

Deep Learning dengan ResNet50 untuk Sistem Rekomendasi Fashion Berbasis Citra

Dewi Rahmawati¹, Kanaya Salsabila Setiawan², Muhammad Fahreza Reynaldy³, M. Rangga Ramadhan Saelan^{4*}

¹²³⁴Sains Data, Universitas Nusa Mandiri

Jl. Margonda No.545, RT.1/RW.7, Pondok Cina, Kec. Beji, Kota Depok, Jawa Barat 16424, (021) 78884819

dewirahmawati9862@gmail.com¹, s.kanayasalsa@gmail.com², fahreza.reynaldy123@gmail.com³, rangga.mgg@nusamandiri.ac.id^{4*}

Abstrak. Perkembangan industri fashion yang pesat menuntut sistem rekomendasi yang tidak hanya akurat, tetapi juga mampu memahami preferensi visual pengguna. Sistem rekomendasi berbasis teks seringkali menghadapi keterbatasan dalam menangkap konteks visual yang kompleks, sehingga pendekatan berbasis citra menjadi solusi yang lebih relevan. Penelitian ini bertujuan untuk membangun sistem *Smart Recommendation Search Engine* berbasis visual dengan memanfaatkan model *Convolutional Neural Network* (CNN) ResNet50 sebagai *feature extractor* dan *dataset DeepFashion*. ResNet50 digunakan untuk mengekstraksi vektor fitur dari gambar produk fashion, yang kemudian dimanfaatkan dalam pencarian gambar serupa menggunakan algoritma *K-Nearest Neighbors* (KNN). Proses penelitian mencakup pra-pemrosesan data, ekstraksi fitur visual, pencarian kemiripan berbasis metrik kemiripan kosinus (*cosine similarity*), serta evaluasi sistem menggunakan metrik *precision* dan *recall* pada top-K results (hasil teratas). Hasil pengujian menunjukkan bahwa metrik *cosine similarity* memberikan performa terbaik dalam menemukan gambar dengan kemiripan visual tinggi, dengan nilai *precision* pada satu hasil teratas (*precision at top-1*) sebesar 0,230. Sistem yang dikembangkan berhasil mengidentifikasi produk fashion serupa secara visual dan mendukung pengalaman belanja yang lebih personal. Temuan ini menegaskan potensi pendekatan berbasis visual dalam meningkatkan akurasi sistem rekomendasi serta mendukung gaya hidup berkelanjutan.

Kata Kunci: *Computer_Vision, Deep_Learning, Fashion, ResNet50, Sistem_Rekomendasi.*

1 Pendahuluan

Seiring dengan tren globalisasi, digitalisasi, dan kebutuhan konsumen akan gaya hidup yang lebih dinamis, industri fashion terus berkembang [1], [2]. Dengan perkembangan ini, para pelaku industri fashion ingin menawarkan konsumen pengalaman berbelanja yang praktis dan personal [3]. Namun, konsumen seringkali menghadapi kesulitan dalam menentukan item fashion yang sesuai dengan preferensi dan kebutuhan mereka di tengah banyaknya merek dan produk yang tersedia. Hal inilah yang menjadi masalah dalam membuat keputusan pembelian yang tepat [4].

Dari perspektif industri, salah satu tantangan utama dalam industri fashion adalah pengelolaan data produk dalam jumlah besar secara efektif [5], [6]. Karena hal tersebut, konsumen sering kali kebingungan dalam memilih item yang paling sesuai seiring bertambahnya variasi produk. Dalam situasi seperti ini, sistem rekomendasi dianggap sebagai solusi teknologi yang dapat menjembatani preferensi pengguna dengan kumpulan produk yang tersedia, sehingga memberikan pengalaman berbelanja yang lebih relevan dan personal [7], [8]. Akan tetapi, sistem rekomendasi konvensional tak jarang menghadapi masalah seperti kelangkaan data, start awal yang dingin, dan kesesuaian visual produk.

Artificial Intelligence (AI), khususnya *machine learning* dan *deep learning*, telah membuka jalan bagi pengembangan sistem rekomendasi fashion yang lebih akurat [5]. Pendekatan ini mampu mempelajari pola preferensi pengguna secara lebih mendalam dan menyesuaikan rekomendasi sesuai dengan perilaku yang terus berubah [9]. Integrasi AI (seperti *machine learning* dan *deep learning*) dalam sistem rekomendasi memungkinkan pendekatan yang lebih kontekstual, termasuk dengan memanfaatkan data tidak terstruktur seperti citra, sehingga meningkatkan akurasi dan pengalaman pengguna [4].

Salah satu cabang AI yang penting dan berkembang dengan pesat dalam sistem rekomendasi fashion adalah *computer vision*. Sebuah teknologi yang memungkinkan computer memahami dan mengekstrak informasi dari sebuah citra [3]. Dengan memanfaatkan pemrosesan citra dan ekstraksi fitur mendalam, sistem dapat menggunakan elemen visual, seperti warna, tekstur, bentuk, dan gaya pakaian untuk dijadikan dasar dalam memberikan rekomendasi yang tepat [10]. Terutama dalam konteks belanja fashion, sistem berbasis visual seperti ini sangat relevan karena tampilan visual menjadi faktor dominan dalam pengambilan keputusan [9].

Seiring perkembangan kebutuhan akan personalisasi, terlahirlah konsep *AI-Styling Assistant*, yaitu sistem cerdas yang mampu memberikan saran berpakaian berbasis data visual kepada pengguna berdasarkan foto pakaian yang dimiliki, preferensi gaya, hingga kondisi cuaca. Namun, dalam penelitian ini sistem masih terbatas pada foto pakaian dan preferensi gaya saja [11]. Sistem ini tidak hanya bertujuan untuk memberikan rekomendasi, tetapi juga untuk mengoptimalkan koleksi pakaian yang sudah ada dengan tujuan mengurangi pembelian impulsif dan mendukung gerakan *circular fashion* demi gaya hidup berkelanjutan.

