

# **LAPORAN AKHIR TUGAS BESAR REKAYASA FITUR**

Diajukan Untuk Memenuhi Tugas Rekayasa Fitur

Dosen Pengampu: Galih Hermawan, S.Kom, M.T



Disusun oleh:

IF 3

Kelompok 3

Fajar Gustiana	10122081
Muhlas Putra Siswaji	10122092
Ryan Bachtiar	10122097
Dyandra Cissy Fadhillah	10122104

**TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK DAN ILMU KOMPUTER**

**UNIVERSITAS KOMPUTER INDONESIA**

**2025**

## **Abstrak**

Pakaian adalah salah satu kebutuhan pokok manusia yang berfungsi untuk melindungi tubuh dari cuaca ekstrim, serangga, penutup aurat. Seiring dengan berjalannya waktu Industri fashion dan e-commerce mengalami pertumbuhan pesat. Platform online membutuhkan sistem otomatis untuk mengkategorikan ribuan produk pakaian yang diunggah setiap hari. Proses manual memakan waktu, tenaga, dan biaya yang besar. Oleh karena itu perlunya kemampuan secara otomatis mengklasifikasikan pakaian menjadi kategori-kategori spesifik seperti celana panjang, gaun, dan jaket, sistem ini dapat membantu dalam otomatisasi proses katalogisasi produk fashion, pencarian produk berdasarkan gambar, dan rekomendasi pakaian. Klasifikasi adalah metode yang digunakan untuk membedakan, mengenali, mengelompokkan dan memahami item ke kelas yang ditentukan. Support Vector Machine (SVM) merupakan salah satu algoritma untuk metode klasifikasi. Sistem ini memanfaatkan kombinasi berbagai teknik ekstraksi fitur visual untuk meningkatkan akurasi klasifikasi. Histogram of Oriented Gradients (HOG) digunakan untuk menangkap informasi bentuk dan tekstur, Local Binary Pattern (LBP) untuk menganalisis pola tekstur lokal, dan Color Histogram untuk mengekstrak informasi warna dari citra pakaian. Hasil menunjukkan bahwa kombinasi HOG + Color Histogram dan HOG + LBP + Color Histogram mencapai akurasi tertinggi sebesar 83.3%. Sebaliknya, kombinasi Color Histogram + LBP menghasilkan akurasi terendah (50.0%),

## **BAB I**

### **PENDAHULUAN**

#### **1.1 LATAR BELAKANG**

Pakaian adalah salah satu kebutuhan pokok manusia yang berfungsi untuk melindungi tubuh dari cuaca ekstrim, serangga, penutup aurat. Seiring dengan berkembangnya zaman, pakaian bisa menentukan status sosial seseorang, adat istiadat, budaya, agama bahkan gender. Perbedaan pakaian perempuan dan laki-laki terdapat pada gaya warna kain. Pakaian laki-laki lebih praktis dibanding perempuan. Tetapi pakaian perempuan lebih luas dalam model dan warna. Pakaian juga dapat digunakan untuk menunjukkan peringkat atau status. Di Roma kuno, misalnya, hanya para senator yang diizinkan untuk memakai pakaian yang dicelup dengan warna ungu Tyrian. Pada awalnya, manusia memanfaatkan kulit pepohonan dan kulit hewan sebagai bahan pakaian, kemudian memanfaatkan benang yang dipintal dari kapas, bulu domba serta

sutera yang kemudian dijadikan kain sebagai bahan dasar pakaian. Kini dikenal berbagai macam jenis kain, di antaranya sutera, katun, nilon, linen, spandex dan lainnya. Pakaian juga banyak jenisnya misalnya baju, kemeja, kaos, rok, celana, jaket [1].

Pada penelitian ini akan melakukan klasifikasi jenis pakaian menggunakan berbagai teknik ekstraksi fitur (HOG, Histogram Warna, dan LBP) dan algoritma SVM. Dataset yang digunakan 3 jenis pakaian yaitu celana panjang, jaket dan gaun masing-masing 50 gambar sehingga total dataset 150 gambar.

