JCTA Journal of Computing Theories and Applications ISSN:3024-9104

Review Article

A Systematic Literature Review of Robustness-Aware Batik Motif Classification: Acquisition Variability, Feature Representation, and Learning Models

Aji Priyambodo [1,2,*], R. Rizal Isnanto [1], and Ridwan Sanjaya [3]

[1] Doctoral Program of Information Systems, Postgraduate School, Diponegoro University, Semarang, 50241, Indonesia; e-mail : ajippro@students.undip.ac.id; rizal_isnanto@yahoo.com
[2] Information Systems and Technology Study Program, Semarang Institute of Technology and Business, Semarang, 50149, Indonesia; e-mail : priyambodo@itbsemarang.ac.id
[3] Department of Information Systems, Faculty of Computer Science, Soegijapranata Catholic University, Semarang, 50234, Indonesia; e-mail : ridwan@unika.ac.id
* Corresponding Author : Aji Priyambodo

Abstract: Batik motif classification has attracted growing attention in visual computing due to its role in cultural heritage preservation, textile informatics, museum documentation, and automated cataloging. Although many studies report high classification accuracy, robustness under real-world acquisition conditions remains insufficiently understood. Batik images are frequently affected by illumination variation, blur, folds, watermark overlays, wearable deformation, scale inconsistency, and background clutter, creating challenges that extend beyond conventional image-noise assumptions. Existing studies largely focus on improving classification performance, while the interactions among acquisition variability, feature representation, evaluation practice, and deployment constraints remain fragmented. This systematic literature review addresses this gap by synthesizing batik classification research through a robustness-aware perspective. Using query expansion, backward and forward citation chaining, relevance screening, and thematic coding, 116 candidate records were identified, resulting in 50 highly relevant studies for detailed analysis. The review reveals that robustness is shaped less by denoising alone than by the combined effects of acquisition conditions, representation design, evaluation realism, and deployment context. Handcrafted descriptors remain competitive for small datasets and structured motifs due to their data efficiency and interpretability, whereas deep learning models achieve the highest reported accuracy when supported by sufficient data diversity and realistic augmentation. Hybrid representations emerge as the most consistently balanced approach, combining local texture stability with higher-level abstraction across heterogeneous acquisition settings. The review further identifies recurring robustness failure patterns, including background dependency, illumination instability, motif-scale inconsistency, wearable deformation, and source-shift vulnerability. Based on these findings, a robustness-oriented research agenda is proposed, emphasizing cross-acquisition evaluation, representation-stability analysis, batik-specific robustness benchmarks, acquisition-aware augmentation, and deployable lightweight or hybrid architectures. The study contributes a domain-specific synthesis that reframes batik motif classification from an accuracy-centric task toward a robustness-aware visual recognition problem.

Keywords: Acquisition variability; Batik motif classification; Cultural heritage preservation; Feature representation; Robustness-aware classification; Textile informatics; Visual recognition; Cross-acquisition generalization.

Received: April, 29th 2026
Revised: May, 30th 2026
Accepted: June, 1st 2026
Published: June, 14th 2026

1. Introduction

Batik is a traditional Indonesian textile art that integrates wax-resist dyeing techniques, regional symbolism, and generations of accumulated craft knowledge. Since its recognition by UNESCO as an Intangible Cultural Heritage of Humanity in 2009, batik has attracted

DOI : 10.62411/jcta.16074

increasing attention not only as a cultural asset but also as an important object of digital preservation and computational analysis. Indonesia hosts a rich diversity of batik motif families, each reflecting distinct historical, philosophical, and socio-cultural traditions. Motifs such as *parang*, *kawung*, *truntum*, *mega mendung*, *nitik*, and numerous coastal variants demonstrate that batik is not a single visual category but a broad collection of texture-rich, repetitive, hierarchical, and often structurally complex patterns produced under diverse artisanal practices [1]–[4]. Consequently, automated batik motif classification has become an increasingly important research topic for cultural archiving, museum documentation, textile informatics, digital cataloging, educational applications, and e-commerce systems.

Despite substantial progress in classification performance, the primary challenge in batik motif recognition extends beyond conventional image noise. In practical settings, batik images are acquired through heterogeneous sources, including mobile photography, flat-layout product imaging, wearable capture, museum digitization, scanned archives, and online repositories. These acquisition scenarios introduce significant variability in image characteristics. Illumination inconsistency affects contrast and color constancy, folds and wearable deformation alter local geometric structures, cropping and scale variation remove contextual information, watermark overlays and compression artifacts distort texture statistics, while background clutter may dominate visual representations intended to capture motif characteristics. Such factors directly influence the stability of co-occurrence statistics, local binary patterns, gradient-based descriptors, and learned deep representations. Consequently, the central challenge is not merely noise reduction but robustness under heterogeneous acquisition conditions, making a robustness-aware perspective more appropriate than a narrowly defined noise-aware perspective when examining batik motif classification [5]–[9].

A second challenge concerns the fragmented methodological landscape of the field. Existing studies vary substantially in their emphasis on preprocessing, handcrafted descriptors, transfer learning, data augmentation, hybrid feature representations, and lightweight deployment-oriented models. However, many studies continue to rely on homogeneous single-source datasets, accuracy-centric evaluation, and limited cross-dataset validation. As a result, reported performance often reflects dataset consistency rather than genuine robustness and generalization capability. The key research question in contemporary batik informatics is therefore no longer limited to maximizing classification accuracy, but rather understanding how robust visual representations can be constructed, maintained, and evaluated under realistic acquisition variability [10]–[19], [20]–[29], [30]–[32].

Motivated by these challenges, this study presents a systematic literature review of robustness-aware batik motif classification through an analytical framework that connects acquisition variability, feature representation, learning models, evaluation realism, and deployment considerations. Rather than treating preprocessing as the primary analytical focus, this review examines robustness as an emergent property arising from the interaction among data acquisition conditions, representational choices, learning architectures, and evaluation protocols. The main contributions of this review are as follows:

- Reframing batik motif classification from a predominantly noise-centric perspective toward a robustness-aware perspective centered on acquisition variability, representation stability, and evaluation realism.
- Synthesizing and critically comparing handcrafted, hybrid, deep, and lightweight learning approaches with respect to robustness trade-offs, data requirements, deployment suitability, and likely failure modes.
- Identifying recurring robustness challenges in batik classification, including illumination instability, background dependency, motif-scale inconsistency, wearable deformation, and source-shift vulnerability, while highlighting limitations of current evaluation practices.
- Formulating a future research agenda focused on cross-acquisition generalization, representation stability, robustness-oriented evaluation, domain-specific benchmarking, and deployment-ready batik informatics systems.

The remainder of this paper is organized as follows. Section 2 describes the review methodology, including literature identification, screening procedures, and thematic coding strategies. Section 3 presents the characteristics of the reviewed corpus. Section 4 synthesizes the literature from the perspectives of acquisition variability, feature representation, and robustness challenges. Section 5 compares major families of learning approaches and discusses their robustness trade-offs and deployment implications. Section 6 outlines key research gaps and future directions. Finally, Section 7 concludes the review and summarizes the principal insights derived from the analysis.

2. Literature Review

2.1. Batik as a Classification Target: Cultural Context, Visual Structure, and Real-World Capture

From a computational perspective, batik presents a distinctive and challenging classification problem for several reasons. First, intra-class variability is inherently high because batik is traditionally handcrafted; two samples belonging to the same motif family may differ in scale, stroke thickness, orientation, color composition, and local regularity. Second, inter-class similarity is often substantial among related motif families, requiring fine-grained recognition systems to distinguish subtle differences in repetitive local structures. Third, the number of recognized batik categories is large and regionally diverse, resulting in multi-class classification problems characterized by class imbalance and partially overlapping visual signatures [3], [18], [19]. At the same time, batik is a culturally significant artifact rather than a generic texture benchmark. Consequently, classification errors may affect digital preservation, museum curation, cultural documentation, and educational applications, extending the implications of model performance beyond purely technical evaluation.

The acquisition environments for batik images are equally diverse. Museum collections are commonly digitized using controlled scanners or professional photography, whereas e-commerce platforms frequently rely on consumer devices operating under inconsistent illumination conditions. Artisan archives may contain folded, partially occluded, or aged textiles, while wearable scenarios introduce deformation, perspective distortion, and pose variation. As a result, the same motif may appear visually stable to human observers while exhibiting substantial variation in measurable image characteristics across acquisition settings. Robust batik classification should therefore be viewed as a problem of preserving representation stability across heterogeneous acquisition conditions rather than merely suppressing pixel-level disturbances. In this context, robustness emerges from the interaction between acquisition conditions and representation design rather than from preprocessing alone [10], [22], [33], [34].

2.2. Acquisition Variability, Robustness, and Batik-Specific Disturbances

Classical batik classification studies frequently employ median filtering, Otsu thresholding, image resizing, region-of-interest extraction, and wavelet decomposition as preprocessing techniques [1], [10], [11], [23], [35]. Although these operations often improve classification performance, their analytical role should be interpreted carefully. Within the batik domain, preprocessing serves not only to suppress Gaussian or impulse noise but also to stabilize visual representations in the presence of illumination inconsistency, background interference, scale variation, and partial deformation. Kusanti et al. [10], for example, demonstrated that resize normalization and ROI cropping reduce the influence of irrelevant backgrounds prior to VGG-16 inference, while wavelet-based approaches [11], [35] suggest that multi-resolution decomposition can preserve salient motif structures under varying acquisition conditions.

The broader computer vision literature further clarifies the distinction between denoising and robustness. Tian et al. [5] and Mao et al. [7] provide comprehensive reviews of image denoising methods, Rawat and Wang [6] survey CNN-based image classification, and Humeau-Heurtier [8], [9] review color and texture descriptors. While these studies offer valuable theoretical foundations, they are not sufficiently domain-specific to explain why batik classifiers may fail under wearable deformation, watermark overlays, illumination shifts, or cross-source acquisition variability. Batik requires a domain-aware interpretation of robustness because motif identity is encoded through repetitive, hierarchical, and texture-rich structures whose visual cues may be disrupted differently by folds, clutter, scale changes, and local occlusion. This observation suggests that robustness in batik classification cannot be fully inferred from generic image-processing literature but must be interpreted in relation to the unique visual characteristics of batik motifs.

The major gap is therefore not simply the absence of stronger denoising algorithms, but rather the lack of controlled and batik-specific robustness analysis. Most reviewed studies adopt a fixed preprocessing configuration while varying feature descriptors or classifiers, making it difficult to determine whether performance improvements originate from input stabilization, feature representation, or learning model selection. A stronger research agenda would compare methods under matched acquisition conditions, explicitly model batik-specific disturbances, and evaluate whether performance gains persist across heterogeneous datasets rather than within a single curated split [1], [10], [11], [23], [36]–[38]. Addressing this gap requires evaluation protocols capable of disentangling the effects of preprocessing, feature representation, and learning models under comparable acquisition settings.

