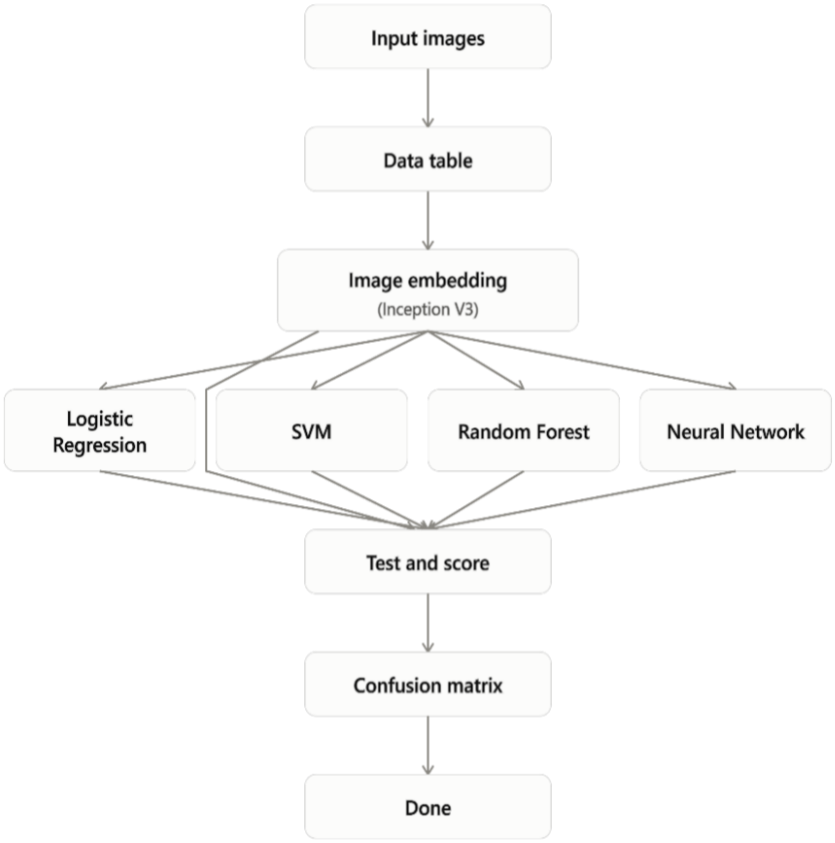


Figure 1. *Songket* Classification Process

RESULTS AND DISCUSSION

Results

This study used machine learning in Orange Data Mining with four classifiers: Logistic Regression, Support Vector Machine (SVM), Random Forest, and Neural Network. These classifiers used image embeddings extracted with Inception-v3. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC), classification accuracy (CA), F1-score, precision, recall, and Matthews correlation coefficient (MCC). The results are presented below.

Table 1
Orange Machine Learning Results

Model	AUC	CA	F1	Prec	Recall	MCC
SVM	0.895	0.675	0.678	0.686	0.675	0.568
Logistic Regression	0.905	0.750	0.750	0.752	0.750	0.667
Random Forest	0.789	0.550	0.552	0.557	0.550	0.401
Neural Network	0.903	0.725	0.726	0.728	0.725	0.633

Table 1 shows that Logistic Regression achieved the best overall classification performance, with an AUC of 0.905, classification accuracy of 0.750, F1-score of 0.750, precision of 0.752, recall of 0.750, and MCC of 0.667. The Neural Network ranked second, with an AUC of 0.903, accuracy of 0.725, F1-score of 0.726, precision of 0.728, recall of 0.725, and MCC of 0.633. SVM achieved an AUC of 0.895, accuracy of 0.675, F1-score of 0.678, precision of 0.686, recall of 0.675, and MCC of 0.568. Random Forest showed the lowest overall accuracy, at 0.550, with an AUC of 0.789, F1-score of 0.552, precision of 0.557, recall of 0.550, and MCC of 0.401.

