

Penerapan Augmentasi Gambar untuk Memperluas Dataset Batik Dayak Kalimantan pada Klasifikasi Pola Berbasis Citra Digital

Implementation Of Image Augmentation To Expand The Kalimantan Dayak Batik Dataset In Digital Image-Based Pattern Classification

Stefany Marchela¹, Muhammad Faqih Dzulqarnain², Irwan Adhi Prasetya³

^{1,2,3}Program Studi D-III Teknologi Informasi Politeknik 'Aisyiyah Pontianak

Jl. Ampera, Sungai Jawi, Kec. Pontianak Kota, Kota Pontianak, Kalimantan Barat 78114, Indonesia

email: ¹stefanymarchela@gmail.com, ²faqih.dzulqarnain@polita.ac.id, ³irwan.adhi@polita.ac.id

ABSTRAK

DOI;
10.30595/jrst.xxxx

Histori Artikel:

Diajukan:
xx/xx/20xx

Diterima:
xx/xx/20xx

Diterbitkan:
xx/xx/20xx

Batik Dayak Kalimantan merupakan warisan budaya dari lima wilayah di Kalimantan yang memiliki keragaman motif khas. Keterbatasan jumlah data latih yang hanya mencakup 600 citra dari lima kelas wilayah menyebabkan model klasifikasi berpotensi mengalami *overfitting* dan menghasilkan akurasi yang tidak optimal, sehingga diperlukan strategi perluasan dataset tanpa pengumpulan data baru di lapangan. Penelitian ini menerapkan teknik augmentasi gambar menggunakan bahasa pemrograman Python dengan pustaka *Keras ImageDataGenerator* melalui empat teknik transformasi, yaitu rotasi hingga 180°, pencerminan horizontal dan vertikal, perbesaran hingga 10%, serta pergeseran warna sebesar $\pm 7,8\%$. Penerapan teknik tersebut berhasil memperluas dataset dari 600 citra menjadi 3.600 citra dengan distribusi yang tetap seimbang antar kelas wilayah. Dataset yang telah diperluas kemudian diklasifikasikan menggunakan aplikasi *Orange Data Mining* melalui proses ekstraksi fitur *Image Embedding* berbasis model CNN pralatih *SqueezeNet (local)* dan algoritma *Neural Network*, serta dievaluasi menggunakan metode *Stratified 5-Fold Cross Validation* berdasarkan nilai AUC, CA, *Precision*, *Recall*, dan *F1-score*. Hasil penelitian menunjukkan peningkatan kinerja klasifikasi yang signifikan setelah augmentasi diterapkan, dengan nilai AUC meningkat dari 0,975 menjadi 0,997, nilai CA meningkat dari 88,7% menjadi 95,3%, serta nilai *Precision*, *Recall*, dan *F1-score* masing-masing meningkat dari 0,886 menjadi 0,953. Hasil ini menunjukkan bahwa augmentasi gambar mampu memperluas variasi dataset dan meningkatkan kinerja klasifikasi motif batik Dayak Kalimantan berdasarkan wilayah asal secara lebih optimal.

Kata Kunci: batik Dayak Kalimantan, augmentasi gambar, *ImageDataGenerator*, *SqueezeNet*, *Neural Network*

ABSTRACT

Dayak Kalimantan batik is a cultural heritage originating from five regions in Kalimantan, characterized by a diverse range of distinctive motifs. The limited amount of training data, comprising only 600 images across five regional classes, causes the classification model to be prone to *overfitting* and results in suboptimal accuracy. This condition necessitates a dataset expansion strategy without collecting new data in the field. This study applies image augmentation techniques using the Python programming language with the *Keras ImageDataGenerator* library through four transformation

techniques, namely rotation up to 180°, horizontal and vertical flipping, zooming up to 10%, and color shifting of $\pm 7.8\%$. These techniques successfully expanded the dataset from 600 images to 3,600 images while maintaining a balanced distribution across regional classes. The expanded dataset was then classified using the *Orange Data Mining* application through an *Image Embedding* feature extraction process based on the pre-trained *SqueezeNet (local)* CNN model and the *Neural Network* algorithm, evaluated using the *Stratified 5-Fold Cross Validation* method based on AUC, CA, *Precision*, *Recall*, and *F1-score* values. The results show a significant improvement in classification performance after augmentation was applied, with the AUC value increasing from 0.975 to 0.997, the CA value increasing from 88.7% to 95.3%, and the *Precision*, *Recall*, and *F1-score* values each increasing from 0.886 to 0.953. These results indicate that image augmentation is able to expand dataset variation and improve the classification performance of Dayak Kalimantan batik motifs based on their region of origin more optimally.