## **1.2 TUJUAN**

Penelitian ini berupaya untuk membangun solusi teknologi yang dapat diaplikasikan pada e-commerce fashion, sistem inventori toko pakaian, aplikasi belanja online, dan asisten virtual untuk fashion. Dengan kemampuan untuk secara otomatis mengklasifikasikan pakaian menjadi kategori-kategori spesifik seperti celana panjang, gaun, dan jaket, sistem ini dapat membantu dalam otomatisasi proses katalogisasi produk fashion, pencarian produk berdasarkan gambar, dan rekomendasi pakaian.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 KONSEP DASAR**

Klasifikasi adalah metode yang digunakan untuk membedakan, mengenali, mengelompokkan dan memahami item ke kelas yang ditentukan. Klasifikasi juga merupakan bagian dari machine learning dimana mesin belajar melalui data training. Setelah mesin dilatih dilakukan pengujian dengan data testing. Klasifikasi dapat didefinisikan sebagai pelatihan kerja/pembelajaran untuk menargetkan fungsi yang memetakan setiap set atribut (fitur) ke salah satu dari sejumlah label kelas yang tersedia [6].

#### **2.2 PENELITIAN TERKAIT**

Penelitian yang dilakukan oleh [2] penelitian membahas klasifikasi jenis kain berdasarkan tekstur dengan metode support vector machine berbasis web flask yang berfokus pada citra tekstil klasifikasi berdasarkan teksturnya. Metode ekstraksi citra yang diimplementasikan adalah Gray Level Co-occurrence Matrix (GLCM).

Penelitian lainnya [4] menggunakan citra ikan laut yang diambil dari public dataset yang memiliki 7 jenis ikan laut dimana setiap jenis ikan laut ada 7.000 gambar yang akan dilakukan tahap segmentasi warna HSV dengan mengambil nilai value sehingga menjadi grayscale yang akan dilanjutkan ke proses HOG dan untuk mengklasifikasikan jenis ikan laut menggunakan SVM.

## **BAB III**

### **METODOLOGI**

#### **3.1 DATASET**

Dataset berupa gambar- gambar pakaian yang diambil dari Carousell, sebuah pasar online, yang secara khusus disusun untuk tugas klasifikasi gambar. Dataset yang digunakan 3 jenis pakaian yaitu celana panjang, jaket dan gaun. Dataset berjumlah 70 gambar setiap jenis pakaian sehingga total ada 210 gambar. Data dibagi menjadi 80% untuk data training dan 20% untuk data testing

#### **3.2 PRA-PEMROSESAN DATA**

Pre-processing dimulai dengan tahap crop image. Crop image bertujuan untuk memfokuskan gambar dari resolusi  $720 \times 1280$  pixel menjadi  $120 \times 120$  pixel. Tahapan kedua adalah konversi format dari BGR ke RGB untuk konsistensi pemrosesan. Selanjutnya data dimasukkan ke dalam struktur Numpy untuk efisiensi komputasi.

Label kategori pakaian di-encode menggunakan `LabelEncoder()` dari scikit-learn untuk mengkonversi label string menjadi representasi numerik yang dapat diproses oleh algoritma machine learning.

#### **3.3 PEMODELAN**

Support Vector Machine (SVM) merupakan salah satu algoritma pada metode klasifikasi. SVM bekerja atas prinsip Structural Risk Minimization (SRM) dengan tujuan menemukan hyperplane terbaik. SVM pada dasarnya didesain untuk klasifikasi 2-kelas, namun dalam beberapa makalah disebutkan bahwa kombinasi K SVM dapat digunakan untuk menyelesaikan masalah klasifikasi K kelas [3]. Alasan dipilih SVM terletak pada kemampuannya untuk menemukan hyperplane terbaik yang memisahkan class pada feature space yang ditunjang oleh strategi Structural Risk Minimization (SRM). SVM juga cocok untuk dataset berdimensi tinggi seperti hasil ekstraksi HOG dan color histogram.

Histogram of Oriented Gradients (HOG) merupakan sebuah metode yang digunakan untuk mengekstraksi fitur bentuk dari sebuah citra dalam bentuk histogram. HOG menjadi unggul dalam merepresentasikan fitur bentuk karena mampu menangkap gradien dari karakteristik lokal pada sebuah citra [4].