Beberapa penelitian sebelumnya telah mengkaji penerapan teknologi *computer vision* dan *machine learning* dalam sistem rekomendasi fashion berbasis citra dengan algoritma yang beragam pula. Penelitian mengembangkan *Adaptive Closet* dengan algoritma *Reinforcement Learning* [4], Kemudian, penelitian [9] menerapkan klasifikasi multilabel dengan *decision tree* dan penelitian [10] meneliti *collaborative filtering* berbasis GNN. Penelitian-penelitian sebelumnya berhasil menunjukkan potensi tinggi dalam menyajikan rekomendasi fashion yang personal. Namun, sebagian besar studi tersebut masih terbatas pada *filtering* berbasis data pengguna atau deskripsi produk, belum banyak yang mengeksplorasi pemrosesan citra visual secara mendalam sebagai fitur utama dalam rekomendasi.

Berdasarkan gap tersebut, penelitian ini hadir dengan tujuan untuk mengembangkan sistem *Smart Recommendation Search Engine* berbasis citra, yang dapat memberikan rekomendasi produk fashion berdasarkan kemiripan visual dari gambar yang diunggah oleh pengguna. Sistem ini memanfaatkan *dataset DeepFashion*, salah satu *dataset* citra *fashion* terbesar dan terlengkap yang mencakup ribuan kelas dan ratusan ribu gambar produk fashion dari berbagai kategori [12]. *Dataset* ini dipilih karena representatif terhadap variasi nyata dalam industri fashion dan sangat sesuai untuk kebutuhan *visual feature extraction*.

Permasalahan utama yang ingin diselesaikan dalam penelitian ini adalah ketidaktepatan sistem rekomendasi konvensional dalam memahami konteks visual, serta keterbatasan pemrosesan citra dalam menangkap kemiripan semantik antar produk fashion secara efektif. Dengan mengimplementasikan arsitektur ResNet50 sebagai *feature extractor* dan algoritma *K-Nearest Neighbors* (KNN) untuk pencarian kemiripan berbasis *embedding* citra, sistem ini dirancang untuk menghasilkan rekomendasi yang lebih kontekstual dan akurat secara visual [13].

Dengan pengembangan sistem ini, *Smart Recommendation Search Engine* diharapkan mampu memberikan rekomendasi produk fashion yang tidak hanya akurat dan personal, tetapi juga mendukung pengalaman belanja yang lebih visual, kontekstual, dan berkelanjutan berdasarkan kemiripan visual dari citra yang diunggah oleh pengguna bukan hanya berdasarkan teks atau metadata produk. Sistem cerdas ini diharapkan dapat menjadi solusi inovatif dalam membantu pengguna mencari produk fashion yang pelanggan sukai dengan lebih cepat dan efisien, sekaligus menjadi alternatif dari sistem berbasis teks yang sering kali kurang akurat dalam memberikan rekomendasi visual tentang apa yang mereka cari.

2 Tinjauan Pustaka

2.1 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan jenis jaringan saraf yang terdiri dari beberapa lapisan konvolusional, di mana struktur dan jumlah lapisannya disesuaikan dengan tujuan dan performa sistem rekomendasi yang ingin dicapai [3]. Fungsi utama CNN adalah mengubah representasi citra dari bentuk piksel menjadi *vector* fitur yang dapat dianalisis oleh sistem rekomendasi. Secara umum, proses konvolusi dapat dirumuskan secara matematis sebagai berikut:

$$y(i, j) = \sum_m \sum_n x(i + m, j + n) \cdot w(m, n) \quad (1)$$

di mana $x(i + m, j + n)$ adalah piksel input, $w(m, n)$ bobot kernel konvolusi, dan $y(i, j)$ hasil dari operasi konvolusi pada titik tertentu.

2.2 ResNet50

ResNet50 merupakan salah satu jenis CNN yang dikenal karena kedalaman dan keefektifannya dalam bidang *computer vision*. Arsitektur ini menggunakan residual connections untuk mengatasi masalah hilangnya gradien saat pelatihan, sehingga sangat unggul dalam mengolah dataset berskala besar seperti dataset *Deep Fashion* Dataset yang digunakan dalam penelitian ini [14]. Alasan penggunaan arsitektur ini adalah kompleksitas komputasionalnya tergolong *medium* dan kemampuan interpretasi nya *moderate*. Model matematis dari residual connection/block dapat dirumuskan sebagai berikut.

$$y = F(x, \{W_i\}) + x \quad (2)$$

di mana x adalah input dari residual block, $F(x, \{W_i\})$ fungsi non-linear yang merepresentasikan operasi konvolusi dan aktivasi, dan y sebagai output dari blok setelah penjumlahan residual. Dengan ini, ResNet50 mampu mempelajari fitur-fitur visual dari citra pakaian dengan lebih efisien tanpa kehilangan informasi penting yang sering kali terjadi pada jaringan konvolusional yang sangat dalam.

2.3 Evaluasi Model

Evaluasi sistem dilakukan dengan menghitung metrik Precision@K dan Recall@K pada nilai $K = 1, 5, 10$. Precision@K mengukur proporsi hasil pencarian teratas yang relevan, sedangkan Recall@K mengukur cakupan item relevan yang berhasil ditemukan [15]. Rumus masing-masing metrik adalah sebagai berikut:

$$Precision@K = \frac{|Rel_K|}{K} \quad (3)$$

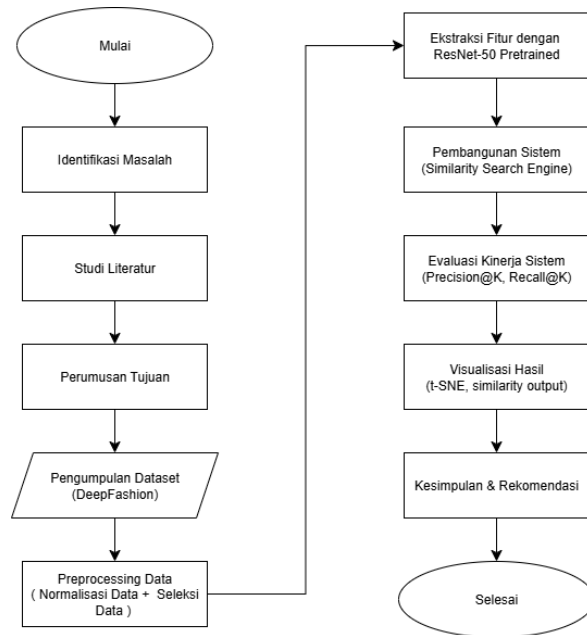
dengan Rel_K adalah jumlah item relevan di antara K item teratas, sedangkan K adalah jumlah item teratas yang dievaluasinya.

