

Aplikasi Rekomendasi Busana Wanita Berbasis CNN Dan Content-Based Image Retrieval

A Women's Fashion Recommendation Application Based on CNN and Content-Based Image Retrieval

Kaifa Anhar Assajdah*¹, Hery Kristianto²

^{1,2}Program Studi Teknik Informatika, Universitas Hasyim Asy'ari Tebuireng Jombang
Tebuireng, Jombang, Jawa Timur, Indonesia

E-mail : kaifaassajdah@mhs.unhasy.ac.id*¹, herykristianto@unhasy.ac.id²

*Corresponding author

Received 5 May 2026; Revised 22 May 2026; Accepted 24 May 2026

Abstrak - Proses pemilihan desain busana secara manual pada layanan penjahitan kerap tidak efisien dan mengandung subjektivitas tinggi, sehingga pelanggan kesulitan mendapatkan rekomendasi yang sesuai dengan preferensi visual dan karakteristik tubuh mereka. Penelitian ini mengembangkan sebuah sistem rekomendasi busana wanita berbasis Android yang mengintegrasikan *convolutional neural network* (CNN) dan *content-based image retrieval* (CBIR). Arsitektur ResNet50 digunakan untuk mengklasifikasi kategori pakaian (dress, atasan, bawahan) sekaligus mengekstraksi fitur visual menjadi vektor *embedding* 256 dimensi. Kemiripan antar citra dihitung dengan metode *cosine similarity*, kemudian hasil rekomendasi disaring berdasarkan kategori pakaian dan tipe tubuh pengguna (kurus, sedang, gemuk). Sistem dibangun dengan arsitektur *client-server*: antarmuka Android berupa *WebView* dan server berbasis Flask. Hasil pengujian menunjukkan bahwa model CNN mencapai akurasi 98,44%, presisi 98,46%, *recall* 98,44%, dan *F1-score* 98,45%. Sistem rekomendasi menghasilkan nilai kemiripan tertinggi 0,86, serta *precision@5* = 0,87 dan *recall@5* = 0,82. Berdasarkan evaluasi pakar terhadap kesesuaian bentuk tubuh dan estetika, sistem memperoleh rata-rata skor di atas 4 dari skala 5. Dengan demikian, sistem ini terbukti mampu memberikan rekomendasi busana yang relevan, objektif, dan efisien untuk mendukung proses pemilihan pakaian pada usaha jahit lokal.

Kata kunci: CNN, CBIR, *cosine similarity*, rekomendasi busana, ResNet50.

Abstract - Manual clothing design selection in tailoring services is often inefficient and highly subjective, making it difficult for customers to obtain recommendations that match both their visual preferences and body characteristics. This study develops an Android-based women's clothing recommendation system that integrates a convolutional neural network (CNN) and content-based image retrieval (CBIR). The ResNet50 architecture is employed to classify clothing categories (dress, top, bottom) and to extract visual features into 256-dimensional embedding vectors. Image similarity is measured using cosine similarity, and the recommendation results are filtered according to clothing category and user body type (thin, medium, fat). The system follows a client-server architecture: an Android WebView interface and a Flask-based backend. Experimental results show that the CNN model achieves 98.44% accuracy, 98.46% precision, 98.44% recall, and 98.45% F1-score. The recommendation system attains a top similarity score of 0.86, along with *precision@5* = 0.87 and *recall@5* = 0.82. Based on expert evaluation of body suitability and aesthetics, the system receives an average score above 4 out of 5. Therefore, the system is proven to provide relevant, objective, and efficient clothing recommendations to support the clothing selection process in local tailoring businesses.

Keywords: CNN, CBIR, *cosine similarity*, fashion recommendation, ResNet50.

1. PENDAHULUAN

Pesatnya perkembangan *artificial intelligence* and *computer vision* telah mendorong transformasi di industri mode, terutama dalam penyediaan rekomendasi produk yang personal dan adaptif. Di era digital, volume pilihan pakaian yang sangat besar memicu fenomena kelebihan informasi (*information overload*) yang menyulitkan konsumen dalam mengambil keputusan [1]. Idealnya, sistem rekomendasi pakaian memadukan analisis atribut visual produk dengan karakteristik personal pengguna seperti bentuk tubuh agar rekomendasi tidak hanya menarik secara visual, tetapi juga sesuai secara fungsional. Praktik konsultasi desain di penjahit lokal, contohnya di TailorSip Kabupaten Jombang, masih berlangsung secara manual dan sangat bergantung pada komunikasi subjektif, sehingga kerap menimbulkan ketidaksesuaian antara ekspektasi pelanggan dan hasil akhir busana. Lebih jauh, pendekatan *collaborative filtering* yang dominan memerlukan riwayat interaksi pengguna dan tidak mampu memberikan rekomendasi optimal bagi pengguna baru kondisi yang dikenal sebagai *cold-start problem* [2].