Gambar Asli



HOG Visualization

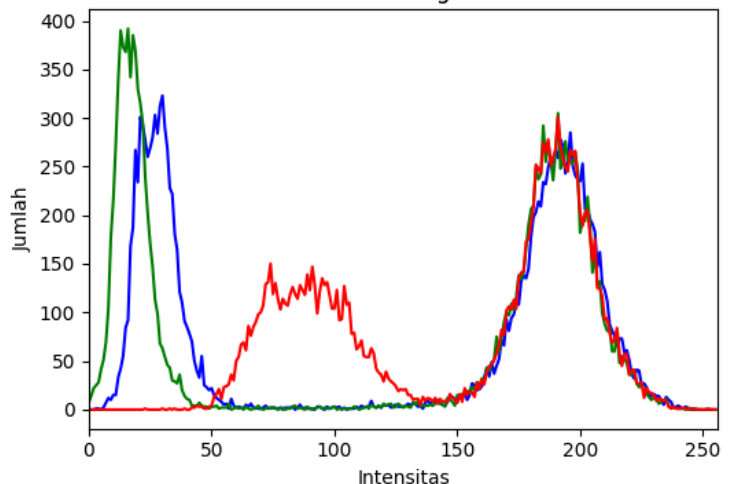


Dalam model warna RGB, warna dibuat dengan mencampurkan tiga warna dasar: merah (Red), hijau (Green), dan biru (Blue) dalam berbagai intensitas. Setiap warna pada gambar digital direpresentasikan oleh kombinasi dari ketiga komponen ini dengan nilai intensitas yang bervariasi dari 0 hingga 255, menghasilkan sekitar 16 juta warna yang berbeda. Ruang warna RGB memainkan peran penting dalam identifikasi produk pertanian karena banyak produk pertanian yang memiliki ciri khas warna yang dapat diidentifikasi secara visual [5].

Gambar Asli



Color Histogram



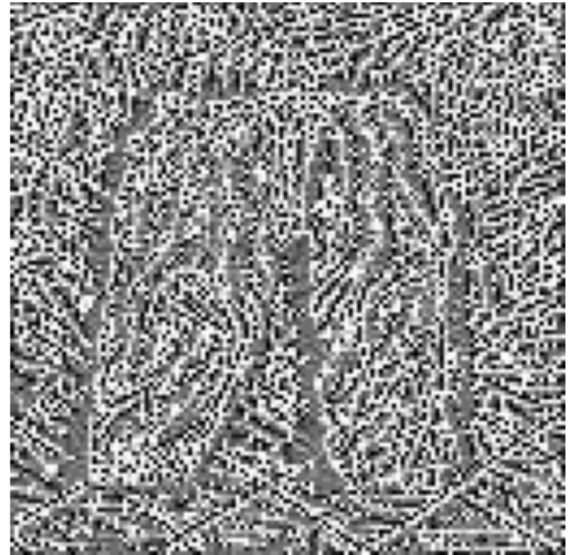
Local Binary Pattern (LBP) adalah operator sederhana dan efisien untuk menggambarkan pola gambar dan mendapatkan hasil yang baik dalam tekstur pengambilan

gambar. Keuntungan dari LBP adalah perubahan grayscale monoatomik tidak berubah, kompleksitas komputasi yang rendah dan ekstensi multiskala yang mudah [6].

Gambar Asli



LBP Visualization



Penelitian ini mengevaluasi empat kombinasi fitur yang berbeda:

1. **HOG + Color Histogram:** Menggabungkan informasi bentuk dan warna
2. **Color Histogram + LBP:** Menggabungkan informasi warna dan tekstur
3. **LBP + HOG:** Menggabungkan informasi tekstur dan bentuk
4. **HOG + LBP + Color Histogram:** Menggabungkan ketiga jenis fitur

### 3.4 EVALUASI

Kinerja model dievaluasi menggunakan beberapa metrik :

