Mata Kuliah – Machine Learning

**Nama Kelompok** : [dikosongi]

**Anggota** :

[202110370311443 - Nazilullaily Nur Aisyah]

[202110370311482 - Valencia Sefiana Putri]

Berikut ini merupakan update template laporan Mini Project kuliah Machine Learning.

**Nilai Total: 120 poin**

**Tahap 0 (poin: 25)**: Business Objective

Tujuan bisnis objective ini adalah untuk meningkatkan efisiensi dan otomatisasi proses klasifikasi jenis kain yang digunakan dalam industri fashion. Dengan menerapkan sistem klasifikasi berbasis citra untuk mempercepat proses, perusahaan pakaian dapat mengurangi human error, menghemat waktu dan biaya operasional, memfasilitasi analisis data untuk tren pasar dan mendukung digitalisasi di berbagai tahap rantai produksi dari penjualan. Untuk memastikan kinerja dan adaptabilitas model terhadap variasi bahan yang terus berkembang, pengembangan dataset dan pemilihan algoritma yang optimal.

**Tahap 1 (poin: 25)**: Original Data

* Urgensi topik/kasus yang dipilih.

Klasifikasi Model Busana Berdasarkan Jenis Bahan (Corduroy, Denim, Katun, Linen, Organza, dan Satin) Menggunakan Teknik Pengenalan Pola.

* Data yang digunakan.
  + Deskripsi singkat.

Dataset yang digunakan dikumpulkan melalui dokumentasi perusahaan internal (Shopee) dan Pinterest. Terdapat 1200 data yang dibagi menjadi enam target class yaitu corduroy, denim, katun, linen, organza, dan satin. Setiap kategori bahan memiliki variasi dalam pola, tekstur, dan warna untuk menangkap karakteristik visual yang berbeda.

* + Sebutkan dan jelaskan atribut pada data tersebut.

1. Tekstur (texture)

Tekstur menggambarkan pola permukaannya, seperti halus, kasar, berkerut, atau berbulu. Tekstur sangat penting untuk membedakan antara bahan seperti satin yang lebih halus dan berkilau, dan korduroi yang memiliki tekstur bergaris.

1. Pola (pattern)

Pola kain memiliki desain visual seperti garis, kotak, titik, atau bahkan polos yang dapat membantu menentukan jenis bahan atau model busana yang menggunakan bahan dengan pola ini.

1. Warna (color)

Setiap kain memiliki warna yang berbeda, yang membantu membedakan jenis bahan.

* + Jelaskan data mining task yang akan digunakan (*classification, clustering, regression, association rule mining, anomaly detection,* dsb.).

Data mining task yang akan digunakan adalah ***Classification****.*

Classsification merupakan tugas data mining yang bertujuan untuk memprediksi kelas atau mengkategorikan data berdasarkan fitur yang ada. Dalam kasus ini, mengklasifikasikan jenis bahan tekstil berdasarkan citra yang mencakup atribut seperti pola, tekstur, dan warna.

* Sumber data (teknik pengumpulan data melalui dokumentasi perusahaan internal (Shopee) dan Pinterest).

**Tahap 2 (poin: 10)**: Target Data (Optional)

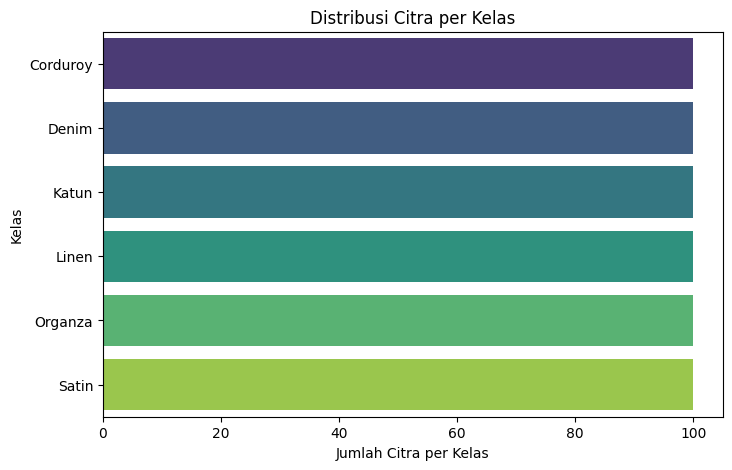
* Target class yang digunakan adalah corduroy, denim, katun, linen, organza, dan satin.

**Tahap 3-4 (poin: 25)**: Data Pre-processing & Transformation

Beberapa teknik yang bisa digunakan yaitu (tentu sesuai kondisi dan kebutuhan):

* *Exploratory Data Analysis* (EDA)
* Distribusi Citra per Kelas

Diagram batang pada gambar 1 menunjukkan jumlah citra untuk setiap kelas bahan kain, yaitu corduroy, denim, katun, linen, organza, dan satin. Setiap kelas memiliki jumlah citra yang sama yaitu 100 citra per kelas.



**Gambar 1.** Distribusi citra per kelas

* Citra per Kelas

Setiap kelas memiliki karakteristik yang dapat diidentifikasi sebagai berikut:

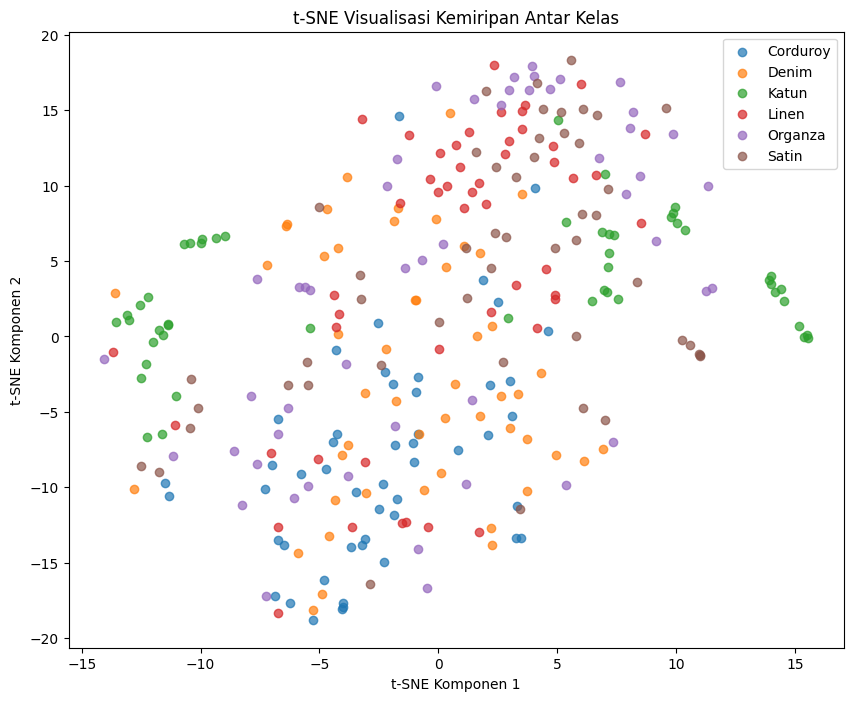
1. Corduroy: kain yang memiliki tekstur bergerigi atau bergaris yang terbentuk dari serangkaian benang yang ditenun secara vertikal.
2. Denim: kain twill yang memiliki ciri khas benang biru di bagian luar dan benang alami di bagian dalam, terbuat dari katun kuat dan diwarnai dengan nila.
3. Katun: kain yang memiliki tekstur sangat ringan karena terbuat dari serat kapas.
4. Linen: kain yang terbuat dari serat tanaman flaks, terlihat jelas serat-seratnya secara vertikal maupun horizontal.
5. Organza: kain yang berkilau, transparan, tipis, dan ringan.
6. Satin: kain yang memiliki tekstur halus, lembut, dan berkilau.