Keywords: Dayak Kalimantan batik, image augmentation, ImageDataGenerator, SqueezeNet, Neural Network

PERPUSTAKAAN

NPP. 6171052A2000001



POLITEKNIK 'AISYIYAH PONTIANAK

1. PENDAHULUAN

Batik merupakan warisan budaya asli Indonesia yang telah diakui oleh UNESCO pada tanggal 2 Oktober 2009 sebagai warisan kemanusiaan yang wajib dilestarikan. Sebagai sebuah karya seni, batik adalah pola pada kain yang dibuat dengan menggunakan lilin atau malam sebagai bahan pelapis untuk menahan warna sehingga menghasilkan motif yang khas dan bernilai tinggi. Keberagaman motif batik di setiap daerah Indonesia mencerminkan kekayaan budaya, filosofi, serta identitas masyarakat setempat, menjadikan batik bukan sekadar kain bermotif, melainkan sebuah karya seni yang sarat akan makna dan nilai historis yang mendalam (Raspati et al., 2023).

Perkembangan sistem klasifikasi pola berbasis citra digital mendorong pemanfaatan teknologi kecerdasan buatan dalam berbagai bidang, termasuk dalam upaya digitalisasi dan pelestarian budaya daerah seperti motif batik Dayak Kalimantan. Tantangan utama dalam pengembangan sistem klasifikasi motif batik Dayak Kalimantan adalah keterbatasan jumlah dan variasi dataset yang tersedia. Dataset awal yang direncanakan dalam penelitian ini mencakup 600 citra batik Dayak Kalimantan yang berasal dari lima wilayah di Kalimantan, yaitu Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kalimantan Timur, dan Kalimantan Utara, dengan masing-masing wilayah diwakili oleh 120 citra. Jumlah tersebut masih sangat terbatas dan belum merepresentasikan keseluruhan keragaman visual batik Dayak yang sesungguhnya tersebar di Kalimantan. Dokumentasi motif batik Dayak belum terstruktur secara menyeluruh, sementara setiap motif memiliki variasi pola, warna, dan orientasi yang cukup beragam. Keterbatasan data pelatihan menyebabkan model klasifikasi sering kali dilakukan dengan jumlah data yang terbatas, sehingga akurasi menurun dan model cenderung mengalami *overfitting* ketika dihadapkan pada variasi citra yang berbeda dari dataset. Jumlah data pelatihan yang terbatas juga dapat menyebabkan model mengalami *underfitting*, yaitu kondisi ketika model belum mampu mempelajari pola penting dalam dataset secara optimal sehingga kinerja klasifikasi menjadi kurang maksimal. Penelitian terdahulu menunjukkan bahwa keterbatasan dataset berpengaruh langsung terhadap penurunan akurasi serta meningkatnya risiko *overfitting* pada model klasifikasi citra (Andzani et al., 2025).

Permasalahan klasifikasi citra semakin kompleks pada motif batik Dayak Kalimantan yang memiliki karakter visual khas dan memiliki detail ornamen yang beragam. Motif yang terinspirasi dari unsur flora, fauna, serta simbol budaya masyarakat Dayak memiliki pola berulang dengan variasi bentuk dan komposisi yang tidak selalu seragam. Keberagaman pola tersebut memerlukan sistem yang mampu mengenali motif secara konsisten pada berbagai kondisi sudut pengambilan gambar, pencahayaan, skala, dan kualitas dokumentasi. Penelitian sebelumnya menunjukkan bahwa variasi kondisi visual seperti perubahan intensitas cahaya dan orientasi objek dapat mempengaruhi kinerja model klasifikasi citra (Rosanti et al., 2024). Keterbatasan jumlah dan variasi dataset menunjukkan bahwa diperlukan strategi yang mampu memperkaya variasi data tanpa harus melakukan pengumpulan data dalam skala besar yang memerlukan waktu dan biaya tinggi.

