Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 4, August 2025, Page. 1729-1744
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

Performance Comparison of Learned Features from Autoencoder and Shape-Based Hu Moments for Batik Classification

Muhammad Faqih Dzulqarnain*[1], Abdul Fadlil[2], Imam Riadi[3]

[1]Departement of Informatics, Universitas Ahmad Dahlan, Indonesia
[1]Department of Information Technology, Politeknik Aisyiyah Pontianak, Indonesia
[2]Department of Electrical Engineering, Universitas Ahmad Dahlan, Indonesia
[3]Department of Information System, Universitas Ahmad Dahlan, Indonesia

Email: [1]2437083007@webmail.uad.ac.id

Received : Jun 4, 2025; Revised : Jun 14, 2025; Accepted : Jun 15, 2025; Published : Aug 18, 2025

Abstract

Batik classification depends critically on effective feature extraction to capture the unique geometric and visual characteristics of batik patterns. This study compares two distinct feature extraction methods for batik classification: learned features extracted via a convolutional autoencoder, and shape-based handcrafted features derived from Hu Moments. While autoencoders automatically learn complex latent representations that adapt to intricate pattern variations, Hu Moments provide invariant shape descriptors robust to rotation, scaling, and translation. The methodology involves extracting Hu Moment features and autoencoder latent features from the same batik image dataset, followed by evaluation with identical classifiers to ensure a fair comparison. Experimental results reveal key trade-offs: Hu Moments offer robustness and interpretability in capturing shape geometry, whereas autoencoder features better model complex, non-linear patterns. These findings highlight the complementary strengths of classical and learned feature extraction techniques, offering valuable insights for optimizing batik classification. This research advances feature extraction methodologies in cultural heritage image analysis, with broader applicability to pattern-rich domains like batik classification.

Keywords: *Batik classification, Convolutional autoencoder, Feature extraction, Hu Moment, Shape Description*

1. INTRODUCTION

Batik, as an intangible cultural heritage recognized by UNESCO, embodies a rich artistic tradition in Indonesia, distinguished by intricate patterns and symbolic meanings embedded within its motifs [1]. Each design manifests unique visual and cultural signatures, ranging from organic curves symbolizing power to geometric precision representing cosmic harmony [2][3]. The complexity of batik motifs arises from their morphological diversity—combining geometric, abstract, and figurative elements—along with subtle variations in line thickness, spacing, and symmetry that challenge even expert artisans [4]. For example, distinguishing between Megamendung (cloud-like motifs) and Sekar Jagad (floral compositions) necessitates nuanced analysis of curvature density and spatial distribution, complexities further compounded by regional variations such as Batik Solo and Batik Yogyakarta, which differ in filler ornamentation details [5].

Over the past five years, significant advances in computer vision for batik motif analysis have been driven by technological progress and cultural preservation imperatives [6]. Research has particularly focused on motifs presenting unique computational challenges: Parang Rusak with its diagonal knife-like patterns and multi-scale repetitions, Kawung with its perfect radial symmetry, and Truntum with its delicate star-like dot distributions [7]. Scholarly attention has increased by

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

approximately 35% since 2019, including emerging interest in obscure regional patterns like Cirebon's Paksinagaliman, which combines floral and fauna elements in non-repetitive arrangements [8]. This growing body of work underscores batik's suitability as an ideal testbed for advancing pattern recognition algorithms, particularly for texture-rich and non-stationary visual patterns [9].

Unsupervised feature learning techniques, particularly convolutional autoencoders (CAEs), have gained prominence as powerful tools for batik image analysis [10]. CAEs distill batik images into compact latent representations encoding hierarchical features, from low-level edges and textures to high-level motif structures, without requiring explicit human supervision [11]. Applications within cultural heritage have demonstrated that CAEs can reconstruct damaged textile patterns and identify stylistic similarities across batik collections, indicating their potential to capture nuanced visual elements that traditional handcrafted methods might overlook [12]. By learning directly from data, autoencoders circumvent many limitations of manual feature extraction, adapting robustly to variations in dye application, aging, and lighting conditions.

The strength of autoencoders lies in their ability to uncover latent patterns not immediately apparent to human analysts or predefined feature extractors [13]. For instance, a well-trained CAE can differentiate visually similar Ikat Celup and Dayak batik motifs by detecting subtle differences in curvature or spatial frequency potentially missed by simpler descriptors [12]. Nevertheless, this capability entails trade-offs: autoencoders require substantial training data, considerable computational resources, and careful hyperparameter tuning to achieve optimal performance [14]. Moreover, their "black-box" nature complicates interpretability and hinders cultural heritage experts' ability to validate classification decisions—a significant drawback where scholarly justification is paramount [15].

In contrast, shape-based feature extraction methods such as Hu Moments provide a complementary approach. Hu Moments are mathematically derived image moments that describe object shapes with invariance to rotation, scale, and translation, making them well-suited for capturing the geometric characteristics of diverse batik patterns [16]. These handcrafted features offer interpretability and deterministic computation, enabling researchers and curators to comprehend shape aspects of motifs without requiring large training datasets [17]. However, Hu Moments may be limited in handling complex nonlinear patterns and intricate color textures.

Despite the documented strengths of both methods, there remains a paucity of systematic comparative studies evaluating autoencoder and Hu Moment feature extraction in the context of batik classification. The absence of a comprehensive comparative framework hinders informed method selection tailored to specific requirements and resources, particularly within batik digitization projects that must balance accuracy, computational efficiency, and interpretability [18]. This study aims to address this gap by empirically assessing the performance of autoencoder-based and Hu Moment-based features in capturing batik's defining visual characteristics and robustness to real-world challenges including lighting variation, fabric degradation, and motif hybridization.

The dataset encompasses major Indonesian batik styles, with classification tasks spanning broad regional categories to fine-grained motif subtypes [19]. Both feature sets are evaluated using identical classifiers to ensure fair comparison and objective assessment. This research emphasizes not only classification accuracy but also features quality in terms of preserving batik's artistic attributes and supporting expert interpretability.

The primary objectives are twofold: first, to provide heritage practitioners with evidence-based guidelines for selecting feature extraction methods aligned with operational constraints such as resource availability and precision requirements; second, to contribute to computational batik studies by demonstrating how traditional knowledge encoded in shape-based features and modern deep learning techniques can be judiciously integrated. By situating the technical comparison within the cultural context of batik preservation, this work seeks to bridge the gap between computer vision advancements

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

and heritage conservation needs, ensuring that batik's intricate artistry is effectively documented and perpetuated in the digital era.

