



Klasifikasi Tingkat Kelunturan Warna Kain Menggunakan KNN, SVM, dan Random Forest

Romindo^{1*}, Triandes Sinaga², Kevin Bastian Sirait³, Arosochi Yosua Daeli⁴, Jepronel Saragih⁵

^{1,2,3,4}Program Studi Sistem Informasi (Kampus Kota Medan), Fakultas *Artificial Intelligence and Data Science*, Universitas Pelita Harapan, Medan, Indonesia

⁵Program Studi Informatika (Kampus Kota Medan), Fakultas *Artificial Intelligence and Data Science*, Universitas Pelita Harapan, Medan, Indonesia

Email: ^{1*}romindo@uph.edu, ²triandes.sinaga@uph.edu, ³kevin.sirait@uph.edu,

⁴arosochi.daeli@uph.edu, ⁵jepronel@lecturer.uph.edu

Abstract

The laundry industry faces challenges in maintaining service quality, particularly regarding fabric color fading after washing. Assessments that are still performed manually tend to be subjective and inconsistent, so a more objective automated classification system is required. This study aims to apply and compare three algorithms, namely KNN, SVM, and Random Forest, to classify the level of fabric color fading based on digital images. The features used comprise color in the RGB and HSV spaces as well as shape in the form of area and shape ratio, all extracted automatically. A total of 300 images were divided into 250 training data and 50 testing data, then mapped into three categories, namely not faded, fairly faded, and faded. The testing results show that Random Forest delivers the best performance with an accuracy of 0.96, followed by SVM at 0.94 and KNN at 0.88. All models faced difficulties in recognizing the minority class due to data imbalance. This study proves that the machine learning approach, particularly Random Forest, is able to assess color fading levels more accurately and consistently than manual evaluation, while supporting quality control in the laundry industry.

Keywords: Classification, Fabric Color Fading, KNN, SVM, Random Forest.

Abstrak

Industri laundry menghadapi tantangan dalam menjaga kualitas layanan, khususnya terkait kelunturan warna kain setelah pencucian. Penilaian yang masih dilakukan secara manual kerap bersifat subjektif dan tidak konsisten, sehingga diperlukan sistem klasifikasi otomatis yang lebih objektif. Penelitian ini bertujuan menerapkan dan membandingkan tiga algoritma, yaitu KNN, SVM, dan Random Forest, untuk mengklasifikasikan tingkat kelunturan warna kain berbasis citra digital. Fitur yang digunakan meliputi warna pada ruang RGB dan HSV serta bentuk berupa area dan rasio bentuk yang diekstraksi secara otomatis. Sebanyak 300 citra dibagi menjadi 250 data latih dan 50 data uji, lalu dipetakan ke dalam tiga kategori, yaitu tidak luntur, cukup luntur, dan luntur. Hasil pengujian menunjukkan Random Forest memberikan kinerja terbaik dengan akurasi 0,96, diikuti SVM 0,94, dan KNN 0,88. Seluruh model menghadapi tantangan dalam mengenali kelas minoritas akibat ketidakseimbangan data. Penelitian ini membuktikan bahwa pendekatan pembelajaran mesin, khususnya Random Forest, mampu menilai tingkat kelunturan warna secara lebih akurat dan konsisten dibandingkan penilaian manual, sekaligus mendukung kendali mutu pada industri laundry.

Kata Kunci: Klasifikasi, Kelunturan Warna Kain, KNN, SVM, Random Forest.

1. PENDAHULUAN

Industri jasa laundry tumbuh pesat seiring meningkatnya kesibukan masyarakat perkotaan dan tuntutan terhadap efisiensi waktu dalam kehidupan keseharian. Banyak individu kini lebih memilih layanan pencucian profesional dibandingkan mencuci secara mandiri, sehingga kualitas hasil pencucian menjadi penentu utama kepuasan sekaligus

loyalitas pelanggan (Arippudin, 2022; Purnamaningtyas et al., 2022). Kualitas tersebut tidak hanya menyangkut ketepatan waktu dan kebersihan hasil cucian, tetapi juga kondisi fisik pakaian setelah proses pencucian, khususnya kestabilan warna kain yang menjadi indikator mutu yang mudah dikenali oleh pelanggan.