Proses identifikasi motif batik Dayak secara manual memiliki berbagai keterbatasan karena banyaknya pola yang memiliki kemiripan visual sehingga sulit dibedakan secara tepat. Perbedaan kecil pada warna, detail ornamen, dan orientasi pola memerlukan pengamatan yang cermat agar identifikasi dapat dilakukan secara akurat. Penelitian terdahulu menunjukkan bahwa identifikasi motif batik secara manual memerlukan keahlian khusus, memakan waktu, dan rentan terhadap perbedaan penilaian (Muria & Dedy, 2025). Pendekatan *Machine Learning* berbasis *Convolutional Neural Network* (CNN) banyak digunakan untuk mengatasi keterbatasan tersebut karena kemampuannya mengenali pola citra secara otomatis tanpa memerlukan perancangan fitur secara manual, sehingga proses klasifikasi motif batik dapat dilakukan secara lebih konsisten dan efisien dibandingkan identifikasi manual (Samidin & Fadjeri, 2024).

Kinerja model *Convolutional Neural Network* (CNN) dalam klasifikasi citra dipengaruhi oleh ketersediaan dataset yang memadai dan bervariasi, dan keterbatasan jumlah serta variasi dataset tidak hanya berdampak pada kinerja sistem klasifikasi berbasis citra digital, tetapi juga mempengaruhi konsistensi proses identifikasi manual pada motif batik Dayak Kalimantan. Penerapan teknik augmentasi citra telah terbukti mampu meningkatkan perluasan dataset dan kemampuan generalisasi model, sebagaimana

perubahan citra seperti rotasi, *flipping*, translasi, dan perubahan skala terbukti mampu meningkatkan keragaman dataset sehingga model menjadi lebih stabil dan adaptif terhadap variasi kondisi visual (Yoo & Kang, 2023). Meskipun demikian, penelitian yang secara khusus mengkaji pengaruh augmentasi citra terhadap klasifikasi motif batik Dayak Kalimantan berdasarkan wilayah asal masih terbatas dan belum dikaji secara mendalam, padahal citra batik Dayak memiliki variasi bentuk, warna, dan detail ornamen yang beragam sehingga membutuhkan dataset yang lebih bervariasi agar model mampu melakukan klasifikasi secara lebih optimal. Kondisi ini menunjukkan adanya celah penelitian pada aspek pengaruh penerapan teknik augmentasi terhadap perluasan dataset dan peningkatan kinerja klasifikasi citra batik Dayak Kalimantan berbasis wilayah asal, yang dalam penelitian ini dijumpai melalui pemanfaatan bahasa pemrograman Python untuk menghasilkan variasi data secara otomatis.

Penelitian ini difokuskan pada satu permasalahan utama yaitu keterbatasan jumlah dan variasi dataset motif batik Dayak Kalimantan dalam pengembangan sistem klasifikasi citra digital, yang berpotensi menyebabkan model mengalami overfitting, menurunkan kemampuan generalisasi, serta menghasilkan tingkat akurasi klasifikasi yang tidak optimal ketika dihadapkan pada data baru. Jumlah dan variasi dataset yang belum memadai menyebabkan variasi pola, warna, dan orientasi motif yang digunakan dalam pengembangan sistem menjadi kurang beragam, sehingga sistem menunjukkan kinerja tidak maksimal ketika mengklasifikasikan citra dengan kondisi sudut pengambilan gambar, pencahayaan, dan skala yang berbeda. Berangkat dari permasalahan tersebut, penelitian ini bertujuan untuk menerapkan teknik augmentasi gambar guna memperluas jumlah dan variasi dataset batik Dayak Kalimantan tanpa mengubah karakteristik pola utama motif batik Dayak, serta membandingkan hasil klasifikasi citra pada Orange Data Mining sebelum dan sesudah augmentasi data berdasarkan nilai AUC, CA, F1-score, Precision, dan Recall.

2. METODE PENELITIAN

Metode penelitian menguraikan cara yang digunakan untuk menyelesaikan masalah, meliputi alat, bahan, dan metode yang digunakan

dalam pemecahan masalah. Penelitian ini menggunakan pendekatan eksperimental untuk membandingkan kinerja klasifikasi citra batik Dayak Kalimantan sebelum dan sesudah augmentasi, menggunakan aplikasi Orange Data Mining. Dataset citra diperoleh dari Kaggle dan penelitian (Dzulqarnain et al., 2025), bersifat legal dan bebas *copyright*, kemudian dikelompokkan berdasarkan wilayah asal. Tahap *preprocessing* dan augmentasi dilakukan dengan Python menggunakan *library Keras ImageDataGenerator* melalui rotasi, *flip*, *zoom*, dan *color shift* tanpa mengubah pola utama motif. Dataset hasil augmentasi selanjutnya diekstraksi fiturnya melalui *Image Embedding* berbasis CNN pralatih dan diklasifikasikan menggunakan algoritma *Neural Network* untuk melihat pengaruh augmentasi terhadap kinerja klasifikasi.