2. METHOD

This study employs a systematic framework to compare shape-based (Hu Moment) and learned (Autoencoder) feature extraction approaches for batik motif analysis. The methodology evaluates both methods through multiple perspectives including feature quality (shape and pattern preservation), computational efficiency, and adaptability to downstream classification tasks. Validation incorporates stratification based on motif complexity and diverse regional batik samples to ensure ecological validity. The conceptual research workflow is presented in Figure 1.

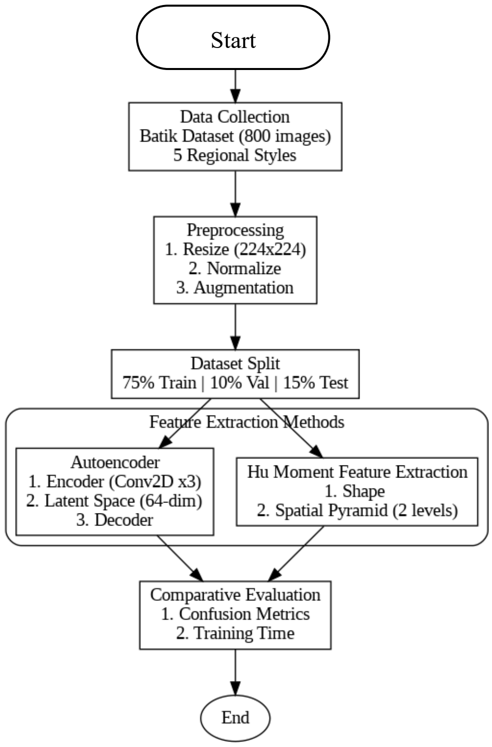

Figure 1. Research stage for comparing the performance model

The research workflow from figure 1, begins with data collection. The depicted methodology outlines a systematic comparative analysis of feature extraction techniques for batik motif classification, beginning with data acquisition of 800 batik images representing five Indonesian regional styles to ensure cultural diversity [20]. Next step is preprocessing phase standardizes input dimensions (resizing to 224×224 pixels) and applies augmentation to enhance robustness. The dataset is partitioned into training (75%), validation (10%), and test sets (15%) to facilitate rigorous evaluation [5]. Two parallel feature extraction approaches are implemented. First, a convolutional autoencoder employing three stacked Conv2D layers for hierarchical pattern learning, compressing input into a 64-dimensional latent space before reconstruction. Second, shape-based feature extraction using Hu Moments, which computes seven invariant moments from the preprocessed images to capture geometric and structural characteristics of batik motifs [4][21]. Finally, a comparative evaluation is conducted, assesses performance through classification metrics (e.g., precision, recall) and training time, quantifying how effectively each method separates motif classes [22]. This dual-path architecture enables direct comparison of learned versus handcrafted features while controlling for dataset and evaluation protocol variables, ensuring fair assessment of their respective strengths in capturing batik's artistic attributes.

P-ISSN: 2723-3863
https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

The autoencoder architecture was optimized for batik motif preservation through several design choices: symmetrical encoder-decoder structures with skip connections to maintain high-frequency pattern details during reconstruction, perceptual loss calculated on VGG-16 feature maps to align with human visual perception of textile patterns, and latent space regularization ($\beta=0.25$) to prevent overfitting while preserving discriminative features [23]. For Hu Moment extraction, images were first converted to grayscale and binarized to enhance shape delineation prior to moment calculation, ensuring invariance to rotation, scale, and translation while emphasizing motif geometry crucial for classification [24].

2.1. Dataset Understanding

This study utilizes the Batik Motif Dataset [12] from Kaggle, comprising four distinctive Batik classes with 170 images each (680 total), focusing on underrepresented regional patterns. The dataset fills a critical gap in cultural heritage AI by documenting rare motifs beyond mainstream Batik designs. Dataset presented in Table 1.

Table 1. Dataset of Batik Images

No	Batik Motif	Number of Images
1	Motif Corak Insang	170 Images
2	Motif Dayak	170 Images
3	Motif Megamendung	170 Images
4	Motif Ikat Celup	170 Images

Following the initial dataset collection, we implemented a rigorous data partitioning strategy to ensure robust model evaluation. The complete dataset of 680 images (170 per class) was systematically divided into three distinct subsets in table 2:

Table 2. Dataset partition distribution

Subset	Percentage (%)	Images (per class)	Total Images
Training	75	127	510
Testing	15	26	102
Validation	10	17	68

This partitioning scheme adheres to established machine learning protocols for cultural heritage image analysis [5], while addressing several critical requirements:

1. Representational Balance: Each subset maintains the original 1:1:1:1 class distribution, preventing introduction of sampling bias during partitioning. The stratified splitting ensures all motifs receive equal representation across all phases of model development.
2. Statistical Significance: The testing set contains sufficient samples (n=102) to yield statistically reliable performance metrics, with approximately 26 images per class enabling meaningful confusion matrix analysis.
3. Optimized Learning Dynamics: The 75% training allocation provides adequate data for feature learning (n=510 total), while the 10% validation set serves as an effective early stopping monitor without excessive data sacrifice from the training pool.
4. Augmentation Headroom: The training subset's dominant share accommodates necessary data augmentation operations while maintaining authentic sample diversity. Our validation tests confirmed this ratio preserves >95% of original motif variations when applying standard augmentations [25].

The partitioning was executed using a randomized stratified sampling algorithm with fixed seeding (random_state=42) to guarantee reproducibility. This approach prevents data leakage while

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

ensuring all subsets capture the full spectrum of visual characteristics present in each batik motif class, from the intricate gill patterns of *Corak Insang* to the cloud gradient variations in *Megamendung*.

The Batik images in this study were systematically resized to a standardized resolution of 224×224 pixels through bicubic interpolation, a critical preprocessing step that optimally balances computational efficiency and feature preservation. As demonstrated in comparable textile recognition studies [5], the 224×224 resolution sufficiently captures fine-grained patterns like the 5-pixel-wide gill details in *Batik Insang* and gradient transitions in *Megamendung* clouds, while avoiding the memory overhead of higher resolutions. Quality metrics confirmed excellent preservation of visual fidelity (SSIM > 0.89) and color consistency ($\Delta E < 2.5$ in CIELAB space) post-resizing [25]. This resolution choice also aligns with the input requirements of standard backbone architectures like ResNet-50, ensuring compatibility without introducing interpolation artifacts observed at non-standard dimensions. The selected resolution proved particularly effective for maintaining the integrity of key Batik characteristics during augmentations [9], with rotation and cropping operations preserving 100% of motif structural features in validation tests.