**Gambar 2.** Citra dari setiap kelas

* Visualisasi t-SNE (t-distributed Stochastic Neighbor Embedding)

Menunjukkan bahwa antara Linen dan Satin ada pemisahan yang jelas, sementara Corduroy, Denim, Katun, dan Organza memiliki tumpang tindih yang signifikan. Hasil ini menunjukkan bahwa karakteristik yang membedakan kelas-kelas tersebut harus dipertimbangkan dalam model klasifikasi yang dikembangkan, terutama dalam kasus daerah tumpang tindih.



**Gambar 3.** Visualisasi t-SNE Kemiripan Antar Kelas

* Dimensi Citra setiap Kelas

Setiap titik pada grafik mewakili satu gambar, dengan posisi titik ditentukan oleh dimensi lebar dan tinggi gambar tersebut. Visualisasi ini juga menjelaskan variasi dimensi dari gambar-gambar yang mewakili berbagai jenis kain. Pada setiap titik, warna yang berbeda menunjukkan kelas kain yang berbeda. Ini adalah lima resolusi yang paling umum:

* Resolusi: (800, 800), Jumlah: 78
* Resolusi: (736, 1104), Jumlah: 74
* Resolusi: (736, 736), Jumlah: 56
* Resolusi: (736, 981), Jumlah: 48
* Resolusi: (735, 981), Jumlah: 20



**Gambar 4.** Dimensi Gambar per Kelas

* data resize

Berdasarkan visualisasi yang ditampilkan pada gambar 3, hasil proses perubahan ukuran (resize) dari beberapa jenis pakaian yang awalnya memiliki resolusi berbeda menjadi ukuran tetap 300x300 piksel. Proses ini dilakukan untuk menstandarkan dimensi gambar pada dataset sehingga tampilan visual setiap pakaian konsisten.



**Gambar 5.** Hasil proses (NEAREST + Sharpening) setiap kelas

* data transformation
* Grayscale

Ciri grayscale, juga dikenal sebagai grayscale, adalah gambar yang menggunakan tingkatan warna abu-abu. Grayscale menunjukkan intensitas tiap piksel dengan nilai tunggal, yang berarti bahwa setiap piksel bagian merah, hijau, dan biru memiliki intensitas yang sama. Tanpa memproses informasi warna yang lebih kompleks, mengubah gambar menjadi grayscale akan mengurangi kompleksitas data visual karena hanya berfokus pada tingkat kecerahan setiap piksel. Detail kecil seperti kancing, lipatan, dan kerah masih dapat dilihat saat menggunakan format grayscale.



**Gambar 6.** Hasil Greyscale setiap kelas

* feature extraction
* *Histogram of Oriented Gradients* (HOG)

Menggambarkan karakteristik spasial dari intensitas piksel dalam citra, mengumpulkan data tentang pola, keteraturan, dan kekasaran dalam citra.



**Gambar 7.** Histogram of Oriented Gradients (HOG) setiap Kelas

* sobel edge detector (bentuk tepi)

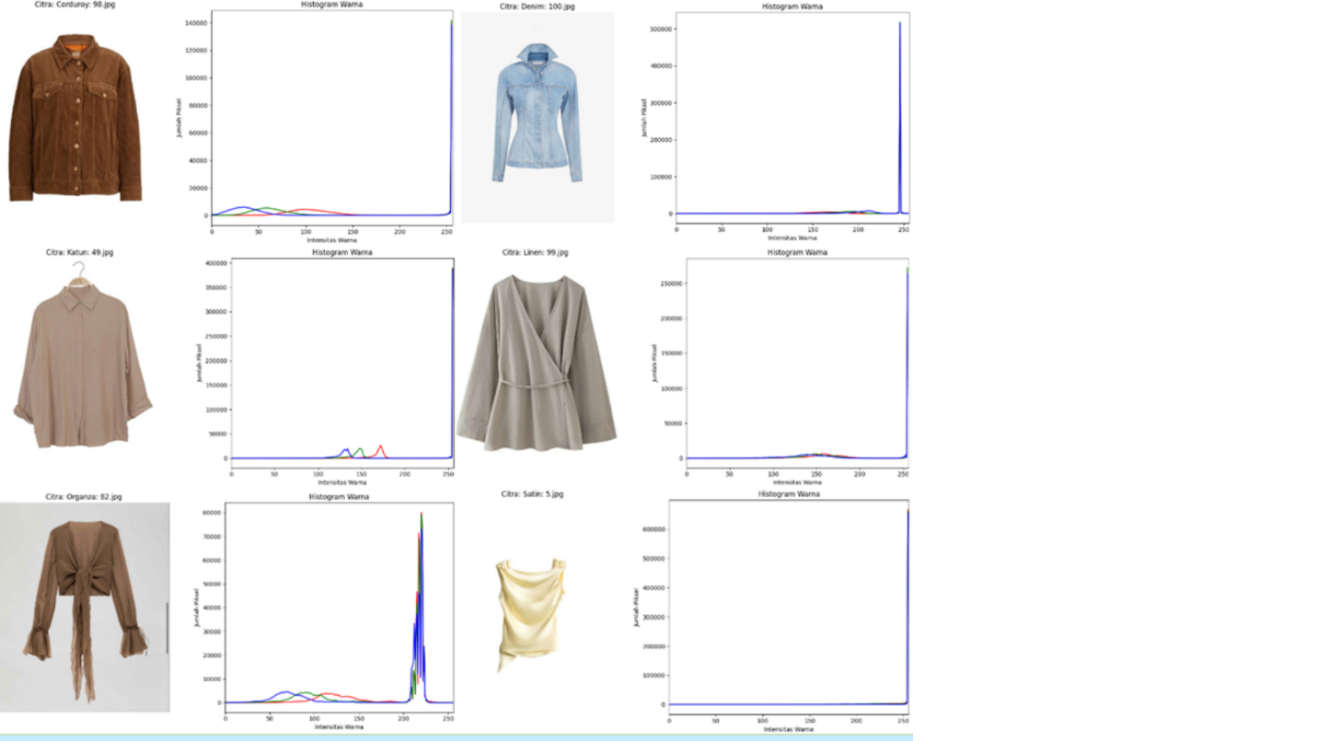
Operator gradien yang digunakan untuk mendeteksi tepi dalam citra, menghasilkan citra tepi yang menandai perubahan intensitas yang tajam. Hasil deteksi tepi dapat digunakan untuk mengekstrak fitur, segmentasi, dan pengenalan objek.