Salah satu persoalan paling krusial dalam layanan ini adalah kelunturan warna kain. Kelunturan menurunkan nilai estetika pakaian sekaligus menimbulkan kerugian bagi pelanggan maupun penyedia jasa karena pakaian dapat menjadi rusak dan tidak layak pakai. Bagi penyedia jasa, insiden kelunturan tidak hanya berdampak pada klaim ganti rugi, tetapi juga menurunkan reputasi usaha dan kepercayaan pelanggan dalam jangka panjang. (Aliei, Carrera-Gallissa, et al., 2024) membuktikan bahwa kondisi pencucian seperti jenis air, suhu, dan komposisi deterjen berpengaruh signifikan terhadap perubahan warna pada kain katun alami. Temuan ini diperkuat oleh kajian lanjutan yang memperlihatkan bahwa pencucian rumah tangga turut mengubah properti kolorimetrik kain katun organik secara nyata (Aliei, Cayuela, et al., 2024). Temuan tersebut menegaskan bahwa pengendalian mutu pada proses pencucian menuntut penilaian yang akurat, objektif, dan dapat direplikasi secara konsisten, terutama ketika volume cucian yang ditangani terus bertambah.

Pada praktiknya, identifikasi tingkat kelunturan warna di sebagian besar usaha laundry masih dilakukan secara manual melalui pengamatan visual oleh karyawan. Pendekatan ini bersifat subjektif karena bergantung pada persepsi tiap individu, sehingga rawan menghasilkan penilaian yang berbeda untuk kondisi kain yang serupa. Selain memerlukan waktu yang relatif lama, metode manual sulit dijaga keseragamannya ketika beban kerja meningkat. Kesalahan persepsi dalam penilaian semacam ini berpotensi memicu keluhan pelanggan sekaligus menghambat proses evaluasi mutu secara menyeluruh. Kondisi inilah yang mendorong kebutuhan akan sistem otomatis yang mampu mengklasifikasikan tingkat kelunturan warna secara objektif berdasarkan karakteristik visual kain.

Kemajuan pengolahan citra digital dan pembelajaran mesin menawarkan solusi yang menjanjikan untuk persoalan tersebut. Pembelajaran mesin memungkinkan model belajar dari data dan menghasilkan prediksi tanpa aturan yang diprogram secara eksplisit, sehingga banyak diterapkan pada analisis prediktif maupun pengolahan citra (Patil (ed) et al., 2024). Perkembangan arsitektur pembelajaran mendalam semakin memperluas cakupan penerapannya pada beragam tugas pengenalan citra (Mienye & Swart, 2024). Kendati demikian, penerapan pembelajaran mesin tetap menghadapi sejumlah tantangan, seperti risiko overfitting, kebutuhan data yang memadai, serta sensitivitas model terhadap ketidakseimbangan kelas (Hasanin et al., 2019; Liang et al., 2020; Park et al., 2023). Dalam ranah tekstil, sejumlah studi telah membuktikan efektivitas pendekatan ini. (Kaya, 2022) memadukan ekstraksi fitur tekstur berbasis GLCM dengan algoritma KNN untuk meningkatkan akurasi klasifikasi kain tenun, sedangkan (Wijaya & Widiarti, 2024) menerapkan KNN berbasis ekstraksi GLCM untuk mengklasifikasikan kain batik menurut pola dan warnanya. (Furferi & Servi, 2023) turut menunjukkan bahwa pendekatan berbasis visi mesin mampu mengklasifikasikan warna kain wol daur ulang dengan akurasi yang kompetitif dibandingkan metode lain. Penerapan serupa juga berkembang pada inspeksi cacat kain berbasis visi komputer (Carrilho et al., 2024; Guo et al., 2023; Kang & Zhang, 2020), maupun pada klasifikasi motif kain tradisional Indonesia seperti batik (Azmi et al., 2023; Ma'ruf et al., 2023; Meranggi et al., 2022).