$$Recall@K = \frac{|Rel_K|}{|Rel_{total}|} \quad (4)$$

dengan Rel_{total} adalah jumlah total item relevan yang tersedia untuk user. Dari rumus-rumus tersebut diproses nilai rata-rata dari Precision dan Recall yang dihitung untuk 100 query acak guna mengevaluasi performa keseluruhan sistem.

2 Metode Penelitian

Secara umum, pendekatan berbasis AI telah menunjukkan kemampuannya dalam meningkatkan efisiensi dan akurasi dalam sistem rekomendasi [16]. Dalam bidang fashion, penggabungan AI khususnya deep learning dengan pemrosesan citra secara digital memungkinkan sistem untuk melakukan personalisasi berdasarkan preferensi visual pengguna secara otomatis [17]. Penelitian ini dilakukan melalui beberapa tahapan seperti yang terlihat pada Gambar 1.



Gambar. 1. Alur Penelitian

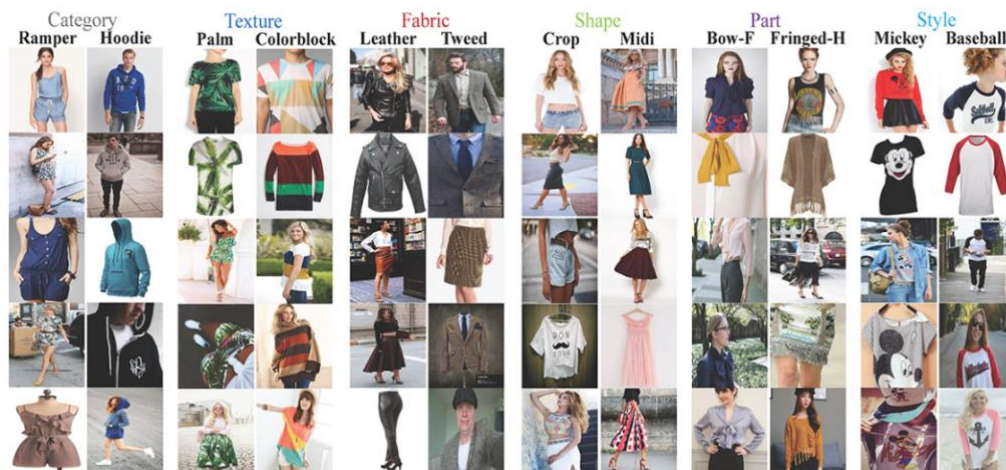
2.1 Identifikasi Masalah

Industri *fashion* yang dinamis dan semakin digital menuntut pengalaman berbelanja yang personal dan praktis. Namun, sering kali pelanggan merasa kesulitan memilih produk di tengah banyaknya pilihan, terlebih lagi dengan adanya tantangan berupa pengelolaan data produk dalam jumlah besar. Selain itu, sistem rekomendasi konvensional juga sering kali gagal karena kelangkaan data, *cold start*, dan ketidaksesuaian visual produk [16]. Untuk mengatasi keterbatasan pencarian berbasis teks dalam memahami nuansa visual, diperlukan pendekatan alternatif yang mampu mengenali karakteristik visual secara langsung dari gambar.

Sebagai solusi, penelitian ini mengembangkan sistem pencarian kemiripan visual berbasis deep learning dengan memanfaatkan ekstraksi fitur dari model ResNet50 terlatih (*pre-trained*) dengan index *Nearest Neighbors* yang efisien. Sistem ini mampu merepresentasikan gambar pakaian dalam bentuk vektor fitur berdimensi 2048, yang kemudian digunakan untuk mencari kemiripan menggunakan berbagai metrik (*Cosine*, *Euclidean*, dan *Manhattan*).

2.2 Pengumpulan Data

Penelitian ini menggunakan DeepFashion Dataset sebagai sumber utama, yang merupakan salah satu dataset paling luas dan komprehensif dalam analisis citra pada bidang fashion [12]. *Dataset DeepFashion* digunakan, berisi 5.621 kelas dan 289.229 gambar dengan distribusi tak seimbang, terdapat kelas dengan jumlah gambar yang sangat sedikit, sementara kelas lain memiliki representasi yang jauh lebih banyak. Ketidakseimbangan ini menjadi pertimbangan krusial dalam tahap *pre-processing* selanjutnya, guna memastikan data yang digunakan relevan dan representatif untuk analisis fitur. Beberapa contoh gambar beserta anotasinya ditunjukkan pada Gambar 2.



Gambar. 2. Contoh gambar dari berbagai kategori dan atribut dalam DeepFashion. Atribut membentuk lima kelompok: *texture*, *Fabric*, *Shape*, *Part*, dan *Style*.