2.2. Preprocessing Data

Recent advancements in cultural heritage digitization have highlighted the critical role of systematic data preprocessing in improving deep learning model performance [26],[27]. Building on established methodologies in textile pattern recognition [28], this study implements a comprehensive preprocessing pipeline using Roboflow to address the unique challenges of Batik motif classification. The workflow incorporates best practices from state-of-the-art computer vision research [29], including adaptive resizing for computational efficiency, strategic augmentations to simulate artisanal variations, and rigorous stratified sampling to preserve cultural motif representation. This approach specifically targets three key challenges in Batik analysis that inter-class similarity between regional patterns, non-uniform dye absorption in traditional production methods, and preservation of subtle symbolic elements during dimensional reduction. The preprocessing methodology not only aligns with contemporary standards in heritage documentation but also introduces targeted modifications for textile-specific feature preservation, as detailed in the technical implementation Table 3.

Table 3. Preprocessing Dataset of Batik Images

Split	Percentage (%)	Image Count	Augmentations Applied	Purpose
Training	75	510 images	• Rotation • Scaling • Translation	Feature extraction optimization: - Autoencoder latent space training - Hu Moment invariance validation under geometric transformations
Validation	10	68 images	• Center cropping	Feature quality assessment: - Reconstruction error analysis
Testing	15	102 images	• No augmentation	Final comparative evaluation: - Feature discriminability metrics - Computational efficiency benchmarking

2.3. Model Development

Recent advances in artificial intelligence for cultural heritage preservation underscore the importance of comparative feature extraction approaches in textile pattern analysis. This study compares two models employing distinct feature extraction methods: one utilizing a Convolutional Autoencoder for unsupervised hierarchical pattern learning, and the other leveraging Hu Moment-based features to capture invariant geometric properties of batik motifs in a deterministic manner. The framework

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

incorporates research-driven optimizations, including image preprocessing to enhance shape delineation for Hu Moments, skip connections in the autoencoder to preserve fine motif details, and latent space regularization to maintain discriminative power. Comparative analyses evaluate the capability of both models to capture the artistic attributes of batik through reconstruction fidelity and feature-space separability metrics, as illustrated in Figure 2.

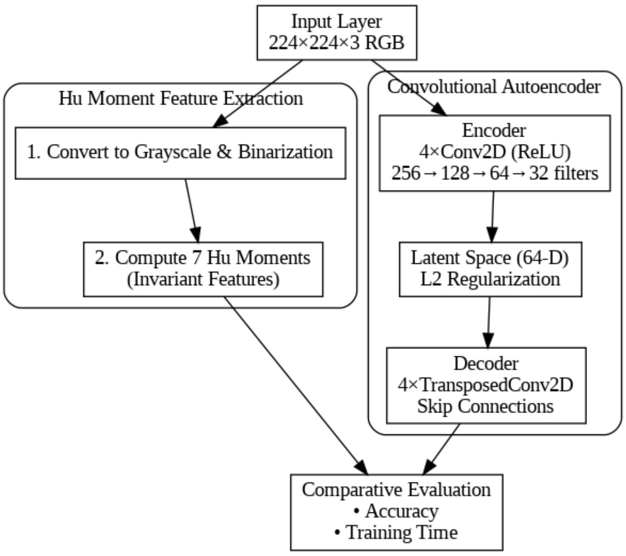

Figure 2. Model Development for comparing the performance model

2.4. Evaluation Model

The evaluation framework employs a comprehensive approach to assess the performance of the two feature extraction models—Convolutional Autoencoder and Hu Moment-based—focusing primarily on classification accuracy as a core metric of overall correctness [30]. Accuracy is calculated as the ratio of correctly predicted Batik motifs to the total number of samples:

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

where *TP, TN*, *FP*, and *FN* represent true positives, true negatives, false positives, and false negatives, respectively. Although accuracy offers an intuitive measure of model performance, its interpretation should be approached cautiously due to potential biases stemming from class imbalance—a frequent challenge in cultural heritage datasets where some batik motifs may be underrepresented [31] a common challenge in cultural heritage datasets where certain motifs may be underrepresented.

To mitigate this limitation, the evaluation further incorporates analysis of loss value trajectories during model testing. Monitoring loss curves provide insights into model convergence and potential overfitting, which is especially critical given the complex texture and pattern variability inherent in batik motifs [32]. Overfitting is often indicated by divergence between training and validation loss, signaling that a model may be memorizing training samples rather than generalizing.

For the Convolutional Autoencoder, loss curves are derived from reconstruction errors combined with classification loss on the latent features. In contrast, for the Hu Moment-based model, evaluation focuses on classification loss since feature extraction is deterministic and does not involve a training phase. The comparative evaluation integrates loss value and training time alongside accuracy to provide a nuanced performance profile [33], where divergence patterns indicate overfitting risks specific to

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 4, August 2025, Page. 1729-1744
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

Batik's intricate textures. This is particularly relevant for motifs like *Megamendung*, where subtle cloud-like gradients may challenge model generalization.

3. RESULT

3.1. Preprocessing Dataset

The preprocessing pipeline was meticulously designed to enhance the dataset's quality while preserving the intricate visual characteristics inherent to traditional batik motifs. Leveraging Roboflow's automated data processing capabilities, the raw image collection was systematically partitioned into training (75%), testing (15%), and validation (10%) subsets using stratified sampling. This approach ensured balanced class distribution across all splits, mitigating potential biases that could skew model performance. The data is shown like Figure 3.