Tiga algoritma klasifikasi yang banyak digunakan dalam konteks ini adalah KNN, SVM, dan Random Forest. KNN dikenal sederhana namun efektif untuk pengenalan pola citra karena bekerja berdasarkan kedekatan jarak antar data dan tidak menuntut asumsi distribusi yang kompleks (Daulay, 2024; Halder et al., 2024). SVM bekerja dengan

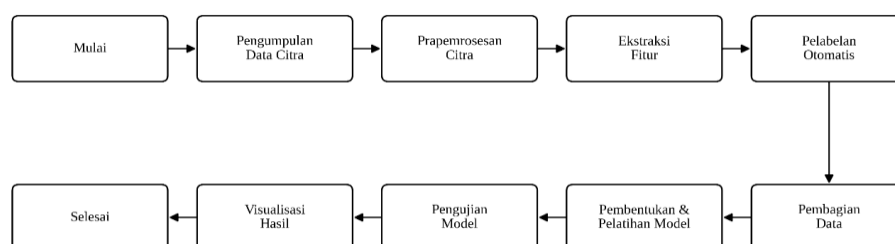
mencari hyperplane optimal yang memisahkan kelas melalui margin maksimum dan telah menunjukkan kinerja tinggi pada beragam persoalan, mulai dari analisis sentimen dengan akurasi mencapai 84,89 persen (Kristiyanti & Sri Hardani, 2023) hingga klasifikasi citra kompleks dengan akurasi 93,34 persen (Mehta et al., 2023). Sementara itu, Random Forest sebagai metode ensemble menggabungkan banyak pohon keputusan untuk memperoleh prediksi yang lebih stabil serta generalisasi yang lebih kuat terhadap data berdimensi tinggi (Ibrahim, 2022; Yuan et al., 2023). Keunggulan lain Random Forest adalah kemampuannya mengevaluasi tingkat kepentingan fitur, sehingga memudahkan identifikasi variabel yang paling berpengaruh dalam proses klasifikasi (Agarwal et al., 2025).

Meskipun ketiga algoritma tersebut telah luas dimanfaatkan, kajian yang secara khusus membandingkan KNN, SVM, dan Random Forest untuk klasifikasi tingkat kelunturan warna kain masih sangat terbatas. Penelitian terdahulu cenderung berfokus pada klasifikasi jenis maupun motif kain, bukan pada perubahan warna sebagai indikator mutu pencucian. Sebagian besar studi tersebut juga hanya mengandalkan fitur tunggal, padahal kombinasi fitur warna dan fitur bentuk berpotensi menghasilkan representasi yang lebih kaya terhadap fenomena kelunturan. Selain itu, tantangan ketidakseimbangan jumlah data antar kelas kelunturan masih jarang dibahas secara mendalam dalam literatur yang ada. Keterbatasan ini menunjukkan perlunya kajian komparatif yang tidak hanya mengukur akurasi keseluruhan, tetapi juga menilai kemampuan tiap model dalam mengenali kelas minoritas.

Berdasarkan celah tersebut, penelitian ini bertujuan mengimplementasikan sekaligus membandingkan kinerja algoritma KNN, SVM, dan Random Forest dalam mengklasifikasikan tingkat kelunturan warna kain berbasis citra digital. Fitur yang digunakan meliputi fitur warna pada ruang RGB dan HSV serta fitur bentuk berupa area dan rasio bentuk yang diperoleh melalui proses ekstraksi otomatis, kemudian dipetakan ke dalam tiga kategori, yaitu tidak luntur, cukup luntur, dan luntur. Kontribusi utama penelitian ini terletak pada penyediaan kerangka klasifikasi otomatis yang objektif serta analisis komparatif yang menilai keunggulan dan keterbatasan tiap algoritma, termasuk dampak ketidakseimbangan data terhadap akurasi. Dengan demikian, temuan penelitian diharapkan dapat mendukung peningkatan efisiensi operasional dan konsistensi mutu pada industri laundry, sekaligus memperkaya literatur pengolahan citra dan pembelajaran mesin pada domain tekstil.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental yang terdiri atas enam tahap utama, yaitu pengumpulan data citra, prapemrosesan, ekstraksi fitur, pelabelan otomatis, pembagian data, pembentukan & pelatihan model, pengujian model, visualisasi hasil, dan selesai. Alur lengkap penelitian ditunjukkan pada Gambar 1. Tujuan tahapan ini adalah menghasilkan model klasifikasi yang mampu menilai tingkat kelunturan warna kain secara otomatis dan objektif. Implementasi dilakukan menggunakan bahasa pemrograman Python dengan dukungan pustaka Scikit Learn, Scikit Image, NumPy, dan OpenCV.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Data berupa citra kain yang diperoleh langsung dari objek penelitian sebelum proses pencucian. Setiap kain difoto menggunakan kamera ponsel beresolusi 64 MP pada jarak sekitar 30 cm dengan pencahayaan yang stabil. Keceragaman jarak dan pencahayaan dijaga untuk meminimalkan variasi warna yang tidak berasal dari kondisi kain itu sendiri. Seluruh citra disimpan dalam format JPEG berukuran seragam 256×256 piksel. Total data yang dikumpulkan sebanyak 300 citra yang mewakili tiga kategori, yaitu tidak luntur, cukup luntur, dan luntur.