2.3 Pra-pemrosesan Data

Tahapan pra-pemrosesan data dalam penelitian ini bertujuan untuk menyiapkan data citra agar memiliki struktur, ukuran, dan kualitas yang seragam sebelum dilakukan proses ekstraksi fitur visual menggunakan model CNN ResNet50. Proses yang dilakukan pada tahap ini meliputi beberapa langkah sebagai berikut:

- 1) **Seleksi Kelas**
Dataset DeepFashion memiliki 5.621 kelas dan total 289.229 gambar. Untuk menjaga efisiensi proses dan mengurangi ketimpangan distribusi kelas (*class imbalance*), dilakukan pemilihan hanya pada 100 kelas dengan jumlah gambar terbanyak. Setiap kelas yang dipilih harus memiliki minimal 75 gambar. Hasil dari proses ini menghasilkan 13.779 gambar yang mewakili 100 kelas terpilih.
- 2) **Resize Gambar**
Seluruh gambar diubah ukurannya menjadi 224×224 piksel menggunakan metode *bilinear interpolation*. Ukuran ini disesuaikan dengan standar input dari model ResNet50, agar kompatibel secara struktur dimensi.
- 3) **Transformasi Tensor (Konversi Format)**
Gambar diubah ke format tensor dan disusun dalam batch untuk efisiensi pemrosesan. Batch size yang digunakan pada tahap ini adalah 32 gambar per batch.

Tahapan pra-pemrosesan dilakukan melalui seleksi kelas teratas berdasarkan jumlah gambar, *resize* seluruh citra ke ukuran 224×224 piksel, serta normalisasi nilai piksel berdasarkan distribusi *ImageNet*. Proses ini memastikan data memiliki struktur yang konsisten, seimbang secara kelas, dan sesuai dengan format *input* model ResNet50.

2.4 Ekstraksi Fitur

Ekstraksi fitur dilakukan menggunakan model ResNet-50 pralatih (*pre-trained*) pada ImageNet. Lapisan akhir klasifikasi dihapus, sehingga keluaran berupa vektor fitur berdimensi 2048. Seluruh citra yang telah diproses dilewatkan melalui model ini untuk menghasilkan representasi fitur yang mendalam (*deep features*), yang kemudian digunakan untuk pencarian kemiripan visual.

2.5 Pemodelan Sistem Similarity Search

Setelah proses ekstraksi fitur selesai dilakukan, sistem pencarian kemiripan visual dibangun menggunakan representasi vektor fitur berdimensi 2048. Untuk mengukur tingkat kemiripan antar citra, sistem memanfaatkan algoritma *K-Nearest Neighbors (KNN)* dengan tiga pilihan metrik: *Cosine Similarity*, *Euclidean Distance*, dan *Manhattan Distance*. Ketiga metrik ini diuji untuk menentukan mana yang menghasilkan performa terbaik dalam mengidentifikasi gambar yang serupa secara visual. Tabel 1 menunjukkan contoh hasil dari pencarian kemiripan menggunakan metrik *Cosine Similarity* pada sebuah gambar kueri. Lima gambar dengan tingkat kemiripan tertinggi ditampilkan dalam urutan peringkat, lengkap dengan nilai jarak (*distance*) dan skor kemiripan (*similarity score*).

Tabel 1. Hasil Top-5 pencarian gambar serupa berdasarkan *Cosine Similarity*

Rank	Nama Produk	Jarak (Dist)	Skor Kemiripan
1	Dotted Chiffon Blouse	0.1481	0.8519
2	Sweet Crochet Blouse	0.1494	0.8506
3	Embroidered Peasant Blouse	0.1542	0.8458
4	Dotted Chiffon Blouse	0.1563	0.8437
5	Tartan Plaid Blouse	0.1622	0.8378

2.6 Sistem Evaluasi

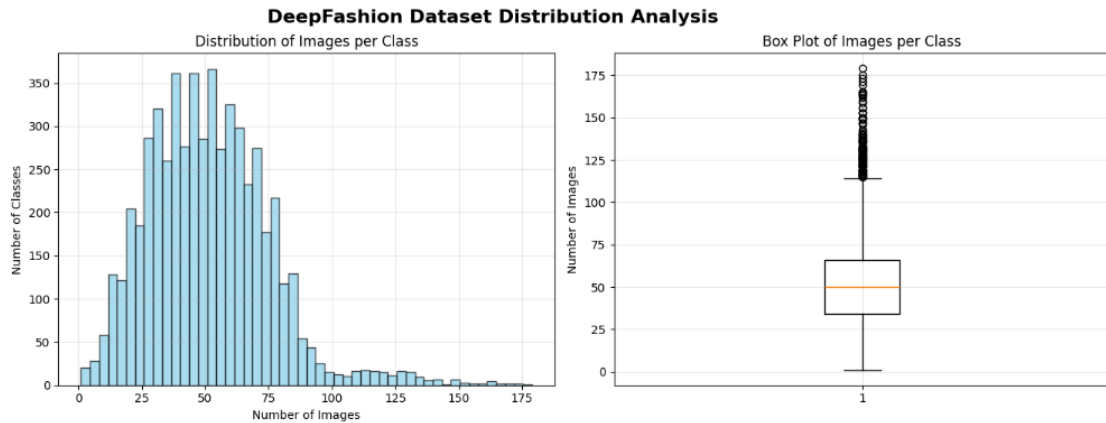
Evaluasi dilakukan dengan mengukur performa pencarian menggunakan metrik *Precision@K* dan *Recall@K* pada $K = 1, 5$, dan 10 . Sebanyak 100 citra dipilih secara acak sebagai *query*. Hasil evaluasi menunjukkan bahwa *Cosine Similarity* memberikan performa terbaik dibandingkan *Euclidean* dan *Manhattan*, dengan nilai *Precision@1* sebesar 0.230.

2.6 Implementasi Model

Setelah proses training dan evaluasi, model yang telah dibangun siap untuk diterapkan ke dalam sistem cerdas. Tahapan ini termasuk konversi model ke format deployment, antarmuka sederhana untuk pengujian model secara langsung, dan integrasi sistem ke dalam aplikasi rekomendasi fashion berbasis web yang siap guna.