The superior performance of Logistic Regression may be explained by the properties of the embeddings and the relatively small dataset. The features were extracted using Inception-v3, a model pretrained on ImageNet, whose final global average pooling layer produces a 2,048-

dimensional representation (Qian et al., 2025; Raghu et al., 2019). If the classes are sufficiently separable in this feature space, a linear decision boundary may be adequate. Logistic Regression can estimate a relatively small number of parameters compared with a more complex neural network, potentially making it suitable for a small dataset. The Neural Network includes a hidden layer whose parameters must be estimated from the same 200 images, increasing the risk of overfitting and potentially explaining its slightly lower performance. Random Forest partitions the feature space using axis-aligned splits. With only 200 samples, the individual trees may be less stable, which may contribute to its lower accuracy. Logistic Regression also demonstrated balanced class-level performance (Table 6), with recall values ranging from 62% to 88%, whereas Random Forest recall ranged from 44% to 66%.

Table 2
SVM Confusion Matrix Results

Actual\ Predicted	Songket Bali	Songket Lombok	Songket Palembang	Songket Riau	Σ
SongketBali	25	12	9	4	50
SongketLombok	19	29	2	0	50
SongketPalembang	9	0	41	0	50
SongketRiau	2	2	6	40	50
Σ	55	43	58	44	200

Table 2 presents the confusion matrix for the Support Vector Machine (SVM) classifier, showing the distribution of predictions across the four *songket* motif categories. This distribution provides insight into the model's ability to distinguish the regional textile categories based on their image embeddings.

The SVM model correctly classified 25 of the 50 Balinese *songket* images. The remaining images were misclassified: 12 were predicted as Lombok *songket*, nine as Palembang *songket*, and four as Riau *songket*. These results indicate that the model had difficulty distinguishing Balinese *songket* from the other categories, particularly Lombok and Palembang *songket*. However, the confusion matrix alone cannot establish whether specific visual similarities, such as fine textures or repetitive geometric patterns, caused these errors. Examining the misclassified images would help determine whether such similarities contributed to the observed pattern.

For Lombok *songket*, SVM correctly classified 29 of the 50 images, while 19 were incorrectly predicted as Balinese *songket* and two as Palembang *songket*. No images were misclassified as Riau *songket*. The frequent confusion between the Balinese and Lombok classes suggests that their image embeddings were not sufficiently distinct for reliable classification in all cases. Further visual analysis would be needed to identify the characteristics associated with these errors.

For Palembang *songket*, SVM correctly classified 41 of the 50 images. The remaining nine were misclassified as Balinese *songket*. This result indicates that the model distinguished Palembang *songket* relatively well, although some images were assigned to another category. The confusion matrix does not reveal which specific visual features contributed to these errors.

For Riau *songket*, SVM correctly classified 40 of the 50 images. Of the remaining images, two were predicted as Balinese *songket*, two as Lombok *songket*, and six as Palembang *songket*. The most frequent error involved the Palembang class. This pattern suggests that the model had some difficulty distinguishing Riau from Palembang *songket*, although the specific visual factors responsible cannot be determined from the confusion matrix alone.

Overall, the confusion matrix indicates that SVM classified Palembang and Riau *songket* more accurately than Balinese and Lombok *songket*. The reciprocal misclassification between the Balinese and Lombok classes was particularly prominent. These results suggest that a linear decision boundary in the embedding space may not fully capture the distinctions between all four categories. Future research could evaluate alternative classifiers, data augmentation, or fine-tuning of the feature extractor to determine whether these approaches improve classification performance.

Table 3
Logistic Regression Confusion Matrix Results

Actual \ Predicted	Songket Bali	Songket Lombok	Songket Palembang	Songket Riau	Σ
SongketBali	31	10	5	4	50
SongketLombok	15	33	2	0	50
SongketPalembang	3	3	44	0	50
SongketRiau	2	1	5	42	50
Σ	51	47	56	46	200

Table 3 shows the classification performance of Logistic Regression for each *songket* motif category. This analysis is relevant to regional motif recognition because the four categories may exhibit overlapping visual characteristics that complicate automated classification.

Palembang *songket* achieved the highest class-level recall, with 44 of 50 images correctly classified. This result indicates that the model distinguished this category relatively well. However, the confusion matrix does not establish which specific visual characteristics account for this performance. Five images were misclassified as Balinese *songket*, and three were misclassified as Lombok *songket*. These errors suggest that some images in the Palembang category were not fully distinguishable from other categories in the embedding space.