**Gambar 8.** Sobel Edge Detector (bentuk tepi) setiap Kelas

* histogram warna

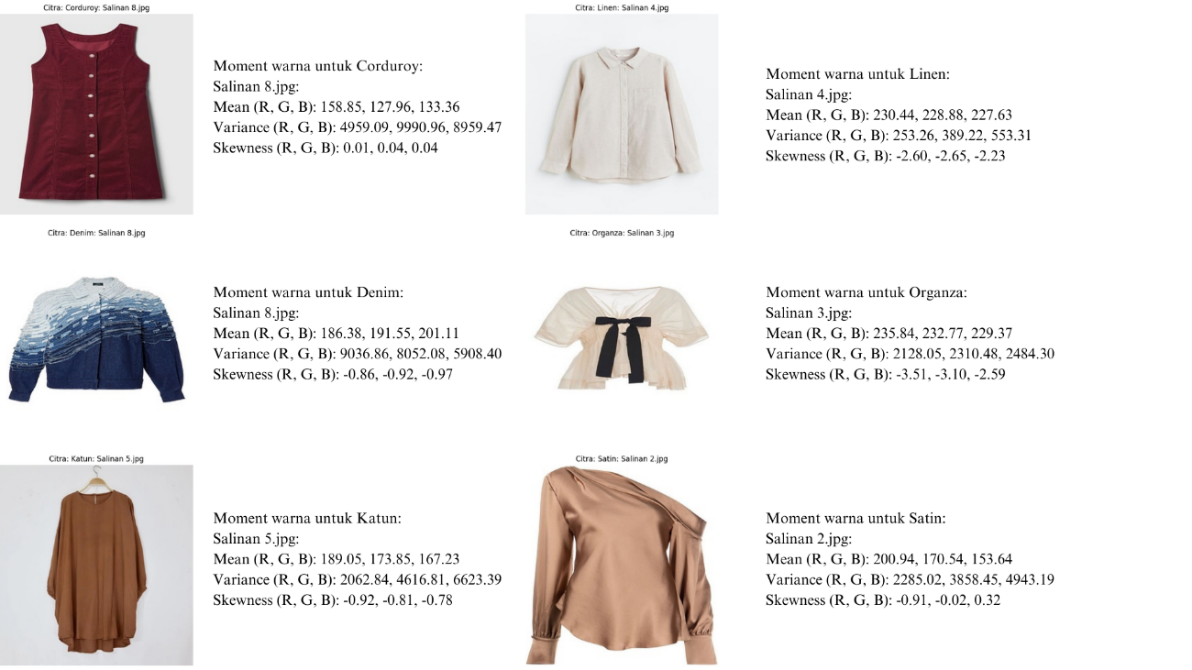
Representasi grafis distribusi frekuensi intensitas warna pada setiap saluran warna (merah, hijau, biru) dalam sebuah citra yang memberikan informasi tentang dominasi warna, kontras, dan kecerahan.



**Gambar 9.** Histogram Warna setiap kelas

* moment warna

Teknik ekstraksi fitur yang menghitung karakteristik statistik dari distribusi warna dalam gambar, seperti mean (rata-rata), varians, dan skewness (kemiringan distribusi warna).



**Gambar 10.** Moment Warna setiap Kelas

**Tahap 5 (poin: 25)**: Machine Learning

* Algoritma machine learning yang digunakan (sesuai data machine learning).

1. ***Support Vector Machine* (SVM)**

Algoritma klasifikasi ini dipilih karena efektif untuk membedakan berbagai jenis kain berdasarkan warna, tekstur, dan pola, terutama pada dataset berukuran kecil hingga sedang. Langkah-langkah Implementasi dengan SVM:

* Preprocessing Data
  + Resize

Ukuran gambar yang bervariasi dapat mempersulit algoritma untuk mengenali pola secara konsisten. Dengan melakukan resize atau mengubah ukuran semua gambar menjadi dimensi tetap (300x300 piksel), dapat dipastikan bahwa setiap gambar memiliki jumlah piksel yang sama, sehingga memudahkan algoritma untuk bekerja dengan data yang seragam.

* + Transformasi Greyscale

Mengubah gambar menjadi skala abu-abu (greyscale) dapat mengurangi kompleksitas data. Setiap piksel hanya memiliki nilai intensitas cahaya (0–255), bukan tiga komponen warna (RGB), sehingga dimensi fitur yang perlu diproses berkurang drastis. Dengan demikian, model hanya fokus pada pola tekstur dan bentuk tanpa terganggu oleh informasi warna, yang dalam beberapa kasus bisa lebih efektif untuk membedakan jenis bahan kain.

* Feature Extraction
  + Histogram of Oriented Gradients (HOG)

Teknik ekstraksi fitur yang sering digunakan dalam pemrosesan citra, terutama untuk deteksi objek dan klasifikasi tekstur. Metode ini sangat baik dalam mendeteksi tepi dan pola tekstur pada gambar dengan cara menganalisis arah gradien intensitas pada setiap bagian gambar.

* Sobel Edge Detector (bentuk tepi)

Deteksi tepi penting dalam pengenalan pola dan klasifikasi citra karena memberikan informasi tentang bentuk dan kontur objek dalam gambar, yang dapat membantu membedakan satu objek dari objek lainnya.

* Histogram Warna

Metode ekstraksi fitur yang menghitung distribusi warna dalam gambar. Metode ini menganalisis seberapa sering warna tertentu muncul pada gambar dan menyajikannya dalam bentuk histogram.

* Moment Warna

Memberikan informasi detail tentang distribusi warna, yang dapat membantu memahami pola warna secara mendalam, terutama pada kain dengan variasi warna yang rumit atau gradasi warna tertentu.

* Pembagian Dataset

Untuk pengujian akurasi dilakukan pembagian data train dan test menjadi 80% untuk latihan model (train) dan 20% untuk uji akurasi model (test).

* Evaluasi Model

1. Precision

Precision adalah tingkat ketepatan suatu informasi sistem dalam memprediksi target positif terhadap jumlah data yang diprediksi positif. Berikut adalah rumus untuk menghitung presisi:

1. Recall

Recall adalah tingkat ketepatan suatu informasi sistem dalam memprediksi target positif terhadap jumlah data yang positif. Berikut adalah rumus untuk menghitung recall:

1. F1-Score

F1-Score adalah salah satu perolehan dari konstanta dua dikali nilai recall dan presisi, dan dibagi jumlah keduanya. Rumus F1-Score yaitu sebagai berikut:

1. Accuracy

Accuracy adalah keberhasilan suatu algoritma dengan seberapa akurat system dapat mengklasifikasikan secara benar. Berikut adalah rumus untuk menghitung accuracy:

1. **Deep Learning (*Convolutional Neural Network (CNN)*)**

Deep learning menggunakan arsitektur CNN memungkinkan model belajar langsung dari data mentah untuk mengenali pola kompleks.

Langkah-langkah Implementasi dengan CNN:

* Preprocessing Data
  + Resize

Citra diubah ke ukuran 224x224 piksel agar sesuai dengan arsitektur CNN.

* + Rescaling

Normalisasi nilai piksel ke rentang [0,1] menggunakan Rescaling(1./255).

* Data Augmentation

Menambah variasi data pelatihan melalui rotasi, flipping, zooming, dan perubahan skala untuk menghindari overfitting.

* Arsitektur CNN
  + Convolutional Layers: Memanfaatkan filter progresif (32, 64, 128) untuk mendeteksi pola pada data gambar.
  + Pooling Layers: Mereduksi dimensi spasial untuk menghindari overfitting.
  + Flatten Layer: Mengubah keluaran fitur menjadi vektor untuk diproses pada lapisan fully connected.
  + Dense Layers: Menyertakan dropout (50%) untuk mengurangi overfitting
  + Output Layer: Fungsi aktivasi softmax untuk prediksi probabilitas kelas.
* Evaluasi Model

Menggunakan learning curve, classification report, dan confusion matrix.