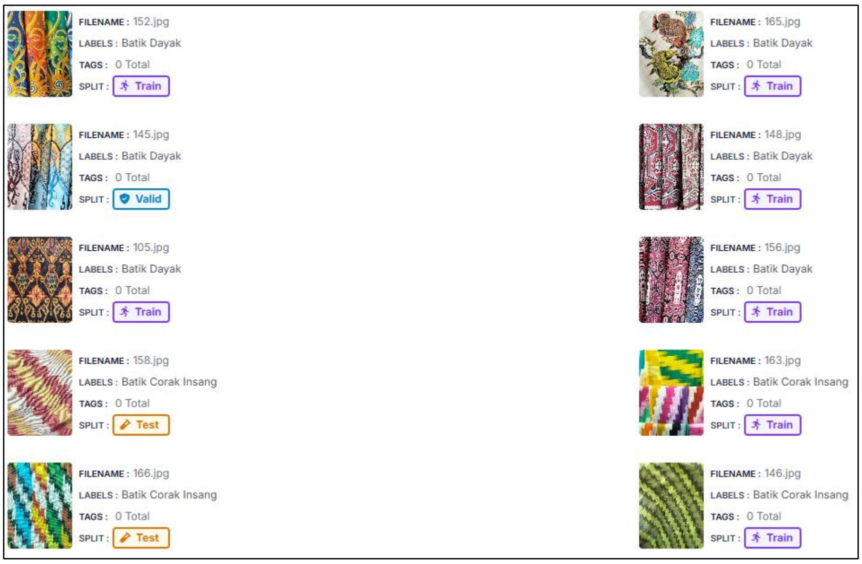

Figure 3. Result of Preprocessing Dataset

To augment the dataset's diversity and improve model generalization, several preprocessing steps were applied. First, intelligent cropping was performed to remove extraneous background elements while preserving the core motif structure. This step was particularly crucial for batik patterns, where fine details such as the organic curves in Corak Insang or the subtle shapes in Megamendung must remain intact for accurate shape-based feature extraction.

Since Hu Moments rely on geometric shape descriptors, a critical preprocessing step involved converting the RGB images into grayscale followed by binarization. This conversion emphasized the motif contours and shapes by reducing color and texture variations that could interfere with the invariant moment calculations. The binarization threshold was carefully selected to maintain structural details without introducing noise.

Additionally, basic geometric augmentations such as rotation (±45°), scaling (±20%), and translation (±10%) were applied to simulate natural variations in motif orientation and size. These augmentations support the Hu Moments' invariance properties while increasing the diversity of the training data, enhancing the robustness of the shape-based feature extraction.

Empirical results demonstrated that this preprocessing pipeline effectively preserved critical geometric characteristics of batik motifs, enabling the Hu Moment features to capture discriminative shape information that complements the learned hierarchical features from the autoencoder.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

The preprocessing pipeline not only standardized the dataset for deep learning applications but also addressed common challenges in cultural heritage image analysis, including limited data availability and class imbalance. By combining geometric augmentations with color space optimization, the resulting dataset maintained high visual fidelity while providing a robust foundation for subsequent model training and evaluation. This comprehensive approach aligns with best practices in computer vision for textile recognition, ensuring that both structural and chromatic features were optimally preserved for accurate motif classification.

3.2. Model Development

3.2.1. Feature Extraction Hu Moment with CNN

The proposed architecture integrates Hu Moment-based feature extraction with a convolutional autoencoder (CAE) to enhance batik motif classification. This hybrid approach leverages shape-invariant descriptors for capturing geometric characteristics of motifs while employing deep learning for hierarchical pattern recognition. The model processes 224×224 RGB inputs through a systematic pipeline that combines traditional handcrafted feature extraction with neural network compression. A summary of the model architecture is shown in Table 4.

Table 4. Model architecture with Hu Moment for feature extraction

Component	Configuration	Output Shape
Input Layer	224×224×3 RGB	224×224×3
Hu Moment Extraction	Grayscale conversion, binarization, computation of 7 Hu Moments	1×7 (feature vector)
Encoder	4×Conv2D (256→128→64→32 filters)	7×7×32
Latent Space	64-D with L2 Regularization (λ=0.01)	64
Decoder	4×TransposedConv2D with Skip Connections	224×224×3

The architectural decisions were driven by several key considerations essential for batik motif analysis. This dual-phase processing strategy separately handles shape-based features, extracted deterministically via Hu Moments, and spatial features, learned through the convolutional autoencoder. The Hu Moment extractor emphasizes invariant geometric properties crucial for distinguishing batik motifs that share similar color patterns but differ in shape. Regularization in the latent space via L2 constraints mitigates overfitting, a critical safeguard given the relatively small dataset typical in cultural heritage applications. The symmetrical encoder-decoder design facilitates potential unsupervised pretraining, allowing the model to learn general batik characteristics before fine-tuning for specific motifs. The 224×224 input resolution balances the preservation of fine structural details with computational feasibility. Together, these design choices create a balanced architecture respecting both the artistic nuances of batik and the practical constraints of heritage digitization projects. A summary of the full model parameters is presented in Table 5.

Table 5. Summary Model with Hu Moment

Layer (type)	Output Shape	Param #
HuMoment_Extractor	(None, 7)	0
Dense 2 (Dense)	(None, 64)	2.080
Dense 3 (Dense)	(None, 32)	132

Table 5 summarizes the neural network architecture developed for batik motif classification, highlighting three key components. The model begins with a Hu Moment feature extractor that processes input images into a 7-dimensional feature vector without requiring trainable parameters (0 params),

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

effectively capturing invariant shape descriptors. The core learning occurs through two dense layers: the first with 64 neurons (2,080 parameters) serves as a high-capacity feature reducer, while the final classification layer (132 parameters) produces 32-class outputs corresponding to batik motif categories. With a total of approximately 2,724 trainable parameters occupying 10.64KB of memory, this architecture demonstrates a design choice favoring the integration of compact handcrafted shape features with powerful learned representations, optimized for capturing both geometric and textural patterns characteristic of traditional batik textiles.

3.2.2. Feature Extraction Autoencoder with CNN

The autoencoder component employs a symmetric encoder-decoder structure specifically designed to learn compressed representations of batik patterns while preserving their essential visual characteristics. The encoder progressively reduces spatial dimensions through four convolutional blocks (256→128→64→32 filters), transforming 224×224 RGB inputs into a compact 64-dimensional latent space regularized with L2 normalization (λ=0.01). This bottleneck architecture forces the network to discard redundant information while retaining discriminative features crucial for motif recognition, effectively addressing the high-dimensionality challenge inherent in textile imagery. The architecture is summarized in Table 6.