Riau *songket* also showed strong classification performance, with 42 of 50 images correctly classified. Of the remaining images, five were predicted as Palembang *songket*, two as Balinese *songket*, and one as Lombok *songket*. These results indicate that Logistic Regression generally distinguished Riau *songket* from the other categories, although some confusion remained, particularly with Palembang *songket*. Further examination of the misclassified images would be needed to determine whether particular visual characteristics contributed to these errors.

The confusion matrix indicates that Balinese and Lombok *songket* were more frequently misclassified as each other than were the other class pairs. Specifically, 10 Balinese *songket* images were predicted as Lombok *songket*, while 15 Lombok *songket* images were predicted as Balinese *songket*. This reciprocal misclassification suggests that the Inception-v3 embeddings did not fully distinguish the visual representations of these two classes. Although similarities in motif composition, texture, or repetitive patterns may contribute to this result, the confusion matrix alone cannot establish which specific visual characteristics account for the classification errors. Further visual analysis is therefore needed to identify features that may distinguish Balinese and Lombok *songket* more reliably.

The remaining errors involving Palembang and Riau *songket* indicate that some images from these categories were also assigned to other classes. The confusion matrix documents these classification errors but does not establish whether they resulted from similarities in ornamental patterns, texture, or other visual characteristics. Overall, the results show that Logistic Regression can classify *songket* motifs using Inception-v3 image embeddings, although its performance varies across the four regional categories. The results do not establish whether the model can reliably distinguish specific textile techniques, textures, or repetitive patterns, particularly between Balinese and Lombok *songket*.

These findings suggest that further research should investigate the visual features associated with classification errors. Possible approaches include fine-tuning a convolutional neural network (CNN), applying data augmentation to increase the diversity of the training dataset, and evaluating attention-based architectures to determine whether they improve the representation of fine-grained visual details. Such investigations could support the development of machine learning-based classification systems and contribute to the digitization, documentation, and preservation of Indonesian textile heritage.

Table 4
Random Forest Confusion Matrix Results

Actual \ Predicted	Songket Bali	Songket Lombok	Songket Palembang	Songket Riau	Σ
SongketBali	24	13	6	7	50
SongketLombok	21	22	2	5	50
SongketPalembang	8	6	31	5	50
SongketRiau	6	4	7	33	50
Σ	59	45	46	50	200

Table 4 presents the confusion matrix for the Random Forest classifier, showing different classification performance across the four *songket* categories. The matrix indicates how effectively the model classified the regional motifs using features extracted by Inception-v3. Overall, the model produced varying classification results, with errors distributed across all four categories.

For Balinese *songket*, Random Forest correctly classified 24 of 50 images. The remaining 26 images were misclassified: 13 as Lombok *songket*, six as Palembang *songket*, and seven as Riau *songket*. The largest number of errors occurred between the Balinese and Lombok classes. These results indicate that the model had difficulty distinguishing Balinese *songket* from the other categories, particularly Lombok *songket*. However, the confusion matrix alone does not establish which specific visual characteristics contributed to these errors.

For Lombok *songket*, the model correctly classified 22 of 50 images. Of the remaining images, 21 were classified as Balinese *songket*, two as Palembang *songket*, and five as Riau *songket*. The most frequent error involved the Balinese class, indicating substantial classification confusion between these two categories. Further examination of the misclassified images would be needed to determine whether particular motif characteristics were associated with this pattern.

For Palembang *songket*, the model correctly classified 31 of 50 images. The remaining 19 images were misclassified: eight as Balinese *songket*, six as Lombok *songket*, and five as Riau *songket*. Palembang *songket* had more correct predictions than either Balinese or Lombok *songket*, although errors occurred across all three other categories. These results describe the model's classification performance but do not, by themselves, establish which visual characteristics account for the observed differences.

For Riau *songket*, the model correctly classified 33 of 50 images. The remaining 17 images were classified as Balinese *songket* (six), Lombok *songket* (four), and Palembang *songket* (seven). The largest number of errors involved Palembang *songket*. This pattern indicates that the model did not consistently distinguish Riau *songket* from the other categories, particularly Palembang *songket*. Image-level examination would be necessary to investigate whether specific visual similarities were associated with these misclassifications.