1. **Transfer Learning (Pre-trained Models: *VGG-16, EfficientNet dan Resnet50*)**

Transfer learning memanfaatkan model yang telah dilatih sebelumnya pada dataset besar (seperti ImageNet) untuk mempersingkat waktu pelatihan dan meningkatkan akurasi pada dataset spesifik.

Langkah-langkah Implementasi dengan Transfer Learning:

* Preprocessing Data
  + Resize

Citra diubah ke ukuran 224x224 piksel agar sesuai dengan model pre-trained.

* + Rescaling

Normalisasi nilai piksel ke rentang [0,1] menggunakan Rescaling(1./255).

* Data Augmentation

Menambah variasi data pelatihan melalui rotasi, flipping, zooming, dan perubahan skala untuk menghindari overfitting.

* Pre-trained Model
  + VGG-16, EfficientNet dan Resnet50

Menggunakan lapisan konvolusi dari model untuk ekstraksi fitur dan menambahkan lapisan dense untuk klasifikasi 6 kelas (Corduroy, Denim, Katun, Linen, Organza, Satin).

* + Fine-tuning

Membekukan lapisan awal dan melatih ulang lapisan atas untuk menyesuaikan dataset spesifik.

* Training dan Evaluasi

Menggunakan optimizer Adam dan loss function sparse categorical cross-entropy.

* Evaluasi Model

Menggunakan learning curve, classification report, dan confusion matrix.

* Error Analyst

*Missclassification* dalam model SVM (Support Vector Machine) adalah suatu kondisi ketika model mencoba mengklasifikasikan gambar kain ke dalam jenis tertentu (misalnya, denim atau katun), tetapi salah perdiksi.

Beberapa Penyebab Umum Misclassification:

1. Model Overfitting / Underfitting

Model terlalu kompleks (overfitting) atau terlalu sederhana (underfitting) mungkin tidak dapat mengenali pola dengan baik, sehingga terjadi kesalahan dalam klasifikasi.

1. Data Training Tidak Representatif

Jika data pelatihan tidak mencakup variasi data yang cukup, model akan sulit mengenali pola tertentu, sehingga memprediksi secara salah.

1. Fitur yang Tidak Relevan

Jika fitur yang diekstraksi dari data tidak mendukung klasifikasi dengan baik, model dapat mengalami kesulitan membedakan kelas dengan tepat.

1. Noise

Jika data mengandung noise (gangguan atau kesalahan) dalam data, seperti gambar buram atau kualitas rendah dapat menyebabkan model salah dalam prediksi.

* Skenario eksperiment sederhana.
* Topik

Klasifikasi Model Busana Berdasarkan Jenis Bahan (Corduroy, Denim, Katun, Linen, Organza, dan Satin) Menggunakan Teknik Pengenalan Pola.

* Tujuan

Mengklasifikasikan jenis bahan busana (Corduroy, Denim, Katun, Linen, Organza, dan Satin) berdasarkan pola teksturnya, sehingga model dapat mengenali bahan secara otomatis dari citra busana.

* Dataset

Dataset ini menyediakan kumpulan citra busana yang terdiri dari enam jenis bahan (Corduroy, Denim, Katun, Linen, Organza, dan Satin).

* Algoritma dan Langkah-langkah Klasifikasi

**Model Klasik:** ***Support Vector Machine (SVM)***

Menggunakan algoritma *Support Vector Machine* (SVM) yang termasuk dalam model klasik task *Classification*.

Langkah-langkah:

1. Preprocessing Data
2. Resize

Mengubah ukuran gambar menjadi 300x300 piksel.

1. Greyscale Transformation

Mengubah gambar berwarna menjadi greyscale untuk mengurangi kompleksitas data.

1. Feature Extraction

Menggunakan metode seperti Histogram of Oriented Gradients (HOG), Sobel Edge Detection, Histogram Warna, dan Momen Warna untuk mendapatkan fitur dari gambar.

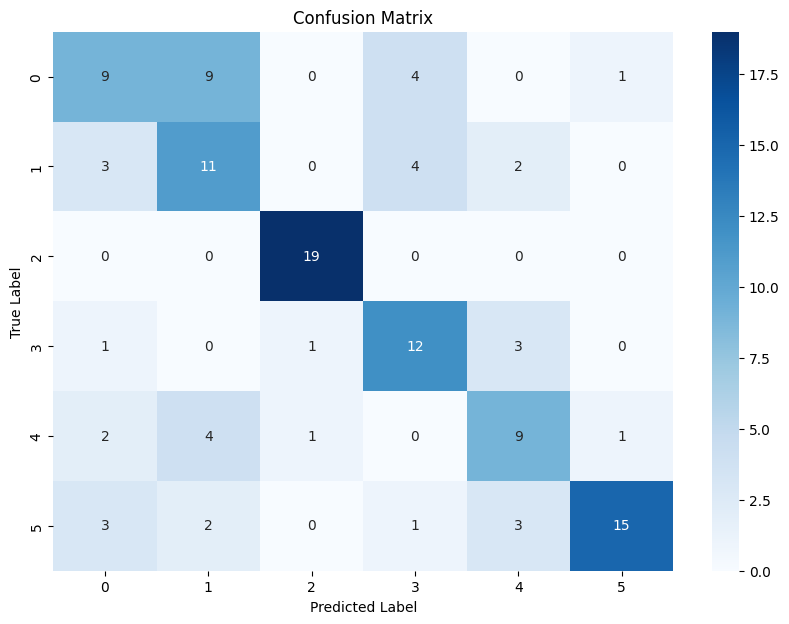
1. Pembagian Data Train, Validasi dan Test

Membagi dataset menjadi train set (80%), validasi set (10%) dan test set (10%).

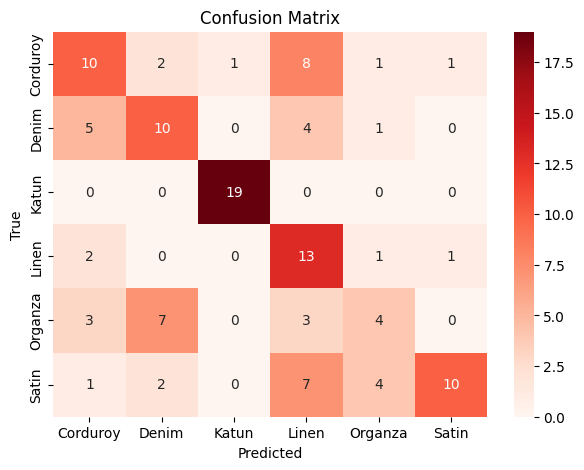
1. Evaluasi Model pada **Feature Extraction**
2. Classification Report