1. Akurasi (Proporsi prediksi yang benar dari total prediksi)  
$$\text{Akurasi} = (TP + TN) / (TP + TN + FP + FN)$$
2. Precision (Proporsi prediksi positif yang benar)  
$$\text{Precision} = TP / (TP + FP)$$
3. Recall (Proporsi sampel positif yang berhasil diidentifikasi)  
$$\text{Recall} = TP / (TP + FN)$$
4. F1-Score (Harmonic mean dari precision dan recall)  
$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

Confusion Matrix:

- True Positive (TP): Prediksi benar untuk setiap kelas
- False Positive (FP): Kesalahan klasifikasi
- False Negative (FN): Kelas yang terlewat

### 3.5 KONTRIBUSI ANGGOTA

Nama	NIM	Kontribusi Utama
Fajar Gustiana	10122081	<ul style="list-style-type: none"><li>• Implementasi ekstraksi fitur HOG (<code>extract_hog_features()</code>)</li><li>• Implementasi ekstraksi fitur Color Histogram (<code>extract_color_histogram_features()</code>)</li><li>• Optimisasi parameter untuk kedua metode ekstraksi fitur</li></ul>
Muhlas Putra Siswaji	10122092	<ul style="list-style-type: none"><li>• Implementasi ekstraksi fitur LBP (<code>extract_lbp_features()</code>)</li><li>• Pengembangan pipeline kombinasi fitur</li><li>• Implementasi dan tuning model SVM</li></ul>
Ryan Bachtiar	10122097	<ul style="list-style-type: none"><li>• Implementasi sistem evaluasi dan metrik performa</li><li>• Pembuatan visualisasi (confusion matrix, bar chart perbandingan)</li><li>• Analisis hasil dan interpretasi performa model</li><li>• Dokumentasi dan pelaporan hasil eksperimen</li></ul>
Dyandra Cissy Fadhillah	10122104	<ul style="list-style-type: none"><li>• Pengumpulan dan organisasi dataset gambar pakaian</li><li>• Implementasi fungsi pra-pemrosesan data (<code>load_images_and_labels()</code>)</li><li>• Dokumentasi struktur data dan pipeline pemrosesan awal</li></ul>
Kolaborasi Tim		<ul style="list-style-type: none"><li>• Brainstorming pendekatan metodologi</li><li>• Review dan debugging kode</li><li>• Validasi hasil eksperimen</li><li>• Penulisan laporan</li></ul>



## BAB IV

### HASIL DAN PEMBAHASAN

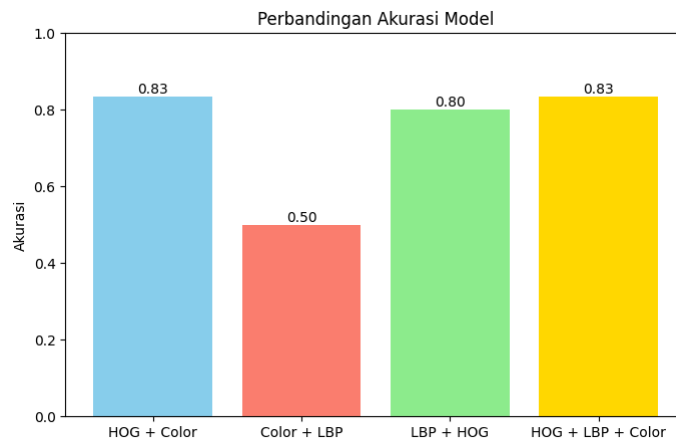
#### 4.1 HASIL EVALUASI MODEL DAN PEMBAHASAN

##### 1. Distribusi dataset

Kategori	Data Train	Data Test	Total
Celana_Panjang	50	20	70
Gaun	50	20	70
Jaket	50	20	70
Total	150	60	210