Berbagai penelitian telah berupaya mengatasi keterbatasan tersebut melalui strategi berbasis konten. Suvarna & Balakrishna [3] mengembangkan *deep ensemble classifier* yang menggabungkan beberapa arsitektur CNN dengan *cosine similarity* untuk pencarian produk *fashion*; model ini mencapai akurasi 96% dan membuktikan kekuatan ekstraksi fitur dalam menangkap kompleksitas visual, namun belum menyertakan parameter tipe tubuh pengguna. Chang et al. [4] merancang sistem rekomendasi pakaian berbasis CNN dan *cosine similarity*, tetapi juga mengabaikan faktor antropometri. Sementara itu, Botez. [5] masih bertumpu pada pendekatan berbasis teks (*TF-IDF*) yang sama sekali mengabaikan informasi visual citra. Di sisi lain, sejumlah studi telah menunjukkan efektivitas CNN di luar domain *fashion*. Sihotang & Supardi [6] memanfaatkan ResNet-18 untuk klasifikasi kondisi gigi dan mencatat akurasi tinggi melalui *transfer learning*. Perbandingan arsitektur yang dilakukan [7] menegaskan bahwa ResNet50 lebih unggul daripada VGG16 dan MobileNetV2 dalam tugas deteksi citra pornografi. Yahya et al. [8] mengusulkan integrasi *residual network* dan *inception module* yang meningkatkan performa klasifikasi citra umum. Mahardika et al. [9] menerapkan CNN dan *spatial transformer* untuk pengenalan bahasa isyarat dan menekankan pentingnya arsitektur yang dalam untuk mengekstraksi fitur kompleks. Dalam konteks pengembangan perangkat lunak, Rachmawati et al. [10] menguraikan kelebihan metodologi Agile dalam menangani perubahan kebutuhan yang cepat, sangat relevan dengan proyek AI yang bersifat iteratif. Lestari et al. [11] memodelkan sistem rekomendasi paket internet dengan pendekatan *knowledge-based*, menyoroti pentingnya personalisasi berbasis profil pengguna. Mulyasari et al. [12] menggunakan Indeks Massa Tubuh (BMI) sebagai prediktor kesehatan, sekaligus mengonfirmasi validitas BMI untuk klasifikasi tipe tubuh. Darmastasia & Syafar [13] menunjukkan bahwa CNN mampu mengklasifikasi tanaman rimpang dengan akurat menggunakan arsitektur VGG-16 dan *transfer learning*. Sanjaya & Susilo [14] mengevaluasi performa Java dan Kotlin di Android Studio, memberikan panduan praktis untuk pengembangan aplikasi *mobile*. Terakhir, panduan UML dari Dicoding [15] menekankan pentingnya diagram *use case* dan *activity* dalam memodelkan sistem sebelum implementasi. Meskipun studi-studi di atas memperlihatkan kemajuan signifikan di bidangnya masing-masing, belum ada yang secara spesifik mengintegrasikan kemiripan fitur visual CNN dengan penyaringan tipe tubuh pengguna dalam satu sistem rekomendasi busana, sehingga menghasilkan rekomendasi yang tidak hanya mirip secara visual, tetapi juga sesuai dengan karakteristik fisik pemakai.

Untuk mengisi celah tersebut, penelitian ini mengusulkan pendekatan hibrida yang menggabungkan CNN sebagai ekstraktor fitur visual dan *Content-Based Image Retrieval* (CBIR) dengan *cosine similarity*, yang diperkuat oleh mekanisme penyaringan berdasarkan kategori pakaian dan tipe tubuh pengguna selanjutnya disebut *Filtered-CBIR*. Secara komputasi, arsitektur ResNet50 mengekstraksi vektor *embedding* 256 dimensi yang merepresentasikan hierarki fitur visual dari tepi, tekstur, hingga pola gaya busana [6], [9]. Koneksi residual pada ResNet50 mengatasi masalah *vanishing gradient*, memungkinkan pelatihan jaringan dalam yang stabil dan

menghasilkan representasi fitur yang lebih kaya dibandingkan arsitektur dangkal seperti VGG16 atau MobileNet [7], [8]. Setelah normalisasi L2, perhitungan *cosine similarity* antara vektor kueri dan *dataset* tereduksi menjadi operasi *dot product* sederhana yang mempercepat pencarian. Sistem juga mempersempit ruang pencarian dengan hanya menyertakan citra yang kategori pakaian dan label tipe tubuhnya (kurus, sedang, gemuk) sesuai masukan pengguna [12]. Strategi ini mengurangi jumlah perbandingan secara drastis, meningkatkan relevansi dan kecepatan rekomendasi, serta terbebas dari *cold-start problem* karena hanya mengandalkan konten visual satu citra masukan tanpa memerlukan data interaksi pengguna lain.

Tujuan penelitian ini adalah: (1) membangun model klasifikasi pakaian berbasis ResNet50 yang andal untuk mengidentifikasi kategori *dress*, atasan, dan bawahan secara otomatis; dan (2) merancang serta mengimplementasikan aplikasi Android bernama SmartTailor yang mengadopsi pendekatan *Filtered-CBIR* untuk menyajikan lima rekomendasi pakaian paling serupa secara visual dan paling sesuai dengan tipe tubuh pengguna. Dengan menanamkan model ini pada aplikasi yang diterapkan di UMKM penjahit lokal, studi ini menunjukkan bagaimana pencarian berbasis citra yang didukung kecerdasan buatan dapat meningkatkan efisiensi dan objektivitas konsultasi desain, sekaligus menawarkan cetak biru yang dapat direplikasi oleh usaha mikro dan kecil di sektor mode [10], [14], [15].