2.3. Feature Representation: Why Handcrafted, Hybrid, and Deep Models Persist

Feature representation constitutes the central mechanism through which acquisition variability is translated into either stable or unstable classification outcomes. Consequently, understanding robustness in batik classification requires examining how different representation strategies preserve motif information under varying acquisition conditions. Handcrafted descriptors remain foundational in batik classification because batik motifs are inherently texture-rich and because carefully designed feature engineering often remains competitive when labeled datasets are limited. The Grey-Level Co-occurrence Matrix (GLCM) [39] captures second-order spatial statistics, including contrast, correlation, energy, and homogeneity, that align closely with the repetitive structural regularities characteristic of many batik motifs. As a result, GLCM appears across a broad range of approaches, including standalone handcrafted pipelines, hybrid CNN systems, and decision-tree-based classifiers [12], [14], [21], [30], [35], [40], [41].

Local Binary Patterns (LBP) [42] and related variants describe local micro-textures through neighborhood intensity comparisons that are relatively robust to monotonic gray-level transformations. Multiple studies have shown that LBP-based representations achieve competitive performance under limited data availability and moderate image degradation [13], [25]–[28], [43]–[49]. Similarly, the Histogram of Oriented Gradients (HOG) [9] emphasizes gradient orientation statistics and is particularly effective for motif families characterized by strong geometric structures and directional regularity. Wavelet-based descriptors [11], [50] provide multi-resolution representations capable of separating coarse structural information from fine-grained details, making them suitable for the inherently multi-scale nature of batik patterns. Gabor filters [51] further contribute orientation-selective frequency analysis that captures directional periodicities often present in traditional motifs. Taken together, these handcrafted descriptors represent different strategies for preserving motif information under acquisition variability. Some emphasize local texture regularity, others capture structural organization or scale-dependent patterns, but all seek to maintain representation stability despite changes in image quality and acquisition conditions.

Deep representations differ fundamentally from handcrafted descriptors because they learn hierarchical abstractions directly from data rather than relying on predefined texture statistics. Their primary advantage lies not merely in higher representational capacity, but in their ability to integrate motif information across multiple spatial scales and levels of abstraction. Elvitaria et al. [14] proposed an ensemble framework combining GLCM descriptors with deep learning predictions, consistently outperforming either component in isolation. Dzulqarnain et al. [15] demonstrated that deep convolutional autoencoders can learn discriminative batik representations without explicit feature engineering, while Anggoro et al. [2] applied ResNet-based architectures to Solo batik and reported classification accuracies exceeding 97% when supported by appropriate preprocessing and augmentation strategies. Broader surveys by Alzubaidi et al. [52] and Khan et al. [53] similarly report that deep models consistently outperform handcrafted baselines on large-scale visual recognition benchmarks, with transfer learning proving particularly effective under domain-shifted conditions. However, the advantages of deep representations are often contingent upon dataset diversity, augmentation realism, and the extent to which training data reflect real-world acquisition variability. Consequently, superior benchmark performance does not necessarily imply stronger robustness when models are deployed beyond the conditions under which they were trained.

2.4. Learning Models, Evaluation Realism, and Deployment Constraints

The landscape of learning models used in batik classification spans both classical machine learning and modern deep learning paradigms. Support Vector Machines (SVM) [54] remain attractive for high-dimensional handcrafted feature spaces due to their strong margin-based generalization properties and robustness in limited-data settings. Random Forest [55] and k-Nearest Neighbor (KNN) [56] classifiers continue to serve as competitive baselines when paired with carefully engineered features, while decision trees [21] offer complete interpretability in applications where human-readable decision rules are desirable. Consequently, classical learning models remain relevant not because they consistently outperform deep architectures, but because they often provide favorable trade-offs among interpretability, computational efficiency, and robustness on relatively small datasets.

Transfer learning has emerged as the dominant strategy for maximizing classification performance. Architectures such as VGG-16, ResNet-50 [57], DenseNet [58], and MobileNetV2 [59] are commonly initialized using ImageNet pretraining and subsequently fine-tuned on batik datasets. This approach leverages generic visual representations learned from large-scale datasets while adapting higher-level features to motif-specific characteristics. MobileNetV2, particularly when combined with quantization-aware training [16], [59], reflects a growing emphasis on deployment feasibility and resource-efficient inference rather than benchmark accuracy alone.

Nevertheless, improved benchmark performance should not automatically be interpreted as stronger robustness, particularly when evaluation is conducted on homogeneous datasets collected under narrowly controlled acquisition conditions. Models may learn acquisition-specific characteristics rather than motif-specific representations, thereby overestimating their ability to generalize across real-world settings.

Evaluation realism remains one of the least developed dimensions of the batik classification literature. Accuracy-only reporting on single train-test splits continues to dominate experimental practice, limiting the ability to assess robustness and generalization. Rawat and Wang [6] and Tian et al. [5] note that high-quality computer vision research increasingly emphasizes comprehensive evaluation protocols incorporating per-class metrics, cross-dataset validation, efficiency analysis, and robustness testing. Such practices have yet to become standard in batik classification studies.

Future research should therefore adopt a minimum evaluation framework that includes per-class precision, recall, and F1-score, macro-F1 metrics, confusion-matrix analysis, at least one cross-dataset validation experiment, and computational efficiency measurements. Without such evaluation protocols, it remains difficult to determine whether reported performance improvements reflect genuine robustness gains or merely improved performance under narrowly controlled acquisition conditions.

3. Review Method

3.1. Review Design and Protocol

This study follows a structured systematic literature review (SLR) design informed by PRISMA 2020 [60], Snyder [61], Okoli [62], Pare et al. [63], and Webster and Watson [64]. The review is guided by the following research question: " How do acquisition conditions, feature representations, and learning models interact to influence robustness in batik motif classification?" To address this question, the review adopts a robustness-oriented analytical framework that integrates three complementary dimensions:

- Acquisition variability and preprocessing strategies, encompassing both image-level stabilization and robustness-oriented data preparation;
- Feature representation families, including handcrafted, hybrid, deep, and lightweight representations;
- Learning-model trade-offs, emphasizing robustness, evaluation realism, generalization capability, and deployment implications rather than classification accuracy alone.

These dimensions are conceptually linked through the framework illustrated in Figure 1, which positions acquisition variability as the primary source of representation instability and robustness challenges in batik motif classification. Because the retrieved literature reports heterogeneous datasets, motif taxonomies, feature pipelines, and evaluation protocols, a quantitative meta-analysis would not yield meaningful pooled effect estimates. Consequently, this study adopts a qualitative and thematic synthesis approach aimed at identifying recurring patterns, robustness challenges, methodological gaps, and future research directions.

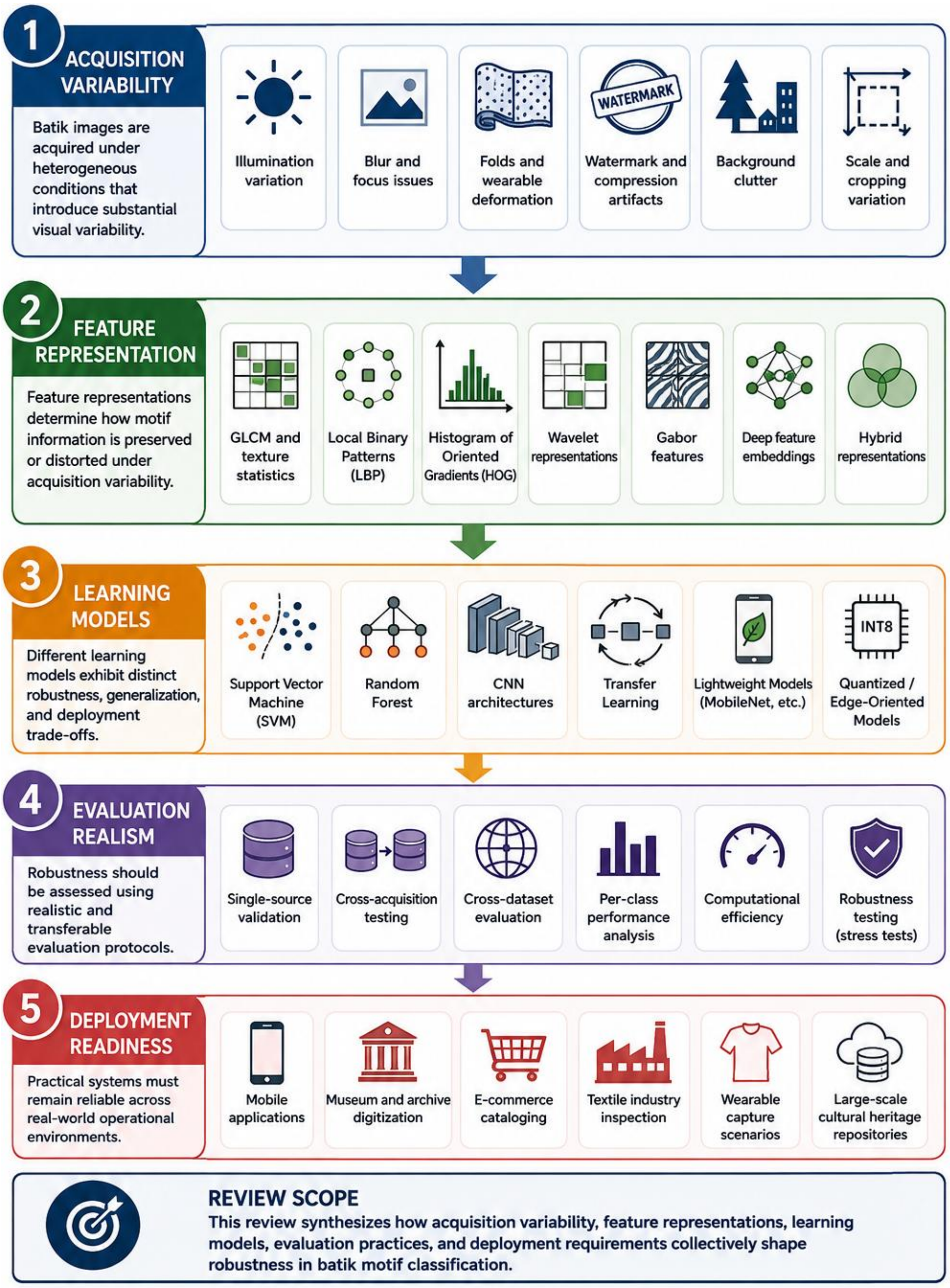


Figure 1. Robustness-oriented conceptual framework linking acquisition variability, representation instability, robustness challenges, evaluation inconsistency, and deployment limitations in batik motif classification.