##### 2. Performa Model dengan Kombinasi Fitur Berbeda

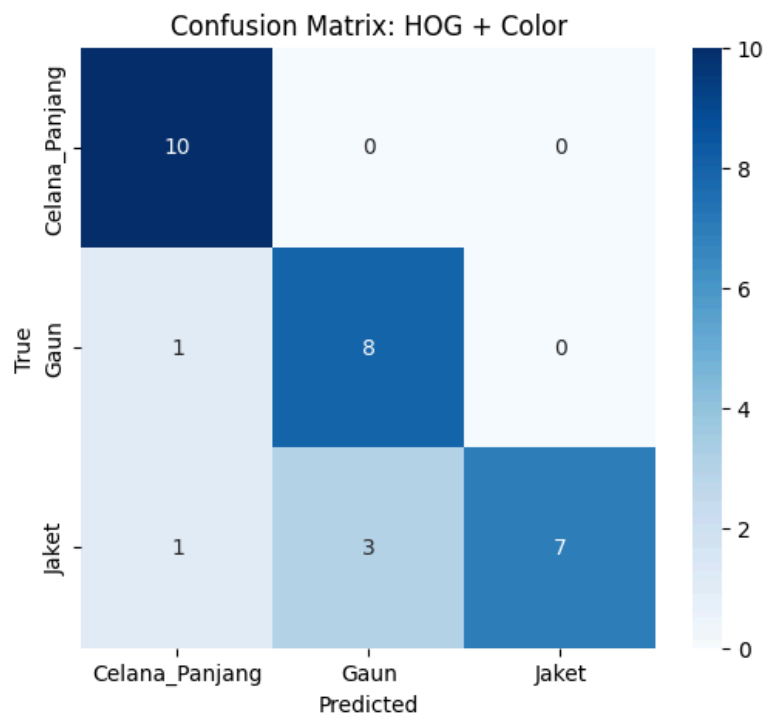
Kombinasi Fitur	Akurasi	(%)
HOG + Color Histogram	0,833	83,3%
Color Histogram + LBp	0,500	50,0%
LBP + HOG	0,800	80,0%
HOG+LBP+Color Histogram	0,833	83.3%



### 3. Analisis Detail per Kombinasi Fitur

**Table Classification Report HOG + Color Histogram**

Kelas	Precision	Recall	F1-Score	Support
Celana Panjang	0.83	1.00	0.91	10
Gaun	0.73	0.89	0.80	9
Jaket	1.00	0.64	0.78	11
<b>Weighted Avg</b>	<b>0.86</b>	<b>0.83</b>	<b>0.83</b>	<b>30</b>



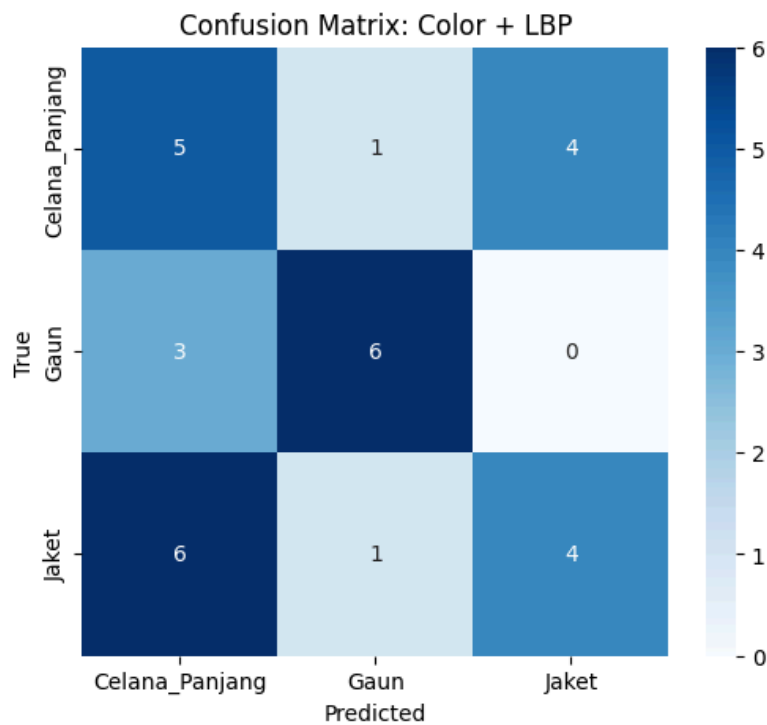
#### Confusion Matrix Analysis:

- Celana Panjang: Prediksi sempurna (10/10 benar)
- Gaun: 8 benar, 1 salah diklasifikasi sebagai Jaket
- Jaket: 7 benar, 3 salah diklasifikasi sebagai Gaun, 1 sebagai Celana Panjang