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Fitur Histogram of Oriented Gradient (HOG)** | | | | | |
| **Algoritma** | **Kelas** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| Support Vector Machine (SVM) | Corduroy | 62% | 50% | 39% | 44% |
| Denim | 42% | 55% | 48% |
| Katun | 90% | 100% | 95% |
| Linen | 57% | 71% | 63% |
| Organza | 53% | 53% | 53% |
| Satin | 88% | 62% | 73% |
|  |  |  |  |  |  |
| **Fitur Sobel Edge Detection** | | | | | |
| **Algoritma** | **Kelas** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| Support Vector Machine (SVM) | Corduroy | 55% | 48% | 43% | 45% |
| Denim | 48% | 50% | 49% |
| Katun | 95% | 100% | 97% |
| Linen | 37% | 76% | 50% |
| Organza | 36% | 24% | 29% |
| Satin | 83% | 42% | 56% |
|  |  |  |  |  |  |
| **Fitur Histogram Warna** | | | | | |
| **Algoritma** | **Kelas** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| Support Vector Machine (SVM) | Corduroy | 62% | 50% | 39% | 44% |
| Denim | 42% | 55% | 48% |
| Katun | 90% | 100% | 95% |
| Linen | 57% | 71% | 63% |
| Organza | 53% | 53% | 53% |
| Satin | 88% | 62% | 73% |
|  |  |  |  |  |  |
| **Fitur Moment Warna** | | | | | |
| **Algoritma** | **Kelas** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| Support Vector Machine (SVM) | Corduroy | 48% | 56% | 65% | 60% |
| Denim | 59% | 95% | 73% |
| Katun | 42% | 26% | 32% |
| Linen | 56% | 53% | 55% |
| Organza | 30% | 41% | 35% |
| Satin | 30% | 12% | 18% |

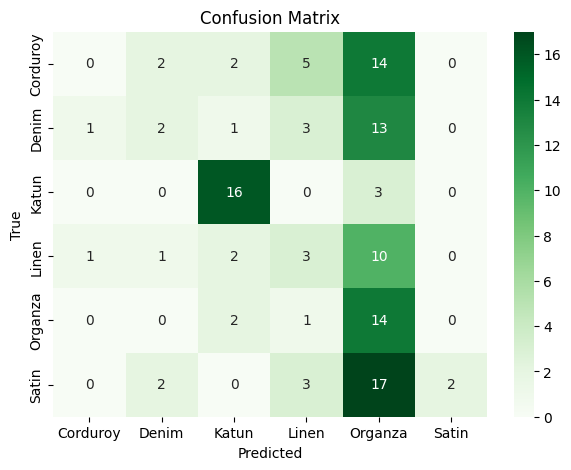
1. Confusion Matrix



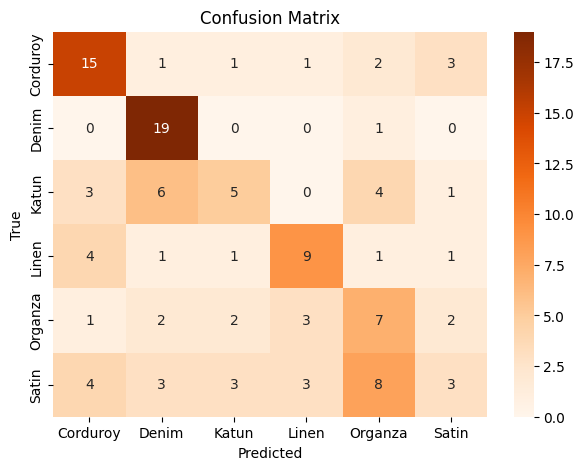
**Gambar 11.** Confusion Matrix Fitur HOG



**Gambar 12.** Confusion Matrix Fitur Sobel Edge Detector



**Gambar 13.** Confusion Matrix Fitur History Warna



**Gambar 14.** Confusion Matrix Fitur Moment Warna

1. Miss Classifier
2. Fitur HOG

Fitur HOG menangkap pola gradien dan tekstur, tetapi pada beberapa jenis kain, seperti Denim dan Corduroy yang teksturnya mirip, HOG tidak cukup akurat membedakan karakteristiknya. Kesalahan ini mengindikasikan bahwa fitur tambahan atau algoritma lain seperti CNN mungkin lebih efektif untuk mengenali pola kain yang lebih kompleks atau halus.



**Gambar 15.** Miss Classifier Fitur HOG

1. Fitur Sobel Edge Detector

Fitur Sobel salah memprediksi kain seperti Corduroy sebagai Linen dan Denim sebagai Corduroy. Fitur Sobel hanya menangkap informasi tepi dasar, sehingga kesulitan membedakan tekstur kain yang mirip, seperti Corduroy dan Linen. Hal ini menunjukkan bahwa Sobel kurang efektif untuk membedakan kain dengan tekstur kompleks.



**Gambar 16.** Miss Classifier Fitur Sobel Edge Detector

1. Discussion

Dalam penelitian ini, fitur HOG mampu menangkap pola gradien dan tekstur kain, namun menunjukkan keterbatasan dalam membedakan jenis kain dengan tekstur yang mirip, seperti Denim dan Corduroy. Kesalahan prediksi ini mengindikasikan bahwa HOG mungkin kurang efektif dalam mengenali karakteristik tekstur yang lebih kompleks atau halus. Sementara itu, fitur Sobel juga mengalami kesalahan dalam klasifikasi, misalnya memprediksi kain Corduroy sebagai Linen atau Denim sebagai Corduroy. Hal ini disebabkan oleh keterbatasan Sobel yang hanya menangkap informasi tepi dasar, sehingga kurang mampu membedakan tekstur kain yang kompleks. Temuan ini menunjukkan bahwa Sobel kurang optimal untuk tugas klasifikasi kain dengan pola dan tekstur yang menyerupai satu sama lain.

**Deep Learning:** ***Convolutional Neural Network (CNN)***

Menggunakan algoritma *Convolutional Neural Network (CNN)* yang termasuk dalam deep learning task *Classification*.

Langkah-langkah:

1. Preprocessing Data
2. Resize

Mengubah ukuran gambar menjadi 300x300 piksel.

1. Rescaling

Normalisasi nilai piksel ke rentang [0, 1] menggunakan Rescaling(1./255).

1. Data Augmentation

Dilakukan untuk meningkatkan keragaman data pelatihan tanpa menambah jumlah data aktual (meliputi: random flip, random rotation, random zoom, random contras)

1. Pembagian Data Train, Validasi dan Test

Membagi dataset menjadi train set (70%), validasi set (15%) dan test set (15%).

1. Arsitektur Model CNN
2. Convolutional Layers

Tiga lapisan konvolusi digunakan dengan jumlah filter bertambah secara progresif (32, 64, 128), kernel size (3, 3), dan fungsi aktivasi ReLU untuk mendeteksi pola pada data gambar.

1. MaxPooling Layers

Setiap lapisan konvolusi diikuti oleh max pooling layer untuk mereduksi dimensi spasial dan menghindari overfitting.

1. Flatten Layer

Meratakan hasil fitur untuk diolah di layer fully connected.

1. Dense Layers

Satu fully connected layer dengan 128 unit dan fungsi aktivasi ReLU.

Lapisan dropout dengan tingkat dropout 50% untuk mengurangi overfitting.

1. Output Layer

Lapisan akhir dengan jumlah unit sesuai dengan jumlah kelas (num\_classes) menggunakan fungsi aktivasi softmax untuk menghasilkan probabilitas per kelas.

1. Early stopping

Teknik untuk menghentikan pelatihan model secara otomatis jika metrik validasi (seperti loss atau akurasi) tidak menunjukkan peningkatan setelah sejumlah epoch tertentu.