3.2. Source Collection, Screening, and Corpus Construction

The literature corpus was constructed through a multi-stage search strategy combining query expansion, backward citation chaining, and forward citation chaining. Search terms were organized into three semantic clusters corresponding to the analytical dimensions of the review:

- Batik classification terms: batik classification, batik motif classification, batik image recognition, batik pattern identification, and traditional textile classification;
- Robustness and acquisition terms: acquisition variability, domain shift, illumination, blur, clutter, watermark, preprocessing, and denoising;

- Feature and learning-model terms: texture descriptors, GLCM, LBP, HOG, CNN, transfer learning, and lightweight models.

Representative search strings included:

- "batik motif classification" AND (LBP OR GLCM OR CNN);
- "batik image recognition" AND (illumination OR blur OR preprocessing);
- "batik pattern classification" AND (texture descriptor OR transfer learning);
- "traditional textile classification" AND (feature extraction OR domain shift OR lightweight model).

Searches were conducted using Google Scholar, IEEE Xplore, Scopus, and the Indonesian academic indexing portal (portal.garuda.kemdikbud.go.id) to capture both international and nationally indexed publications. The review prioritized peer-reviewed journal and conference papers published between 2015 and 2025, written in either English or Indonesian, and containing sufficient methodological detail to identify datasets, feature representations, learning models, and evaluation settings. Studies were excluded when they:

- represented duplicate records,
- lacked accessible full text,
- focused primarily on batik generation, retrieval, segmentation, or craft-process analysis rather than motif classification,
- contained only peripheral batik relevance,
- or lacked sufficient methodological information to support thematic coding.

The initial direct search yielded 35 candidate studies. Backward citation chaining identified an additional 49 studies, while forward citation chaining contributed 32 studies, resulting in 116 candidate records. After title, abstract, and duplicate screening, 114 studies were retained as broadly relevant to the review topic. Subsequent full-text assessment identified 50 highly relevant studies suitable for detailed thematic synthesis and coding.

Within this final synthesis corpus, batik-specific empirical studies constitute the analytical core of the review. Broader computer-vision, denoising, and texture-analysis surveys were retained only when they provided transferable concepts related to robustness, feature representation, or evaluation methodology. Maintaining this distinction is important because the objective of the review is not to reproduce a generic computer-vision survey, but to develop a domain-specific understanding of robustness challenges in batik motif classification.

3.3. Synthesis Procedure and Coding Scheme

Each of the 50 highly relevant studies was coded according to:

- acquisition conditions and preprocessing strategies;
- feature-representation family (handcrafted, hybrid, deep, or lightweight);
- learning model and classification approach;
- evaluation metrics and validation protocols;
- dataset provenance and sample characteristics;
- reported performance outcomes;
- stated limitations, failure modes, and future research directions.

The resulting codes were aggregated through iterative thematic analysis to identify higher-level patterns across the literature. The synthesis was organized around five analytical themes: (1) acquisition robustness, (2) representation stability, (3) evaluation realism, (4) deployment feasibility, (5) methodological maturity.

When multiple studies converged on similar observations, the resulting pattern was treated as a synthesized finding. Conversely, where evidence remained fragmented, conclusions were explicitly presented as analytical interpretations rather than universally established facts. This distinction is particularly important because robustness-related evidence remains unevenly distributed across the batik literature. Many studies report classification performance without explicitly analyzing acquisition variability, representation stability, or deployment constraints. Consequently, the present review emphasizes synthesis and interpretation rather than simple aggregation of reported accuracies. To position the contribution of the present study, Table 1 compares the review against adjacent survey streams and representative literature categories.

Table 1. Positioning of the present review relative to adjacent survey streams and batik-specific empirical literature.

Literature Stream	Representative References	Primary Focus	Limitation Relative to This Review
General denoising surveys	[5], [7]	Image denoising taxonomies and restoration methods	Do not address batik-specific acquisition variability, motif semantics, representation stability, or evaluation realism.
General CNN and image-classification surveys	[6], [52], [53], [65]	Deep architectures and large-scale image-classification practices	Not centered on batik datasets, texture-specific robustness challenges, or cultural-domain deployment.
General texture and descriptor surveys	[8], [9]	Texture descriptors, feature families, and visual pattern analysis	Provide theoretical foundations but do not offer batik-specific robustness synthesis.
Present review	Batik-focused corpus [1], [2], [16]–[23], [29], [30], [3], [33], [35], [4], [10]–[15] plus adjacent surveys	Acquisition variability, representation stability, evaluation realism, and deployment implications in batik motif classification.	Provides the robustness-oriented synthesis and research agenda missing from adjacent survey streams.

Table 1 highlights that the novelty of the present review does not lie in introducing a new classification algorithm, but in synthesizing how acquisition variability, feature representation, learning models, evaluation practices, and deployment considerations collectively shape robustness in batik motif classification. This robustness-oriented perspective forms the analytical foundation for the subsequent synthesis and discussion sections.

4. Findings and Thematic Synthesis

4.1. Descriptive Overview of the Reviewed Corpus

The 50 highly relevant studies reveal that batik motif classification has evolved from predominantly handcrafted texture-analysis pipelines toward increasingly diverse combinations of hybrid and deep-learning approaches. However, the literature has not converged on a single dominant paradigm because the relative strengths of different methods remain highly dependent on acquisition conditions, dataset characteristics, deployment constraints, and robustness requirements. As illustrated in Figure 2 and Figure 3, the field continues to exhibit substantial methodological diversity, reflecting the broader challenge of balancing accuracy, representation stability, and practical deployment considerations.

More importantly, the reviewed literature indicates that robustness-aware batik classification remains a relatively specialized research area. While classification accuracy is frequently reported, explicit investigations of acquisition variability, cross-acquisition generalization, deployment readiness, and robustness evaluation remain comparatively uncommon. This imbalance suggests that methodological development has progressed more rapidly than robustness assessment, creating a gap between benchmark performance and real-world applicability.

The composition of the corpus also provides important context for interpreting the findings of this review. Approximately two-thirds of the detailed synthesis corpus consists of batik-specific empirical studies, while the remaining studies provide complementary theoretical foundations in areas such as image denoising, texture representation, machine learning, and evaluation methodology. These supporting studies are not intended to replace the batik-focused evidence base; rather, they provide transferable concepts that help interpret robustness, representation stability, and evaluation realism within the batik domain.

This evidentiary structure establishes the scope and limitations of the review. Conclusions regarding the strengths, weaknesses, robustness characteristics, and deployment implications of different approaches are derived primarily from batik-specific studies and subsequently interpreted using broader computer-vision literature. Accordingly, statements concerning the relative stability, transferability, or robustness of particular method families should be understood as synthesized observations across heterogeneous studies rather than universally established causal relationships.

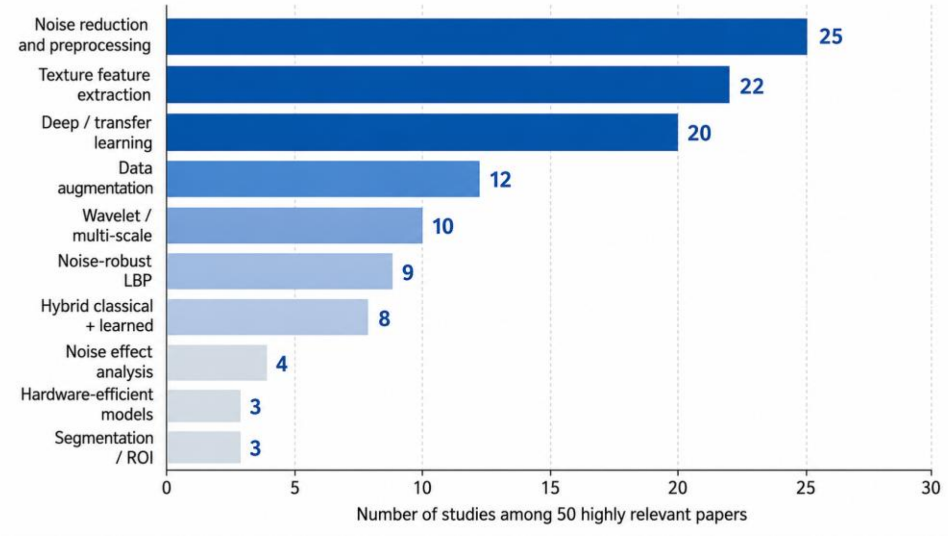


Figure 2. Thematic prevalence across the 50 highly relevant studies included in the synthesis corpus.

Figure 2 further illustrates the thematic distribution of the reviewed literature. Acquisition variability and preprocessing appear in 25 of the 50 studies (50%), making them the most frequently discussed theme and highlighting the continued importance of image quality, capture conditions, and input stabilization. Feature-representation research remains equally prominent, with texture descriptors appearing in 22 studies (44%), indicating that handcrafted approaches continue to play a significant role despite the increasing adoption of deep learning. Deep and transfer-learning methods appear in 20 studies (40%), reflecting substantial growth in data-driven approaches, although not complete displacement of classical techniques. In contrast, deployment-oriented studies, robustness evaluation, and cross-acquisition analysis remain comparatively rare, suggesting that the field continues to emphasize performance optimization more strongly than robustness validation.

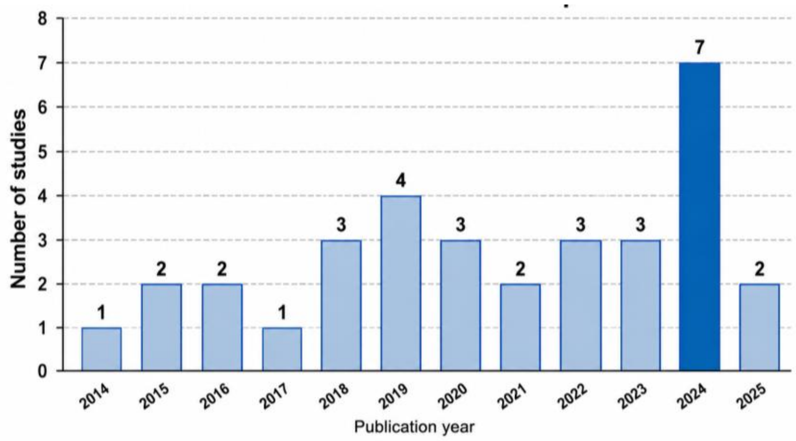


Figure 3. Publication-year distribution of representative batik classification studies (2014–2025).

