

# Klasifikasi Citra Pakaian Menggunakan Convolution Neural Network (CNN) Berbasis MobileNetV2

Rycho Febrian Nalendra Putra<sup>1</sup>, Peniel Manurung<sup>2</sup>, Muzaki Syifauz Zain Arrifki<sup>3</sup>, Dewi Oktafiani<sup>4</sup>

<sup>1234</sup>STMIK Amikom Surakarta

<sup>1234</sup>Sukoharjo - Indonesia

Email: <sup>1</sup>[rhyco.10418@mhs.amikomsolo.ac.id](mailto:rhyco.10418@mhs.amikomsolo.ac.id),

<sup>2</sup>[peniel.10420@mhs.amikomsolo.ac.id](mailto:peniel.10420@mhs.amikomsolo.ac.id), <sup>3</sup>[muzaki.10422@mhs.amikomsolo.ac.id](mailto:muzaki.10422@mhs.amikomsolo.ac.id),

<sup>4</sup>[dewioktafiani@dosen.amikomsolo.ac.id](mailto:dewioktafiani@dosen.amikomsolo.ac.id)

## Abstract

*Clothing image classification is a vital application of computer vision that significantly facilitates the automation of e-commerce systems and digital inventory management. A primary challenge remains the scarcity of training data, which can hinder the model's generalization capabilities. This study aims to design and evaluate a Convolutional Neural Network (CNN) model utilizing Transfer Learning with the MobileNetV2 architecture to enhance the efficiency and accuracy of clothing image classification on small-scale datasets. The dataset consists of 2,000 clothing images distributed across four classes: tops, bottoms, outerwear, and shoes. The research stages include image preprocessing, data augmentation (rotation, flipping, zooming), model training for 250 epochs using the Adam optimizer, and performance evaluation based on accuracy, precision, recall, and F1-score metrics. The results demonstrate a training accuracy of 98.05% and a validation accuracy of 91.75%, with a macro F1-score of 0.91. These findings confirm that the MobileNetV2 architecture offers an optimal balance between computational efficiency and high accuracy, even with limited datasets. This research contributes to the development of lightweight, flexible image classification systems suitable for implementation in e-commerce applications and other visual recognition systems.*

**Keywords:** Convolutional Neural Network (CNN), Image Classification, Clothing, MobileNetV2, Transfer Learning.

## Abstraksi

*Klasifikasi citra pakaian merupakan salah satu penerapan computer vision yang sangat membantu untuk otomatisasi sistem perdagangan daring dan pengelolaan inventaris secara digital. Permasalahan utama yang masih dihadapi adalah kurangnya jumlah data pelatihan, yang dapat menghalangi efektivitas model dalam melakukan generalisasi. Penelitian ini bertujuan untuk merancang dan mengevaluasi model Convolutional Neural Network (CNN) yang menggunakan Transfer Learning dengan arsitektur MobileNetV2, guna meningkatkan efisiensi dan akurasi klasifikasi citra pakaian pada dataset yang kecil. Dataset yang digunakan terdiri dari 2.000 citra pakaian yang terbagi dalam empat kelas, yaitu tops, bottoms, outerwear, dan shoes. Tahapan penelitian ini meliputi prapemrosesan citra, augmentasi data (rotasi, flipping, zooming), pelatihan model selama 250 epoch menggunakan optimizer Adam, serta evaluasi performa berdasarkan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan akurasi pelatihan sebesar 98.05% dan akurasi validasi sebesar 91.75%,*

dengan nilai *macro F1-score* mencapai 0.91. Temuan ini menegaskan bahwa arsitektur *MobileNetV2* dapat memberikan kombinasi yang baik antara efisiensi komputasi dan akurasi tinggi, meskipun dengan dataset yang terbatas. Penelitian ini berkontribusi pada pengembangan sistem klasifikasi citra yang ringan, fleksibel, dan dapat diimplementasikan pada aplikasi berbasis *e-commerce* maupun sistem pengenalan visual lainnya.