2.2 Prapemrosesan dan Ekstraksi Fitur

Setiap citra dikonversi menjadi ruang warna grayscale dan HSV, kemudian intensitas piksel dinormalisasi ke rentang 0 sampai 1. Fitur warna diperoleh dari nilai rerata kanal R, G, B pada ruang RGB serta kanal H, S, V pada ruang HSV. Penggunaan dua ruang warna dimaksudkan agar representasi warna lebih kuat, karena RGB sensitif terhadap perubahan pencahayaan sementara HSV memisahkan komponen warna dari kecerahan. Fitur bentuk diperoleh melalui segmentasi Otsu yang mengubah citra grayscale menjadi citra biner untuk memisahkan objek dari latar, lalu dihitung nilai area dan rasio bentuk dari objek tersegmentasi. Dengan demikian, tiap citra direpresentasikan oleh delapan fitur numerik.

2.3 Pelabelan Otomatis

Label kelas ditentukan secara otomatis melalui skor kelunturan yang menggabungkan beberapa metrik fitur dengan bobot tertentu, dirumuskan sebagai berikut:

$$\text{skor} = (0,30 \times (1 - V)) + (0,20 \times (1 - \mu_{\text{warna}})) + (0,15 \times \sigma_{\text{warna}}) + (0,05 \times (1 - \text{areaVal})) + (0,05 \times |1 - \text{rasio}|) \quad (1)$$

dengan V sebagai brightness, μ_{warna} sebagai rerata warna, dan σ_{warna} sebagai variasi warna. Pendekatan berbasis aturan ini menghilangkan kebutuhan anotasi manual sehingga penilaian menjadi konsisten dan dapat direplikasi. Penentuan kelas berdasarkan skor disajikan pada Tabel 1.

Tabel 1. Aturan pelabelan otomatis berdasarkan skor kelunturan

Rentang Skor	Kelas
$\text{skor} \leq 0,4$	Tidak Luntur
$0,4 < \text{skor} \leq 0,6$	Cukup Luntur
$\text{skor} > 0,6$	Luntur

2.4 Pembagian Data

Pembagian data merupakan tahap penting untuk memastikan bahwa model dievaluasi secara objektif pada data yang belum pernah dipelajari. Dataset dibagi menjadi data latih dan data uji dengan rasio 83,33 persen berbanding 16,67 persen, yakni 250 citra untuk pelatihan dan 50 citra untuk pengujian. Data latih digunakan sebagai dasar pembentukan model, sehingga algoritma dapat mempelajari pola hubungan antara fitur warna dan bentuk dengan tingkat kelunturan kain. Proporsi data latih yang lebih besar dimaksudkan agar model memperoleh contoh yang memadai untuk membangun batas keputusan yang stabil.

2.5 Pemodelan

Tiga algoritma dipilih karena mewakili paradigma yang berbeda, yaitu pembelajaran berbasis instans pada KNN, pemisahan margin pada SVM, dan pendekatan ensemble pada Random Forest. Pada KNN, jumlah tetangga ditentukan menggunakan $k = \sqrt{N}$, sedangkan kedekatan antar data dihitung dengan jarak Euclidean:

$$d = \sqrt{(\sum (x_i - y_i)^2)} \quad (2)$$

SVM dibangun menggunakan kernel RBF untuk memperoleh hyperplane optimal yang memaksimalkan margin antar kelas, dan Random Forest dibentuk dari kumpulan pohon keputusan dengan penentuan kelas melalui voting mayoritas. Ringkasan parameter disajikan pada Tabel 2.