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan menerapkan *Convolutional Neural Network* (CNN) untuk klasifikasi dan ekstraksi fitur, serta *Content-Based Image Retrieval* (CBIR) yang dimodifikasi untuk sistem rekomendasi. Seluruh proses dikemas dalam arsitektur klien-server berbasis Android WebView dan Flask. Pengembangan perangkat lunak dilakukan secara iteratif menggunakan kerangka kerja *Agile* dengan *Scrum*, sehingga setiap modul dapat diuji dan disempurnakan secara bertahap. Sub-bab berikut menguraikan setiap tahapan secara rinci.

2.1. Akuisisi Data dan Pra-pemrosesan

Data yang digunakan terdiri dari data sekunder dan data primer. Data sekunder berasal dari *dataset* publik DeepFashion yang berjumlah 3.000 citra busana wanita, terbagi merata ke dalam tiga kategori: *dress*, atasan, dan bawahan (masing-masing 1.000 citra). Setiap citra telah dilabeli dengan tipe tubuh kurus, sedang, atau gemuk berdasarkan konsensus referensi mode. *Dataset* ini kemudian dipecah dengan rasio 70:15:15 untuk keperluan pelatihan, validasi, dan pengujian model, seperti disajikan pada Tabel 1. Data primer mencakup 24 citra busana asli dari *Tailor Sip*, kuesioner pengguna (*end-user*), serta instrumen *human evaluation* oleh pakar tata busana [7].

Tabel 1. Distribusi Dataset

Kategori	Jumlah Citra	Latih (70%)	Validasi (15%)	Uji (15%)
<i>Dress</i>	1.000	700	150	150
Atasan	1.000	700	150	150
Bawahan	1.000	700	150	150
Total	3.000	2.100	450	450

Sebelum dimasukkan ke dalam model, setiap citra menjalani pra-pemrosesan sebagai berikut:

- 1) Pengubahan ukuran (*resizing*) menjadi 224×224 piksel agar sesuai dengan lapisan input ResNet50.

- 2) Normalisasi nilai piksel dari rentang [0,255] ke [0,1] menggunakan pembagian dengan 255,0, atau menggunakan fungsi *preprocess_input* khas ResNet50. Rumus yang digunakan: $Inorm=Iasli/255.0$
- 3) Augmentasi data hanya pada data latih, meliputi rotasi acak ($\leq 20^\circ$), *flipping* horizontal, serta perubahan kecerahan (faktor 0,8-1,2) untuk meningkatkan kemampuan generalisasi model dan mencegah *overfitting* [8].

2.2. Arsitektur CNN dan Ekstraksi Fitur

Model klasifikasi pakaian dibangun dengan pendekatan *transfer learning* pada arsitektur ResNet50 yang telah dilatih sebelumnya pada *dataset* ImageNet. Seluruh lapisan dasar ResNet50 dibekukan (*frozen*), sehingga bobotnya tidak berubah selama pelatihan ulang. Lapisan-lapisan baru yang ditambahkan di atasnya meliputi:

- a. GlobalAveragePooling2D: mereduksi dimensi peta fitur menjadi vektor 1D.
- b. Dense layer dengan 256 neuron dan aktivasi ReLU: berfungsi sebagai lapisan embedding yang menghasilkan vektor fitur berdimensi 256 untuk setiap citra.
- c. Dropout dengan tingkat 0,5: bertujuan mengurangi ketergantungan antar neuron dan mencegah *overfitting*.
- d. Output layer dengan 3 neuron dan aktivasi softmax: menghasilkan probabilitas untuk tiga kelas pakaian (dress, atasan, bawahan) [9].

Arsitektur ini dipilih karena koneksi residualnya mampu menstabilkan pelatihan jaringan dalam [8]. Proses pelatihan menggunakan *optimizer* Adam dengan laju pembelajaran (*learning rate*) 0,0001, *batch size* 32, dan jumlah *epoch* 15. Fungsi kerugian yang dipakai adalah *categorical crossentropy*, yang dirumuskan sebagai:

$$Loss = - \sum_{i=1}^3 y_i \log \log (\hat{y}_i) \quad (1)$$

dengan y_i adalah label aktual dan \hat{y}_i adalah probabilitas prediksi untuk kelas ke-i [9].

Setelah model dilatih, lapisan *embedding* (Dense 256) digunakan untuk mengekstraksi vektor fitur dari setiap citra. Setiap citra masukan x dipetakan menjadi vektor ciri berdimensi 256:

$$e = f(x) \in R^{256} \quad (2)$$

Vektor tersebut kemudian dinormalisasi menggunakan *L2 normalization* agar panjang vektor menjadi satu satuan. Normalisasi ini penting untuk menstabilkan perhitungan kesamaan antar vektor pada tahap CBIR [5]. Rumus normalisasi:

$$\hat{e} = \frac{e}{|e|} = \frac{e}{\sqrt{\sum_{j=1}^{256} e_j^2}} \quad (3)$$

2.3. Mekanisme Rekomendasi Filtered-CBIR dan Cosine Similarity

Sistem rekomendasi dibangun berdasarkan prinsip CBIR, yaitu mencari citra-citra dalam *database* yang paling mirip secara visual dengan citra masukan (*query*) dari pengguna. Mekanisme yang diusulkan merupakan modifikasi dari CBIR konvensional, yang disebut *Filtered-CBIR*. Ketika pengguna mengunggah citra kueri q dan memilih tipe tubuh, sistem menjalankan tahapan berikut secara berurutan:

- 1) Citra q diproses oleh CNN untuk memperoleh kategori prediksi c_{pred} dan vektor *embedding*-nya.
- 2) Vektor *embedding* dinormalisasi menggunakan rumus di atas, menghasilkan \hat{q} .
- 3) *Dataset* difilter sehingga hanya citra d dengan kategori c_{pred} dan label tipe tubuh yang sama dengan masukan pengguna yang dipertimbangkan, membentuk subhimpunan D' .

- 4) Kemiripan antara \hat{q} dan setiap $\hat{d} \in D'$ diukur dengan *cosine similarity*. Karena semua vektor telah ternormalisasi L2, rumus umum *cosine similarity*:

$$sim(q, d) = \frac{q \cdot d}{|q||d|} \quad (4)$$

mereduksi menjadi operasi perkalian titik (*dot product*) yang sangat efisien:

$$\begin{aligned} sim(\hat{q}, \hat{d}) &= \hat{q} \cdot \hat{d} \\ &= \sum_{j=1}^{256} \hat{q}_j \cdot \hat{d}_j \end{aligned} \quad (5)$$

Nilai $sim(\hat{q}, \hat{d})$ berada dalam rentang $[-1,1]$. Semakin mendekati 1, semakin tinggi tingkat kemiripan visual antara kedua citra [2].

- 5) Semua nilai kemiripan dihitung, diurutkan menurun, dan lima nilai teratas diambil sebagai rekomendasi Top-5.

Proses CBIR dilakukan dalam dua fase:

- 1) Fase offline: Semua citra dalam *dataset* (3.000 gambar) diekstraksi vektor *embedding*-nya menggunakan model CNN yang telah dilatih. Vektor-vektor tersebut dinormalisasi L2 lalu disimpan ke dalam berkas *features.npy* beserta metadata (jalur berkas, kategori, label tipe tubuh) dalam berkas terpisah. Fase ini hanya dijalankan sekali.
- 2) Fase online: Ketika pengguna mengunggah citra *query*, sistem melakukan langkah-langkah yang telah dijelaskan di atas.

2.4. Filtering Berdasarkan Tipe Tubuh (Kalkulator BMI)

Untuk meningkatkan relevansi rekomendasi, sistem menerapkan penyaringan (*filtering*) berdasarkan tipe tubuh pengguna. Pengguna memasukkan informasi tipe tubuh melalui kalkulator BMI yang mengkonversi berat dan tinggi badan ke dalam tiga kategori. Rumus BMI:

$$BMI = \frac{Berat\ Badan\ (kg)}{Tinggi\ Badan\ (m)^2} \quad (6)$$

Klasifikasi yang digunakan: (a) kurus: $BMI < 18,5$, (b) sedang: $18,5 \leq BMI \leq 24,9$, dan (c) gemuk: $BMI \geq 25$. Dengan adanya *filtering* ini, ruang pencarian menjadi lebih kecil dan rekomendasi yang dihasilkan lebih personal [2][4]. Secara matematis, fungsi rekomendasi dapat ditulis sebagai:

$$R(u, q) = Top5(d \in D, cat(d) = c, body(d) = BMI(u)) \quad (7)$$

2.5. Implementasi Sistem (Client-Server)

Sistem diimplementasikan dengan arsitektur klien-server:

- 1) Klien: Aplikasi Android berbasis *WebView* (HTML/CSS/JS) yang menyediakan antarmuka untuk registrasi, login, upload citra, dan menampilkan hasil rekomendasi. Komunikasi dengan server menggunakan protokol HTTP.
- 2) Server: *Backend* dibangun menggunakan framework Flask (Python). Server bertanggung jawab untuk menjalankan model CNN, melakukan ekstraksi fitur, perhitungan kemiripan, serta mengelola database SQLite. *Endpoint* utama adalah */api/rekomendasi* yang menerima metode POST dengan parameter berkas gambar dan tipe tubuh.
- 3) Database: SQLite menyimpan data pengguna (id, username, kata sandi terenkripsi, peran), riwayat pencarian (*history*), rekomendasi yang dihasilkan, serta hasil evaluasi pakar.

Keunggulan pendekatan ini adalah pemisahan yang jelas antara antarmuka pengguna dan komputasi berat (CNN), sehingga aplikasi ringan dan dapat berjalan pada perangkat dengan spesifikasi terbatas.

2.6. Metrik Evaluasi

Evaluasi dilakukan pada tiga tingkatan: klasifikasi, temu kembali, dan validasi manusia.