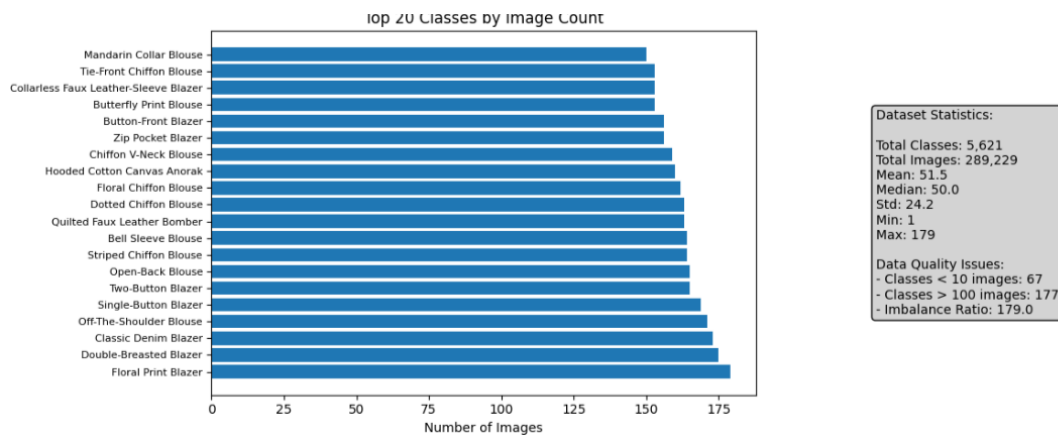
3 Hasil dan Pembahasan

Tahap awal dilakukan dengan menganalisis distribusi jumlah citra pada dataset *DeepFashion*. Gambar 3 menunjukkan distribusi jumlah gambar per kelas, yang memperlihatkan ketimpangan distribusi (*data imbalance*). Terdapat kelas dengan jumlah citra sangat sedikit, sementara beberapa kelas memiliki ratusan citra.



Gambar. 3. Distribusi Awal Dataset Deep Fashion

Setelah dilakukan seleksi dengan ambang minimal 75 citra per kelas, distribusi data menjadi lebih seimbang. Dari total 5.621 kelas dan 289.229 gambar, dataset difilter menjadi 100 kelas teratas, menghasilkan 13.779 gambar yang siap diproses lebih lanjut. Berikut top n-20 kelas gambar terbanyak yang terlihat pada Gambar 4.



Gambar. 4. Top n-20 classes

Setelah data diseleksi, dilakukan proses ekstraksi fitur menggunakan model ResNet50 yang telah dilatih sebelumnya pada dataset *ImageNet*. *Layer* klasifikasi terakhir dari model dihilangkan, sehingga fitur yang dihasilkan berupa representasi vektor berdimensi 2048 per citra. Proses ekstraksi dilakukan pada seluruh gambar hasil pra-pemrosesan, dengan konfigurasi sebagai berikut.

Tabel 2. Feature Extractor Initialized

Parameter	Nilai
Ukuran Citra Input	224x224 piksel
Ukuran Batch	32
Total Gambar	13.779
Jumlah kelas	100
Dimensi Fitur Output	13779,2048

Hasil dari proses ini menghasilkan matriks fitur (feature.npy) yang berisi representasi fitur dari setiap citra, kemudian label kelas (labels.npy), daftar path gambar masing-masing (image_paths.pkl) dengan nama kelasnya class_names.pkl. Yang akan digunakan dalam tahap pembangunan aplikasi sistem pencarian citra berbasis kemiripan visual.

Sistem pencarian dibangun menggunakan algoritma *K-Nearest Neighbors* (KNN) dengan tiga metrik, yaitu *cosine*, *euclidean*, dan *manhattan*. Model KNN memanfaatkan vektor fitur untuk mencari gambar serupa berdasarkan kedekatan di ruang fitur berdimensi 2048. Model KNN dibangun untuk setiap metrik dan disimpan sebagai model terpisah dalam bentuk file *.pkl*, sehingga sistem siap digunakan tanpa perlu pelatihan ulang. Proses ini menghasilkan sistem pencarian yang mampu mengembalikan daftar gambar paling mirip terhadap gambar kueri berdasarkan jarak terdekat dalam ruang fitur berdimensi tinggi.

Tabel 3. Hasil Top 5 Query-42 Berdasarkan *Cosine Similarity*

Rank	Nama Produk	Jarak (Dist)	Skor Kemiripan
1	Dotted Chiffon Blouse	0.1481	0.8519
2	Sweet Crochet Blouse	0.1494	0.8506
3	Embroidered Peasant Blouse	0.1542	0.8458
4	Dotted Chiffon Blouse	0.1563	0.8437
5	Tartan Plaid Blouse	0.1622	0.8378

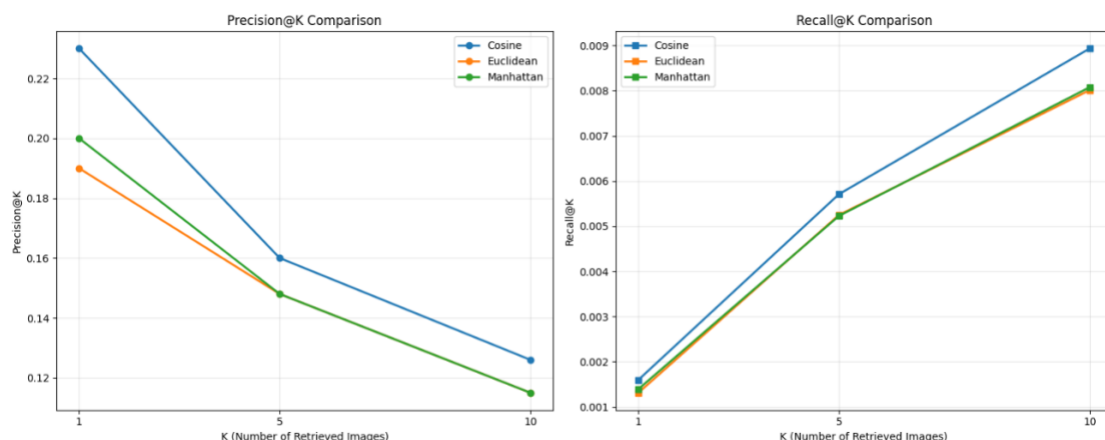
Kemudian sebagai contoh, dilakukan pencarian gambar dengan indeks ke-42 (*query_index*) menggunakan metrik *cosine* dan jumlah top-5 terdekat. Sistem berhasil mengembalikan lima gambar dengan kelas yang relevan, dengan nilai jarak dan kemiripan yang cukup tinggi. Hasil ranking yang ditunjukkan pada Tabel 2 menunjukkan item-item *fashion* dari kelas serupa, dengan skor kemiripan mendekati 1, menandakan tingkat kemiripan visual yang tinggi.