**Kata Kunci:** *Convolutional Neural Network (CNN)*, *Klasifikasi Citra*, *Pakaian*, *MobileNetV2*, *Transfer Learning*.

## 1. PENDAHULUAN

Perkembangan teknologi *computer vision* telah mendorong pemanfaatan sistem pengenalan citra dalam berbagai sektor, terutama pada perdagangan daring (*e-commerce*) [1]. Salah satu penerapannya adalah klasifikasi citra pakaian, yang dapat mengidentifikasi jenis pakaian berdasarkan gambar yang ada. Sistem ini berperan penting dalam mengotomatisasi katalog produk, memberikan rekomendasi busana, serta mengelola inventaris digital. Namun, permasalahan utama yang masih dihadapi adalah kurangnya jumlah data pelatihan, yang dapat menghalangi efektivitas model dalam melakukan generalisasi [2].

Berbagai penelitian sebelumnya telah menggunakan arsitektur *Convolutional Neural Network (CNN)* seperti *VGG16*, *ResNet50*, dan *InceptionV3*. Meskipun mampu mencapai akurasi tinggi pada dataset besar, model-model tersebut seringkali tidak efisien secara komputasi dan rentan mengalami *overfitting* ketika diterapkan pada dataset kecil [3]. Hal ini menciptakan celah penelitian (*research gap*) dalam mengembangkan model *CNN* yang mampu memberikan keseimbangan antara efisiensi komputasi dan akurasi yang tinggi meskipun dengan jumlah data yang terbatas.

Pada penelitian ini, digunakan dataset berukuran kecil yang terdiri dari 2.000 citra pakaian dengan empat kelas utama: *tops*, *bottoms*, *outerwear*, dan *shoes*. Pemilihan jumlah 2.000 citra bertujuan untuk melakukan simulasi kondisi nyata di lingkungan *e-commerce* berskala kecil dan menengah (UMKM), di mana sering kali data yang diberi label sangat terbatas. Oleh karena itu, dibutuhkan model yang ringan, cepat, dan tetap memberikan hasil yang akurat.

Arsitektur *MobileNetV2* dipilih karena memiliki karakteristik *depthwise separable convolution* dan *inverted residual block* yang mampu menekan jumlah parameter model tanpa menurunkan akurasi secara signifikan [4]. Dengan memanfaatkan pendekatan *Transfer Learning*, model dapat menggunakan bobot yang telah dilatih sebelumnya dari dataset besar seperti *ImageNet*, sehingga proses pelatihan menjadi lebih efisien dan konvergensi lebih cepat tercapai meskipun data pelatihan terbatas [5].

Tujuan penelitian ini adalah untuk merancang dan mengevaluasi model *CNN* yang menggunakan *MobileNetV2* dalam melakukan klasifikasi citra pakaian dengan dataset yang terbatas. Evaluasi dilakukan secara menyeluruh dengan menggunakan metrik seperti akurasi, presisi, *recall*, dan *F1-score*. Kontribusi utama dari penelitian ini meliputi: (1) Implementasi arsitektur *MobileNetV2* untuk klasifikasi citra pakaian skala kecil

menggunakan pendekatan *transfer learning*.; (2) Analisis kinerja model secara komprehensif melalui metrik *macro-F1*, *weighted-F1*, presisi, dan *recall* per kelas.; dan (3) Visualisasi kesalahan klasifikasi serta pembahasan faktor-faktor penyebab *overfitting* untuk memberikan wawasan terhadap peningkatan performa model.

## 2. TINJAUAN PUSTAKA

Penelitian yang membahas klasifikasi citra pakaian melalui *deep learning* menunjukkan perkembangan yang signifikan dalam lima tahun terakhir. Arsitektur *Convolutional Neural Network (CNN)* menjadi metode utama karena kemampuannya dalam mengekstrak fitur visual dengan cara yang terstruktur dan efisien. Menurut Kayed et al. (2020), penerapan *CNN* pada dataset *Fashion-MNIST* mampu mencapai tingkat akurasi yang tinggi, namun model yang kompleks seringkali memerlukan sumber daya komputasi yang besar [6].