Tabel 2. Parameter tiap algoritma

Algoritma	Parameter Utama
KNN	k = 3, jarak Euclidean
SVM	kernel RBF
Random Forest	n_estimators = 200, voting mayoritas

2.6 Evaluasi

Kinerja model dinilai menggunakan confusion matrix beserta empat metrik, yaitu accuracy, precision, recall, dan skor F1, yang dihitung sebagai berikut:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (3)$$

$$\text{Precision} = TP / (TP + FP) \quad (4)$$

$$\text{Recall} = TP / (TP + FN) \quad (5)$$

$$\text{Skor F1} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (6)$$

dengan TP, TN, FP, dan FN secara berturut menyatakan true positive, true negative, false positive, dan false negative. Selain nilai keseluruhan, metrik dihitung pula per kelas serta dalam bentuk rerata makro dan rerata terbobot untuk mengukur kinerja model pada kelas mayoritas maupun minoritas.

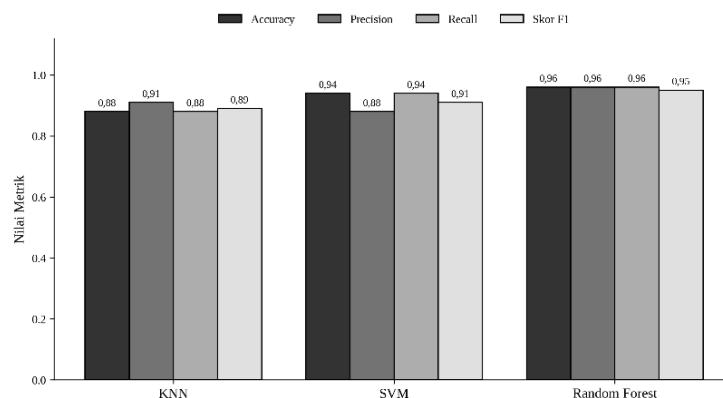
3. HASIL DAN PEMBAHASAN

3.1 Perbandingan Kinerja Keseluruhan Model

Tahap awal analisis difokuskan pada perbandingan kinerja menyeluruh ketiga algoritma sebelum ditelaah secara lebih rinci. Evaluasi dilakukan terhadap 50 citra data uji untuk menilai kinerja tiga algoritma, yaitu KNN, SVM, dan Random Forest, dalam mengklasifikasikan tingkat kelunturan warna kain. Rekapitulasi kinerja keseluruhan disajikan pada Tabel 3, sedangkan perbandingannya divisualisasikan pada Gambar 2. Secara umum, Random Forest memperoleh kinerja tertinggi dengan akurasi 0,96, diikuti SVM dengan akurasi 0,94, dan KNN dengan akurasi 0,88.

Tabel 3. Perbandingan kinerja ketiga model

Model	Accuracy	Precision	Recall	Skor F1
KNN	0,88	0,91	0,88	0,89
SVM	0,94	0,88	0,94	0,91
Random Forest	0,96	0,96	0,96	0,95



Gambar 2. Perbandingan kinerja ketiga model

3.2 Analisis Confusion Matrix dan Perhitungan Akurasi

Untuk memahami sumber perbedaan akurasi tersebut, analisis dilanjutkan pada struktur confusion matrix tiap model. Nilai akurasi diturunkan dari confusion matrix tiap model yang dirangkum pada Tabel 4. Dalam evaluasi ini, kelas Cukup Luntur diperlakukan sebagai kelas positif dan kelas Tidak Luntur sebagai kelas negatif.

Tabel 4. Confusion matrix ketiga model

Model	TP	TN	FP	FN
KNN	1	43	4	2
SVM	0	47	0	3
Random Forest	1	47	0	2

Akurasi dihitung menggunakan persamaan $Accuracy = (TP + TN) / Total$. Untuk KNN diperoleh $(1 + 43) / 50 = 0,88$, untuk SVM $(0 + 47) / 50 = 0,94$, dan untuk Random Forest $(1 + 47) / 50 = 0,96$. Hasil ini menegaskan keunggulan Random Forest dalam mengenali pola kelunturan secara keseluruhan.

3.3 Evaluasi Metrik per Kelas (Precision, Recall, dan Skor F1)

Selain akurasi, kualitas klasifikasi perlu ditinjau melalui metrik per kelas agar perilaku model pada tiap kategori dapat dipahami secara lebih utuh. Sebagai ilustrasi perhitungan metrik kelas positif pada Random Forest, diperoleh $Precision = TP / (TP + FP) = 1 / (1 + 0) = 1,00$, $Recall = TP / (TP + FN) = 1 / (1 + 2) = 0,33$, dan $Skor F1 = 2 \times (1,00 \times 0,33) / (1,00 + 0,33) \approx 0,50$.