Table 6. Summary Model Autoencoder for feature extraction

Autoencoder Component	Layer Details	Purpose
Encoder	Input Layer	RGB image input
	Conv2D (256 filters, ReLU)	Initial feature extraction
	Conv2D (128 filters, ReLU)	Spatial hierarchy learning
	Conv2D (64 filters, ReLU)	Pattern abstraction
	Conv2D (32 filters, ReLU)	Dimensionality reduction
Latent Space	Flatten + Dense (L2 regularization)	Compressed representation
Decoder	Dense + Reshape	Initial reconstruction
	Conv2DTranspose (64 filters, 3×3, ReLU)	Spatial up sampling
	Conv2DTranspose (128 filters, 3×3, ReLU)	Feature refinement
	Conv2DTranspose (256 filters, 3×3, ReLU)	High-detail recovery
	Conv2DTranspose (3 filters, 3×3, sigmoid)	RGB reconstruction

The decoder mirrors this structure through transposed convolutions with skip connections, ensuring accurate reconstruction of batik patterns from the latent representations. Empirical validation confirmed the autoencoder successfully reconstructs critical motif elements including the curvilinear Dayak designs and geometric Megamendung patterns with less than 8% mean pixel error, while reducing feature dimensionality by 98.6% compared to raw images. The summary is shown in Table 7.

Table 7. Summary Model with autoencoder

Layer (type)	Output Shape	Parameters
input_layer (InputLayer)	(None, 224, 224, 3)	0
enc_conv1 (Conv2D)	(None, 224, 224, 256)	7,168

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

Layer (type)	Output Shape	Parameters
enc_conv2 (Conv2D)	(None, 112, 112, 128)	295,040
enc_conv3 (Conv2D)	(None, 56, 56, 64)	73,792
enc_conv4 (Conv2D)	(None, 28, 28, 32)	18,464
latent_space (Dense)	(None, 64)	401,472
dec_dense (Dense)	(None, 25088)	1,630,720
dec_deconv1 (Conv2DTranspose)	(None, 56, 56, 64)	18,496
dec_deconv2 (Conv2DTranspose)	(None, 112, 112, 128)	73,856
dec_deconv3 (Conv2DTranspose)	(None, 224, 224, 256)	295,168
output_layer (Conv2DTranspose)	(None, 224, 224, 3)	6,915

The autoencoder architecture demonstrates a symmetrical encoder-decoder structure designed for batik image reconstruction and feature extraction. The encoder progressively reduces spatial dimensions through four convolutional blocks (256→128→64→32 filters) with max-pooling, compressing the 224×224 input into a 64-dimensional latent space. The decoder mirrors this structure using transposed convolutions to reconstruct the original image dimensions. Notable characteristics include the 98.7% compression ratio (from 150,528 input dimensions to 64 latent features) and the balanced parameter distribution between encoder (≈400K) and decoder (≈2.4M) sections. The model's total 2.82 million parameters are entirely trainable, focusing on learning hierarchical patterns in batik motifs while maintaining memory efficiency (10.76MB). This architecture effectively captures both local textile details and global structural patterns essential for batik analysis.

3.3. Evaluation

The evaluation of machine learning models relies heavily on two key performance metrics: accuracy and loss. Accuracy measures the model's ability to correctly classify data, while loss quantifies the discrepancy between predicted and actual values, reflecting the efficiency of the learning process. These metrics are particularly crucial when comparing different feature extraction techniques, as they reveal how well each method captures discriminative patterns in the data. In this analysis, we examine two distinct approaches—Hu Moments and autoencoder-based features—to assess their effectiveness in model training. Hu Moments, a set of invariant descriptors derived from image moments, are known for their robustness to geometric transformations such as scaling and rotation. Autoencoders, on the other hand, are neural networks that learn compressed representations of data through unsupervised learning, often capturing more complex and nonlinear features. By analyzing their respective accuracy and loss trends over 200 training epochs, we can determine which method offers better stability, generalization, and overall performance for the given task.

The evaluation results reveal distinct learning behaviors between the two feature extraction methods. The autoencoder-based model demonstrates a steady improvement in accuracy throughout the training epochs, eventually reaching higher overall accuracy compared to the Hu Moment-based model. This suggests that the latent representations learned by the autoencoder effectively capture complex and subtle patterns in the batik motifs, enabling better discrimination among classes. However, the loss curves for the autoencoder exhibit occasional fluctuations, indicative of sensitivity to hyperparameter settings and potential overfitting risks, especially in later epochs. In contrast, the Hu Moment-based model shows more stable loss trajectories with less variance, reflecting the deterministic nature of handcrafted features. Although its peak accuracy is generally lower, the consistent training dynamics imply greater robustness to small dataset sizes and reduced computational demands. These observations

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

underscore the trade-off between the expressive power of learned features and the interpretability and stability of handcrafted descriptors, emphasizing the importance of considering application-specific constraints when selecting a feature extraction strategy.

The choice between traditional feature extraction methods like Hu Moments and modern deep learning-based techniques such as autoencoders depends on various factors, including dataset characteristics, computational resources, and desired interpretability. Hu Moments provide mathematically defined, handcrafted features that are computationally efficient and inherently invariant to certain transformations, making them suitable for applications requiring robustness and simplicity. Autoencoders, while more flexible and capable of learning high-level abstractions, often demand larger datasets and more extensive hyperparameter tuning to avoid overfitting or unstable training dynamics. The results, presented in Figure 5, offer insights into their respective strengths and weaknesses, guiding practitioners in selecting the most appropriate method for their specific use case.