2.1 Jenis Citra

Citra digital tersusun dari piksel yang menyimpan informasi warna dan kecerahan, di mana perbedaan nilai piksel membentuk pola dan objek pada gambar. Representasi berbasis piksel ini mendukung dokumentasi digital serta analisis dan klasifikasi motif batik.

Tabel 2.1 Data Jumlah Citra Batik Dayak

No	Wilayah Asal	Jumlah Citra
1	Kalimantan Barat	120
2	Kalimantan Selatan	120
3	Kalimantan Tengah	120
4	Kalimantan Timur	120
5	Kalimantan Utara	120
Total		160

Dataset berjumlah 600 citra terbagi merata ke dalam 5 kelas wilayah (Tabel 2.1), masing-masing 120 citra, sebagai data masukan pada tahap pelatihan model klasifikasi.

2.2 Proses Preprocessing Citra

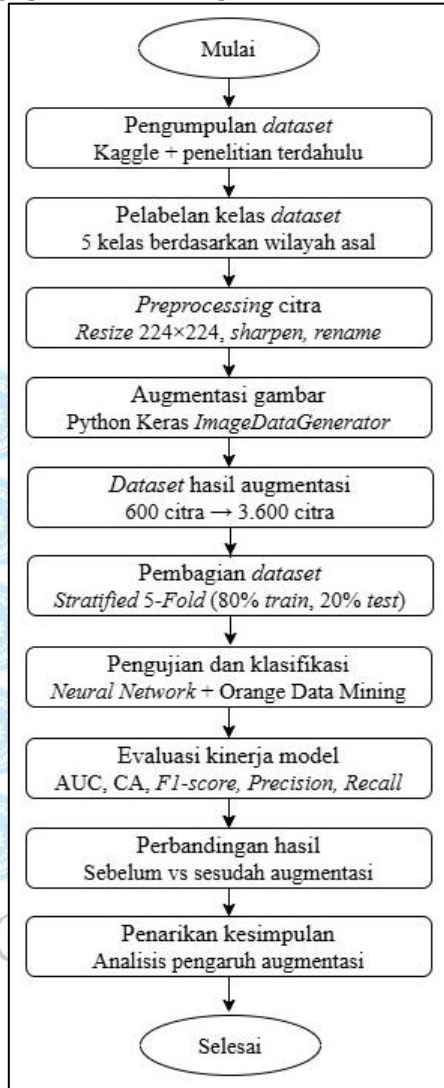
Preprocessing dilakukan agar seluruh citra seragam sebelum augmentasi, melalui dua tahap: (1) *resize* ke ukuran 224×224 piksel dengan *crop* tengah dan algoritma *Lanczos* untuk menghindari *blur* dan distorsi, sesuai ukuran masukan model *SqueezeNet*; dan (2) penyeragaman ketajaman gambar (*sharpen* 1,5) serta penamaan ulang file berdasarkan *prefix* kelas wilayah untuk memudahkan pengelolaan data.

2.3 Proses Augmentasi Gambar

Augmentasi diterapkan pada 600 citra awal menggunakan Python dan *library Keras ImageDataGenerator* untuk menghasilkan variasi

citra baru tanpa mengubah karakteristik motif batik.

Gambar 1 menunjukkan alur kerja penelitian yang menguraikan tahapan-tahapan yang dilakukan secara berurutan dalam penelitian ini, mulai dari pengumpulan *dataset* hingga penarikan kesimpulan.



Gambar 1 Diagram alir penelitian

Tabel 2.2 Parameter Teknik Augmentasi Citra

No	Teknik	Parameter
1	Rotasi	180°
2	Zoom	10%
3	Flip	Horizontal & Vertical
4	Color Shift	$channel_shift_range = 20,0 (\pm 7,8\%)$

Sumber: (Akhmad et al., 2023)

Teknik-teknik pada Tabel 2.2 menghasilkan variasi sudut, skala, orientasi, dan

warna citra tanpa mengubah karakteristik visual utama motif batik.

2.4 Proses Klasifikasi

Klasifikasi dilakukan melalui *workflow* Orange Data Mining: *Import Images* untuk mengimpor dataset, *Image Embedding* dengan model CNN pralatih *SqueezeNet (local)* untuk mengekstraksi fitur citra menjadi vektor numerik, lalu *Neural Network* untuk klasifikasi berdasarkan vektor tersebut. Evaluasi menggunakan *Stratified 5-Fold Cross Validation* (80:20) melalui *widget Test and Score* dan divisualisasikan dengan *Confusion Matrix*. Pengujian dilakukan pada dataset sebelum (600 citra) dan sesudah augmentasi (3.600 citra).