Nilai precision yang sempurna menunjukkan bahwa setiap prediksi Cukup Luntur memang benar, tetapi recall yang rendah memperlihatkan bahwa sebagian besar sampel Cukup Luntur belum berhasil dikenali. Sebagai pembandingan, perhitungan pada KNN menghasilkan $Precision = 1 / (1 + 4) = 0,20$, $Recall = 1 / (1 + 2) = 0,33$, dan $Skor F1 \approx 0,25$. Empat sampel Tidak Luntur yang terprediksi sebagai Cukup Luntur menjadi penyebab utama rendahnya precision KNN pada kelas minoritas.

3.4 Analisis Ketidakseimbangan Kelas

Evaluasi per kelas mengungkap kesenjangan kinerja yang dipengaruhi oleh distribusi data antar kategori. Pola serupa muncul pada ketiga model ketika dianalisis per kelas. Pada kelas mayoritas Tidak Luntur, seluruh model bekerja sangat baik, dengan Skor F1 berkisar antara 0,93 hingga 0,98. Sebaliknya, pada kelas minoritas Cukup Luntur kinerjanya menurun tajam. KNN hanya mencapai recall 0,33, sedangkan SVM gagal sepenuhnya mengenali kelas tersebut sehingga precision, recall, dan Skor F1 bernilai 0. Random Forest relatif paling baik karena tetap mampu mengidentifikasi sebagian sampel minoritas dengan precision 1,00 meskipun recall tetap rendah. Perbedaan ini tampak jelas ketika membandingkan rerata makro dan rerata terbobot. Rerata terbobot tetap tinggi karena didominasi kelas mayoritas, sedangkan rerata makro yang memberi bobot setara pada tiap kelas menurun signifikan, terutama pada SVM. Hal ini menunjukkan bahwa akurasi keseluruhan saja tidak cukup untuk menggambarkan kinerja pada data yang tidak seimbang. Kesenjangan tersebut bersumber dari komposisi data uji yang timpang, yaitu hanya tiga sampel Cukup Luntur berbanding 47 sampel Tidak Luntur.

3.5 Pembahasan Komparatif Antar Algoritma

Perbedaan karakteristik tiap algoritma turut menjelaskan variasi kinerja yang teramati. Pada SVM, hyperplane cenderung bergeser ke arah kelas mayoritas sehingga seluruh sampel minoritas terklasifikasi sebagai Tidak Luntur. KNN yang berbasis kedekatan jarak juga rentan karena tetangga terdekat didominasi kelas mayoritas.

Random Forest lebih tahan terhadap kondisi tersebut karena struktur ensemble menggabungkan keputusan banyak pohon sehingga menghasilkan generalisasi yang lebih stabil. Temuan ini sejalan dengan studi terdahulu. Adewale et al. (2023) menunjukkan kinerja tinggi SVM pada data yang relatif seimbang, sementara Yuan et al. (2023) menegaskan keandalan Random Forest pada data citra berdimensi tinggi, dan Furferi serta Servi (2023) membuktikan efektivitas pendekatan berbasis fitur warna untuk klasifikasi kain.

3.6 Implikasi Praktis dan Keterbatasan

Temuan tersebut memiliki sejumlah implikasi praktis sekaligus keterbatasan yang perlu dicatat. Secara praktis, Random Forest paling layak diterapkan sebagai mesin klasifikasi otomatis tingkat kelunturan warna kain karena memberikan akurasi tinggi dan prediksi yang konsisten, sehingga dapat mempercepat pemeriksaan hasil cucian sekaligus mengurangi subjektivitas penilaian manual. Akan tetapi, rendahnya recall pada kelas minoritas menjadi keterbatasan utama penelitian ini. Penambahan sampel pada kelas Cukup Luntur serta penerapan teknik penyeimbangan data berpotensi meningkatkan kemampuan model dalam mengenali kategori yang jarang muncul.