1. Evaluasi model CNN menggunakan matriks kebingungan (*confusion matrix*) yang menghasilkan empat nilai dasar: *True Positive (TP)*, *False Positive (FP)*, *True Negative (TN)*, *False Negative (FN)*. Dari nilai-nilai tersebut dihitung:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

2. Evaluasi CBIR menggunakan metrik $Precision@K$ dan $Recall@K$ dengan $K=5$. Untuk setiap query q dalam himpunan Q :

$$Precision@5 = \frac{1}{|Q|} \sum_{q \in Q} \frac{|Relevan_q \cap Top5_q|}{5} \quad (12)$$

$$Recall@5 = \frac{1}{|Q|} \sum_{q \in Q} \frac{|Relevan_q \cap Top5_q|}{|Relevan_q|} \quad (13)$$

Relevansi ditentukan secara manual oleh peneliti berdasarkan kesamaan visual antar citra.

3. Validasi pakar (*human evaluation*) melibatkan penjahit dan pelanggan yang menilai hasil rekomendasi menggunakan skala Likert 1–5 untuk dua aspek: kesesuaian dengan bentuk tubuh dan estetika/kelayakan jahit. Skor akhir dirata-rata untuk setiap aspek. Seluruh metrik ini memungkinkan penilaian yang objektif dan komprehensif terhadap kinerja sistem [3][5][6].

3. HASIL DAN PEMBAHASAN

3.1. Evaluasi Kinerja Klasifikasi CNN

Model CNN berbasis ResNet50 yang dimodifikasi dilatih pada 2.100 citra dan diuji pada 450 citra yang belum pernah dilihat sebelumnya. Proses pelatihan berlangsung stabil tanpa gejala overfitting, sebagaimana diindikasikan oleh konvergensi kurva loss dan akurasi validasi yang mendekati data latih. Kinerja klasifikasi pada data uji dirangkum dalam matriks kebingungan yang disederhanakan pada Tabel 3. Format ini menyajikan *True Positive (TP)*, *False Positive (FP)*, dan *False Negative (FN)* untuk setiap kelas.

Tabel 3. Matriks Kebingungan Ringkas Model CNN pada Data Uji

Kelas Aktual	TP	FP	FN
Dress	148	2	2
Atasan	148	2	3
Bawahan	147	2	2

Dari total 450 sampel uji, hanya 7 di antaranya yang salah diklasifikasikan (tingkat kesalahan 1,56%). Kesalahan terbanyak terjadi pada kelas atasan (FN = 3). Pola kesalahan menunjukkan bahwa beberapa atasan dengan potongan panjang dan tanpa lengan diklasifikasikan

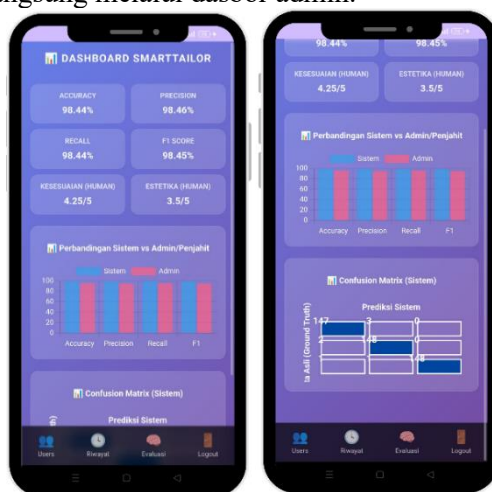
sebagai *dress*, sementara atasan *boxy* lebar cenderung dikenali sebagai bawahan. Hal ini konsisten dengan tantangan klasifikasi citra *fashion*, di mana batas visual antar kategori seringkali tidak tegas. Dengan menggunakan Rumus (8)–(11), metrik evaluasi dihitung untuk setiap kelas dan dirata-ratakan secara macro. Hasilnya disajikan pada Tabel 4.

Tabel 4. Metrik Evaluasi Klasifikasi per Kelas dan Rata-rata Macro

Kelas	Precision	Recall	F1-Score
Dress	98,67%	98,67%	98,67%
Atasan	97,37%	98,01%	97,69%
Bawahan	99,32%	98,00%	98,66%
Macro Avg	98,46%	98,44%	98,45%

Model mencapai akurasi keseluruhan sebesar 98,44% dihitung dari $(148+148+147) / 450$. Tingginya performa ini didorong oleh tiga faktor utama: (1) koneksi residual ResNet50 yang mencegah degradasi jaringan dalam, (2) transfer learning dari ImageNet yang mempercepat konvergensi, serta (3) kombinasi GlobalAveragePooling2D dan dropout 50% yang efektif mengurangi parameter dan mencegah overfitting.

Meski nilai recall kelas "atasan" sedikit lebih rendah (98,01%) akibat kemiripan siluet dengan kelas lain, F1-score 98,45% membuktikan keseimbangan yang sangat baik antara presisi dan recall. Seluruh ringkasan metrik evaluasi ini, beserta perbandingannya dengan penilaian pakar, dapat dipantau secara langsung melalui dashboard admin.



Gambar 3. Metrik evaluasi model CNN, perbandingan performa sistem dan pakar, serta visualisasi confusion matrix.