Tabel 2.3 Alur Kerja Klasifikasi

No	Widget	Fungsi
1	<i>Import Images</i>	Impor dataset
2	<i>Image Embedding</i>	Ekstraksi fitur
3	<i>Neural Network</i>	Klasifikasi
4	<i>Test and Score</i>	Evaluasi model
5	<i>Confusion Matrix</i>	Visualisasi hasil

Widget-widget pada Tabel 2.3 membentuk alur kerja klasifikasi motif batik secara berurutan, mulai dari impor dataset, ekstraksi fitur citra, pelatihan model klasifikasi, hingga evaluasi dan visualisasi hasil prediksi.

2.5 Metrik Evaluasi

Kinerja model diukur menggunakan lima metrik evaluasi, yaitu AUC, CA, *Precision*, *Recall*, dan *F1-score*.

AUC (*Area Under the Curve*) mengukur luas area di bawah kurva ROC (*Receiver Operating Characteristic*), yaitu kurva yang menggambarkan hubungan antara *True Positive Rate* dan *False Positive Rate* pada berbagai ambang batas klasifikasi. Nilai AUC berkisar 0-1, semakin mendekati 1 menunjukkan model semakin baik membedakan antarkelas.

1. Classification Accuracy (CA)

Pertama, *Confusion Matrix* dihasilkan dari hasil klasifikasi pada setiap *fold*. Jumlah data yang berhasil diklasifikasikan dengan benar, diperoleh dari diagonal utama *Confusion Matrix*, kemudian dibagi dengan total data pengujian, sebagaimana dihitung pada Persamaan (1).

$$\text{Akurasi (CA)} = \frac{\sum \text{Data Benar}}{\text{Total Data Pengujian}} \times 100\%$$

dengan $\sum D$ merupakan jumlah total data yang berhasil diklasifikasikan dengan benar (diagonal utama *Confusion Matrix*), dan N_{uji} merupakan jumlah total data pengujian.

2. Precision

Selanjutnya, *Precision* dihitung untuk setiap kelas guna mengukur seberapa banyak data yang diprediksi sebagai kelas tertentu yang benar-benar tepat, sebagaimana ditunjukkan pada Persamaan (2).

$$recision = \frac{TP}{TP + FP}$$

dengan *TP*(*True Positive*) merupakan jumlah data yang diprediksi dengan benar sebagai kelas tertentu, dan *FP*(*False Positive*) merupakan jumlah data yang salah diprediksi sebagai kelas tersebut.

3. Recall

Kemudian, *Recall* dihitung untuk setiap kelas guna mengukur seberapa banyak data aktual pada suatu kelas yang berhasil dikenali oleh model, sebagaimana ditunjukkan pada Persamaan (3).

$$Recall = \frac{TP}{TP + FN}$$

dengan *TP* merupakan jumlah data yang diprediksi dengan benar sebagai kelas tertentu, dan *FN*(*False Negative*) merupakan jumlah data yang sebenarnya termasuk kelas tersebut namun diprediksi sebagai kelas lain.

4. F1-score

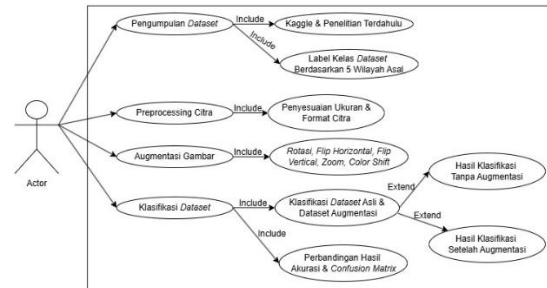
Terakhir, *F1-score* dihitung sebagai rata-rata harmonik antara *Precision* dan *Recall* untuk menggambarkan keseimbangan antara kedua metrik tersebut, sebagaimana ditunjukkan pada Persamaan (4).

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

dengan *Precision* dan *Recall* merupakan nilai yang diperoleh dari Persamaan (2) dan (3).