Perkembangan selanjutnya menekankan pada arsitektur yang lebih ringan seperti *MobileNet* dan *EfficientNet* untuk meningkatkan efisiensi. Sandler et al. (2018) memperkenalkan *MobileNetV2* dengan pendekatan *inverted residuals* dan *linear bottlenecks* yang memungkinkan peningkatan performa sambil mengurangi jumlah parameter secara drastis dibandingkan arsitektur tradisional [4].

Selain itu, penelitian terbaru juga menekankan signifikansi dari pengembangan arsitektur *CNN* yang lebih ringan untuk perangkat dengan kemampuan komputasi yang terbatas. Model-model seperti *L-Net* (2024) dan *EdgeFace* (2023) telah menunjukkan peningkatan yang signifikan dalam efisiensi dan kecepatan inferensi tanpa mengorbankan tingkat akurasi, sehingga membuatnya cocok untuk diimplementasi pada sistem yang berbasis *edge computing* [7][8].

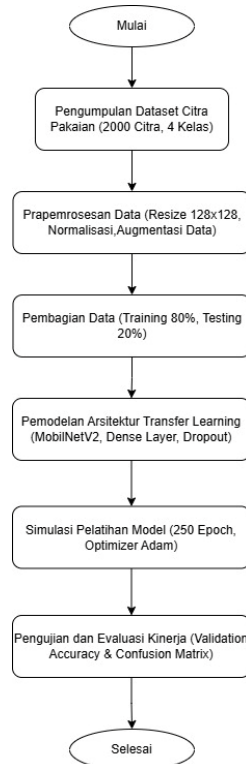
Selain aspek arsitektur, augmentasi data menjadi elemen penting dalam meningkatkan kemampuan generalisasi model, terutama pada dataset berukuran kecil. Shorten dan Khoshgoftaar (2019) mengungkapkan bahwa variasi augmentasi seperti rotasi, *flipping*, dan *zooming* sangat efektif dalam mengurangi risiko *overfitting* dan memperkaya variasi fitur yang dipelajari oleh model [9].

Dari tinjauan literatur tersebut, terlihat adanya gap penelitian di mana sebagian besar studi sebelumnya berfokus pada dataset besar (seperti *DeepFashion*) atau dataset resolusi rendah (seperti *Fashion-MNIST*). Penelitian yang secara spesifik mengevaluasi *MobileNetV2* pada dataset citra pakaian berwarna dengan resolusi menengah dan jumlah sampel terbatas (simulasi lingkungan *e-commerce* kecil) masih perlu diperdalam. Oleh karena itu, penelitian ini menempatkan penggunaan *CNN* berbasis *MobileNetV2* dengan metode *transfer learning* pada dataset 2.000 citra sebagai solusi untuk menjembatani kesenjangan antara akurasi tinggi dan efisiensi komputasi.

Sejalan dengan hal tersebut, penelitian terbaru seperti LitE-SNN (2024) juga mengungkapkan jalur baru dalam pengembangan *lightweight neural network* yang berfokus pada efisiensi energi dan waktu untuk melakukan inferensi, yang memiliki kemungkinan untuk diterapkan dalam klasifikasi citra pada skala kecil [10].

### 3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen berbasis Pemodelan dan Simulasi sistem deep learning. Metodologi penelitian disusun dalam empat tahap utama yang saling berurutan, yaitu persiapan data, perancangan model *transfer learning*, simulasi pelatihan dan pengujian, serta evaluasi kinerja model. Tahapan keseluruhan penelitian divisualisasikan dalam *flowchart* pada Gambar 1.



Gambar 1. Flowchart Metode Penelitian

#### 3.1. Persiapan Data

Dataset yang digunakan terdiri dari 2.000 citra pakaian yang dikategorikan ke dalam empat kelas utama, yaitu *tops*, *bottoms*, *outerwear*, dan *shoes*, dengan distribusi merata masing-masing 500 citra per kelas. Tahap prapemrosesan dimulai dengan mengubah ukuran seluruh citra (*resize*) menjadi  $128 \times 128$  piksel untuk memastikan konsistensi input model. Selanjutnya, dilakukan normalisasi nilai piksel ke rentang 0, 1 dengan membagi nilai piksel asli dengan 255.0. Hal ini bertujuan untuk mempercepat proses konvergensi gradien selama pelatihan.