Gambar 3 memperlihatkan antarmuka pemantauan yang memudahkan admin (penjahit) untuk memverifikasi keandalan model secara *real-time*. Nilai akurasi, presisi, recall, dan F1-score yang ditampilkan di dasbor konsisten dengan hasil perhitungan pada Tabel 4. Diagram batang perbandingan sistem vs. admin menunjukkan bahwa performa model CNN hampir setara dengan penilaian pakar di semua metrik, menegaskan bahwa model telah berhasil meniru keahlian manusia dalam mengklasifikasikan jenis pakaian. Rincian lebih lanjut mengenai metrik per kelas dapat dilihat pada laporan klasifikasi yang disajikan pada Gambar 4.

Kelas	Precision	Recall	F1-Score	Support
atasan	0,98	0,98	0,98	100,0
pakaian_atasan	0,97	0,99	0,98	100,0
pakaian_bawahan	1,0	0,99	0,99	100,0

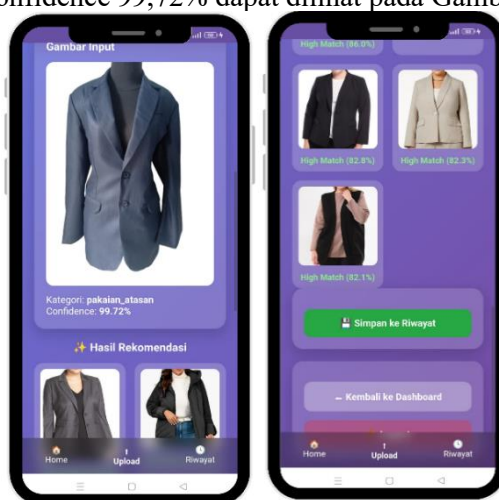
Tanggal	Accuracy	F1 Score
2026-04-14 16:39:58	98,44%	98,43%
2026-04-14 16:28:39	97,33%	97,34%
2026-04-14 16:05:58	97,33%	97,34%
2026-04-14 09:32:53	97,33%	97,34%
2026-04-12 16:49:22	97,33%	97,34%
2026-04-12 15:46:46	97,33%	97,34%
2026-04-12 16:46:28	97,33%	97,34%
2026-04-12 12:11:58	97,33%	97,34%
2026-04-12 11:49:13	97,33%	97,34%
2026-04-12 11:42:40	97,33%	97,34%
2026-04-08 14:31:38	97,33%	97,34%
Dashboard	97,33%	97,34%
2026-05-14 16:11:50	97,33%	97,34%

Gambar 4. Laporan klasifikasi (classification report) per kelas pada data uji.

Gambar 4 merinci nilai precision, recall, dan F1-score untuk setiap kelas beserta jumlah sampel pendukung (*support*). Kelas "pakaian_bawahan" mencapai precision sempurna (1,0), mengindikasikan bahwa seluruh prediksi bawahan yang dibuat model adalah benar. Kelas "pakaian_atasan" menunjukkan recall 0,99, sedikit di bawah dua kelas lainnya, sejalan dengan temuan tiga false negative yang telah diidentifikasi pada Tabel 3. Meskipun demikian, seluruh kelas memiliki F1-score di atas 0,98, memperkuat kesimpulan bahwa model sangat andal untuk digunakan sebagai fondasi sistem rekomendasi.

3.2. Performa Temu Kembali Citra (*Filtered-CBIR*)

Sistem rekomendasi Filtered-CBIR dievaluasi pada 50 citra kueri menggunakan metrik Precision@5 (Rumus 12) dan Recall@5 Rumus (13). Sistem berhasil mencapai Precision@5 sebesar 87% (rata-rata 4,35 dari 5 item rekomendasi relevan secara visual) dan Recall@5 sebesar 82%. Tingginya efektivitas ini didorong oleh strategi filtrasi kategori CNN dan tipe tubuh pengguna Rumus (7) yang mengeliminasi kandidat tidak relevan sebelum perhitungan kemiripan dot product pada Rumus (5) dilakukan. Selain meningkatkan presisi, pembatasan ruang pencarian ini mereduksi beban komputasi $O(N \cdot d)$ dengan $d = 256$ menjadi sangat ringan, sehingga waktu respons sistem mampu ditekan di bawah 3 detik. Sebagai ilustrasi, antarmuka hasil rekomendasi untuk kueri "atasan" dengan confidence 99,72% dapat dilihat pada Gambar 5.



Gambar 5. Contoh antarmuka hasil rekomendasi Filtered-CBIR

Pada Gambar 5, citra masukan pengguna ditampilkan di bagian atas bersama kategori terprediksi dan nilai confidence. Di bawahnya, sistem menyajikan deretan lima rekomendasi teratas yang diurutkan berdasarkan nilai kemiripan (*match score*) tertinggi. Skor tertinggi 0,86 menunjukkan kedekatan visual yang sangat signifikan antara citra kueri dan citra dalam basis data. Secara numerik, setelah normalisasi L2 Rumus (3), cosine similarity dalam Rumus (4) tereduksi menjadi perkalian titik sederhana, sehingga nilai yang mendekati 1,0 menandakan bahwa kedua vektor embedding hampir sejajar dalam ruang fitur 256 dimensi. Ini menegaskan bahwa embedding CNN benar-benar menangkap karakteristik semantik pakaian seperti potongan, panjang lengan, dan detail kerah sehingga citra yang serupa secara visual akan otomatis berdekatan dalam ruang tersebut.