2.6 Metode Pengujian

Pengujian dilakukan dengan membandingkan hasil klasifikasi antara *dataset* citra asli dan *dataset* hasil augmentasi menggunakan aplikasi *Orange Data Mining*, guna menganalisis pengaruh penerapan teknik augmentasi gambar terhadap peningkatan kinerja klasifikasi pola batik Dayak Kalimantan. Skenario pengujian mencakup proses augmentasi gambar untuk menambah jumlah dan variasi citra, serta proses klasifikasi menggunakan algoritma *Neural Network* untuk mengevaluasi kinerja model berdasarkan nilai *AUC*, *CA*, *Precision*, *Recall*, dan *F1-score*, yang divisualisasikan melalui *Confusion Matrix*. Diagram *use case* penelitian yang menggambarkan interaksi antara Peneliti dan sistem dalam empat proses utama disajikan pada Gambar 2.

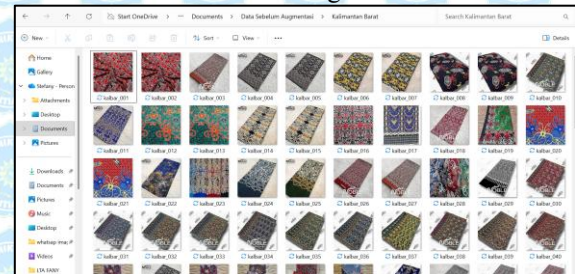


Gambar 2 Diagram use case sistem

3. HASIL DAN PEMBAHASAN

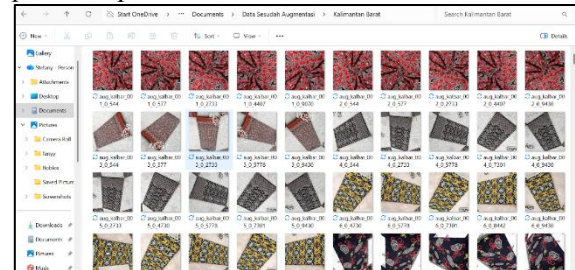
3.1 Hasil Augmentasi Gambar

Penerapan augmentasi gambar pada dataset batik Dayak Kalimantan menggunakan bahasa pemrograman Python dengan *library* Keras *ImageDataGenerator* berhasil memperluas dataset dari 600 menjadi 3.600 citra melalui kombinasi empat teknik transformasi, yaitu rotasi hingga 180 derajat, *flip* horizontal dan vertikal, *zoom* hingga 10%, serta pergeseran warna ($\pm 7,8\%$). Setiap citra asli menghasilkan 5 variasi baru sehingga tiap kelas wilayah bertambah dari 120 menjadi 720 citra. Gambar 1 menampilkan contoh citra batik Kalimantan Barat sebelum augmentasi.



Gambar 3 Citra Sebelum augmentasi

Gambar 4 menampilkan contoh hasil citra yang sama setelah proses augmentasi diterapkan, dengan variasi sudut, orientasi, skala, dan warna pada setiap citra hasil transformasi.



Gambar 4 Citra Sesudah augmentasi

Perbandingan Gambar 3 dan Gambar 4 menunjukkan bahwa augmentasi berhasil menghasilkan variasi citra yang beragam melalui rotasi, *flip*, *zoom*, dan pergeseran warna tanpa mengubah struktur pola ornamen khas batik Dayak

Kalimantan, sehingga identitas kelas tetap terjaga sekaligus model dapat mempelajari motif dari berbagai sudut pandang dan kondisi pencahayaan. Pemilihan keempat teknik ini sesuai dengan karakteristik objek: rotasi dan *flip* relevan karena batik tidak memiliki orientasi baku, *zoom* dibatasi agar detail motif tidak terpotong, dan *color shift* mensimulasikan variasi pencahayaan yang wajar.

3.2 Hasil Klasifikasi

Untuk mengukur pengaruh perluasan *dataset* ini terhadap kinerja klasifikasi, pengujian dilakukan menggunakan *widget Test and Score* pada *Orange Data Mining* dengan metode *Stratified 5-Fold Cross Validation*. Gambar 3 menampilkan hasil evaluasi model *Neural Network* sebelum augmentasi, dengan AUC 0,975, CA 0,887 (88,7%), F1-score 0,886, Precision 0,886, dan Recall 0,887, serta 532 dari 600 data pengujian berhasil diklasifikasikan dengan benar.