The temporal distribution shown in Figure 3 reveals a clear acceleration of research activity after 2019, culminating in the highest publication density in 2024. Early studies were dominated by handcrafted texture descriptors and conventional machine-learning classifiers. The subsequent period introduced transfer learning, data augmentation, and deeper convolutional architectures, reflecting broader trends in computer vision research. More recent studies demonstrate increasing methodological diversification through hybrid representations, lightweight architectures, FPGA deployment, and discussions of robustness-related challenges.

Despite this methodological expansion, evaluation practices have not evolved at the same pace. Cross-dataset validation, acquisition-aware robustness testing, failure-mode analysis, and batik-specific benchmarking protocols remain relatively uncommon. Consequently, the literature appears to be advancing more rapidly in terms of model diversity than evaluation maturity. This observation reinforces one of the central findings of the present review: robustness remains a substantially less explored dimension than classification accuracy within current batik motif classification research. The principal themes identified from the reviewed corpus and their analytical significance are summarized in Table 2.

Table 2. High-level themes identified in the reviewed corpus and their analytical interpretations.

Theme	Frequency (of 50)	Analytical Interpretation
Acquisition variability and preprocessing	25	Most frequently discussed theme, indicating that image quality, acquisition conditions, and representation stabilization remain central concerns
Texture feature representation	22	Handcrafted descriptors remain important because batik motifs are inherently texture-rich and many datasets remain relatively small
Deep and transfer learning	20	Achieve the highest reported performance, but robustness remains strongly dependent on data diversity and evaluation realism
Data augmentation	12	Frequently used to mitigate limited training data, although often without explicit robustness-oriented validation
Wavelet and multi-scale representation	10	Important for capturing motif structures that span multiple spatial and frequency scales
Noise-robust LBP variants	9	Reflect continued efforts to improve representation stability under image degradation and acquisition variability
Hybrid handcrafted–deep representations	8	Emerging strategy that combines complementary strengths of classical descriptors and learned representations
Noise-effect analysis	4	Formal investigation of robustness under controlled disturbances remains relatively uncommon
Hardware-efficient models	3	Indicates growing interest in deployment-oriented and resource-constrained applications
Segmentation and ROI extraction	3	Applied selectively to improve motif localization and reduce background interference

Taken together, the thematic distribution highlights a field in transition. While deep learning and hybrid approaches are increasingly prominent, handcrafted representations, preprocessing strategies, and texture-aware feature engineering continue to play a significant role. More importantly, the comparatively low frequency of robustness-oriented evaluation themes suggests that future research should prioritize acquisition-aware validation, representation stability analysis, and deployment-oriented assessment to complement ongoing advances in classification accuracy.

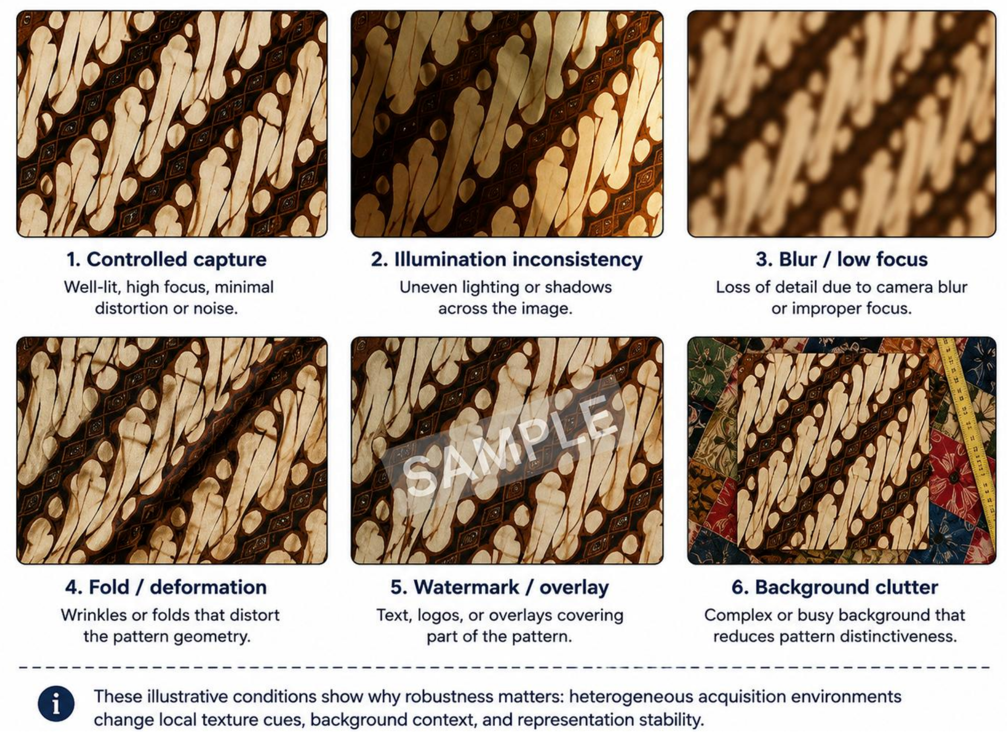
4.2. Thematic Finding 1: Acquisition Stabilization Improves Performance but Does Not Guarantee Robustness

The first major finding is that preprocessing plays a stabilizing role, but it is rarely sufficient to ensure robustness on its own. Median filtering, wavelet decomposition, Otsu thresholding, image normalization, and region-of-interest extraction are frequently employed because they reduce nuisance variation before feature extraction or model inference [1], [10], [11], [23], [35]. However, the reviewed literature does not support the stronger claim that denoising or preprocessing alone can solve the batik classification problem. Rather, the central challenge lies in maintaining stable motif representations when the same motif is observed under heterogeneous acquisition conditions.

This distinction is important because robustness emerges from multiple interacting components rather than from preprocessing alone. While input stabilization can reduce unwanted variability, classification performance ultimately depends on whether the underlying representation remains discriminative despite residual changes in illumination, blur, deformation, scale variation, and background interference. In this sense, preprocessing should be viewed as one element within a broader robustness pipeline that also includes data curation, feature representation, learning-model design, augmentation strategy, and evaluation protocol.

A useful conceptual distinction can therefore be made between input-level stabilization and representation-level robustness. Input-level stabilization seeks to reduce unwanted variation before feature extraction, whereas representation-level robustness concerns the ability of a descriptor or learned feature to preserve motif identity despite remaining variability. The reviewed literature increasingly combines both perspectives. ROI extraction and resizing reduce background interference, robust local descriptors mitigate sensitivity to localized perturbations, and augmentation or transfer learning helps learned representations tolerate greater appearance variation across acquisition settings.

The importance of this distinction becomes particularly evident in real-world batik imagery, where the same motif may be captured under substantially different conditions while retaining its semantic identity. Consequently, robustness should be understood as the preservation of representation stability under acquisition variability rather than merely the suppression of image noise. Figure 4 illustrates representative batik-like acquisition conditions that motivate this broader robustness-oriented perspective.



These illustrative conditions show why robustness matters: heterogeneous acquisition environments change local texture cues, background context, and representation stability.

Figure 4. Illustrative batik-like acquisition conditions highlighting the influence of illumination inconsistency, blur, deformation, watermark overlays, and background clutter on representation stability.

4.3. Thematic Finding 2: Representational Diversity Reflects Different Robustness Trade-offs

The second major finding concerns the central role of feature representation in mediating robustness under acquisition variability. Across the reviewed literature, differences in classification performance are often explained not only by the choice of learning model but also by how motif information is represented and preserved across changing acquisition conditions. Batik motifs are inherently hierarchical visual structures composed of local stroke primitives, repeated micro-textures, meso-scale repetitions, and larger compositional arrangements. Consequently, feature representation cannot be reduced to a simple distinction between handcrafted and deep approaches. Instead, it must be understood in terms of how effectively different representations capture local versus global information, preserve multiscale motif characteristics, and maintain representation stability when confronted with illumination variation, blur, deformation, cropping, and background interference.

This perspective helps explain the continued relevance of handcrafted descriptors despite the rapid expansion of deep learning. Local descriptors such as LBP, GLCM, and patch-based texture statistics remain effective because they directly encode repetitive micro-patterns, edge transitions, and short-range structural regularities that are characteristic of many batik motifs. Their strengths are particularly evident in relatively small datasets where data-efficient representations often outperform more complex models. However, the same locality that gives these descriptors their discriminative power can also become a source of vulnerability. Changes in illumination, blur, crease boundaries, and localized deformation may alter neighborhood relationships and texture statistics even when the semantic identity of the motif remains unchanged. As a result, the robustness of handcrafted representations depends not only on descriptor design but also on how effectively preprocessing and acquisition conditions preserve the visual structures upon which those descriptors rely.

At the opposite end of the representational spectrum, global and learned representations are generally better suited to capturing motif layout, long-range repetition, and broader structural organization. Deep feature hierarchies can aggregate information across multiple spatial scales and model nonlinear relationships that are difficult to encode using handcrafted statistics alone. This capacity becomes particularly valuable for motifs whose class identity emerges from interactions among multiple visual regions rather than from isolated texture elements. Nevertheless, higher-level representations introduce their own robustness challenges. Because they often rely on broader contextual information, they may become sensitive to cropping, partial visibility, wearable deformation, background clutter, or acquisition-specific characteristics that are only indirectly related to motif identity. In homogeneous datasets, deep models may therefore achieve impressive classification accuracy while implicitly learning source-dependent cues such as color balance, imaging conditions, camera characteristics, or recurring backgrounds rather than genuinely transferable motif representations.

A recurring pattern across the reviewed studies is that neither local nor global representations are sufficient in isolation because batik motifs simultaneously exhibit fine-grained texture detail and large-scale structural organization. Fine wax-line patterns contribute to motif identity, but so do repeated compositional units, directional rhythm, and inter-region arrangement. This observation highlights the importance of multi-scale representation as a robustness strategy. Representations that aggregate evidence across multiple levels of visual granularity are generally more resilient because they do not depend exclusively on a single set of visual cues that may be disrupted under varying acquisition conditions. In this regard, representation stability emerges as a more informative concept than classification accuracy alone, since stable representations are more likely to preserve motif identity when acquisition conditions change.

This interpretation also helps explain the growing interest in hybrid representations. Their value extends beyond simply combining handcrafted and learned features. Rather, hybrid approaches integrate complementary robustness mechanisms that operate at different representational levels. Handcrafted descriptors preserve domain-relevant local texture evidence, while learned representations capture broader structural abstractions and higher-order semantic relationships. As a result, hybrid systems frequently offer a favorable balance between data efficiency, representation stability, and abstraction capacity, particularly in datasets characterized by limited scale or heterogeneous acquisition conditions. The more meaningful analytical question is therefore not which representation family achieves the highest average performance, but which representation strategy is best aligned with a specific robustness objective, motif hierarchy, dataset scale, and deployment context.