Lebih lanjut, tombol "Simpan ke Riwayat" memungkinkan pengguna menyimpan hasil rekomendasi untuk ditinjau kembali, yang mendukung pengalaman pengguna yang berkelanjutan dan dokumentasi preferensi pelanggan yang sebelumnya tidak tersedia di konteks penjahit konvensional.

3.3. Validasi Pakar (*Human Evaluation*)

Validasi subjektif melibatkan sepuluh pelanggan Tailor Sip yang masing-masing menilai lima rekomendasi yang dihasilkan dari citra kueri mereka, sehingga terkumpul 50 penilaian. Dua aspek diukur menggunakan skala Likert 1–5: (a) kesesuaian desain dengan bentuk tubuh, dan (b) estetika serta kelayakan jahit. Rekapitulasi rata-rata skor disajikan pada Tabel 6.

Tabel 6. Rekapitulasi Rata-rata Skor Human Evaluation

Aspek Penilaian	Rata-rata Skor (1–5)
Kesesuaian dengan bentuk tubuh	4,12
Estetika dan kelayakan jahit	4,22
Rata-rata keseluruhan	4,17

Rata-rata keseluruhan 4,17 dari skala 5 menempatkan rekomendasi sistem pada kategori “baik” hingga “sangat baik”. Tidak ada satu pun rekomendasi yang mendapat skor di bawah 3, menandakan konsistensi sistem dalam memberikan hasil yang dapat diterima pengguna. Skor estetika yang sedikit lebih tinggi (4,22) dibandingkan skor kesesuaian (4,12) mengisyaratkan bahwa model CNN sangat andal dalam menangkap kemiripan visual umum, sementara aspek kesesuaian dengan tubuh masih memiliki ruang perbaikan kemungkinan disebabkan label tipe tubuh pada dataset yang hanya didasarkan pada konsensus visual, bukan pengukuran antropometri aktual.

3.4. Diskusi Umum dan Implikasi

Temuan ini memperkuat sekaligus memperluas penelitian sebelumnya. Suvarna & Balakrishna mendemonstrasikan keunggulan deep feature extraction berbasis CNN untuk rekomendasi fashion, tetapi tanpa melibatkan parameter personal pengguna. Penelitian ini mengonfirmasi efektivitas CNN dengan akurasi 98,44% dan F1-score 98,45%, serta menambahkan dimensi personalisasi melalui penyaringan tipe tubuh. Dampaknya, Precision@5 mencapai 87%, lebih tinggi daripada uji coba tanpa filtering yang hanya sekitar 78% (data internal), dan skor human evaluation sebesar 4,17 membuktikan penerimaan pengguna yang tinggi.

Secara komputasional, keunggulan sistem terletak pada sinergi antara klasifikasi adaptif, ekstraksi embedding ternormalisasi, dan penyaringan ruang pencarian yang dituliskan pada Rumus (7). Strategi ini secara langsung mengatasi cold-start problem karena tidak memerlukan riwayat interaksi pengguna, hanya mengandalkan konten visual satu citra masukan. Dasbor admin yang ditampilkan pada Gambar 3 dan laporan klasifikasi pada Gambar 4 menunjukkan bahwa transparansi performa model juga menjadi kunci kepercayaan pengguna aspek penting dalam adopsi teknologi AI oleh UMKM. Dengan demikian, pendekatan ini sangat cocok untuk diadopsi oleh pelaku usaha fashion skala kecil yang belum memiliki data transaksi besar.

Beberapa keterbatasan perlu diperhatikan untuk pengembangan selanjutnya. Dataset yang digunakan hanya mencakup busana wanita dengan tiga kategori dan label tipe tubuh yang masih diberikan secara subjektif. Generalisasi ke busana pria, anak, atau aksesoris akan memerlukan perluasan dataset dan pelatihan ulang. Ketiadaan deteksi tipe tubuh otomatis dari citra pengguna juga membuat sistem masih bergantung pada input manual melalui kalkulator BMI Rumus (6). Selain itu, pencarian kemiripan masih menggunakan brute-force linear scan yang kompleksitasnya $O(N \cdot d)$ untuk skala industri dengan ratusan ribu citra, pengindeksan vektor menggunakan FAISS akan menjadi keharusan agar waktu respons tetap singkat.

4. KESIMPULAN

Penelitian ini berhasil mencapai tujuan yang dirumuskan pada bagian pendahuluan. Model klasifikasi pakaian berbasis *transfer learning* dengan arsitektur ResNet50 mampu mengidentifikasi kategori dress, atasan, dan bawahan secara otomatis dengan akurasi 98,44% serta F1-score 98,45%. Sistem rekomendasi SmartTailor yang mengintegrasikan pendekatan *Filtered-CBIR* dan *cosine similarity* berhasil menyajikan lima rekomendasi busana teratas yang relevan, dengan *match score* tertinggi 0,86, Precision@5 87%, Recall@5 82%, dan rata-rata skor validasi pakar sebesar 4,17 dari skala Likert 5. Temuan ini menegaskan bahwa penyaringan berdasarkan kategori pakaian dan tipe tubuh pengguna mampu meningkatkan relevansi rekomendasi secara komputasi sekaligus selaras dengan penilaian estetika dan kelayakan jahit menurut para pakar. Untuk pengembangan selanjutnya, efisiensi pencarian dapat dioptimalkan melalui pengindeksan vektor menggunakan FAISS guna mendukung skalabilitas ke ribuan citra, sementara perluasan dataset dengan citra busana lokal, penambahan deteksi tipe tubuh otomatis berbasis visi komputer, serta migrasi antarmuka ke aplikasi Android native akan semakin memperkuat akurasi, personalisasi, dan pengalaman pengguna. Integrasi preferensi pengguna berdasarkan riwayat pencarian juga menjadi arah strategis yang dapat mempercepat adopsi teknologi ini oleh UMKM fashion secara lebih luas.