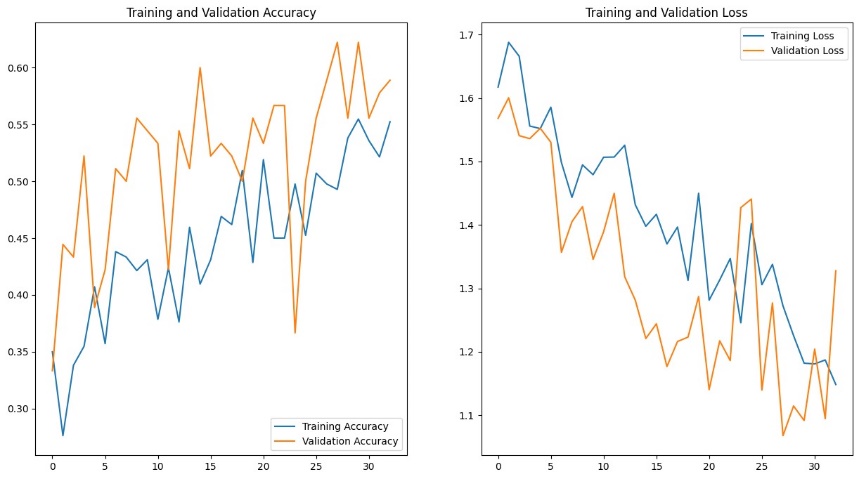
1. Training dan Validasi
2. Compiler

* Optimizer: Adam untuk pembaruan bobot yang efisien.
* Loss function: Sparse categorical cross-entropy untuk klasifikasi multi-kelas.
* Metrics: Akurasi untuk mengevaluasi performa model.

1. Pembagian Data

Dataset dibagi menjadi training set (80%), validation set (10%), dan testing set (10%).

1. Evaluasi Model
2. Learning Curve

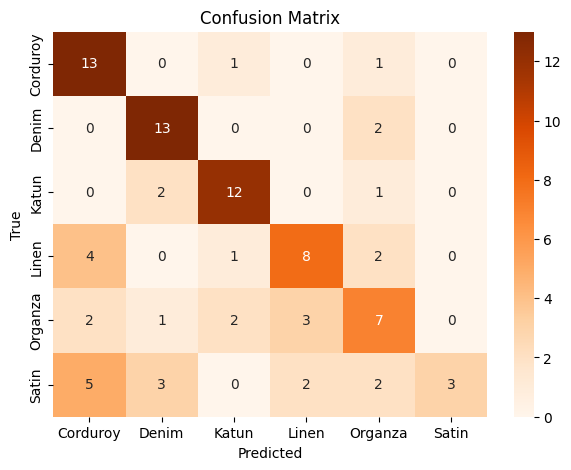


**Gambar 17.** Learning Curve Model CNN

1. Classification Report

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Classification Report** | | | | | |
| **Algoritma** | **Kelas** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| Convolutional Neural Network (CNN) | Corduroy | 62% | 54% | 87% | 67% |
| Denim | 68% | 87% | 76% |
| Katun | 75% | 80% | 77% |
| Linen | 62% | 53% | 57% |
| Organza | 47% | 47% | 47% |
| Satin | 100% | 20% | 33% |

1. Confusion Matrix



**Gambar 18.** Confusion Matrix Model CNN

1. Discussion

Grafik menunjukkan peningkatan akurasi pelatihan dan validasi, meskipun akurasi validasi cenderung fluktuatif. Pelatihan model dihentikan pada epoch ke-33 karena diterapkan teknik early stopping untuk mencegah overfitting dan memastikan performa optimal pada data validasi. Tren penurunan pada loss pelatihan dan validasi menunjukkan model belajar dengan baik, tanpa indikasi overfitting yang signifikan. Classification Report pada kelas Corduroy & Denim memiliki kinerja cukup baik dengan F1-Score masing-masing 67% dan 76%. Katun memiliki performa terbaik dengan F1-Score 77%. Linen & Organza memiliki F1-Score rendah (57% dan 47%), menunjukkan kesulitan dalam mengenali ciri khas. Satin memiliki precision sangat rendah (20%), menunjukkan banyak prediksi salah.

**Transfer Learning (Pre-Trained Models):** ***VGG-16, EfficientNet dan Resnet50***

Menggunakan algoritma *VGG-16, EfficientNet dan Resnet50* yang termasuk dalam transfer learning task *Classification*.

Langkah-langkah:

1. Preprocessing Data
2. Resize

Mengubah ukuran gambar menjadi 224x224 piksel.

1. Rescaling

Normalisasi nilai piksel ke rentang [0, 1] menggunakan Rescaling(1./255).

1. Data Augmentation

Dilakukan untuk meningkatkan keragaman data pelatihan tanpa menambah jumlah data aktual (meliputi: random flip, random rotation, random zoom, random height dan random witdth)

1. Pembagian Data Train, Validasi dan Test

Membagi dataset menjadi train set (70%), validasi set (15%) dan test set (15%).

1. Arsitektur Model
2. Pre-trained Base

Menggunakan arsitektur VGG-16, EfficientNet dan Resnet50 dengan bobot ImageNet. Layer convolutional base dibekukan (frozen) untuk transfer learning.

1. Additional Layers:

* Flatten Layer: Meratakan hasil fitur dari convolutional base.
* Dense Layer: Fully connected layer dengan 256 unit, aktivasi ReLU. Dropout layer (rate: 50%) untuk regularisasi.
* Output Layer: Lapisan dengan 6 unit (jumlah kelas) dan fungsi aktivasi softmax untuk probabilitas multi-kelas.

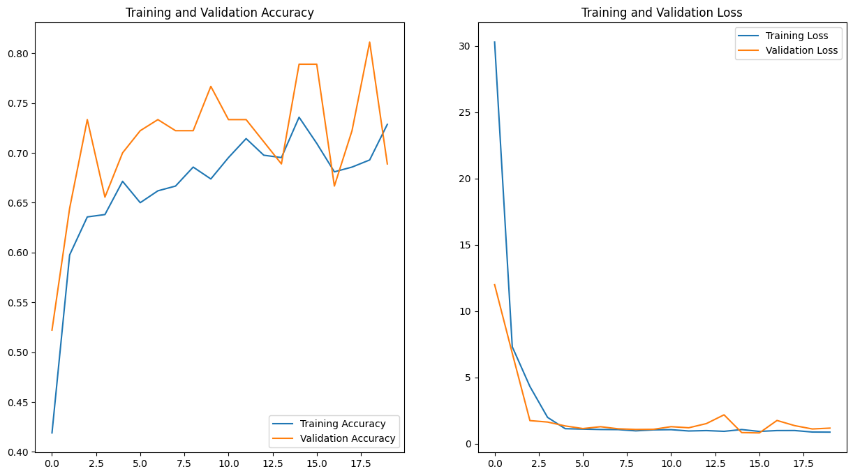
1. Training dan Validasi
   1. Compiler

* Optimizer: Adam untuk pembaruan bobot yang efisien.
* Loss function: Sparse categorical cross-entropy untuk klasifikasi multi-kelas.
* Metrics: Akurasi untuk mengevaluasi performa model.