Pada evaluasi performa sistem pencarian visual, dilakukan pengujian terhadap 100 query acak menggunakan tiga jenis metrik kesamaan, yaitu *Cosine*, *Euclidean*, dan *Manhattan*. Evaluasi dilakukan menggunakan metrik *Precision@K* dan *Recall@K* pada nilai $K = \{1, 5, 10\}$.

Tabel 4. Hasil Evaluasi Precision dan Recall

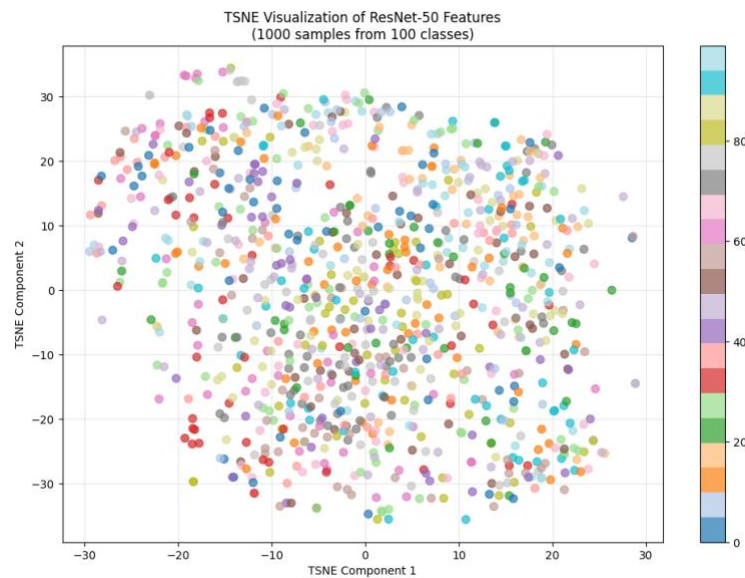
Matrik	Precision			Recall		
	@1	@5	@10	@1	@5	@10
Cosine	0.230	0.160	0.126	0.002	0.006	0.009
Euclidean	0.190	0.148	0.115	0.001	0.005	0.008
Manhattan	0.200	0.148	0.115	0.001	0.005	0.008

Berdasarkan tabel tersebut, metrik *cosine similarity* menunjukkan kinerja terbaik dibandingkan dua metrik lainnya, baik dalam hal precision maupun recall pada seluruh nilai K . Hal ini mengindikasikan bahwa representasi fitur dari model lebih sejajar secara arah dibandingkan secara jarak absolut. Gambar 5 menyajikan grafik *Precision@K* dan *Recall@K* untuk masing-masing metrik kemiripan.



Gambar. 5. Grafik Evaluasi Sistem

Untuk mengevaluasi representasi fitur antar kelas yang dihasilkan oleh model *backbone* (ResNet50), dilakukan visualisasi dua dimensi menggunakan algoritma t-SNE. Sebanyak 1.000 sampel acak dari 100 kelas dianalisis pada tahap ini. Hasil reduksi dimensi ditampilkan pada Gambar 1, yang memperlihatkan sebaran fitur cukup merata di ruang dua dimensi. Walau ditemukan beberapa pengelompokan lokal, secara umum distribusi masih saling tumpang tindih, mencerminkan adanya kedekatan visual antar kelas.



Gambar. 6. Visualisasi Algoritma t-SNE

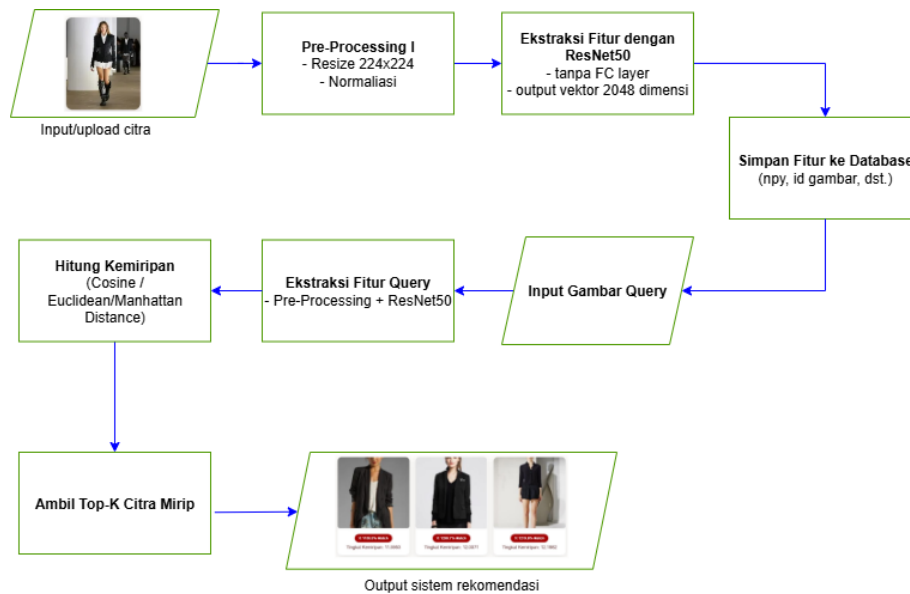
Sebagai pelengkap visualisasi distribusi fitur, dilakukan juga eksperimen pencarian gambar serupa menggunakan metrik *cosine similarity*. Untuk setiap kueri, sistem menghasilkan lima gambar teratas dengan tingkat kemiripan tertinggi. Gambar 6 dan 7 dibawah ini menampilkan contoh hasil pencarian untuk kueri 'Tweed Boyfriend Blazer' dan 'Sheer Pleated-Front Blouse'.