Overall, the Random Forest confusion matrix shows varying classification performance across the four *songket* categories. Riau *songket* had the highest number of correct predictions (33 of 50), followed by Palembang *songket* (31 of 50), Balinese *songket* (24 of 50), and Lombok *songket* (22 of 50). The most prominent reciprocal confusion occurred between the Balinese and Lombok classes. These findings identify the categories that were more difficult for the model to distinguish, but they do not establish the visual causes of the errors. Future work could examine misclassified images and evaluate alternative feature representations or data augmentation strategies to investigate potential improvements.

Table 5
Neural Network Confusion Matrix Results

Actual \ Predicted	Songket Bali	Songket Lombok	Songket Palembang	Songket Riau	Σ
SongketBali	29	12	4	5	50
SongketLombok	16	33	1	0	50
SongketPalembang	6	2	40	2	50
SongketRiau	1	3	3	43	50

Actual \ Predicted	Songket Bali	Songket Lombok	Songket Palembang	Songket Riau	Σ
Σ	52	50	48	50	200

The confusion matrix shows the classification performance of the Neural Network model across the four *songket* categories. The results indicate that the model classified some categories more accurately than others, with the greatest difficulties occurring between categories that were frequently misclassified as each other.

Discussion

The Neural Network correctly classified 29 of 50 images in the Balinese *songket* category. The remaining images were misclassified as follows: 12 as Lombok *songket*, four as Palembang *songket*, and five as Riau *songket*. The relatively high number of errors involving Lombok *songket* indicates that the model had difficulty distinguishing these two categories. However, the confusion matrix alone cannot establish whether similarities in repetitive patterns, basic textures, or other visual characteristics caused the errors.

The model correctly classified 33 of 50 Lombok *songket* images. Of the remaining images, 16 were predicted as Balinese *songket* and one as Palembang *songket*. No images were misclassified as Riau *songket*. These results indicate that the model distinguished Lombok *songket* from Riau *songket* in this dataset, while confusion with Balinese *songket* remained substantial. Further examination of the misclassified images would be needed to investigate the visual characteristics associated with this pattern.

For Palembang *songket*, the model correctly classified 40 of 50 images. Six images were predicted as Balinese *songket*, two as Lombok *songket*, and two as Riau *songket*. These results indicate that the model distinguished Palembang *songket* relatively well. However, the confusion matrix alone cannot establish whether specific ornamental patterns or other visual characteristics explain its performance.

The model achieved its highest number of correct predictions for Riau *songket*, with 43 of 50 images correctly classified. The remaining images were misclassified as Balinese *songket* (one), Lombok *songket* (three), and Palembang *songket*(three). This result indicates that Riau *songket* was the best-classified category for the Neural Network. Examining the misclassified images could help determine whether particular visual characteristics were associated with the remaining errors.

Overall, the Neural Network confusion matrix shows that Palembang and Riau *songket* were classified more accurately than Balinese and Lombok *songket*. The model correctly classified 40 of 50 Palembang images and 43 of 50 Riau images, compared with 29 of 50 Balinese images and 33 of 50 Lombok images. These results indicate that the main classification difficulties involved Balinese and Lombok *songket*. Further research could examine dataset expansion, image augmentation, and alternative network architectures to improve the distinction between visually similar classes.

Table 6

Per-class recall and precision of the four classifiers, calculated from the confusion matrices in Tables 2–5

Class	LR Recall	LR Precision	SVM Recall	SVM Precision	RF Recall	RF Precision	NN Recall	NN Precision
Balinese	62%	61%	50%	45%	48%	41%	58%	56%
Lombok	66%	70%	58%	67%	44%	49%	66%	66%
Palembang	88%	79%	82%	71%	62%	67%	80%	83%
Riau	84%	91%	80%	91%	66%	66%	86%	86%

Note: recall = correct predictions ÷ images of the class (row total); precision = correct predictions ÷ images predicted as the class (column total). LR, Logistic Regression; SVM, Support Vector Machine; RF, Random Forest; NN, Neural Network.