Importantly, the comparative observations presented in this review should not be interpreted as universal rules applicable to every batik dataset. Rather, they represent recurring patterns synthesized from heterogeneous evidence across multiple studies. The conclusion that hybrid representations often provide a productive middle ground, or that deep models may overfit homogeneous acquisition environments, emerges consistently when acquisition variability, representation stability, dataset diversity, and evaluation realism are considered together. Viewed collectively, the literature suggests that future progress in batik motif classification is likely to depend less on incremental increases in model complexity and more on the development of representations that remain stable, transferable, and semantically meaningful across diverse real-world acquisition conditions.

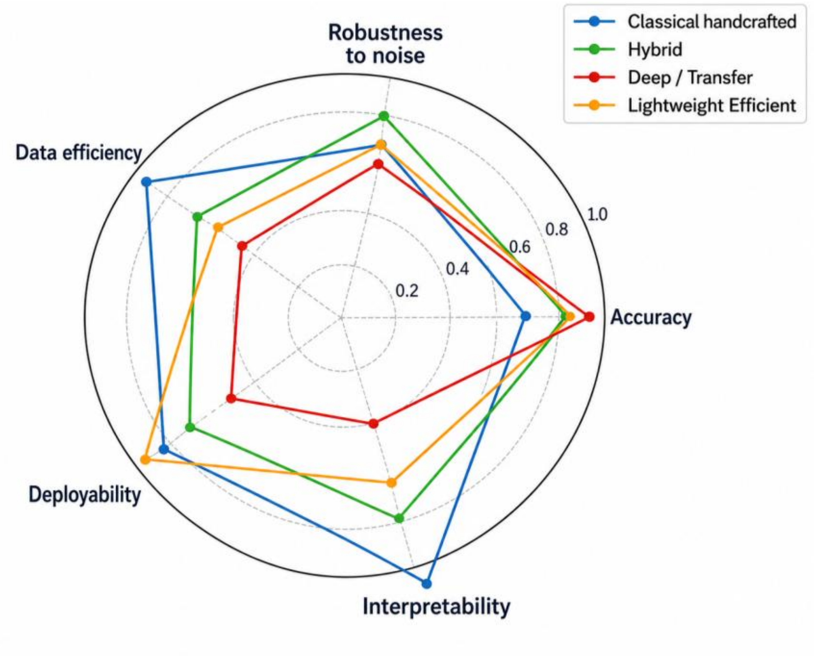


Figure 5. Multi-criteria comparison of major representation and learning-approach families across robustness, data efficiency, interpretability, deployment suitability, and evaluation realism dimensions.

4.4. Thematic Finding 3: Hybrid Representations Provide a Balanced Robustness–Complexity Trade-off

The third major finding is that hybrid representations consistently emerge as one of the most promising directions within the reviewed literature. Batik motifs combine repetitive local structures, larger compositional patterns, and culturally specific texture regularities that are not always captured effectively by a single representation family. Hybrid approaches address this challenge by combining complementary sources of information rather than relying exclusively on handcrafted or learned representations.

This complementary behavior helps explain why handcrafted descriptors remain relevant despite the rapid growth of deep learning. Handcrafted features often preserve domain-specific texture characteristics with high data efficiency, while deep representations provide richer abstraction and improved scalability. By integrating both perspectives, hybrid systems can simultaneously exploit local motif evidence and higher-level structural information.

The growing popularity of hybrid approaches also reflects practical considerations. Many batik datasets remain relatively small, imbalanced, or acquisition-dependent, conditions under which purely end-to-end deep learning may not fully realize its theoretical advantages. Consequently, hybrid systems frequently represent a pragmatic compromise between representation stability, classification performance, computational complexity, and data availability. Taken together, the reviewed studies suggest that the value of hybridization extends beyond performance improvement alone. Hybrid systems can be interpreted as an attempt to balance robustness, interpretability, and representational flexibility in environments characterized by substantial acquisition variability.

4.5. Thematic Finding 4: Evaluation Realism Remains the Main Methodological Weakness

The fourth major finding is that evaluation realism remains the most persistent methodological weakness in the current batik classification literature. Although many studies report high classification accuracy, evaluation protocols often rely on homogeneous datasets and single-source train-test splits that provide limited evidence regarding robustness and generalization. From a robustness perspective, the critical question is not whether a model performs well on a particular dataset, but whether that performance remains stable across variations in acquisition conditions, motif presentation, and dataset provenance. Robustness-aware evaluation should therefore extend beyond overall accuracy and incorporate per-class performance metrics, cross-acquisition validation, cross-dataset testing, computational efficiency measures, and explicit robustness analysis.

Without such evaluation protocols, it remains difficult to determine whether a model is genuinely learning motif identity or merely exploiting acquisition-specific regularities associated with a particular collection process. This issue is particularly important for batik classification because acquisition variability often exceeds the controlled conditions under which benchmark datasets are constructed.

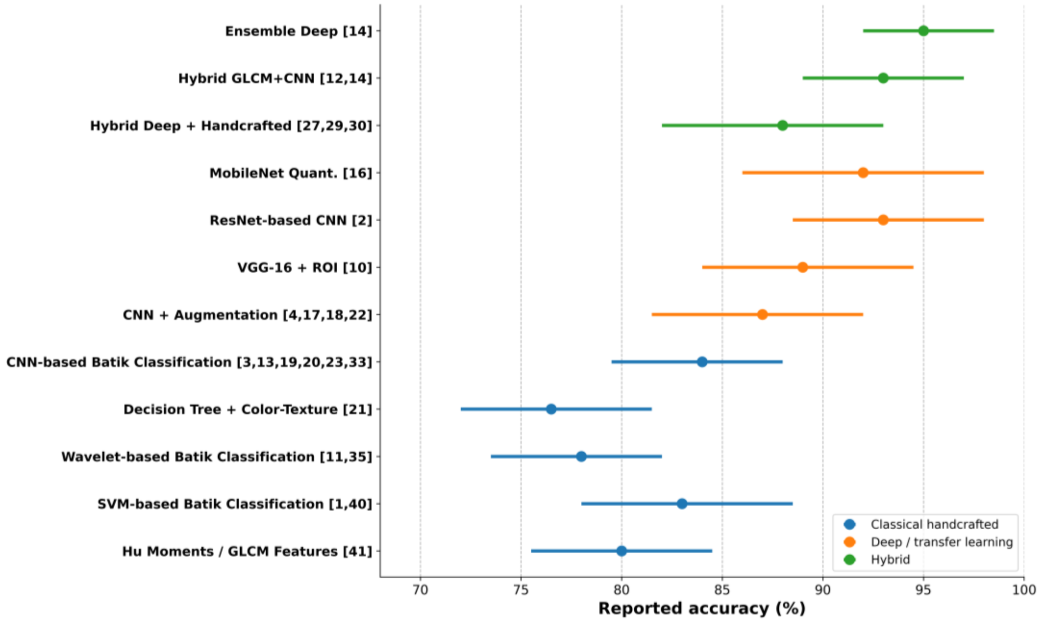


Figure 6. Representative reported accuracy values from selected batik-classification studies, presented at the study/method level; supporting citations were updated to match the revised bibliography.

Figure 6 provides a study-level visual summary of representative reported accuracies from selected batik-classification studies. It is included as an evidence-level illustration only. The broader comparative analysis across approach families, including robustness trade-offs and family-level accuracy ranges, is presented in Table 3.

5. Comparative Analysis of Learning Approaches and Robustness Trade-offs

5.1. Classical Handcrafted Pipelines

Classical handcrafted pipelines remain relevant under three recurring conditions identified across the reviewed literature: limited dataset size, constrained computational resources, and applications where interpretability remains important. Their continued use is not merely a legacy of earlier computer vision practice, but reflects the visual characteristics of batik motifs themselves. Because batik is inherently texture-rich, descriptors based on second-order statistics, local binary patterns, gradients, and multi-resolution analysis can often capture meaningful motif structure without requiring large quantities of labeled data. This data efficiency is particularly valuable given that many publicly available batik datasets remain relatively small and are often collected under homogeneous acquisition settings.

However, the robustness of handcrafted pipelines is strongly influenced by acquisition variability. Since many descriptors explicitly encode local contrast relationships, gradient distributions, or texture statistics, their performance can degrade when illumination conditions change or when motifs are affected by blur, folds, background clutter, or partial occlusion. Consequently, the effectiveness of handcrafted approaches frequently depends on careful preprocessing, acquisition control, or descriptor engineering. Their primary advantage therefore lies not in maximizing classification accuracy, but in providing interpretable and computationally efficient representations that remain useful when data availability or deployment resources are limited.

5.2. Hybrid Representation Systems

Hybrid systems occupy an increasingly important position within the reviewed literature because they attempt to reconcile the complementary strengths and weaknesses of handcrafted and learned representations. Their principal advantage is not simply higher classification accuracy, but the diversification of representational evidence across multiple levels of visual abstraction. Domain-specific handcrafted descriptors can preserve motif-relevant local texture cues that might otherwise be overlooked, while learned representations contribute broader structural and semantic information that is difficult to encode manually.

This characteristic is particularly relevant in batik classification because motif identity is distributed across multiple spatial scales. Local texture statistics contribute important discriminative information, yet larger compositional structures, repeated ornament blocks, and motif organization also play a significant role. A purely handcrafted pipeline may fail to capture higher-order relationships, whereas a purely deep representation trained on limited data may inadvertently absorb acquisition-specific artifacts. Hybrid systems reduce these risks by distributing representational responsibility across complementary sources of evidence.

Viewed from a robustness perspective, hybrid approaches frequently provide a balanced compromise between data efficiency, representation stability, and abstraction capacity. Their success does not arise from combining features indiscriminately, but from integrating representations that respond differently to acquisition variability. This explains why hybrid systems repeatedly emerge as strong performers in heterogeneous batik datasets, particularly when robustness rather than benchmark accuracy is treated as the primary objective.

5.3. Deep and Transfer Learning Models

Deep and transfer-learning models currently achieve the highest reported performance across many batik classification studies, particularly when supported by data augmentation, transfer learning, and sufficiently diverse training data. Their primary advantage lies in the ability to learn hierarchical feature representations that capture complex motif structures, long-range dependencies, and nonlinear relationships that are difficult to model through handcrafted descriptors alone. This capability is particularly beneficial for motifs whose class identity emerges from interactions among multiple visual regions rather than isolated texture elements.

Nevertheless, the reviewed literature suggests that the principal robustness challenge associated with deep learning is not overfitting in the conventional machine-learning sense, but overfitting to the acquisition regime itself. When datasets are collected under relatively homogeneous conditions, deep models may learn source-specific characteristics such as illumination style, background composition, capture distance, camera properties, or color balance together with motif identity. As a result, impressive benchmark performance may not necessarily translate into reliable generalization when acquisition conditions change.