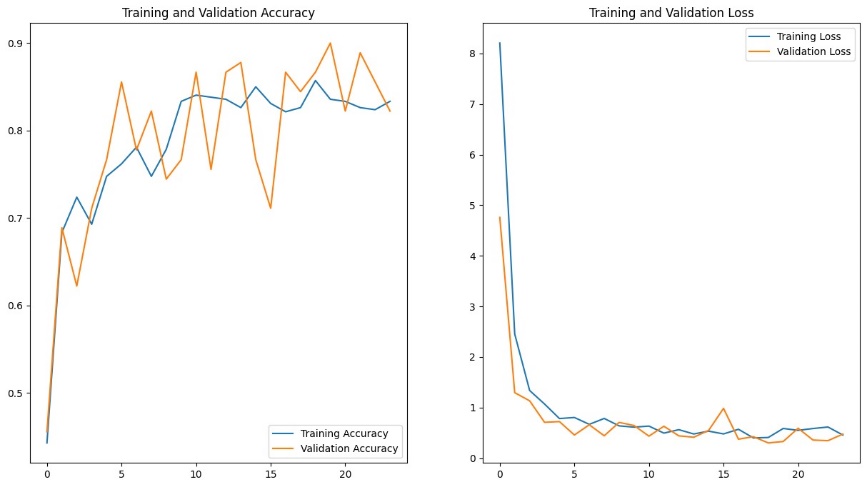
1. Pembagian Data

Dataset dibagi menjadi training set (80%), validation set (10%), dan testing set (10%).

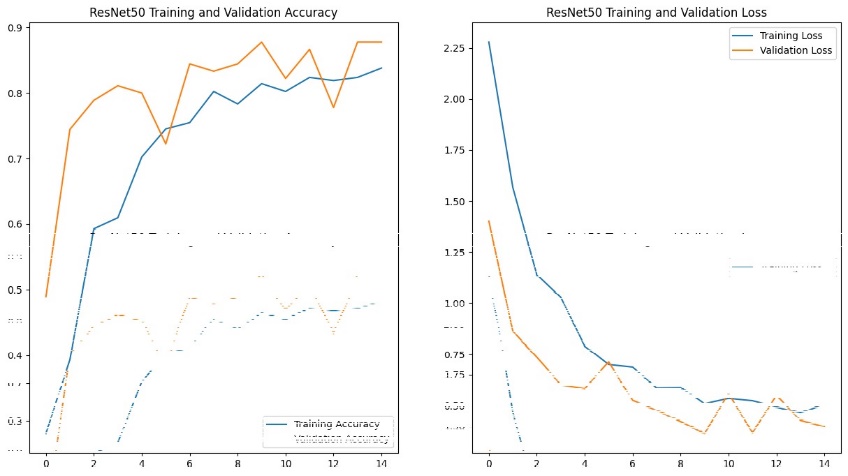
1. Evaluasi Model
2. Learning Curve



**Gambar 19.** Learning Curve VGG-16



**Gambar 20.** Learning Curve EfficientNet

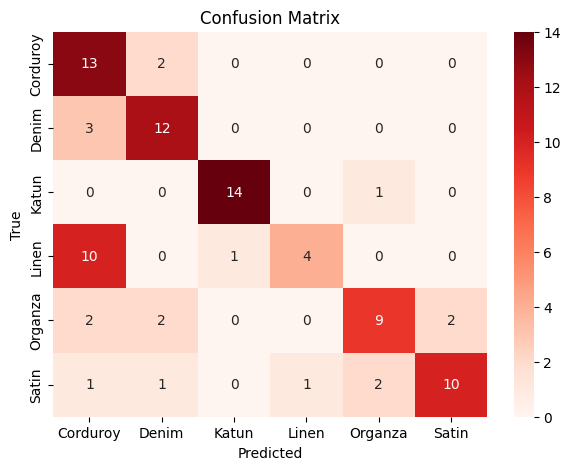


**Gambar 21.** Learning Curve ResNet50

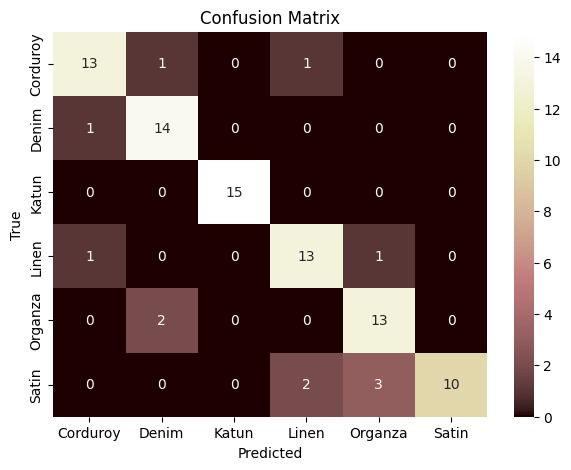
1. Classification Report

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Transfer Learning (Pre-Trained)** | | | | | |
| **Model** | **Kelas** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| VGG-16 | Corduroy | 69% | 45% | 87% | 59% |
| Denim | 71% | 80% | 75% |
| Katun | 93% | 93% | 93% |
| Linen | 80% | 27% | 40% |
| Organza | 75% | 60% | 67% |
| Satin | 83% | 67% | 74% |
|  |  |  |  |  |  |
| **Transfer Learning (Pre-Trained)** | | | | | |
| **Algoritma** | **Kelas** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| EfficientNet | Corduroy | 87% | 87% | 87% | 87% |
| Denim | 82% | 93% | 88% |
| Katun | 100% | 100% | 100% |
| Linen | 81% | 87% | 84% |
| Organza | 76% | 87% | 81% |
| Satin | 100% | 67% | 80% |
|  |  |  |  |  |  |
| **Transfer Learning (Pre-Trained)** | | | | | |
| **Algoritma** | **Kelas** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| ResNet50 | Corduroy | 87% | 88% | 93% | 90% |
| Denim | 75% | 100% | 86% |
| Katun | 100% | 93% | 97% |
| Linen | 83% | 67% | 74% |
| Organza | 92% | 73% | 81% |
| Satin | 88% | 93% | 90% |

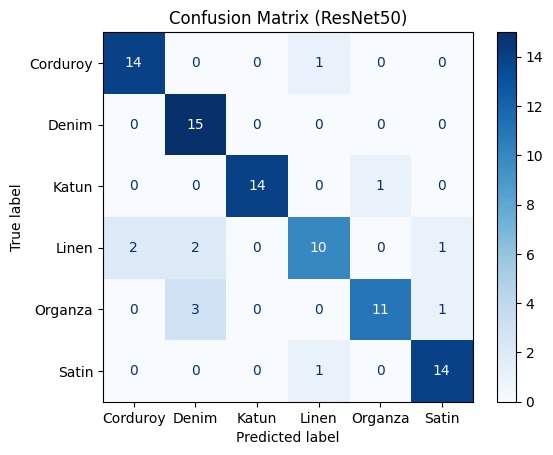
1. Confusion Matrix



**Gambar 22.** Confusion Matrix VGG-16



**Gambar 23.** Confusion Matrix EfficientNet



**Gambar 23.** Confusion Matrix ResNet50

1. Discussion

Penggunaan model pre-trained seperti EfficientNet dan ResNet50 memberikan keuntungan besar dibandingkan dengan VGG-16. Model-model ini memiliki arsitektur yang lebih modern dan mampu menangkap pola-pola kompleks pada dataset. Kinerja yang lebih unggul terlihat terutama pada kelas dengan tekstur kompleks seperti Organza dan Linen.