4. KESIMPULAN

Penelitian ini menerapkan dan membandingkan tiga algoritma klasifikasi, yaitu KNN, SVM, dan Random Forest, untuk mengklasifikasikan tingkat kelunturan warna kain berbasis citra digital menggunakan fitur warna (RGB dan HSV) serta fitur bentuk (area dan rasio bentuk) yang diekstraksi secara otomatis ke dalam tiga kategori, yaitu tidak luntur, cukup luntur, dan luntur. Pengujian terhadap 50 data uji menghasilkan akurasi tertinggi pada Random Forest sebesar 0,96, diikuti SVM 0,94, dan KNN 0,88. Akan tetapi, perlu ditegaskan bahwa akurasi keseluruhan ini bersifat menyesatkan dan tidak dapat menjadi tolok ukur tunggal kinerja, karena terbias oleh distribusi kelas yang sangat timpang, yaitu dominasi kelas Tidak Luntur (47 sampel) terhadap Cukup Luntur (3 sampel). Angka tersebut lebih mencerminkan keberhasilan mengenali kelas mayoritas, sedangkan evaluasi per kelas memperlihatkan seluruh model lemah pada kelas minoritas, bahkan SVM gagal sepenuhnya dengan recall dan skor F1 bernilai nol. Oleh karena itu, kinerja sebenarnya lebih tepat dinilai melalui metrik per kelas, recall kelas minoritas, dan rerata makro. Random Forest tetap paling tahan terhadap ketidakseimbangan data, namun interpretasi hasil harus dilakukan dengan cermat. Penelitian selanjutnya disarankan menambah sampel kelas minoritas, menerapkan teknik penyeimbangan seperti SMOTE, serta mengevaluasi model dengan metrik yang sensitif terhadap ketidakseimbangan kelas.

REFERENSI

- Agarwal, A., Kenney, A. M., Tan, Y. S., Tang, T. M., & Yu, B. (2025). *Integrating Random Forests and Generalized Linear Models for Improved Accuracy and Interpretability*.
- Aliei, H., Carrera-Gallisa, E., & Cayuela, D. (2024). Evaluating the Impact of Washing Conditions on the Color Changes of Naturally Colored Cotton Fabrics: A Focus on Detergents, Water Types, and Temperature. *Materials*, 17(23), 5777. <https://doi.org/10.3390/ma17235777>
- Aliei, H., Cayuela, D., Capdevila, F. J., & Carrera-Gallissà, E. (2024). Influence of Household Washing on the Colorimetric Properties of Intrinsically Natural Organic Cotton Fabrics. *Journal of Natural Fibers*, 21(1). <https://doi.org/10.1080/15440478.2024.2351160>
- Arippudin, A. (2022). LAUNDRY SERVICE BUSINESS MANAGEMENT. *Neo Journal of Economy and Social Humanities*, 1(1), 9–16. <https://doi.org/10.56403/nejesh.v1i1.3>