**Table Classification Report Color Histogram + LBP**

Kelas	Precision	Recall	F1-Score	Support
-------	-----------	--------	----------	---------

Celana Panjang	0.36	0.50	0.42	10
Gaun	0.75	0.67	0.71	9
Jaket	0.50	0.36	0.42	11
<b>Weighted Avg</b>	<b>0.53</b>	<b>0.50</b>	<b>0.51</b>	<b>30</b>



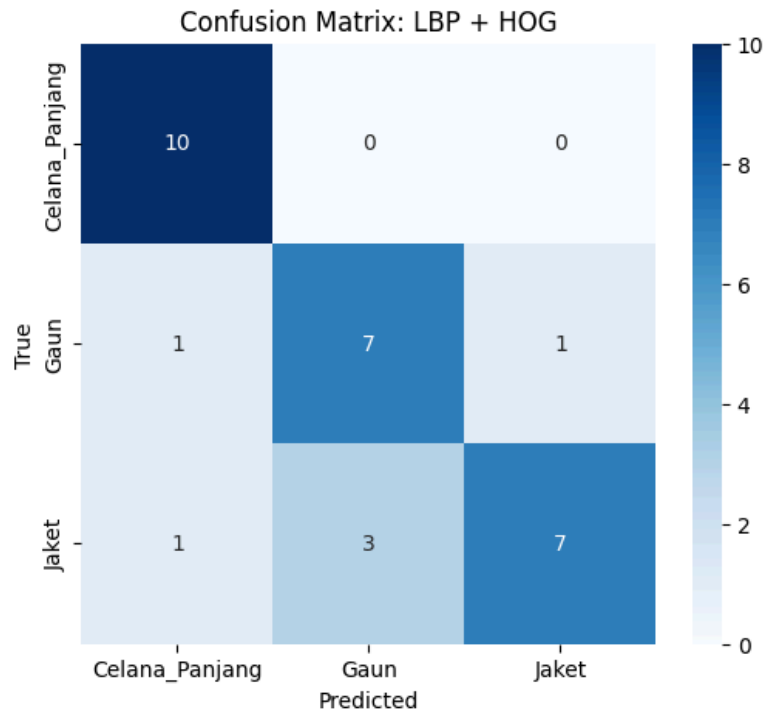
### Confusion Matrix Analysis:

- Performa terburuk dengan banyak misklasifikasi
- Celana Panjang: 5 benar, 1 sebagai Gaun, 4 sebagai Jaket
- Gaun: 6 benar, 3 sebagai Celana Panjang
- Jaket: 4 benar, 6 sebagai Celana Panjang, 1 sebagai Gaun

### Table Classification Report LBP + HOG

Kelas	Precision	Recall	F1-Score	Support
Celana Panjang	0.83	1.00	0.91	10
Gaun	0.70	0.78	0.74	9