Data augmentation partially addresses this limitation, although its effectiveness depends strongly on realism rather than quantity alone. Simple rotations, flips, and generic photometric perturbations improve invariance only when they approximate the types of variability encountered during deployment. In the context of batik classification, acquisition-aware augmentation strategies that simulate folds, blur, watermark overlays, illumination imbalance, cropping, and scale inconsistency are likely to contribute more directly to robustness than generic augmentation pipelines.

Deployment constraints introduce an additional layer of complexity. A highly accurate model may still be unsuitable for mobile heritage applications, museum digitization platforms, or embedded industrial systems if it requires excessive memory, computational resources, or inference time. Consequently, robustness-aware batik classification requires model selection criteria that jointly consider representation quality, acquisition variability, computational efficiency, and deployment feasibility rather than benchmark accuracy alone.

5.4. Lightweight and Deployment-Oriented Models

Lightweight and deployment-oriented architectures represent an emerging but increasingly important direction within batik informatics. As applications expand beyond laboratory settings toward mobile cataloging systems, museum kiosks, cultural heritage repositories, and industrial inspection platforms, computational efficiency becomes a practical requirement rather than a secondary consideration. Architectures such as MobileNet and quantized deep-learning models demonstrate that meaningful classification performance can be achieved under constrained hardware environments.

However, the current evidence base remains relatively limited. While existing studies suggest that compact models can achieve favorable accuracy–efficiency trade-offs, there is still insufficient evidence to conclude that lightweight architectures maintain robustness under heterogeneous real-world acquisition conditions. Most deployment-oriented studies continue to evaluate performance under relatively controlled settings, leaving questions regarding cross-acquisition stability, source-shift resilience, and long-term operational reliability largely unresolved. Consequently, efficiency should not be evaluated independently of robustness, but rather as one component of a broader deployment-readiness framework.

Table 3. Analytical comparison of major learning-approach families across robustness-oriented evaluation dimensions.

Approach Family	**Accuracy Range (%)**	**Sensitivity to Acquisition Variability**	**Data Efficiency**	**Deployability**	**Typical Failure Mode**	**Representative Refs**
Classical handcrafted	78–88	Moderate–High (sensitive to illumination, blur, and clutter)	High (effective on small datasets)	High (CPU/mobile feasible)	Descriptor instability under illumination variation, blur, and local deformation	[1], [21], [31], [32], [35], [44], [50], [51], [66]
Hybrid	89–96	Moderate (complementary cues improve stability)	Medium	Medium	Fusion strategy remains unresolved; performance depends on balanced integration	[14], [27], [29], [30], [67]
Deep / Transfer Learning	86–98	Variable (strong with diverse data, vulnerable under source bias)	Low (requires larger datasets)	Low–Medium	Overfitting to background, acquisition conditions, or source-specific characteristics	[2], [3], [19], [20], [33], [57]–[59], [68], [4], [12]–[18]
Lightweight Efficient Models	83–92	Moderate	Medium	High (edge, mobile, FPGA)	Reduced robustness under heterogeneous field conditions	[16], [59]

Table 3 functions not only as a comparative summary but also as a robustness-oriented decision framework. Classical handcrafted pipelines are most appropriate when dataset size is limited, interpretability is required, and deployment resources are constrained. However, their effectiveness remains highly dependent on acquisition stability because many descriptors encode explicit assumptions about texture contrast and local structure. Hybrid approaches become particularly attractive when the objective is to balance representation stability with abstraction capacity, especially in heterogeneous datasets where neither handcrafted nor learned representations are individually sufficient.

Deep and transfer-learning models offer the greatest representational flexibility and highest best-case performance, but their robustness should be viewed as conditional rather than automatic. Their success depends strongly on dataset diversity, augmentation realism, and evaluation design. Lightweight architectures provide an additional deployment-oriented alternative, although evidence regarding their long-term robustness under realistic acquisition variability remains limited.

Viewed collectively, the central trade-off among approach families concerns how robustness is achieved. Handcrafted methods obtain efficiency at the cost of sensitivity to disturbance; deep models obtain abstraction at the cost of greater data dependency and potential source bias; hybrid systems obtain stability through complementary evidence but require more sophisticated representation integration. These trade-offs reinforce a recurring theme throughout the reviewed literature: model selection should be guided by acquisition conditions, robustness requirements, deployment constraints, and failure tolerance rather than by headline accuracy alone.

5.5. Common Robustness Failure Patterns in Batik Classification

Across the reviewed literature, several robustness-related failure patterns recur frequently enough to warrant explicit synthesis. The first is background dependency, whereby models partially learn contextual information such as cloth presentation, table texture, mannequin configuration, or acquisition setup rather than motif content itself. This issue is particularly problematic because such contextual cues may disappear when the model is deployed under different acquisition conditions.

A second recurring challenge is illumination instability. Uneven lighting conditions alter color consistency, local contrast, and texture visibility, causing both handcrafted descriptors and learned representations to drift away from their intended motif-specific characteristics. A third failure pattern concerns motif-scale inconsistency, where motifs observed at one spatial scale during training are not represented similarly when captured at different distances, cropped, partially visible, or presented in wearable form.

The fourth recurring challenge is wearable deformation, including folds, bending, non-planar presentation, and self-occlusion. These effects disrupt both local texture regularity and larger compositional structures, making them particularly difficult to address through preprocessing alone. Finally, many studies implicitly reveal source-shift vulnerability, whereby performance degrades when images originate from different institutions, devices, acquisition protocols, or preprocessing pipelines. In such cases, the learned decision boundary often proves narrower than the task itself.

These recurring failure patterns help explain why a robustness-aware perspective is especially important for batik motif classification. The fundamental challenge extends beyond recognizing texture categories under controlled conditions; it involves preserving motif identity across acquisition environments that vary substantially in quality, presentation, and context. Explicitly identifying these failure modes also clarifies why different learning approaches exhibit distinct robustness characteristics, even when their reported classification accuracies appear similar.

6. Methodological Gaps and Future Research Directions

The future research agenda emerging from the reviewed literature is not simply a call for more sophisticated models, but for more mature robustness-oriented methodology. Across the reviewed studies, methodological progress has generally outpaced robustness evaluation, resulting in a literature that increasingly reports strong classification performance yet often provides limited evidence regarding stability, transferability, and deployment readiness. The central challenge therefore lies in aligning batik-specific acquisition realities, feature representation strategies, evaluation protocols, and deployment requirements within a coherent robustness framework. The following gaps summarize the principal areas where such methodological maturation remains incomplete.

6.1. Gap 1: Limited Cross-Acquisition Robustness Analysis

The most significant methodological weakness identified in the reviewed corpus is the limited use of controlled robustness analysis across acquisition conditions. Most studies evaluate performance within a single dataset or acquisition environment and rarely examine whether the same model remains reliable when confronted with variations in illumination, cropping, wearable presentation, background complexity, watermark contamination, camera characteristics, or image quality. As a result, reported performance often reflects within-source optimization rather than genuine robustness.

Future research should move beyond conventional train–test validation toward explicit cross-acquisition evaluation protocols designed to separate memorization of acquisition conditions from genuine motif recognition. Shared robustness benchmarks, controlled condition-shift experiments, and systematic reporting of performance degradation under acquisition variability would substantially strengthen the evidentiary basis of the field. Ultimately, cross-acquisition performance should be treated as a primary evaluation outcome rather than an optional extension.

6.2. Gap 2: Incomplete Understanding of Representation Stability

A second major gap concerns the limited understanding of representation stability under different disturbance conditions. Handcrafted descriptors are frequently assumed to be fragile, while deep representations are often assumed to be inherently robust, yet relatively few studies directly validate these assumptions. Most comparisons remain performance-oriented and provide limited insight into why a particular representation succeeds or fails under specific acquisition conditions.

Future work should therefore shift attention from model-level comparisons toward disturbance-specific representation analysis. Questions such as which representations remain stable under illumination shifts, which are resilient to scale variation, which fail under wearable deformation, and which depend strongly on color consistency remain largely unanswered. Addressing these questions will require controlled ablation studies, local-versus-global representation analysis, feature-attribution techniques, and representation-level robustness evaluation. Such investigations would contribute not only to higher performance but also to a deeper understanding of how motif information is preserved across heterogeneous acquisition environments.

6.3. Gap 3: Evaluation Practice Remains Too Narrow

The review also reveals that evaluation practice remains substantially narrower than the robustness challenges faced by real-world batik classification systems. Accuracy on a single train–test split continues to dominate experimental reporting despite providing only a partial view of model behavior. This practice obscures class-specific weaknesses, source-shift failures, calibration issues, and deployment-related limitations that often become visible only under more comprehensive evaluation protocols.

Future studies should adopt a minimum robustness-oriented evaluation framework that includes per-class precision, recall, and F1-score; macro-level performance measures; confusion-matrix analysis; cross-dataset or cross-source validation; calibration analysis where appropriate; and deployment-oriented metrics such as inference latency, memory consumption, and computational efficiency. Evaluation realism should be regarded as a component of methodological quality rather than a reporting detail. Without such practices, it remains difficult to determine whether performance gains reflect genuine robustness improvements or merely improved fitting to a particular dataset.

6.4. Gap 4: Batik-Specific Visual and Dataset Resources Remain Limited

The niche status of robustness-aware batik classification is reinforced by the limited availability of datasets explicitly designed to capture acquisition variability. Existing datasets frequently emphasize motif diversity while providing only limited coverage of the conditions under which batik is encountered in practice. As a consequence, many robustness-related challenges remain underrepresented during both model development and evaluation.

Future datasets should include a broader range of acquisition conditions, including flat and wearable presentation, controlled and uncontrolled illumination, blur, folds, watermark overlays, cluttered backgrounds, partial occlusion, and scale variation. Equally important, dataset documentation should explicitly describe source provenance, acquisition procedures, and preprocessing pipelines to facilitate meaningful cross-source evaluation. Such resources would enable more realistic benchmarking and help establish batik classification as a distinctive robustness problem rather than a generic texture-recognition task.

6.5. Gap 5: Deployment-Oriented Research Remains Underdeveloped

Although deployment feasibility is increasingly discussed in the literature, deployment-oriented research remains comparatively underdeveloped. Practical batik informatics applications frequently operate under conditions that differ substantially from laboratory benchmarks, including constrained hardware resources, variable illumination, intermittent connectivity, and real-time processing requirements. These constraints introduce robustness challenges that are rarely captured by conventional performance metrics.