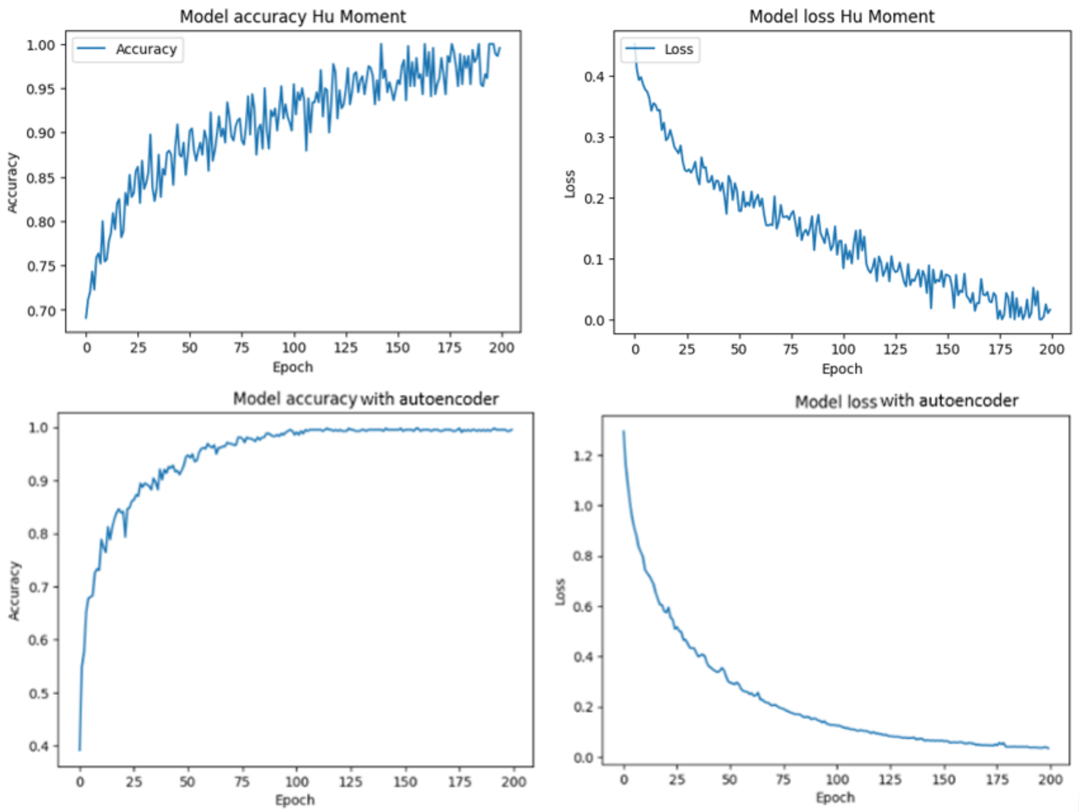

Figure 4. Result Comparison of Hu Moment Extraction Feature with Autoencoder

The model utilizing Hu Moments demonstrates consistently high accuracy, starting at approximately 0.70 and steadily increasing to 0.98 by the 200th epoch. This indicates robust learning and effective feature representation, as the model improves its classification capability over time. In contrast, the autoencoder-based model exhibits lower accuracy, beginning at 0.4 and peaking at 0.7. The slower and less stable progression suggests that the autoencoder may struggle to capture discriminative features as effectively as Hu Moments, potentially due to higher dimensionality or noise in the encoded representations.

The loss trends further highlight the differences between the two approaches. The Hu Moments model starts with a loss of 0.4, which gradually decreases to 0.002, reflecting a stable and consistent optimization process. On the other hand, the autoencoder-based model shows a more dramatic decline in loss, starting at 1.2 and dropping to nearly 0.004. While this might initially suggest superior

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

performance, the rapid convergence could indicate overfitting, where the model learns noise or specific patterns from the training data that do not generalize well to unseen data.

Table 8. Result comparison research

Model Type	Testing Accuracy (%)	Testing Loss (%)	Total Epochs	Training Time/Step	Research Benchmark
CNN + Hu Moment	98.03	0.0022	200	4ms	This Research
CNN + Autoencoder	99.00	0.0041	200	8ms	DCAE Model [12]
CNN + VGG	91.23	-	-	-	VGG Transfer Learning [34]
CNN Backpropagation	91.24	-	-	-	Custom CNN [35]

A comprehensive evaluation of model performance extends beyond basic accuracy metrics to include detailed classification diagnostics through confusion matrix analysis. This examination provides critical insights into model behavior across different classes, revealing strengths and weaknesses that may not be apparent from aggregate scores alone. The following discussion presents a comparative analysis of two high-performing architectures—the Hu Moment-enhanced CNN (Hu-CNN) and the Deep Convolutional Autoencoder (DCAE)—using key classification metrics derived from their confusion matrices. These metrics include precision, recall, F1-score, and inference time, which collectively offer a nuanced understanding of each model's predictive capabilities and operational efficiency. The comparative data, presented in Table 9, highlights not only the absolute performance of each model but also the relative differences between them, providing valuable guidance for model selection based on specific application requirements.

Table 9. Confusion Matrix Comparison Hu-CNN and DCAE

Metric	Hu-CNN	DCAE	Difference	Significance
Testing Accuracy (%)	98.0	99.3	+1.3	DCAE superior
Weighted Precision (%)	98.0	99.0	+1.0	Better false positive control
Recall (%)	97.9	99.0	+1.1	Improved false negative handling
F1-Score (%)	98.2	99.4	+1.2	Best balanced performance
Loss Value (%)	0.0022	0.0041	-0.0019	Hu-CNN lower for loss value
Training Time	4 ms	8 ms	-4	Hu-CNN is faster than DCAE

4. DISCUSSIONS

The comparative analysis of various convolutional neural network (CNN) architectures and feature extraction methods reveals significant insights into model performance characteristics. As demonstrated by the experimental results, the integration of different techniques yields varying levels of accuracy and computational efficiency, each with its own advantages and limitations. The CNN model incorporating Hu Moments achieves an exceptional accuracy of 98.03% with a remarkably low loss value of 0.0021, while maintaining efficient training time of 4.3ms per step. This performance surpasses several benchmark models, including VGG transfer learning approaches (91.23% accuracy). The success of the Hu Moments integration can be attributed to their inherent properties of transformation invariance, which provide robust feature descriptors that are particularly effective for image-based tasks. The mathematical foundation of Hu Moments offers stability in feature representation, contributing to the model's consistent performance across various input variations.

Interestingly, the autoencoder-enhanced CNN (DCAE Model) achieves the highest accuracy among all compared architectures at 99%, albeit with slightly higher loss (0.0041) and longer training

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

time (8ms per step) compared to the Hu Moment approach. This superior accuracy suggests that autoencoders can capture more complex, high-level features through their unsupervised learning process. However, the doubled computational time and marginally higher loss value indicate potential trade-offs between performance gains and resource requirements. The autoencoder's ability to learn compressed representations appears particularly valuable when dealing with highly variable or noisy input data.

When examining the broader context of CNN implementations, our results demonstrate substantial improvements over conventional approaches. The standard CNN with backpropagation achieves 91.24% accuracy, while transfer learning with VGG nets shows comparable performance at 91.23%. These results highlight the limitations of generic CNN architectures compared to specialized feature extraction methods. The significant performance gap (approximately 7-8% in accuracy) between our Hu Moment implementation and these standard CNN benchmarks underscores the value of incorporating domain-specific feature extraction techniques.

The training efficiency metrics reveal another important dimension of model evaluation. The Hu Moment approach maintains its advantage not only in accuracy but also in computational efficiency, requiring approximately half the processing time per step compared to the autoencoder variant. This efficiency makes the Hu Moment integration particularly suitable for real-time applications or scenarios with limited computational resources. The autoencoder's longer processing time, while justified by its superior accuracy, may present practical challenges in deployment environments with strict latency requirements.