This study contributes to three areas. In computer vision, it shows that frozen, ImageNet-pretrained Inception-v3 embeddings combined with a simple linear classifier can distinguish four

regional textile classes with 75% accuracy, providing a reproducible baseline for comparison with fine-tuning and data augmentation approaches (Kornblith et al., 2019; Nayman et al., 2024). In digital heritage preservation, it demonstrates that a no-code workflow in Orange Data Mining can support the cataloging of regional textiles by documentation staff and cultural practitioners, addressing documentation needs identified for *songket* collections (Bongon et al., 2024). In traditional textile classification, the class-level confusion analysis identifies Balinese and Lombok *songket* as the most frequently confused categories, providing a concrete direction for dataset expansion and model refinement (Anggoro et al., 2024; Ariessaputra et al., 2024). These contributions are summarized in Table 7.

Table 7
Positioning of this study relative to CNN-based studies of Indonesian textile motifs

Study	Textile and motif classes	Approach	Contribution relative to this study
Anggoro et al. (2024)	Solo batik patterns	Deep learning CNN classification	Single batik tradition; no songket classes
Ariessaputra et al. (2024)	Lombok songket and Sasambo batik motifs	CNN classification	Lombok songket is the only songket class
This study	Balinese, Lombok, Palembang, and Riau songket (50 images per class)	Inception V3 embeddings with Logistic Regression, SVM, Random Forest, and Neural Network in Orange	Four classifiers on the same embedding with class-level confusion analysis

Across the four classifiers, Balinese and Lombok *songket* presented greater classification difficulties than Palembang and Riau *songket*, particularly because of reciprocal misclassification between the Balinese and Lombok classes. The confusion matrices establish this difference in classification performance, but they do not independently identify the visual characteristics responsible for the errors. Therefore, explanations involving motif composition, texture, symmetry, or repetitive patterns should be interpreted as possible explanations rather than confirmed causes. Future research could examine misclassified images directly, document their relevant visual characteristics, and evaluate whether these characteristics are associated with classification performance. Such analysis could provide a stronger basis for improving the dataset and refining the classification approach.

CONCLUSION

This study classified four regional *songket* motif categories—Balinese, Lombok, Palembang, and Riau—using Inception-v3 image embeddings and four machine-learning classifiers in Orange Data Mining. Logistic Regression was the best-performing classifier, achieving an accuracy of 75%, an AUC of 0.905, and an MCC of 0.667. These findings address the research question concerning comparative model performance. The confusion matrices showed that Balinese and Lombok *songket* were the categories most frequently confused with each other, whereas Palembang and Riau *songket* were generally classified more accurately.

Future work should expand and balance the dataset, document image acquisition and validation procedures, evaluate fine-tuned CNNs and data augmentation for the most frequently confused classes, and assess a mobile or web application for cultural practitioners. The proposed workflow may contribute to the digitization, cataloging, and documentation of traditional textiles, supporting the preservation of *songket* heritage and potentially contributing to the economic development of weaving communities.

ACKNOWLEDGEMENT

The author would like to thank the Master's Program in Human Resource Development and the Department of Information and Library Science, Faculty of Social and Political Sciences, Universitas Airlangga, for the support provided during this research. The author also thanks Roboflow Universe for providing access to the songket image dataset used in this research.

AUTHOR CONTRIBUTION STATEMENT

Maya Fitriana Primasari contributed to the research conceptualization, data collection and curation, data analysis, and writing of the initial manuscript. Imam Yuadi contributed to research supervision, methodology validation, and manuscript review and editing. Both authors have read and approved the final published manuscript.

REFERENCES

Anggoro, D. A., Marzuki, A. A. T., & Supriyanti, W. (2024). Classification of Solo Batik patterns using deep learning convolutional neural networks algorithm. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *22*(1), 232–240.

Ariessaputra, S., Vidiasari, V. H., Al Sasongko, S. M., Darmawan, B., & Nababan, S. (2024). Classification of Lombok Songket and Sasambo Batik motifs using the Convolution Neural Network (CNN) algorithm. *JOIV: International Journal on Informatics Visualization*, *8*(1), 38–44.

Bao, Y., Sivanandan, S., & Karaletsos, T. (2024). Channel vision transformers: An image is worth 1 x 16 x 16 words. *International Conference on Learning Representations*, *2024*, 50651–50677.