- Azmi, K., Defit, S., & Sumijan, S. (2023). Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat. *JURNAL UNITEK*, 16(1), 28–40. <https://doi.org/10.52072/unitek.v16i1.504>
- Carrilho, R., Yaghoubi, E., Lindo, J., Hambarde, K., & Proença, H. (2024). Toward Automated Fabric Defect Detection: A Survey of Recent Computer Vision Approaches. *Electronics*, 13(18), 3728. <https://doi.org/10.3390/electronics13183728>
- Daulay, R. S. (2024). Analisis Kritis dan Pengembangan Algoritma K-Nearest Neighbor (KNN): Sebuah Tinjauan Literatur. *Jurnal Pendidikan Sains Dan Komputer*, 4(02), 131–141. <https://doi.org/10.47709/jpsk.v4i02.5055>
- Furferi, R., & Servi, M. (2023). A Machine Vision-Based Algorithm for Color Classification of Recycled Wool Fabrics. *Applied Sciences*, 13(4), 2464. <https://doi.org/10.3390/app13042464>
- Guo, Y., Kang, X., Li, J., & Yang, Y. (2023). Automatic Fabric Defect Detection Method Using AC-YOLOv5. *Electronics*, 12(13), 2950. <https://doi.org/10.3390/electronics12132950>
- Halder, R. K., Uddin, M. N., Uddin, Md. A., Aryal, S., & Khraisat, A. (2024). Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. *Journal of Big Data*, 11(1), 113. <https://doi.org/10.1186/s40537-024-00973-y>
- Hasanin, T., Khoshgoftaar, T. M., Leevy, J. L., & Bauder, R. A. (2019). Severely imbalanced Big Data challenges: investigating data sampling approaches. *Journal of Big Data*, 6(1), 107. <https://doi.org/10.1186/s40537-019-0274-4>
- Ibrahim, S. (2022). Improving Land Use/Cover Classification Accuracy from Random Forest Feature Importance Selection Based on Synergistic Use of Sentinel Data and Digital Elevation Model in Agriculturally Dominated Landscape. *Agriculture*, 13(1), 98. <https://doi.org/10.3390/agriculture13010098>
- Kang, X., & Zhang, E. (2020). A Universal and Adaptive Fabric Defect Detection Algorithm Based on Sparse Dictionary Learning. *IEEE Access*, 8, 221808–221830. <https://doi.org/10.1109/ACCESS.2020.3041849>
- Kaya, J. U. (2022). *Klasifikasi Motif Kain Tenun menggunakan K-Nearest Neighbor Berdasarkan Gray Level Co-occurrence Matrix*. 3(2). <https://ejournalunsam.id/index.php/jicom/>
- Kristiyanti, D. A., & Sri Hardani. (2023). Sentiment Analysis of Public Acceptance of Covid-19 Vaccines Types in Indonesia using Naïve Bayes, Support Vector Machine, and Long Short-Term Memory (LSTM). *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(3), 722–732. <https://doi.org/10.29207/resti.v7i3.4737>
- Liang, X. W., Jiang, A. P., Li, T., Xue, Y. Y., & Wang, G. T. (2020). LR-SMOTE — An improved unbalanced data set oversampling based on K-means and SVM. *Knowledge-Based Systems*, 196, 105845. <https://doi.org/10.1016/j.knosys.2020.105845>
- Ma'ruf, M. T., Putra, E. D., Reswan, Y., & Juhardi, U. (2023). Classification Of Besurek Batik Fabrics Using Gray Level Co-Occurrence Matrix (GLCM) Features Extraction. *Jurnal Komputer, Informasi Dan Teknologi*, 3(2). <https://doi.org/10.53697/jkomitek.v3i2.1211>
- Mehta, A., Sengupta, P., Garg, D., Singh, H., & Diamand, Y. S. (2023). *Benchmarking the Effectiveness of Classification Algorithms and SVM Kernels for Dry Beans*.
- Meranggi, D. G. T., Yudistira, N., & Sari, Y. A. (2022). Batik Classification Using Convolutional Neural Network with Data Improvements. *JOIV : International Journal on Informatics Visualization*, 6(1), 6. <https://doi.org/10.30630/joiv.6.1.716>
- Mienye, I. D., & Swart, T. G. (2024). A Comprehensive Review of Deep Learning: Architectures, Recent Advances, and Applications. *Information*, 15(12), 755. <https://doi.org/10.3390/info15120755>

- Park, J., Kwon, S., & Jeong, S.-P. (2023). A study on improving turnover intention forecasting by solving imbalanced data problems: focusing on SMOTE and generative adversarial networks. *Journal of Big Data*, 10(1), 36. <https://doi.org/10.1186/s40537-023-00715-6>
- Patil (ed), D., Rane (ed), N. L., Desai (ed), P., & Rane (ed), J. (2024). Trustworthy Artificial Intelligence in Industry and Society. In *Trustworthy Artificial Intelligence in Industry and Society*. Deep Science Publishing. <https://doi.org/10.70593/978-81-981367-4-9>
- Purnamaningtyas, I., Fajria, H., & Bakti, S. (2022). *Survey of Maintenance Implementation on Laundry Business in Indonesia, Jakarta Region*. <https://ieomsociety.org/proceedings/2022istanbul/112.pdf>
- Wijaya, D., & Widiarti, A. R. (2024). Batik classification using KNN algorithm and GLCM features extraction. *E3S Web of Conferences*, 475, 02012. <https://doi.org/10.1051/e3sconf/202447502012>
- Yuan, X., Liu, S., Feng, W., & Dauphin, G. (2023). Feature Importance Ranking of Random Forest-Based End-to-End Learning Algorithm. *Remote Sensing*, 15(21), 5203. <https://doi.org/10.3390/rs15215203>