Gambar. 7. Query (a) Tweed Boyfriend Blazer, (b) Sheer Pleated-Front Blouse

Sebagaimana dapat dilihat, sistem mampu mengambil gambar-gambar yang memiliki kemiripan visual dan semantik tinggi dengan kueri, seperti pola kain yang serupa, potongan blazer yang mirip, serta warna dan gaya yang konsisten. Nilai cosine similarity antara kueri dan hasil teratas berkisar antara 0.87 hingga 0.95, yang menunjukkan performa yang baik dalam pencarian berbasis konten visual.

Penerapan model sistem rekomendasi *fashion* ini didasarkan pada pendekatan *transfer learning* menggunakan arsitektur CNN ResNet50. Pendekatan ini dipilih untuk memanfaatkan kemampuan ResNet50 yang telah terlatih pada dataset skala besar sehingga memungkinkan ekstraksi fitur visual yang mendalam dan robust dari citra *fashion*. Model ini dikembangkan menggunakan Google Colab sepenuhnya



Gambar. 8. Alur kerja sistem

Proses sistem diawali dengan input berupa gambar fashion, yang kemudian diproses melalui tahapan *pre-processing*, termasuk pengubahan ukuran dan normalisasi. Setelah itu, fitur visual diekstraksi menggunakan model ResNet-50 tanpa lapisan klasifikasi akhir, menghasilkan vektor berdimensi 2048. Vektor ini disimpan sebagai representasi citra untuk keperluan pencarian.

Saat pengguna mengunggah gambar kueri, sistem menjalankan proses serupa untuk mengekstrak fitur, lalu menghitung tingkat kemiripan terhadap seluruh data dalam basis menggunakan metrik seperti *Cosine*, *Euclidean*, atau *Manhattan*. Hasil pencarian disusun berdasarkan skor kemiripan, dan sistem menampilkan Top-K gambar yang paling relevan, lengkap dengan visualisasi hasil serta evaluasi performa pencarian

Proses pelatihan dan deployment dilakukan di platform Google Colab dan Visual Studio Code, yang mendukung efisiensi eksperimen dan komputasi. ResNet50 dipilih karena keunggulannya dalam mengekstrak hierarki fitur visual yang kompleks, seperti tekstur dan pola pakaian, sehingga cocok digunakan dalam sistem rekomendasi berbasis citra. Hasil embedding dari model ini digunakan untuk menghitung kemiripan antar gambar, memungkinkan sistem merekomendasikan produk fashion secara visual serupa dengan preferensi pengguna.

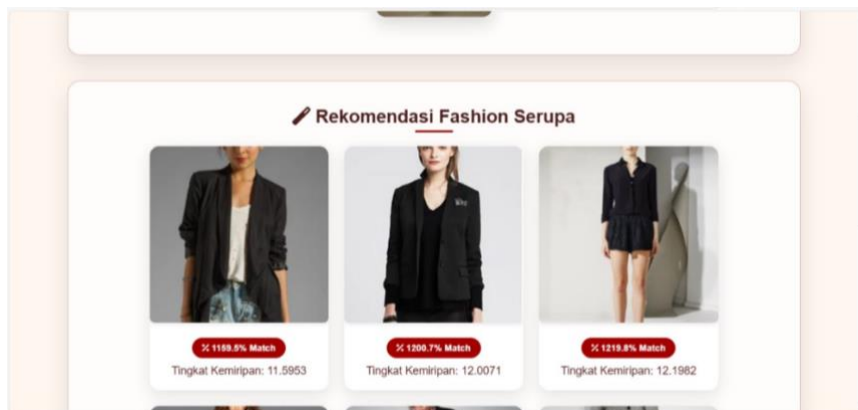
Sistem rekomendasi fashion ini dirancang untuk membantu pengguna menemukan item fashion yang serupa secara visual dengan preferensi mereka. Proses deployment model ini tercermin melalui antarmuka pengguna yang intuitif, yang memandu pengguna dari pemilihan gambar hingga menampilkan hasil rekomendasi.

Area besar dengan teks "Drag & Drop File" atau tombol "Browse file" jelas menunjukkan fungsi utama aplikasi: menerima gambar sebagai query awal. Ini adalah titik masuk bagi pengguna untuk berinteraksi dengan model ResNet-50 yang telah terlatih untuk mengekstrak fitur visual kompleks.



Gambar. 9. Input Citra yang Ingin Dicari Mirip

Pada Gambar 9 merupakan tampilan ketika pengguna sudah berhasil mengenakan sebuah jaket berwarna coklat) menjadi masukan bagi model ResNet-50. Setelah gambar ini diunggah, sistem secara otomatis akan memulai proses prapemrosesan (pengubahan ukuran dan normalisasi) dan ekstraksi fitur visual untuk menghasilkan vektor berdimensi 2048. Keberadaan tombol "Cari" atau "Search" di dekat gambar mengkonfirmasi bahwa setelah gambar kueri siap, pengguna dapat memicu proses pencarian kemiripan terhadap basis data embedding citra yang telah disimpan, memulai fase komputasi metrik kemiripan seperti *Cosine*, *Euclidean*, atau *Manhattan*.



Gambar. 10. Output Fashion yang Mirip dengan yang Diunggah

Dapat dilihat pada Gambar 10 yang menerangkan bahwa Top-K, adalah gambar yang paling relevan dan secara visual mirip dengan gambar kueri yang diunggah pengguna. Setiap gambar hasil rekomendasi disajikan dengan skor kemiripan, memungkinkan pengguna untuk mengevaluasi seberapa dekat relevansi visual antara kueri dan hasil yang direkomendasikan. Tampilan ini merupakan puncak dari alur kerja sistem, di mana embedding yang dihitung oleh ResNet-50 dan perhitungan kemiripan antar gambar diterjemahkan menjadi rekomendasi produk fashion yang nyata. Visualisasi hasil ini juga bisa mencakup evaluasi performa pencarian, memberikan transparansi dan validasi terhadap efektivitas sistem dalam memenuhi preferensi visual pengguna.