Untuk meningkatkan generalisasi model, diterapkan teknik Data *Augmentation* menggunakan pustaka *ImageDataGenerator*. Parameter augmentasi meliputi: rotasi acak (*rotation range*) sebesar 20 derajat, pergeseran *horizontal* (*width shift*) 0.2, pergeseran vertikal (*height shift*) 0.2, *shear range* 0.2, *zoom range* 0.2, serta pembalikan *horizontal* (*horizontal flip*). Teknik ini memungkinkan model mempelajari fitur pakaian dari berbagai orientasi dan skala, sehingga lebih tangguh terhadap variasi data baru. Dataset kemudian dibagi dengan rasio 80:20, menghasilkan 1.600 data latih dan 400 data uji.

### 3.2. Perancangan Arsitektur Model

Penelitian ini menerapkan *Transfer Learning* menggunakan arsitektur *MobileNetV2* sebagai base model. Bobot model diinisialisasi menggunakan bobot *pre-trained* dari *ImageNet*, dan lapisan dasar *MobileNetV2* dibekukan (*frozen*) agar bobot fitur dasar tidak berubah selama pelatihan awal.

Lapisan klasifikasi baru (*Custom Head*) ditambahkan di atas base model untuk menyesuaikan dengan tugas klasifikasi 4 kelas. Struktur *custom head* terdiri dari lapisan *GlobalAveragePooling2D*, diikuti oleh *Dense layer* dengan 128 *neuron* dan fungsi aktivasi *ReLU*. Untuk mencegah *overfitting*, ditambahkan lapisan *Dropout* dengan rasio 0.5. Lapisan terakhir adalah *Dense layer* dengan 4 neuron dan fungsi aktivasi *Softmax* untuk menghasilkan probabilitas prediksi tiap kelas. Arsitektur model yang diterapkan pada penelitian ini ditampilkan pada Gambar 2, yang menunjukkan penggabungan antara base model *MobileNetV2* dan *custom head* yang dirancang khusus untuk klasifikasi empat kelas.

***	block_16_expand (Conv2D)	(None, 4, 4, 960)	153,600	block_15_add[0][...]
	block_16_expand_BN (BatchNormalizatio...	(None, 4, 4, 960)	3,840	block_16_expand[...]
	block_16_expand_re... (ReLU)	(None, 4, 4, 960)	0	block_16_expand_...
	block_16_depthwise (DepthwiseConv2D)	(None, 4, 4, 960)	8,640	block_16_expand_...
	block_16_depthwise... (BatchNormalizatio...	(None, 4, 4, 960)	3,840	block_16_depthwi...
	block_16_depthwise... (ReLU)	(None, 4, 4, 960)	0	block_16_depthwi...
	block_16_project (Conv2D)	(None, 4, 4, 320)	307,200	block_16_depthwi...
	block_16_project_BN (BatchNormalizatio...	(None, 4, 4, 320)	1,280	block_16_project...
	Conv_1 (Conv2D)	(None, 4, 4, 1280)	409,600	block_16_project...
	Conv_1_bn (BatchNormalizatio...	(None, 4, 4, 1280)	5,120	Conv_1[0][0]
	out_relu (ReLU)	(None, 4, 4, 1280)	0	Conv_1_bn[0][0]
	global_average_poo... (GlobalAveragePool...	(None, 1280)	0	out_relu[0][0]
	dense (Dense)	(None, 128)	163,968	global_average_p...
	dropout (Dropout)	(None, 128)	0	dense[0][0]
	dense_1 (Dense)	(None, 5)	645	dropout[0][0]
Total params: 2,422,597 (9.24 MB)				
Trainable params: 164,613 (643.02 KB)				
Non-trainable params: 2,257,984 (8.61 MB)				