Test and Score Thu May 21 26, 20:14:55							
Settings							
Sampling type: Stratified 5-fold Cross validation Target class: None, show average over classes							
Scores							
Model	AUC	CA	F1	Prec	Recall	MCC	
Neural Network	0.975	0.887	0.886	0.886	0.887	0.858	

Gambar 5 Citra Sesudah augmentasi

Gambar 5 menampilkan hasil evaluasi model sesudah augmentasi menggunakan metode pengujian yang sama, dengan AUC meningkat menjadi 0,997, CA menjadi 0,953 (95,3%), F1-score, Precision, dan Recall yang seragam pada 0,953, serta jumlah data yang berhasil diklasifikasikan dengan benar meningkat menjadi 3.432 dari 3.600 data.

Test and Score Thu May 21 26, 20:20:58							
Settings							
Sampling type: Stratified 5-fold Cross validation Target class: None, show average over classes							
Scores							
Model	AUC	CA	F1	Prec	Recall	MCC	
Neural Network	0.997	0.953	0.953	0.953	0.953	0.942	

Gambar 6 Citra Sesudah augmentasi

Peningkatan CA sebesar 6,6% dan AUC sebesar 0,022, disertai keseragaman nilai F1-score, Precision, dan Recall, mengindikasikan bahwa peningkatan kinerja terjadi secara merata pada seluruh kelas tanpa kecenderungan bias terhadap kelas wilayah tertentu. Temuan ini sejalan dengan teori bahwa augmentasi data memperluas variasi pola pelatihan sehingga model dapat membangun representasi fitur yang lebih generalis.

Distribusi hasil klasifikasi per kelas wilayah divisualisasikan melalui *Confusion Matrix*. Gambar

5 menunjukkan hasil sebelum augmentasi dari 600 data pengujian, dengan kelas Kalimantan Utara memiliki tingkat kesalahan tertinggi (95 dari 120 data benar), terutama karena kemiripannya dengan kelas Kalimantan Selatan.

Confusion Matrix for Neural Network (showing number of instances) Thu May 21 26, 20:33:37							
Actual \ Predicted	Predicted					Σ	
	Kalimantan Barat	Kalimantan Selatan	Kalimantan Tengah	Kalimantan Timur	Kalimantan Utara		
Kalimantan Barat	109	1	4	4	2	120	
Kalimantan Selatan	0	107	1	0	12	120	
Kalimantan Tengah	0	0	113	3	4	120	
Kalimantan Timur	5	0	4	108	3	120	
Kalimantan Utara	4	15	0	0	120	120	
Σ	118	123	122	121	116	600	

Gambar 6 Confusion Matrix Sebelum Augmentasi

Gambar 6 menunjukkan hasil sesudah augmentasi dari 3.600 data pengujian. Kelas Kalimantan Utara masih menjadi kelas dengan tingkat kesalahan tertinggi (666 dari 720 data benar), namun mengalami peningkatan signifikan dibandingkan sebelum augmentasi.

Confusion Matrix for Neural Network (showing number of instances) Thu May 21 26, 20:37:52							
Actual \ Predicted	Predicted					Σ	
	Kalimantan Barat	Kalimantan Selatan	Kalimantan Tengah	Kalimantan Timur	Kalimantan Utara		
Kalimantan Barat	689	5	7	14	5	720	
Kalimantan Selatan	4	678	2	0	36	720	
Kalimantan Tengah	0	0	696	6	6	720	
Kalimantan Timur	8	3	4	702	3	720	
Kalimantan Utara	11	35	3	5	686	720	
Σ	721	725	712	727	715	3.600	

Gambar 7 Confusion Matrix Sebelum Augmentasi

Secara per kelas, peningkatan akurasi paling signifikan terjadi pada kelas Kalimantan Utara, yaitu dari 79,2% menjadi 92,5% (+13,3%), karena kelas ini memiliki akurasi terendah sebelum augmentasi. Namun, kelas Kalimantan Utara dan Kalimantan Selatan tetap menunjukkan kesalahan klasifikasi silang yang konsisten baik sebelum maupun sesudah augmentasi, menunjukkan bahwa kemiripan visual antarkelas akibat kedekatan geografis dan budaya merupakan tantangan yang berasal dari karakteristik data itu sendiri. Sebaliknya, kelas Kalimantan Timur menjadi kelas dengan akurasi terbaik setelah augmentasi (97,5%), menggantikan posisi Kalimantan Tengah yang sebelumnya tertinggi (94,2%).