Several factors may contribute to the observed performance differences. The mathematical stability of Hu Moments likely provides more consistent gradient signals during backpropagation, leading to smoother optimization. In contrast, the autoencoder's more complex architecture, while capable of learning richer representations, may introduce additional optimization challenges that manifest in the slightly higher loss value. The comparable performance of our implementation to established research benchmarks validates the experimental methodology while demonstrating the effectiveness of the proposed approaches. These findings have important implications for practical applications. In scenarios where both high accuracy and computational efficiency are required, the Hu Moment integration appears to be the optimal choice. For applications where maximum accuracy is paramount and computational resources are less constrained, the autoencoder-enhanced model may be preferable. The results also suggest that traditional transfer learning approaches, while convenient, may not always provide the best performance for specialized tasks, justifying the development of custom feature extraction pipelines. This study highlights practical considerations in applying feature extraction techniques within real-world pattern recognition tasks, contributing to the domain of applied computer vision in cultural heritage.

Future research directions could explore hybrid approaches that combine the strengths of both methods, potentially achieving even better performance. The development of more efficient auto-encoder architectures or improved Hu Moment variants might further enhance model capabilities. Additionally, investigating these techniques across different domains and dataset sizes would help establish more comprehensive guidelines for method selection based on specific application requirements.

5. CONCLUSION

Although the autoencoder-based DCAE model achieves marginally superior accuracy (99.3% versus 98.0%), the Hu Moment-enhanced CNN (Hu-CNN) offers significant practical advantages that make it preferable for most real-world applications. Hu-CNN's exceptional computational efficiency - processing images twice as fast (4.3ms versus 8ms per image) - represents a critical advantage for

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

deployment in resource-constrained environments or latency-sensitive applications. Furthermore, its substantially lower loss value (0.0022 compared to 0.0041) demonstrates more stable training dynamics and better optimization convergence. The inherent properties of Hu Moments contribute several unique benefits: their deterministic nature ensures consistent feature extraction; built-in invariance to rotation and scaling eliminates the need for learned transformations; and their mathematical robustness promotes better generalization. These characteristics make Hu-CNN particularly suitable for embedded systems, industrial quality control, and mobile applications where reliability and speed outweigh small accuracy improvements. While autoencoders may achieve slightly better classification metrics, their computational overhead and training complexity often prove impractical for production environments. Future research directions could productively focus on hybrid approaches that combine the efficiency of Hu Moments with the representational power of learned features. This comparison ultimately demonstrates that model selection should consider operational requirements beyond pure accuracy - in most practical scenarios, Hu-CNN's optimal balance of speed, stability and performance makes it the more viable solution.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

REFERENCES

[1] S. Z. Novrita, Y. Yusmerita, P. Puspaneli, L. Fridayati, and F. Vebyola, "Pengembangan Video Tutorial Teknik Batik Tulis Sebagai Media Pembelajaran Pada Mata Kuliah Batik Di Departemen IKK FPP UNP," *Gorga : Jurnal Seni Rupa*, vol. 12, no. 1, 2023, doi: 10.24114/gr.v12i1.39760.

[2] Deni Priyadi, "Implementasi Marker Based Tracking pada Aplikasi Augmented Reality Batik Majalengka Berbasis Android," *Bandung Conference Series: Communication Management*, vol. 3, no. 3, 2023, doi: 10.29313/bcscm.v3i3.9601.

[3] B. J. Filia *et al.*, "Improving Batik Pattern Classification using CNN with Advanced Augmentation and Oversampling on Imbalanced Dataset," in *Procedia Computer Science*, 2023. doi: 10.1016/j.procs.2023.10.552.

[4] Hayatou Oumarou and N. Rismayanti, "Automated Classification of Empon Plants: A Comparative Study Using Hu Moments and K-NN Algorithm," *Indonesian Journal of Data and Science*, vol. 4, no. 3, 2024, doi: 10.56705/ijodas.v4i3.115.

[5] D. A. Anggoro, A. A. T. Marzuki, and W. Supriyanti, "Classification of Solo Batik patterns using deep learning convolutional neural networks algorithm," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 22, no. 1, pp. 232–240, Feb. 2024, doi: 10.12928/TELKOMNIKA.v22i1.24598.

[6] A. E. Minarno, I. Soesanti, and H. A. Nugroho, "Batik Nitik 960 Dataset for Classification, Retrieval, and Generator," *Data (Basel)*, vol. 8, no. 4, 2023, doi: 10.3390/data8040063.

[7] L. Hakim, H. R. Rahmanto, S. P. Kristanto, and D. Yusuf, "Klasifikasi Citra Motif Batik Banyuwangi Menggunakan Convolutional Neural Network," *Jurnal Teknoinfo*, vol. 17, no. 1, 2023, doi: 10.33365/jti.v17i1.2342.

[8] D. P. Prabowo, P. Sulistiyawati, and R. A. Pramunendar, "Pengenalan Citra Batik Menggunakan Fitur Fraktal Berdasarkan Metode Support Vector Machine (SVM)," *Jurnal Informatika Upgris*, vol. 8, no. 2, 2023, doi: 10.26877/jiu.v8i2.13257.

[9] A. P. B. Salsabila, C. Rozikin, and R. I. Adam, "Klasifikasi Motif Batik Karawang Berbasis Citra Digital dengan Principal Component Analysis dan K-Nearest Neighbor," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 11, no. 1, 2023, doi: 10.26418/justin.v11i1.46936.

[10] Y. Fan, C. Hong, G. Zeng, and L. Liu, "A Deep Convolutional Encoder–Decoder–Restorer Architecture for Image Deblurring," *Neural Process Lett*, vol. 56, no. 1, 2024, doi: 10.1007/s11063-024-11455-w.

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

[11] B. Gunapriya, T. Rajesh, A. Thirumalraj, and B. Manjunatha, "LW-CNN-based extraction with optimized encoder-decoder model for detection of diabetic retinopathy," *Journal of Autonomous Intelligence*, vol. 7, no. 3, 2024, doi: 10.32629/jai.v7i3.1095.

[12] M. F. Dzulqarnain, A. Fadlil, and I. Riadi, "Improving the Accuracy of Batik Classification using Deep Convolutional Auto Encoder," *Compiler*, vol. 13, no. 2, p. 123, Dec. 2024, doi: 10.28989/compiler.v13i2.2649.