Jaket	0.88	0.64	0.74	11
<b>Weighted Avg</b>	0.81	0.80	0.79	30

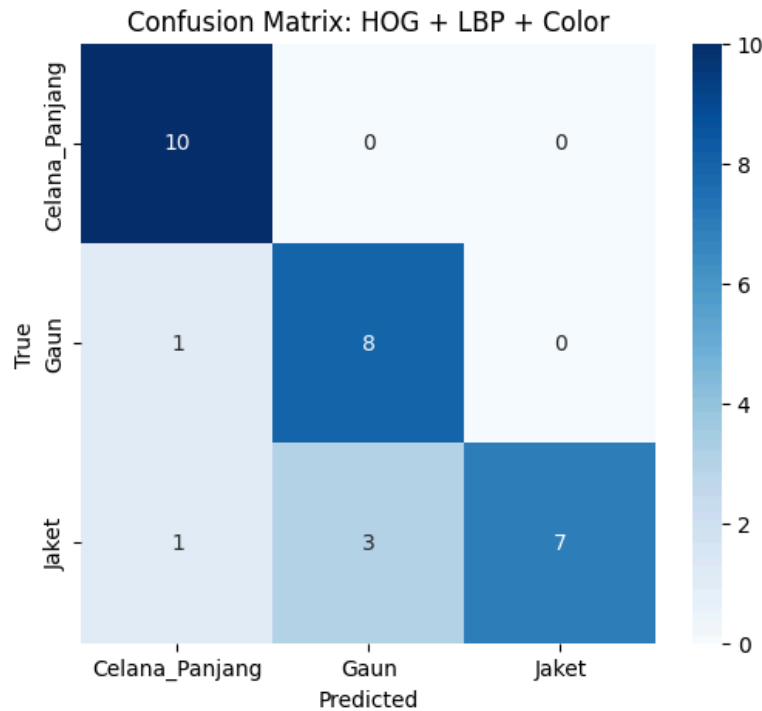


### Confusion Matrix Analysis:

- Celana Panjang: Prediksi sempurna (10/10 benar)
- Gaun: 7 benar, 1 sebagai Celana Panjang, 1 sebagai Jaket
- Jaket: 7 benar, 1 sebagai Celana Panjang, 3 sebagai Gaun

### Table Classification Report HOG + LBP + Color Histogram

Kelas	Precision	Recall	F1-Score	Support
Celana Panjang	0.83	1.00	0.91	10
Gaun	0.73	0.89	0.80	9
Jaket	1.00	0.64	0.78	11
<b>Weighted Avg</b>	0.86	0.83	0.83	30



Berdasarkan hasil eksperimen, beberapa temuan penting dapat diidentifikasi, kedua kombinasi **HOG + Color Histogram** dan **HOG + LBP + Color Histogram** mencapai akurasi tertinggi sebesar 83.3%. Hal ini menunjukkan bahwa **Fitur HOG** memberikan kontribusi signifikan dalam pengenalan bentuk dan struktur pakaian. **Color Histogram** efektif dalam membedakan kategori pakaian berdasarkan karakteristik warna. Penambahan fitur LBP pada kombinasi lengkap tidak meningkatkan performa, menunjukkan kemungkinan redundansi informasi tekstur

Sebaliknya, kombinasi **Color Histogram + LBP** menghasilkan akurasi terendah (50.0%), yang mengindikasikan tanpa informasi bentuk dari HOG, klasifikasi menjadi sangat sulit. Tekstur (LBP) dan warna saja tidak cukup untuk membedakan kategori pakaian secara efektif serta terjadi overfitting atau informasi yang tidak cukup diskriminatif

Kelas "Celana Panjang" mencapai recall sempurna (1.00) pada tiga dari empat kombinasi fitur. Celana panjang memiliki karakteristik visual yang paling distinktif. Bentuk dan struktur celana panjang juga mudah dikenali oleh fitur HOG. Konsistensi warna pada celana panjang berpengaruh dalam mendukung klasifikasi yang akurat

## 4.2 KETERBATASAN PENELITIAN

- **Ukuran Dataset Terbatas:** Salah satu keterbatasan utama dalam penelitian ini adalah ukuran dataset yang relatif kecil. Evaluasi model mungkin belum sepenuhnya representatif terhadap performa model pada skala yang lebih besar dan kondisi real-world yang lebih beragam.
- **Kategori Pakaian Terbatas:** Penelitian ini hanya fokus pada tiga kategori pakaian (Celana Panjang, Gaun, Jaket). Dalam aplikasi nyata, sistem klasifikasi pakaian perlu mampu menangani kategori yang jauh lebih beragam seperti sepatu, aksesoris, dan berbagai subcategory dalam setiap jenis pakaian.
- **Variasi Visual Terbatas:** Dataset yang digunakan mungkin belum mencakup variasi yang cukup dalam hal pencahayaan, sudut pengambilan gambar, latar belakang, pose model, dan kondisi lingkungan yang berbeda-beda yang akan ditemui dalam aplikasi real-world.
- **Dependensi pada Kualitas Gambar:** Sistem ini mengasumsikan input berupa gambar dengan kualitas yang memadai dan objek pakaian yang terlihat jelas. Dalam kondisi nyata, gambar mungkin blur, terpotong, atau memiliki noise yang dapat mempengaruhi performa klasifikasi.
- **Keterbatasan Fitur Manual:** Pendekatan ekstraksi fitur manual yang digunakan mungkin tidak dapat menangkap kompleksitas visual yang lebih tinggi dibandingkan dengan pendekatan deep learning yang dapat mempelajari representasi fitur secara otomatis.