Future research should therefore examine efficient architectures, model compression, quantization, pruning, knowledge distillation, and edge deployment strategies within realistic acquisition environments rather than under highly controlled conditions alone. Deployment-oriented studies should evaluate not only model size and computational efficiency, but also robustness under field variability. A mature research agenda will require closer integration among algorithm design, robustness evaluation, deployment constraints, and application-specific requirements, ensuring that efficiency improvements contribute to practical reliability rather than merely reducing computational cost. The principal robustness-oriented methodological gaps identified in this review and their corresponding research priorities are summarized in Table 4.

Table 4. Summary of major robustness-oriented research gaps and recommended future directions.

Gap	Why It Matters	Recommended Direction	Priority
Limited cross-acquisition testing	Within-source performance may substantially overestimate real-world robustness	Develop cross-source and cross-condition benchmarks with explicit robustness protocols	High
Limited representation-stability analysis	The robustness characteristics of different feature representations remain poorly understood	Conduct disturbance-specific ablations, representation analysis, and feature-level stability studies	High
Narrow evaluation practice	Accuracy alone obscures class-level weaknesses and source-shift failures	Report macro-F1, confusion matrices, cross-dataset validation, calibration measures, and efficiency metrics	High
Limited batik-specific datasets and visual resources	Current datasets insufficiently capture real acquisition variability	Build richer corpora including wearable, cluttered, blurred, watermarked, and multi-condition acquisitions	High
Underdeveloped deployment-oriented research	Practical applications require robustness under constrained hardware environments	Evaluate compression, quantization, distillation, and efficient architectures under realistic field conditions	Medium

7. Conclusions

This review synthesized 50 highly relevant studies on batik motif classification through a robustness-oriented perspective that integrates acquisition variability, feature representation, learning models, evaluation realism, and deployment considerations. The principal conclusion is not that a single approach family universally dominates, but that robustness emerges from the interaction among representation design, acquisition conditions, motif hierarchy, evaluation protocols, and deployment constraints. Handcrafted descriptors remain relevant because of their data efficiency and ability to capture repetitive local texture structures, particularly in small datasets and geometrically regular motifs. Deep learning models achieve the highest reported performance when supported by sufficient data diversity and effective augmentation strategies, although their apparent superiority can be overstated when evaluation is confined to homogeneous acquisition settings. Hybrid representations appear particularly promising because they combine complementary local and global evidence, frequently providing a favorable balance between representation stability, abstraction capacity, and practical feasibility.

More broadly, the reviewed literature suggests that the central challenge in contemporary batik classification is no longer simply increasing benchmark accuracy. Rather, the challenge is ensuring that motif representations remain stable, transferable, and interpretable under heterogeneous real-world acquisition conditions. The recurring failure patterns identified in this review—including background dependency, illumination instability, scale inconsistency, wearable deformation, and source-shift vulnerability—demonstrate that strong classification performance does not automatically imply robustness. In many cases, models that perform well under controlled conditions may degrade substantially when confronted with realistic acquisition variability.

Accordingly, the next stage of batik informatics research should prioritize cross-acquisition benchmarking, disturbance-specific robustness analysis, representation-stability studies, richer batik-specific datasets, and deployment-ready architectures evaluated under realistic operating conditions. Progress should therefore be assessed less by isolated best-case accuracy and more by the ability of classification systems to maintain reliable motif recognition across the diverse acquisition environments that characterize batik in practical use. From this perspective, robustness is not merely an auxiliary evaluation criterion, but the central requirement for transforming batik classification from a laboratory task into a dependable real-world technology for cultural heritage preservation, documentation, and digital textile informatics.

Author Contributions: A.P., R.R.I., and R.S.; Methodology: A.P. and R.R.I.; Software: A.P.; Validation: A.P., R.R.I., and R.S.; Formal analysis: A.P. and R.S.; Investigation: A.P.;

Resources: R.R.I. and R.S.; Data curation: A.P.; Writing original draft: A.P.; Writing review & editing: R.R.I. and R.S.; Visu-alization: A.P.; Supervision: R.R.I. and R.S.; Project administration: A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new dataset was created. The article synthesizes published studies cited in the reference list.

Acknowledgments: The authors thank the Doctoral Program of Information Systems, Postgraduate School, Diponegoro University, Semarang and Department of Information Systems, Faculty of Computer Science, Soegijapranata Catholic University, Semarang for academic support. All literature screening, selection, interpretation, synthesis, and final scholarly responsibility remain with the authors. Generative artificial intelligence tools were used only to assist with English language refinement and the creation of illustrative conceptual figures. The authors independently conducted and verified all scientific analyses, interpretations, syntheses, and conclusions presented in this review.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] R. Wiryadinata, M. R. Adli, R. Fahrizal, and R. Alfanz, "Klasifikasi 12 Motif Batik Banten Menggunakan Support Vector Machine," *J. EECCIS (Electrics, Electron. Commun. Control. Informatics, Syst.*, vol. 13, no. 1, pp. 60–64, Apr. 2019, doi: 10.21776/jeeccis.v13i1.570.

[2] D. A. Anggoro, A. A. T. Marzuki, and W. Supriyanti, "Classification of Solo Batik patterns using deep learning convolutional neural networks algorithm," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 22, no. 1, p. 232, Aug. 2023, doi: 10.12928/telkomnika.v22i1.24598.

[3] M. W. Ilahi, C. N. Apriyani, A. Desiani, N. Gofar, Y. Andriani, and M. R. Halim, "Classification of Geometric Batik Motif Typical of Indonesian Using Convolutional Neural Network," *J. Tek. Informtika*, vol. 15, no. 1, pp. 91–100, Jun. 2022, doi: 10.15408/jti.v15i1.24968.

[4] S. Aras, A. Setyanto, and Rismayani, "Classification of Papuan Batik Motifs Using Deep Learning and Data Augmentation," in *2022 4th International Conference on Cybernetics and Intelligent System (ICORIS)*, Oct. 2022, pp. 1–5. doi: 10.1109/ICORIS56080.2022.10031320.

[5] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C.-W. Lin, "Deep learning on image denoising: An overview," *Neural Networks*, vol. 131, pp. 251–275, Nov. 2020, doi: 10.1016/j.neunet.2020.07.025.

[6] W. Rawat and Z. Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review," *Neural Comput.*, vol. 29, no. 9, pp. 2352–2449, Sep. 2017, doi: 10.1162/neco_a_00990.

[7] J. Mao, L. Sun, J. Chen, and S. Yu, "Overview of Research on Digital Image Denoising Methods," *Sensors*, vol. 25, no. 8, p. 2615, Apr. 2025, doi: 10.3390/s25082615.

[8] A. Humeau-Heurtier, "Color Texture Analysis: A Survey," *IEEE Access*, vol. 10, pp. 107993–108003, 2022, doi: 10.1109/ACCESS.2022.3213439.

[9] A. Humeau-Heurtier, "Texture Feature Extraction Methods: A Survey," *IEEE Access*, vol. 7, pp. 8975–9000, 2019, doi: 10.1109/ACCESS.2018.2890743.

[10] J. Kusanti, E. Noersasongko, M. A. Soeleman, F. Alzami, Purwanto, and Z. A. Hasibuan, "Pre-Processing Of Resize And Region Of Interest (ROI) To Improve The Accuracy Of Batik Detection Based On VGG-16," in *2023 10th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, Sep. 2023, pp. 242–247. doi: 10.1109/EECSI59885.2023.10295666.

[11] A. Farkhan, Z. Arif, A. A. Murtopo, and Gunawan, "Penerapan Metode Wavelet Daubechies untuk Klasifikasi Content Based Image Retrieval Batik," *PROSISKO J. Pengemb. Ris. dan Obs. Sist. Komput.*, vol. 11, no. 2, pp. 211–220, Sep. 2024, doi: 10.30656/prosisko.v11i2.8615.

[12] A. R. Dani and I. Handayani, "Klasifikasi Motif Batik Yogyakarta Menggunakan Metode GLCM dan CNN," *J. Teknol. Terpadu*, vol. 10, no. 2, pp. 142–156, Dec. 2024, doi: 10.54914/jtt.v10i2.1451.

[13] A. Oktarino, Y. Desnita-Tasri, and A. Efendi, "Identification of Batik Motif Based Deep Learning-Convolutional Neural Network Approach," *J. Ocean. Mech. Aerosp. -science Eng.*, vol. 68, no. 3, pp. 139–147, Nov. 2024, doi: 10.36842/jomase.v68i3.381.

[14] L. Elvitaria, E. F. A. Shaubari, N. A. Samsudin, S. K. A. Khalid, S. -, and Z. Indra, "A Proposed Batik Automatic Classification System Based on Ensemble Deep Learning and GLCM Feature Extraction Method," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 10, 2024, doi: 10.14569/IJACSA.2024.0151058.

[15] 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.

[16] N. Setyawan, C.-C. Sun, W.-K. Kuo, and M.-H. Hsu, "FPGA-Based Batik Classification Using Quantization Aware Training of MobileNet and Data-flow Implementation," in *2024 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, Nov. 2024, pp. 306–310. doi: 10.1109/APCCAS62602.2024.10808389.

[17] C. Uswatun Khasanah, E. Utami, and S. Raharjo, "Implementation of Data Augmentation Using Convolutional Neural Network for Batik Classification," in *2020 8th International Conference on Cyber and IT Service Management (CITSM)*, Oct. 2020, pp. 1–5. doi: 10.1109/CITSM50537.2020.9268890.

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

[19] D. G. T. Meranggi, N. Yudistira, and Y. A. Sari, "Batik Classification Using Convolutional Neural Network with Data Improvements," *JOIV Int. J. Informatics Vis.*, vol. 6, no. 1, p. 6, Mar. 2022, doi: 10.30630/joiv.6.1.716.

[20] R. F. Alya, M. Wibowo, and P. Paradise, "Classification of Batik Motif Using Transfer Learning on Convolutional Neural Network (CNN)," *J. Tek. Inform.*, vol. 4, no. 1, pp. 161–170, Feb. 2023, doi: 10.52436/1.jutif.2023.4.1.564.

[21] A. Tejawati *et al.*, "Batik Pattern Classification Using Decision Tree Based on Color-Texture Features," in *2023 Sixth International Conference on Vocational Education and Electrical Engineering (ICVEE)*, Oct. 2023, pp. 347–352. doi: 10.1109/ICVEE59738.2023.10348243.

[22] S. Suprapto, M. N. Tentua, and A. R. Maulana, "Optimizing nitik batik classification through comparative analysis of image augmentation," *IAES Int. J. Artif. Intell.*, vol. 14, no. 5, p. 3970, Oct. 2025, doi: 10.11591/ijai.v14.i5.pp3970-3981.