Secara teoritis, *dataset* awal yang kecil (120 citra per kelas) berisiko menyebabkan *overfitting* pada model berbasis *deep learning*; nilai CA 88,7% sebelum augmentasi mengindikasikan kondisi tersebut. Peningkatan metrik secara merata setelah augmentasi memberikan indikasi bahwa risiko tersebut berkurang. Perlu dicatat, augmentasi dilakukan secara luring (*offline*) sebelum pembagian *fold*, sehingga berpotensi menimbulkan kebocoran data (*data leakage*). Namun, konsistensi pola kesalahan antarkelas mirip sebelum dan sesudah augmentasi mengindikasikan peningkatan performa lebih mencerminkan variasi visual yang nyata, bukan model menghafal data.

4. KESIMPULAN

Augmentasi gambar (rotasi, *flip*, *zoom*, *color shift*) berhasil memperluas *dataset* batik Dayak Kalimantan dari 600 menjadi 3.600 citra tanpa mengubah karakteristik visual motifnya, dan terbukti meningkatkan kinerja klasifikasi model *Neural Network*: AUC naik dari 0,975 menjadi 0,997 dan CA dari 88,7% menjadi 95,3%, dengan peningkatan yang merata pada seluruh kelas wilayah. Peningkatan terbesar terjadi pada kelas Kalimantan Utara (+13,3%), meski kelas ini bersama Kalimantan Selatan tetap saling tertukar akibat kemiripan visual motif yang tidak sepenuhnya teratasi oleh augmentasi. Penelitian ini membuktikan augmentasi data efektif meningkatkan performa klasifikasi pada *dataset* terbatas tanpa perlu pengumpulan data baru, dengan saran penelitian lanjutan menerapkan augmentasi secara *online* atau membagi *dataset* sebelum augmentasi untuk menghindari potensi kebocoran data.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada Politeknik 'Aisyiyah Pontianak beserta Program Studi D-III Teknologi Informasi atas fasilitas dan bimbingan yang telah diberikan selama penelitian ini dilaksanakan.

DAFTAR PUSTAKA

- Akhmad, I., Jasril, J., Sanjaya, S., Handayani, L., & Yanto, F. (2023). Klasifikasi Citra Daging Sapi dan Babi Menggunakan CNN Alexnet dan Augmentasi Data. *Journal of Information System Research (JOSH)*, 4(4), 1176–1185. <https://doi.org/10.47065/josh.v4i4.3702>
- Andzani, S., Arie Wijaya, Y., Faqih, A., Kurniawan, R., & Studi Manajemen Informatika, P. (2025). Perbandingan Teknik Optimasi Cnn Untuk Klasifikasi Gambar Sampah Berbasis Dataset Kecil Dan Tidak Seimbang. <https://doi.org/10.32485/Jcsai>
- Dzulqarnain, M. F., Fadlil, A., & Riadi, I. (2025). Performance Comparison of Learned Features from Autoencoder and Shape-Based Hu Moments for Batik Classification. *Jurnal Teknik Informatika (Jutif)*, 6(4), 1729–1744. <https://doi.org/10.52436/1.jutif.2025.6.4.4827>
- Muria, A. A., & Dedy, K. (2025). Implementasi Arsitektur Resnet50 Pada Klasifikasi Motif Batik Indonesia Menggunakan Metode Convolutional Neural Network (Cnn). *Jurnal Rekayasa Sistem Informasi Dan Teknologi*,

2(4).

<https://doi.org/10.70248/Jrsit.V2i4.2375>

- Raspati, M. R., Sarwenty, P. N., Aksaningtyas, L. L., Tondang, B. A., Chevin, V. A., Saraswati, R. M. D., & Aldo, D. (2023). Sistem Informasi Ragam Batik Di Indonesia Sebagai Media Pembelajaran Menggunakan Multimedia Interaktif. *Idealis : Indonesia Journal Information System*, 6(2). <https://doi.org/10.36080/Idealis.V6i2.3032>
- Rosanti, N., Latifah, R., Munir, S., & Maududi, I. A. Q. (2024). Pengaruh Jarak Objek Citra pada Model Deteksi dan Klasifikasi Botol Plastik menggunakan YOLO. *Jurnal Teknologi Terpadu*, 10(1). <https://doi.org/10.54914/jtt.v10i1.1247>
- Samidin, S., & Fadjeri, A. (2024). Klasifikasi Gambar Batu-Kertas-Gunting Menggunakan Convolutional Neural Network dengan Fungsi Callback untuk Mencegah Overfitting. *Jurnal Penelitian Inovatif*, 4(2). <https://doi.org/10.54082/jupin.413>
- Yoo, J., & Kang, S. (2023). Class-Adaptive Data Augmentation for Image Classification. *IEEE Access*, 11, 26393–26402. <https://doi.org/10.1109/ACCESS.2023.3258179>