

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan *Artificial Intelligence* (AI) telah mendorong transformasi digital di berbagai sektor melalui pemanfaatan analisis data berbasis pembelajaran mesin yang adaptif dan efisien (Hughes et al., 2021). Transformasi digital ditandai dengan meningkatnya penggunaan sistem cerdas untuk mengotomatisasi proses analisis data yang sebelumnya dilakukan secara manual. Teknologi AI memungkinkan sistem komputer mempelajari pola dari data historis sehingga mampu menghasilkan prediksi dan rekomendasi yang lebih akurat. Kemajuan perangkat keras komputasi dan ketersediaan *big data* turut mempercepat adopsi AI di berbagai bidang. Salah satu cabang AI yang berkembang pesat adalah computer vision, yaitu bidang yang berfokus pada kemampuan sistem komputer dalam memahami dan menginterpretasikan informasi visual dari citra digital secara otomatis. Oleh karena itu, computer vision menjadi teknologi kunci dalam berbagai aplikasi analisis citra modern (Zhao et al., 2024).

Convolutional Neural Network (CNN) telah menjadi pendekatan dominan dalam klasifikasi citra karena kemampuannya mengekstraksi fitur visual secara hierarkis tanpa memerlukan rekayasa fitur manual (Rodríguez et al., 2021). CNN bekerja dengan memanfaatkan lapisan konvolusi untuk menangkap pola lokal pada citra, seperti tepi, tekstur, dan bentuk, yang kemudian dikombinasikan menjadi representasi fitur tingkat tinggi. Pendekatan ini terbukti lebih efektif dibandingkan metode tradisional yang bergantung pada fitur buatan manusia. Metode berbasis *machine learning* mampu mengekstraksi fitur secara otomatis

dan efektif dalam menangkap pola kompleks pada citra, sedangkan metode ekstraksi fitur manual lebih unggul dalam merepresentasikan karakteristik bentuk yang bersifat invarian terhadap rotasi dan skala (Dzulqarnain et al., 2025). Efektivitas pendekatan berbasis CNN dalam mengenali pola kompleks juga telah dibuktikan pada berbagai penelitian klasifikasi tekstil dan citra budaya (Ji et al., 2024). Keberhasilan model CNN sangat dipengaruhi oleh kualitas dan pengelolaan *dataset* yang digunakan, karena *dataset* yang baik akan meningkatkan kemampuan generalisasi dan akurasi prediksi model (Minarno et al., 2023). Penyusunan *dataset* yang representatif menjadi fondasi penting dalam pengembangan sistem klasifikasi citra.

Salah satu objek citra yang menarik perhatian peneliti adalah batik. Batik sebagai warisan budaya Indonesia telah diakui oleh UNESCO sebagai warisan budaya takbenda dunia dan sedang diupayakan pelestariannya melalui dokumentasi digital yang sistematis (Arciszewski et al., 2022). Digitalisasi penting tidak hanya untuk tujuan arsip budaya, tetapi juga untuk mendukung penelitian berbasis teknologi yang dapat membantu identifikasi dan klasifikasi motif secara otomatis, mengingat setiap daerah memiliki karakteristik motif yang

unik, termasuk Batik Dayak Kalimantan yang kaya akan simbol flora dan fauna khas serta nilai filosofis budaya lokal (Aras et al., 2022). Kompleksitas visual pada motif Batik Dayak Kalimantan menjadikannya objek yang menantang sekaligus potensial untuk dianalisis menggunakan pendekatan *machine learning*. Sayangnya, dokumentasi *dataset* citra Batik Dayak Kalimantan saat ini masih terbatas dan belum tersusun secara sistematis untuk mendukung kebutuhan analisis berbasis *machine learning* (Manap et al., 2024). Keterbatasan

dokumentasi *dataset* citra Batik Dayak Kalimantan berpotensi menghambat pengembangan model klasifikasi yang optimal. Secara konkret, ketersediaan *dataset* Batik Dayak Kalimantan yang siap digunakan untuk *machine learning* masih sangat terbatas. Berbeda dengan *dataset* Batik Jawa yang telah memiliki ratusan hingga ribuan citra terkurasi, *dataset* citra Batik Dayak Kalimantan yang terdokumentasi secara publik masih terbatas. Sebagian besar *dataset* yang digunakan dalam penelitian sebelumnya dikembangkan untuk kebutuhan penelitian masing-masing, memiliki jumlah citra yang relatif sedikit, dan belum dikelola secara terstruktur untuk mendukung pengembangan model *machine learning*. Penelitian ini melakukan penggunaan ulang sebanyak 1.200 citra yang terbagi ke dalam lima kelas berdasarkan provinsi asal di Kalimantan, yaitu Kalimantan Barat, Kalimantan Timur, Kalimantan Selatan, Kalimantan Utara, dan Kalimantan Tengah, sebagai upaya untuk membangun *dataset* yang lebih representatif.