#### 4.3 POTENSI PERBAIKAN

Perbesar dataset menjadi minimal 1000 gambar per kelas serta tambahkan variasi pencahayaan, latar belakang, dan orientasi. Implementasikan data augmentation (rotasi, flipping, scaling). Pilih dataset publik seperti Fashion-MNIST atau DeepFashion. Eksplorasi fitur tekstur lanjutan (GLCM, Wavelet) dan kernel SVM lain (RBF, polynomial). Lebih dalam lagi bisa mengembangkan interface user-friendly yang terdapat implementasi real-time classification terintegrasi dengan mobile application. Dengan saran perbaikan ini, penelitian selanjutnya dapat mencapai performa yang lebih tinggi dan aplikabilitas yang lebih luas dalam klasifikasi pakaian otomatis.

## **BAB V**

### **KESIMPULAN**

Penelitian ini telah mengembangkan sistem klasifikasi otomatis untuk pakaian menggunakan algoritma Support Vector Machine (SVM) dengan berbagai kombinasi teknik ekstraksi fitur. Dari empat kombinasi fitur yang dievaluasi, diperoleh Performa kombinasi HOG + Color Histogram dan HOG + LBP + Color Histogram mencapai akurasi tertinggi sebesar 83.3%. Hasil ini menunjukkan bahwa fitur HOG memberikan kontribusi signifikan dalam pengenalan bentuk dan struktur pakaian, sementara Color Histogram efektif dalam membedakan kategori berdasarkan karakteristik warna. Sebaliknya Kombinasi Color Histogram + LBP menghasilkan akurasi terendah (50.0%), mengindikasikan bahwa tanpa informasi bentuk dari HOG, klasifikasi menjadi sangat sulit dan tidak akurat.

Kelas "Celana\_Panjang" menunjukkan performa terbaik dengan recall sempurna (1.00) pada tiga dari empat kombinasi fitur, menunjukkan bahwa celana panjang memiliki karakteristik visual yang paling distinktif dan mudah dikenali oleh sistem.

Disarankan perbesar dataset menjadi minimal 1000 gambar per kelas serta tambahkan variasi pencahayaan, latar belakang, dan orientasi. Implementasikan data augmentation (rotasi, flipping, scaling). Pilih dataset publik seperti Fashion-MNIST atau DeepFashion. Eksplorasi fitur tekstur lanjutan (GLCM, Wavelet) dan kernel SVM lain (RBF, polynomial).

## DAFTAR PUSTAKA

- [1] Pakaian. (2024, Juli 30). Di Wikipedia, Ensiklopedia Bebas. Diakses pada 08:04, Juli 30, 2024, dari <https://id.wikipedia.org/w/index.php?title=Pakaian&oldid=26100937>
- [2] Nadhiroh, A. Y. (2024). Sistem Klasifikasi Jenis Kain Berdasarkan Tekstur Menggunakan Metode Support Vector Machine Berbasis Web Flask. *Jurnal Ilmiah Informatika dan Komputer*, 1(1), 56-60.
- [3] Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003). Support vector machine. *Proceeding Indones. Sci. Meeting Cent. Japan*.
- [4] Rachmat, N., Yohannes, Y., & Mahendra, A. (2021). Klasifikasi jenis ikan laut menggunakan metode SVM dengan fitur HOG dan HSV. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(4), 2235-2247.
- [5] Irma, I., Muchtar, M., Adawiyah, R., & Sarimuddin, S. (2024). Klasifikasi Tingkat Kematangan Cabai Merah Keriting Menggunakan Svm Multiclass Berdasarkan Ekstraksi Fitur Warna. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(3).
- [6] Sari, C. A., & Rachmawanto, E. H. (2021). Fitur Esktraksi LBP dan Naive Bayes dalam Klasifikasi Jenis Pepaya Berdasarkan Citra Daun. *Jurnal Masyarakat Informatika*, 12(2), 102-113.