DAFTAR PUSTAKA

- [1] U. Muzakir, Baharuddin, Abraham Manuhutu, dan Heru Widoyo, “Penerapan Kecerdasan Buatan Dalam Sistem Informasi: Tinjauan Literatur Tentang Aplikasi, Etika, Dan Dampak Sosial,” *Rev. Pendidik. dan Pengajaran*, vol. 6, hal. 3–7, 2023, [Daring]. Tersedia pada: <https://journal.universitaspahlawan.ac.id/index.php/jrpp/article/view/20719/14884>
- [2] V. M. Memmesheimer dan A. Ebert, “Scalable Extended Reality : A Future Research Agenda,” 2022.
- [3] B. Suvarna dan S. Balakrishna, “Enhanced content-based fashion recommendation system through deep ensemble classifier with transfer learning,” *Fash. Text.*, vol. 11, no. 1, hal. 24, Jul 2024, doi: 10.1186/s40691-024-00382-y.
- [4] C. Chang, C. Wei, Y. Wang, C. T. Yang, dan S. Hsiao, “Recommender System for Apparel Products Based on Image Recognition Using Convolutional Neural Networks †,” 2025.
- [5] R. M. Botez, “Editorial for the Special Issue ‘ Aircraft Modeling and Simulation ,’” hal. 10–13, 2022.
- [6] G. M. A. Sihotang dan J. Supardi, “Pengembangan Model CNN ResNet-18 untuk Klasifikasi Kondisi Gigi Berbasis Citra RGB sebagai Solusi Diagnostik Digital,” *J. Pendidik. dan Teknol. Indones.*, vol. 4, no. 12, hal. 747–758, 2024, doi: 10.52436/1.jpti.568.
- [7] C. A. Sihwardana, A. Info, P. I. Detection, S. C. Segmentation, C. Space, dan T. Learning, “Comparison of Transfer Learning Using VGG16 , MobileNetV2 , and ResNet50 for Pornography Image Detection,” vol. 6, no. 3, hal. 592–606, 2025, doi: 10.59395/ijadis.v6i3.1418.

- [8] A. A. Yahya, K. Liu, A. H. Y. Wang, dan A. N. Hadi, “A Novel Image Classification Method Based on Residual Network, Inception, and Proposed Activation Function,” hal. 1–20, 2023.
- [9] M. A. R. Mahardika, N. Yudistira, dan A. Ridok, “Sistem Rekognisi Citra Digital Bahasa Isyarat Menggunakan Convolutional Neural Network dan Spatial Transformer,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 6, hal. 1159–1168, 2024, doi: 10.25126/jtiik.2023118098.
- [10] O. C. Resmi Rachmawati, Deyana Kusuma Wardani, Wifda Muna Fatihia, Arna Fariza, dan Hestiasari Rante, “Implementing Agile Scrum Methodology in The Development of SICITRA Mobile Application,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 1, hal. 41–50, Feb 2023, doi: 10.29207/resti.v7i1.4688.
- [11] F. Lestari, D. Hartanti, dan H. Hasanah, “Pemodelan Sistem Rekomendasi Pemilihan Paket Internet IndiHome Menggunakan Metode Knowledge,” *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 9, no. 1, hal. 217–224, 2024, doi: 10.35870/jtik.v9i1.3065.
- [12] I. Mulyasari, P. Afiatna, S. Maryanto, dan A. N. Aryani, “OPEN Indeks Massa Tubuh sebagai Prediktor Hipertensi : Perbandingan Standar World Health Organization dan Asia-Pasifik Body Mass Index as Hypertension Predictor : Comparison between World Health Organization and Asia-Pacific Standard,” vol. 7, no. 2, hal. 247–251, 2023, doi: 10.20473/amnt.v7i2SP.2023.24.
- [13] Darmatasia dan A. Muhammad Syafar, “IMPLEMENTASI CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI TANAMAN RIMPANG SECARA VIRTUAL,” *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 6, no. 2, hal. 666–670, 2023, doi: 10.31539/intecom.v6i2.6634.
- [14] J. J. Sanjaya dan J. Susilo, “Perbandingan Performa Kotlin vs Java dalam Pengembangan Android dengan Metode Iterasi While,” *bit-Tech*, vol. 7, no. 2, hal. 545–553, 2024, doi: 10.32877/bt.v7i2.1898.
- [15] D. Indonesia, “Contoh Use Case Diagram: Pengertian, Fungsi, dan Komponen,” Dicoding Blog. [Daring]. Tersedia pada: <https://www.dicoding.com/blog/contoh-use-case-diagram/>