Gambar 2. Arsitektur Model MobileNetV2 dengan Custom Head

### 3.3. Simulasi Pelatihan dan Evaluasi

Model dikompilasi menggunakan *optimizer Adam* dengan *learning rate* 0.0001 dan fungsi kerugian *categorical cross-entropy*. Pelatihan dilakukan selama 250 *epoch* dengan ukuran *batch* 32. Jumlah epoch yang cukup besar ini dipilih untuk mengamati perilaku konvergensi model dan mendeteksi titik *overfitting*. Kinerja model dievaluasi menggunakan data uji (20%) yang tidak pernah dilihat model selama pelatihan. Metrik evaluasi yang digunakan meliputi *Accuracy* (akurasi), serta analisis mendalam menggunakan *Confusion Matrix* untuk menghitung Presisi, *Recall*, dan *F1-Score* guna memastikan model tidak bias pada kelas tertentu.

## 4. HASIL DAN PEMBAHASAN

### 4.1. Hasil Simulasi Pelatihan Model

Simulasi pelatihan model dilaksanakan selama 250 *epoch*. Ringkasan dari hasil pelatihan dan validasi model ditampilkan dalam Tabel 1, yang menunjukkan nilai *accuracy*, *loss*, serta skor F1 untuk setiap tahap pelatihan.

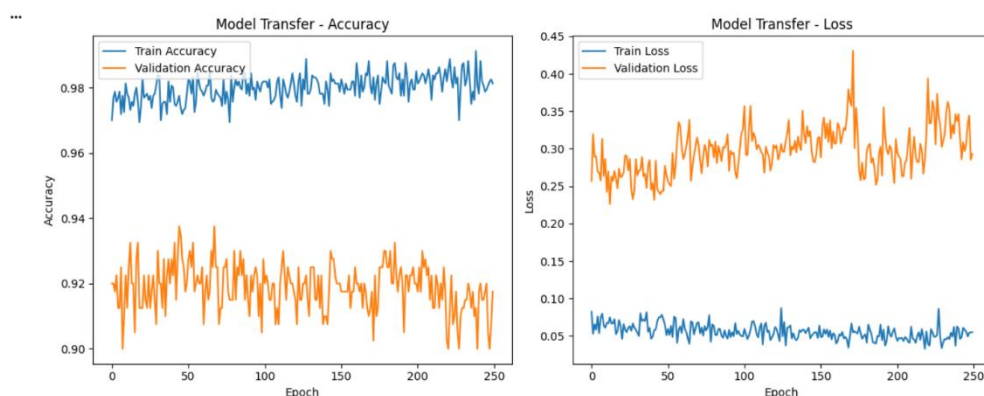
Tabel 1. Hasil Simulasi Pelatihan Model CNN

Metrik Kinerja	Nilai Akhir (Epoch 250)
Training Accuracy	98.05%
Validation Accuracy	91.75%
Training Loss	0.0547
Validation Loss	0.2483
Macro-F1 Score	0.91
Weighted-F1 Score	0.92

Nilai *Validation Accuracy* sebesar 91.75% menunjukkan bahwa model mampu menggeneralisasi data baru dengan sangat baik. Namun, terdapat selisih sekitar 6.3% antara akurasi pelatihan (98.05%) dan validasi, serta perbedaan yang cukup signifikan antara *Training Loss* (0.0547) dan *Validation Loss* (0.2483). Hal ini mengindikasikan terjadinya *overfitting* ringan, di mana model mulai menghafal pola data latih secara spesifik setelah *epoch* tertentu.

### 4.2. Analisis Grafik Kinerja

Kinerja model dianalisis lebih lanjut melalui Gambar 3 yang menampilkan grafik perubahan *accuracy* seiring bertambahnya *epoch*, serta grafik *loss* memperlihatkan perubahan *loss* selama proses pelatihan dan validasi.