Bongon, M. S., Raymundo, C. M., Santelices, L. V, Sumayo, R. O., Vibandor, D. B., Villafuerte, M. R. C., Sumayo, R. M., Setyawati, N. K. A., Rideng, I. W., Mahendrawati, N. L. M., Mardika, I. M., Ekayani, N. N. S., Muliana, I. N., Wesna, P. A. S., & Susanti, A. A. D. (2024). Mobilizing stakeholders in preserving and revitalizing the traditional craft of Balinese songket textile of the women weavers. *Community Service Journal*, *7*(1), 47–58.

Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, *106*, 249–259.

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1), 6.

Das, S., Mallik, D., Ali, A., & Chaudhuri, S. S. (2023). Invo-ResNet: Rethinking the Most Used Backbone in Computer Vision. *International Conference on Frontiers in Computing and Systems*, 551–561.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. *International Conference on Learning Representations (ICLR)*. https://doi.org/10.48550/arXiv.2010.11929

Duggirala, N., & Kollipara, S. (2025). *Utilizing Densely Connected Convolutional Network Architectures with Effective Feature Reuse for the Enhanced Early Detection of Pancreatic Neuroendocrine Malignancies*.

Fernández, A., García, S., Herrera, F., & Chawla, N. V. (2018). SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence Research*, *61*, 863–905.

Gao, X., Xie, D., Zhang, Y., Wang, Z., Chen, C., He, C., Yin, H., & Zhang, W. (2026). A comprehensive survey on imbalanced data learning. *Frontiers of Computer Science*, *20*(11), 2011622.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

Hill, J. C., LaBonte, T., Zhang, X., & Muthukumar, V. (2025). On the unreasonable effectiveness of last-layer retraining. *ArXiv Preprint ArXiv:2512.01766*.

Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4708.

Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, *6*, Article 27.

Kornblith, S., Shlens, J., & Le, Q. V. (2019). Do better ImageNet models transfer better? *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Miró, J. P., Mohsin, M. A., Alkamil, A., & Perera, D. G. (2025). FPGA-based Hardware Accelerator for Bottleneck Residual Blocks of MobileNetV2 Convolutional Neural Networks. *2025 IEEE 68th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 847–851.

Nayman, N., Golbert, A., Noy, A., & Zelnik-Manor, L. (2024). Diverse imagenet models transfer better. *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 1903–

1914.
Qian, Y., An, R., Liu, G., Tang, H., Xiao, G., & Bavirisetti, D. P. (2025). TransFusion: Transfer learning-driven adaptive fusion network for infrared and visible image. *Infrared Physics & Technology*, *150*, 105906.
Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding transfer learning for medical imaging. *Advances in Neural Information Processing Systems*, *32*.
Roboflow. (2023). *Roboflow Universe: Open-source computer vision datasets*. https://universe.roboflow.com/
Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510–4520.
Sharma, S., & Gosain, A. (2025). Addressing class imbalance in remote sensing using deep learning approaches: a systematic literature review. *Evolutionary Intelligence*, *18*(1), 23.
Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*, Article 60.
Sun, C., Shrivastava, A., Singh, S., & Gupta, A. (2017). Revisiting unreasonable effectiveness of data in deep learning era. *Proceedings of the IEEE International Conference on Computer Vision*, 843–852.
Szabó, S., Holb, I. J., Abriha-Molnár, V. É., Szatmári, G., Singh, S. K., & Abriha, D. (2024). Classification Assessment Tool: A program to measure the uncertainty of classification models in terms of class-level metrics. *Applied Soft Computing*, *155*, 111468.
Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2818–2826. https://doi.org/10.1109/CVPR.2016.308
Tamura, J., Itaya, Y., Hayashi, K., & Yamamoto, K. (2026). Statistical inference of the Matthews correlation coefficient for multiclass classification. *Journal of Applied Statistics*, 1–26.
Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). *A survey on deep transfer learning*. Springer.
Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, *17*(1), 168–192.
Xu, G., Wang, X., Wu, X., Leng, X., & Xu, Y. (2025). Development of residual learning in deep neural networks for computer vision: A survey. *Engineering Applications of Artificial Intelligence*, *142*, 109890.
Yu, F., Xiu, X., & Li, Y. (2022). A survey on deep transfer learning and beyond. *Mathematics*, *10*(19), 3619.
Zeng, W. (2024). Image data augmentation techniques based on deep learning: A survey. *Mathematical Biosciences and Engineering*, *21*(6), 6190–6224.