4 Kesimpulan

Permasalahan klasik seperti kelangkaan data, kesesuaian visual, dan keterbatasan sistem rekomendasi berbasis teks berhasil diatasi melalui pemanfaatan arsitektur CNN ResNet50 yang mampu mengekstraksi fitur visual mendalam dari citra. Penelitian ini juga menunjukkan bahwa pendekatan visual memiliki potensi kuat dalam meningkatkan pengalaman pengguna dan mendukung efisiensi pencarian produk fashion yang lebih personal dan kontekstual. Proses ekstraksi fitur dengan ResNet50 dan implementasi similarity search berbasis *K-Nearest Neighbors* (KNN) membuktikan efektivitas sistem dalam mengidentifikasi kemiripan visual antargambar. Hasil evaluasi menggunakan metrik *Precision* dan *Recall* pada *top-K Results* (hasil teratas) menunjukkan bahwa metrik *Cosine Similarity* memberikan performa terbaik dalam pencarian visual. Visualisasi t-SNE juga memperkuat temuan bahwa representasi fitur yang dihasilkan cukup baik dalam mengelompokkan kelas fashion. Secara keseluruhan, sistem ini tidak hanya memberikan rekomendasi yang akurat, tetapi juga berpotensi mendukung gerakan *circular fashion* dan pengurangan pembelian impulsif melalui optimalisasi koleksi pakaian yang sudah dimiliki pengguna.

Referensi

- [1] A. Etuk, J. A. Anyadighibe, E. E. James, and M. U. Ukpe, "Sociological factors and consumer buying behaviour towards fashion clothing," *International Journal of Applied Research in Social Sciences*, vol. 4, no. 2, pp. 21–34, 2022.
- [2] A. Balabanyts, I. Kyianytsia, T. Bielialov, S. Mosiiuk, O. Nedzvetska, and O. Verbytska, "The impact of tourism and digital media on the dynamics of the fashion industry in the global creative economy," *Multidisciplinary Reviews*, vol. 7, 2024.
- [3] S. Chakraborty, M. S. Hoque, N. Rahman Jeem, M. C. Biswas, D. Bardhan, and E. Lobaton, "Fashion recommendation systems, models and methods: A review," in *Informatics*, MDPI, 2021, p. 49.
- [4] O. K. Majeed, Z. ul Abideen, U. Arshad, R. H. Ali, A. Habib, and R. Mustafa, "Adaptive closet: Reinforcement learning in personalized clothing recommendations," in *2023 18th International Conference on Emerging Technologies (ICET)*, IEEE, 2023, pp. 305–309.
- [5] Q. Zhang, J. Lu, and Y. Jin, "Artificial intelligence in recommender systems," *Complex & intelligent systems*, vol. 7, no. 1, pp. 439–457, 2021.
- [6] M. Mohiuddin Babu, S. Akter, M. Rahman, M. M. Billah, and D. Hack-Polay, "The role of artificial intelligence in shaping the future of agile fashion industry," *Production Planning & Control*, vol. 35, no. 15, pp. 2084–2098, 2024.
- [7] S. Silvester and S. Kurian, "Recommendation systems: enhancing personalization and customer experience," in *2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, IEEE, 2023, pp. 1–6.
- [8] D.-N. Nguyen, V.-H. Nguyen, T. Trinh, T. Ho, and H.-S. Le, "A personalized product recommendation model in e-commerce based on retrieval strategy," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 10, no. 2, p. 100303, 2024, doi: <https://doi.org/10.1016/j.joitmc.2024.100303>.
- [9] B. Yang, "Clothing design style recommendation using decision tree algorithm combined with deep learning," *Comput Intell Neurosci*, vol. 2022, no. 1, p. 5745457, 2022.
- [10] C. Gao, X. Wang, X. He, and Y. Li, "Graph neural networks for recommender system," in *Proceedings of the fifteenth ACM international conference on web search and data mining*, 2022, pp. 1623–1625.
- [11] M. Princessilia, J. Wijaya, W. Fernando, Y. Indrayudha, C. Martin, and R. Y. Ningsih, "Perancangan Aplikasi Berbasis AI untuk Rekomendasi Pakaian Harian sebagai Solusi Pembelian Impulsif," *Jurnal Sosial Teknologi*, vol. 5, no. 1, pp. 112–122, 2025.
- [12] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "Deepfashion: Powering robust clothes recognition and retrieval with rich annotations," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1096–1104. Accessed: Jul. 28, 2025. [Online]. Available: <https://mmlab.ie.cuhk.edu.hk/projects/DeepFashion.html>
- [13] M. K. Syiam, A. T. Wibowo, and E. B. Setiawan, "Fashion Recommendation System using Collaborative Filtering," *Building of Informatics Technology and Science (BITS)*, vol. 5, no. 2, 2023.
- [14] C.-C. Chang, C.-H. Wei, Y.-H. Wang, C.-H. T. Yang, and S. Hsiao, "Recommender System for Apparel Products Based on Image Recognition Using Convolutional Neural Networks," *Engineering Proceedings*, vol. 89, no. 1, p. 38, 2025.
- [15] A. Çınar, M. Yıldırım, and Y. Eroğlu, "Classification of pneumonia cell images using improved ResNet50 model," *Traitement du Signal*, vol. 38, no. 1, pp. 165–173, 2021.

- [16] B. E. Jin and D. C. Shin, “The power of 4th industrial revolution in the fashion industry: what, why, and how has the industry changed?,” *Fashion and Textiles*, vol. 8, no. 1, p. 31, 2021.
- [17] Z. Lu, Y. Hu, Y. Chen, and B. Zeng, “Personalized outfit recommendation with learnable anchors,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 12722–12731.