Dalam pengembangan model klasifikasi citra, distribusi kelas yang tidak seimbang dapat menyebabkan bias model dan menurunkan performa klasifikasi (Ciaglia et al., 2022). Untuk mengatasi permasalahan tersebut, pemanfaatan

atribut visual citra seperti warna dan pola geometris terbukti mampu meningkatkan akurasi sistem klasifikasi visual (Bianconi et al., 2021). Konteks pada Batik Dayak Kalimantan, atribut-atribut tersebut memiliki peran krusial karena setiap motif memiliki ciri khas tersendiri, misalnya dominasi warna tanah dan merah pada motif tertentu, serta pola geometris berulang yang merepresentasikan simbol flora dan fauna Dayak. Pendekatan berbasis atribut ini terbukti efektif dalam studi klasifikasi tekstil tradisional dengan kompleksitas

visual tinggi (Quan et al., 2024). Teknik data *augmentation* digunakan untuk meningkatkan variasi *dataset* dan mengurangi risiko *overfitting* dan *underfitting* pada model *machine learning* (Zeng., 2024). Melalui kombinasi strategi tersebut, model diharapkan memiliki kemampuan generalisasi yang lebih baik terhadap data baru.

Dataset batik yang terstruktur dengan metadata lengkap mampu meningkatkan performa model klasifikasi secara signifikan (Minarno et al., 2023).

Penelitian klasifikasi batik berbasis CNN telah banyak dilakukan di Indonesia, meskipun sebagian besar berfokus pada motif Batik Jawa seperti Batik Solo dan

Batik Yogyakarta (Sari et al., 2025). Implementasi VGG16 dalam klasifikasi batik

berbasis *web* juga menunjukkan hasil akurasi tinggi pada *dataset* yang telah terkurasi dengan baik (Shafira et al., 2024). Optimasi CNN melalui augmentasi

dan *preprocessing* meningkatkan stabilitas model dalam klasifikasi citra (Alin et al., 2023). Penelitian yang secara khusus membangun dan mengelola *dataset* citra

Batik Dayak Kalimantan secara sistematis berdasarkan atribut visual masih terbatas. *Dataset* Batik Dayak Kalimantan yang belum terstruktur dan terlabel

dengan baik, serta belum siap pakai untuk pelatihan model, menjadi celah

penelitian yang perlu segera dilakukan. Penelitian ini hadir untuk mengisi celah

penelitian tersebut dengan membangun dan mengelola *dataset* citra Batik Dayak

berbasis atribut menggunakan platform Roboflow.

Pemanfaatan platform seperti Roboflow memungkinkan integrasi proses anotasi, augmentasi, dan pengelolaan *dataset* dalam satu sistem terstandarisasi.

Beberapa penelitian menunjukkan bahwa Roboflow digunakan secara luas dalam proses pembuatan dan peningkatan kualitas *dataset* citra untuk model *machine*

learning (Yauna et al., 2025). Adanya platform Roboflow memungkinkan proses pengelolaan *dataset* menjadi lebih konsisten dan mudah direproduksi. Penelitian ini berfokus pada pengelolaan *dataset* citra Batik Dayak Kalimantan berbasis atribut menggunakan Roboflow untuk mendukung analisis klasifikasi motif batik berbasis *machine learning* secara lebih optimal.

1.2 Rumusan Masalah

1. Bagaimana menerapkan pengelolaan *dataset* citra Batik Dayak Kalimantan berdasarkan atribut visual yang digunakan sebagai dasar deskripsi/analisis pendukung secara terstruktur berdasarkan kelas citra menggunakan platform Roboflow?
2. Bagaimana kualitas *dataset* yang telah dikelola dapat dievaluasi melalui performa model klasifikasi CNN berdasarkan metrik akurasi?

1.3 Tujuan Penelitian

1. Mengelola *dataset* citra Batik Dayak Kalimantan secara terstruktur melalui tahapan pengumpulan, *preprocessing*, anotasi, augmentasi, dan pembagian *dataset* menggunakan platform *Roboflow*.
2. Pelabelan kelas dilakukan berdasarkan provinsi asal, sedangkan atribut visual digunakan sebagai dasar deskripsi/analisis pendukung.
3. Mengevaluasi kualitas *dataset* yang dihasilkan melalui pelatihan model CNN dan pengukuran metrik akurasi menggunakan data validasi.

1.4 Ruang Lingkup

Batasan penelitian ini disusun agar pembahasan tetap terarah dan tidak menyimpang dari tujuan yang telah ditetapkan. Adapun batasan dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini berfokus pada **pengelolaan *dataset* citra Batik Dayak**

Kalimantan yang mencakup proses pengumpulan, *preprocessing*, anotasi, pelabelan, augmentasi dan pembagian *dataset* menggunakan platform Roboflow.

2. *Dataset* yang digunakan dalam penelitian ini berupa citra motif Batik Dayak Kalimantan dengan jumlah target sebanyak 1.200 citra yang terbagi ke dalam 5 kelas berdasarkan provinsi, yaitu Kalimantan Barat, Kalimantan Timur, Kalimantan Selatan, Kalimantan Utara, dan Kalimantan Tengah.

3. Proses pelabelan *dataset* dilakukan berdasarkan provinsi asal motif Batik Dayak Kalimantan.

4. Augmentasi digunakan sebagai tahap pendukung dalam pengelolaan *dataset*.

5. Penelitian ini dilaksanakan dalam rentang waktu Januari hingga Juni 2026.