[23] A. Kusuma Whardana and U. Alfaruq Abdul Aziz, "Klasifikasi Motif Batik Menggunakan Convolutional Neural Network (CNN) dengan Otsu thresholding dan Median filter," *J. SISKOM-KB (Sistem Komput. dan Kecerdasan Buatan)*, vol. 8, no. 3, pp. 205–217, May 2025, doi: 10.47970/siskom-kb.v8i3.844.

[24] H. Hadizadeh, "Noise-resistant and rotation-invariant texture description and representation using local Gabor wavelets binary patterns," in *2015 The International Symposium on Artificial Intelligence and Signal Processing (AISP)*, Mar. 2015, pp. 30–34. doi: 10.1109/AISP.2015.7123521.

[25] Y. Liang, S. Lu, R. Weng, C. Han, and M. Liu, "Unsupervised noise-robust feature extraction for aerial image classification," *Sci. China Technol. Sci.*, vol. 63, no. 8, pp. 1406–1415, Aug. 2020, doi: 10.1007/s11431-020-1600-9.

[26] Q. Wang, B. Li, X. Chen, J. Luo, and Y. Hou, "Random Sampling Local Binary Pattern Encoding Based on Gaussian Distribution," *IEEE Signal Process. Lett.*, vol. 24, no. 9, pp. 1358–1362, Sep. 2017, doi: 10.1109/LSP.2017.2728122.

[27] Fawad, M. A. Asghar, A. Saeed, M. J. Khan, M. Zahid, and M. ur Rehman, "Texture Classification using a Hybrid Deep and Handcrafted Features," in *2019 UK/ China Emerging Technologies (UCET)*, Aug. 2019, pp. 1–4. doi: 10.1109/UCET.2019.8881836.

[28] Fawad *et al.*, "Texture Representation Through Overlapped Multi-Oriented Tri-Scale Local Binary Pattern," *IEEE Access*, vol. 7, pp. 66668–66679, 2019, doi: 10.1109/ACCESS.2019.2918004.

[29] M. F. Dzulqarnain, A. Fadlil, and I. Riadi, "Performance Comparison of Learned Features from Autoencoder and Shape-Based Hu Moments for Batik Classification," *J. Tek. Inform.*, vol. 6, no. 4, pp. 1729–1744, Aug. 2025, doi: 10.52436/1.jutif.2025.6.4.4827.

[30] A. E. Minarno, I. Soesanti, and H. A. Nugroho, "Batik Image Representation using Multi Texton Co-occurrence Histogram," *JOIV Int. J. Informatics Vis.*, vol. 8, no. 3–2, p. 1582, Nov. 2024, doi: 10.62527/joiv.8.3-2.3095.

[31] M. K. Nazir, M. A. N. U. Ghani, A. Ashraf, S. M. Alam, R. U. Farooq, and Z. Latif, "A Textile Image Classification based on Texture and Shape Features," in *2021 International Conference on Innovative Computing (ICIC)*, Nov. 2021, pp. 1–6. doi: 10.1109/ICIC53490.2021.9693067.

[32] A. C. Siregar and B. C. Octariadi, "Classification of Sambas traditional fabric 'kain lunggi' using texture feature," *Indones. J. Comput. Cybern. Syst*, vol. 13, no. 4, pp. 389–398, 2019, doi: 10.22146/IJCCS.49782.

[33] L. Elvitaria *et al.*, "A Data-Driven Approach for Batik Pattern Classification Using Convolutional Neural Networks (CNN)," *Semarak Int. J. Electron. Syst. Eng.*, vol. 4, no. 1, pp. 22–30, Dec. 2024, doi: 10.37934/sijese.4.1.2230.

[34] M. A. Rasyidi, R. Handayani, and F. Aziz, "Identification of batik making method from images using convolutional neural network with limited amount of data," *Bull. Electr. Eng. Informatics*, vol. 10, no. 3, pp. 1300–1307, Jun. 2021, doi: 10.11591/eei.v10i3.3035.

[35] A. E. Minarno, I. Soesanti, and H. A. Nugroho, "Batik Classification using Microstructure Co-occurrence Histogram," *JOIV Int. J. Informatics Vis.*, vol. 8, no. 1, p. 134, Mar. 2024, doi: 10.62527/joiv.8.1.2152.

[36] T. D. Pham, "Noise-Added Texture Analysis," in *Lecture Notes in Computer Science*, 2017, pp. 93–100. doi: 10.1007/978-3-319-52277-7_12.

[37] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995, doi: 10.1109/18.382009.

[38] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pp. 839–846. doi: 10.1109/ICCV.1998.710815.

[39] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Trans. Syst. Man. Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.

[40] E. Winarno, W. Hadikurniawati, A. Septiarini, and H. Hamdani, "Analysis of color features performance using support vector machine with multi-kernel for batik classification," *Int. J. Adv. Intell. Informatics*, vol. 8, no. 2, p. 151, Jul. 2022, doi: 10.26555/ijain.v8i2.821.

[41] B. P. H. K. M. D. Senarathna and R. M. T. P. Rajakaruna, "Feature Descriptor for Sri Lankan Batik Patterns Using Hu Moment Invariants and GLCM," in *2021 10th International Conference on Information and Automation for Sustainability (ICIAfS)*, Aug. 2021, pp. 197–202. doi: 10.1109/ICIAfS52090.2021.9606106.

[42] M. Pietikäinen and G. Zhao, "Two decades of local binary patterns," in *Advances in Independent Component Analysis and Learning Machines*, Elsevier, 2015, pp. 175–210. doi: 10.1016/B978-0-12-802806-3.00009-9.

[43] B. Xiao, K. Wang, X. Bi, W. Li, and J. Han, "2D-LBP: An Enhanced Local Binary Feature for Texture Image Classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 9, pp. 2796–2808, Sep. 2019, doi: 10.1109/TCSVT.2018.2869841.

[44] I. D. Mienye and T. G. Swart, "A Comprehensive Review of Deep Learning: Architectures, Recent Advances, and Applications," *Information*, vol. 15, no. 12, p. 755, Nov. 2024, doi: 10.3390/info15120755.

[45] M. H. Shakoor and R. Boostani, "Noise robust and rotation invariant texture classification based on local distribution transform," *Multimed. Tools Appl.*, vol. 80, no. 6, pp. 8639–8666, Mar. 2021, doi: 10.1007/s11042-020-10084-4.

[46] A. Saeed *et al.*, "Robustness-Driven Hybrid Descriptor for Noise-Deterrent Texture Classification," *IEEE Access*, vol. 7, pp. 110116–110127, 2019, doi: 10.1109/ACCESS.2019.2932687.

[47] T. P. Nguyen and A. Manzanera, "Noise tolerant descriptor for texture classification," in *2015 International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Nov. 2015, pp. 237–241. doi: 10.1109/IPTA.2015.7367137.

[48] S. Hu, J. Li, H. Fan, S. Lan, and Z. Pan, "Scale and pattern adaptive local binary pattern for texture classification," *Expert Syst. Appl.*, vol. 240, p. 122403, Apr. 2024, doi: 10.1016/j.eswa.2023.122403.

[49] S. Karanwal and M. Diwakar, "Triangle and orthogonal local binary pattern for face recognition," *Multimed. Tools Appl.*, vol. 82, no. 23, pp. 36179–36205, Sep. 2023, doi: 10.1007/s11042-023-15072-y.

[50] N. Anantrasirichai, J. Burn, and D. R. Bull, "Robust texture features based on undecimated dual-tree complex wavelets and local magnitude binary patterns," in *2015 IEEE International Conference on Image Processing (ICIP)*, Sep. 2015, pp. 3957–3961. doi: 10.1109/ICIP.2015.7351548.

[51] A. W. Muzaffar *et al.*, "Gabor Contrast Patterns: A Novel Framework to Extract Features From Texture Images," *IEEE Access*, vol. 11, pp. 60324–60334, 2023, doi: 10.1109/ACCESS.2023.3280053.

[52] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[53] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5455–5516, Dec. 2020, doi: 10.1007/s10462-020-09825-6.

[54] D. Mustafa Abdullah and A. Mohsin Abdulazeez, "Machine Learning Applications based on SVM Classification A Review," *Qubahan Acad. J.*, vol. 1, no. 2, pp. 81–90, Apr. 2021, doi: 10.48161/qaj.v1n2a50.

[55] J. Xiang, "A Review of Machine Learning Classification Based on Random Forest Algorithm," *Artif. Intell. Robot. Res.*, vol. 13, no. 01, pp. 143–152, 2024, doi: 10.12677/AIRR.2024.131016.

[56] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, "Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction," *Sci. Rep.*, vol. 12, no. 1, p. 6256, Apr. 2022, doi: 10.1038/s41598-022-10358-x.

[57] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, vol. 2016-Decem, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[58] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 2261–2269. doi: 10.1109/CVPR.2017.243.

[59] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.

[60] M. J. Page *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, p. n71, Mar. 2021, doi: 10.1136/bmj.n71.

[61] H. Snyder, "Literature review as a research methodology: An overview and guidelines," *J. Bus. Res.*, vol. 104, pp. 333–339, Nov. 2019, doi: 10.1016/j.jbusres.2019.07.039.

[62] C. Okoli, "A Guide to Conducting a Standalone Systematic Literature Review," *Commun. Assoc. Inf. Syst.*, vol. 37, 2015, doi: 10.17705/1CAIS.03743.

[63] G. Paré, M.-C. Trudel, M. Jaana, and S. Kitsiou, "Synthesizing information systems knowledge: A typology of literature reviews," *Inf. Manag.*, vol. 52, no. 2, pp. 183–199, Mar. 2015, doi: 10.1016/j.im.2014.08.008.

[64] J. Webster and R. T. Watson, "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Q.*, vol. 26, no. 2, pp. xiii–xxiii, Jun. 2002, doi: 10.2307/4132319.

[65] R. Archana and P. S. E. Jeevaraj, "Deep learning models for digital image processing: a review," *Artif. Intell. Rev.*, vol. 57, no. 1, p. 11, Jan. 2024, doi: 10.1007/s10462-023-10631-z.

[66] S. Tiecheng, F. Jie, L. Shuang, and Z. Tianqi, "Color Context Binary Pattern Using Progressive Bit Correction for Image Classification," *Chinese J. Electron.*, vol. 30, no. 3, pp. 471–481, May 2021, doi: 10.1049/cje.2021.03.010.

[67] M. Alimoussa, A. Porebski, N. Vandenbroucke, S. El Fkihi, and R. Oulad Haj Thami, "Compact Hybrid Multi-Color Space Descriptor Using Clustering-Based Feature Selection for Texture Classification," *J. Imaging*, vol. 8, no. 8, p. 217, Aug. 2022, doi: 10.3390/jimaging8080217.

[68] H.-C. Shin *et al.*, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/TMI.2016.2528162.