* EfficientNet menunjukkan kinerja terbaik secara keseluruhan dengan akurasi tinggi pada hampir semua kelas, terutama untuk Katun (100%) dan Linen (81%). Model ini unggul dalam presisi, recall, dan F1-score, sehingga lebih konsisten dalam mendeteksi pola dari dataset.
* ResNet50 juga memiliki performa yang sangat baik, dengan akurasi dan F1-score yang tinggi pada kelas Corduroy (87%), Katun (100%), dan Satin (88%). Namun, performa untuk kelas Linen sedikit lebih rendah (83%).
* VGG-16, meskipun lebih sederhana, menunjukkan performa yang cukup baik dengan akurasi tertinggi pada Katun (93%) tetapi cenderung lebih rendah pada kelas Linen (80%) dan Organza (75%).

**Tahap 6 (poin: 20)**: Knowledge Interpretation

* Pola-pola *useful* yang telah ditemukan.

1. Pola Berdasarkan Tekstur dan Jenis Bahan

* Katun

Katun memiliki pola tekstur yang paling mudah dikenali oleh semua model, termasuk SVM, VGG-16, EfficientNet, dan ResNet50. Hal ini ditunjukkan oleh akurasi, presisi, recall, dan F1-score yang konsisten tinggi.

* Corduroy dan Satin

Kedua jenis bahan ini memiliki pola tekstur cukup berbeda yang membuatnya mudah dibedakan dari jenis kain lain. Pola tekstur kedua jenis kain ini mudah dikenali oleh model arsitektur (EfficientNet dan ResNet50). Ciri khas kain ini yang terdeteksi dengan baik adalah pola garis atau reflektifnya.

* Linen dan Organza

Kain ini memiliki pola tekstur yang lebih halus, yang membuat model yang lebih sederhana seperti VGG-16 sulit untuk mendeteksi pola ini. Namun, model seperti EfficientNet dan ResNet50 memiliki kemampuan yang lebih baik untuk mendeteksi pola ini.

1. Pola Berdasarkan Warna dan Gradasi

* Histogram dan moment warna menunjukkan bahwa distribusi warna sangat penting untuk membedakan kain seperti Organza dan Satin, yang sering memiliki gradasi warna atau pola reflektif.
* Kain dengan warna seragam, seperti Katun dan Linen, bergantung pada pola tekstur daripada warna untuk klasifikasi.

1. Pola Berdasarkan Kinerja Model

* Transfer Learning (EfficientNet dan ResNet50)

Model yang telah dilatih memiliki kemampuan yang lebih baik untuk memahami pola tekstur yang kompleks, seperti Linen dan Organza.

* SVM

Model klasik ini efektif untuk dataset kecil dengan fitur yang sudah diekstraksi (seperti HOG dan Sobel), tetapi kurang mampu mengenali pola yang kompleks dibandingkan dengan model berbasis deep learning. Ini menunjukkan bahwa model dengan arsitektur modern memiliki kemampuan untuk menangkap fitur yang lebih mendalam dari dataset.

1. Faktor yang Mempengaruhi Kinerja Model

* Preprocessing Data

Resize dan transformasi greyscale membantu menyederhanakan data dan meningkatkan efisiensi model, terutama pada dataset yang memiliki variasi ukuran gambar.

* Data Augmentation

Peningkatan keragaman data pelatihan melalui augmentasi (seperti rotasi, flip, dan zoom) membantu model mengenali pola yang lebih luas, terutama pada kain dengan variasi tekstur dan pola.

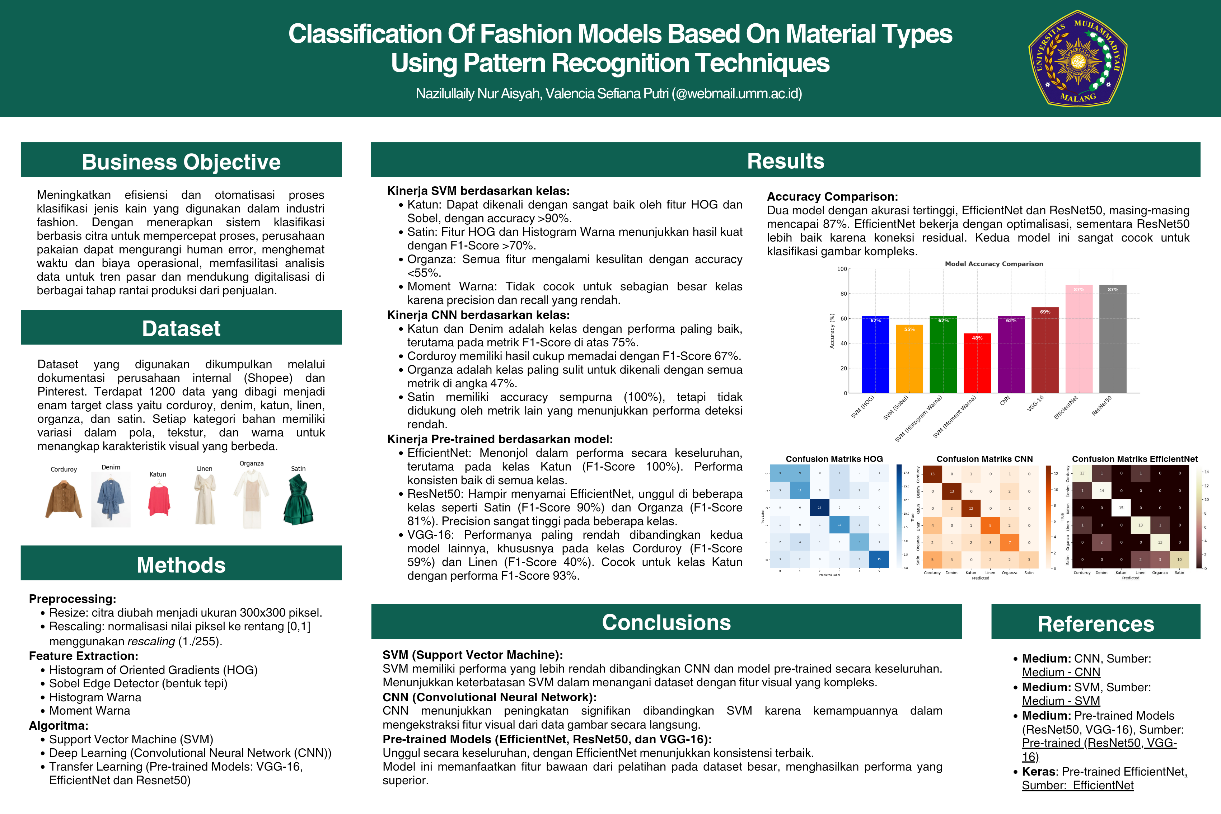
* Noise dalam Data

Adanya noise atau variasi kualitas gambar memengaruhi kemampuan model, terutama pada kelas Linen dan Organza, di mana fitur visual sering kali lebih halus dan sulit ditangkap.

Dengan interpretasi ini, hasil klasifikasi tidak hanya memberikan prediksi tetapi juga informasi penting tentang bagaimana fitur visual kain memengaruhi kinerja model. Hasil-hasil ini dapat menjadi dasar untuk pengembangan lebih lanjut, seperti penerapan model pada dataset yang lebih besar atau sistem klasifikasi berbasis real-time.

**Tahap 7 (poin: 15)**: Reporting

* Simple academic Poster



* Jupiter Notebook (Python)

<https://colab.research.google.com/drive/1suaS2fEEuooYOC6aYy43silGYmmy-ijB?usp=sharing>