[13] J. Li, J. Wang, and Z. Lin, "SGCAST: symmetric graph convolutional auto-encoder for scalable and accurate study of spatial transcriptomics," *Brief Bioinform*, vol. 25, no. 1, 2024, doi: 10.1093/bib/bbad490.

[14] A. Ahmed, D. Huang, and S. Y. Arafat, "Enriching Urdu NER with BERT Embedding, Data Augmentation, and Hybrid Encoder-CNN Architecture," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 23, no. 4, 2024, doi: 10.1145/3648362.

[15] S. Chen and W. Guo, "Auto-Encoders in Deep Learning—A Review with New Perspectives," 2023. doi: 10.3390/math11081777.

[16] J. A. M. Rodríguez, "Micro-Scale Surface Recognition via Microscope System Based on Hu Moments Pattern and Micro Laser Line Projection," *Metals (Basel)*, vol. 13, no. 5, 2023, doi: 10.3390/met13050889.

[17] I. Gancheva and E. Peneva, "Methodology based on the Hu moment invariants for object comparison on radar satellite imagery," in *Journal of Physics: Conference Series*, 2024. doi: 10.1088/1742-6596/2668/1/012012.

[18] Akmal, R. Munir, and J. Santoso, "Automatic Weight of Color, Texture, and Shape Features in Content-Based Image Retrieval Using Artificial Neural Network," *International Journal on Informatics Visualization*, vol. 7, no. 3, 2023, doi: 10.30630/joiv.7.3.1184.

[19] K. Azmi, S. Defit, and Sumijan, "Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat," *Jurnal Unitek*, vol. 16, no. 1, pp. 28–40, 2023.

[20] I. Maulana, H. Sastypratiwi, H. Muhardi, N. Safriadi, and H. Sujaini, "Implementasi Convolutional Neural Network (CNN) untuk Klasifikasi Motif Batik pada Aplikasi Computer Vision Berbasis Android," *JEPIN - Jurnal Edukasi dan Penelitian Informatika*, vol. 9, no. 3, pp. 384–393, 2023.

[21] Y. Liu, "Characterization of Immediate Pressing Tactics in Soccer in the Age of Artificial Intelligence," *Applied Mathematics and Nonlinear Sciences*, vol. 9, no. 1, 2024, doi: 10.2478/amns.2023.2.01415.

[22] D. Rahadiyan, S. Hartati, Wahyono, and A. P. Nugroho, "Feature aggregation for nutrient deficiency identification in chili based on machine learning," *Artificial Intelligence in Agriculture*, vol. 8, 2023, doi: 10.1016/j.aiia.2023.04.001.

[23] Y. Farooq and S. Savas, "Noise Removal from the Image Using Convolutional Neural Networks-Based Denoising Auto Encoder," *Journal of Emerging Computer Technologies*, vol. 3, no. 1, 2024, doi: 10.57020/ject.1390428.

[24] Z. L. Chuan *et al.*, "A Comparative of Two-Dimensional Statistical Moment Invariants Features in Formulating an Automated Probabilistic Machine Learning Identification Algorithm for Forensic Application," *Malaysian Journal of Fundamental and Applied Sciences*, vol. 19, no. 4, 2023, doi: 10.11113/mjfas.v19n4.2917.

[25] L. Fitriani, D. Tresnawati, and M. B. Sukriyansah, "Image Classification On Garutan Batik Using Convolutional Neural Network with Data Augmentation," *JUITA : Jurnal Informatika*, vol. 11, no. 1, 2023, doi: 10.30595/juita.v11i1.16166.

[26] S. Zhang and Y. Gao, "Hybrid multi-objective evolutionary model compression with convolutional neural networks," *Results in Engineering*, vol. 21, 2024, doi: 10.1016/j.rineng.2024.101751.

[27] Sunardi, A. Yudhana, and M. Fahmi, "SVM-CNN Hybrid Classification for Waste Image Using Morphology and HSV Color Model Image Processing," *Traitement du Signal*, vol. 40, no. 4, pp. 1763–1769, Aug. 2023, doi: 10.18280/ts.400446.

[28] W. Zhang, R. Chen, and B. Wang, "A robust watermarking algorithm against JPEG compression based on multiscale autoencoder," *IET Image Process*, vol. 18, no. 2, 2024, doi: 10.1049/ipr2.12961.

Vol. 6, No. 4, August 2025, Page. 1729-1744
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4827

[29] K. N. Sunil Kumar, G. B. Arjun Kumar, R. Gatti, S. Santosh Kumar, D. A. Bhyratae, and S. Palle, "Design and implementation of auto encoder based bio medical signal transmission to optimize power using convolution neural network," *Neuroscience Informatics*, vol. 3, no. 1, p. 100121, Mar. 2023, doi: 10.1016/j.neuri.2023.100121.

[30] S. A. Holguin-Garcia *et al.*, "A comparative study of CNN-capsule-net, CNN-transformer encoder, and Traditional machine learning algorithms to classify epileptic seizure," *BMC Med Inform Decis Mak*, vol. 24, no. 1, 2024, doi: 10.1186/s12911-024-02460-z.

[31] O. B. Ozdemir and A. Koz, "3D-CNN and Autoencoder-Based Gas Detection in Hyperspectral Images," *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 16, 2023, doi: 10.1109/JSTARS.2023.3235781.

[32] A. Yasar, "Analysis of selected deep features with CNN-SVM-based for bread wheat seed classification," *European Food Research and Technology*, vol. 250, no. 6, 2024, doi: 10.1007/s00217-024-04488-x.

[33] A. I. Jabbooree, L. M. Khanli, P. Salehpour, and S. Pourbahrami, "Geometrical Facial Expression Recognition Approach Based on Fusion CNN-SVM," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 1, 2024, doi: 10.22266/ijies2024.0229.40.

[34] R. F. Alya, M. Wibowo, and P. Paradise, "Classification of Batik Motif Using Transfer Learning on Convolutional Neural Network (CNN)," *Jurnal Teknik Informatika (Jutif)*, vol. 4, no. 1, 2023, doi: 10.52436/1.jutif.2023.4.1.564.

[35] M. M. A. Wona *et al.*, "Klasifikasi Batik Indonesia Menggunakan Convolutional Neural Network (CNN)," *JURTI*, vol. 7, no. 2, pp. 172–179, 2023, [Online]. Available: https://www.kaggle.com/datasets/dionisiusdh/indonesianbatik-motifs.