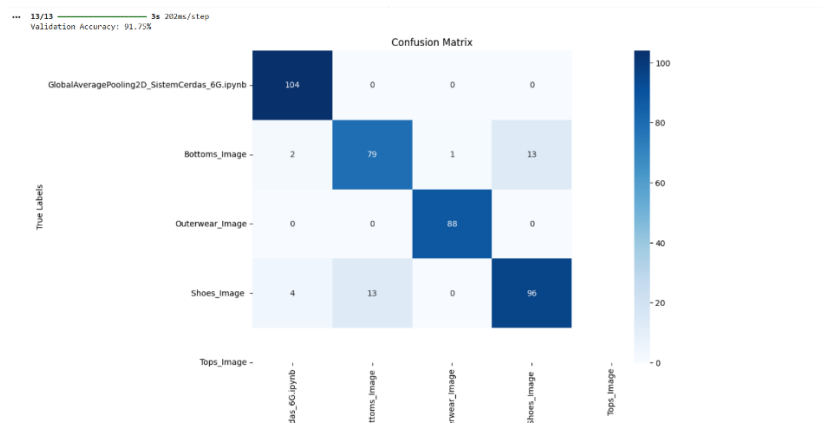
Gambar 3. Grafik Accuracy Model terhadap Epoch dan Grafik Loss Model terhadap Epoch

Berdasarkan Gambar 3, grafik akurasi pelatihan meningkat konsisten dan stabil mendekati angka 98%. Sementara itu, pada Grafik *Loss* terlihat pola yang menarik di mana *Training Loss* terus menurun hingga mendekati nol, sedangkan *Validation Loss* mulai berfluktuasi dan cenderung meningkat perlahan setelah *epoch* ke-150. Kenaikan *validation loss* ini mengonfirmasi analisis sebelumnya mengenai *overfitting*. Meskipun demikian, stabilitas akurasi validasi di angka 91% menunjukkan bahwa model masih

mempertahankan kemampuan klasifikasi yang handal meskipun *loss*-nya sedikit meningkat.

### 4.3. Evaluasi Kinerja Klasifikasi (Confusion Matrix)

Evaluasi mendalam dilakukan menggunakan *Confusion Matrix* untuk membedah kesalahan prediksi antar kelas. Visualisasi hasil klasifikasi dapat dilihat pada Gambar 4.



Gambar 4. Visualisasi Confusion Matrix

Berdasarkan *Confusion Matrix*, performa per kelas dapat dianalisis sebagai berikut.

1. Kelas *Outerwear*: Mencapai performa sempurna dengan prediksi benar sebanyak 88 dari 88 data uji. Model sangat efektif mengenali fitur jaket atau luaran yang distingtif.
2. Kelas *Tops*: Menunjukkan performa sangat baik dengan 104 prediksi benar dari 107 sampel.
3. Kelas *Shoes*: Terdapat 96 prediksi benar, namun terjadi kesalahan prediksi signifikan di mana 4 citra diprediksi sebagai *Tops* dan 13 citra diprediksi sebagai *Bottoms*.
4. Kelas *Bottoms*: Merupakan kelas dengan tingkat kesalahan tertinggi. Dari 93 sampel, hanya 79 yang diprediksi benar. Kesalahan terbesar adalah 13 citra *Bottoms* yang keliru diprediksi sebagai *Shoes*.

Kesalahan klasifikasi yang paling menonjol terjadi antara kelas *Bottoms* (bawahan/celana) dan *Shoes* (sepatu), di mana kedua kelas ini saling tertukar (*misclassified*) dengan jumlah yang signifikan (masing-masing 13 kesalahan). Hal ini kemungkinan disebabkan oleh kemiripan fitur visual tertentu pada citra hasil augmentasi atau pengambilan sudut pandang (*angle*) foto. Misalnya, foto sepatu yang diambil dari atas (*top-down*) bisa memiliki siluet memanjang yang menyerupai celana lipat. Selain itu, tekstur bahan antara celana dan sepatu jenis tertentu bisa memiliki kemiripan warna dan pola bagi model *CNN*.

### 4.4. Visualisasi Kesalahan Klasifikasi

Agar pemahaman tentang kinerja model, dilakukan visualisasi dari beberapa contoh kesalahan prediksi. Gambar 5 menunjukkan salah satu kesalahan klasifikasi yang paling sering muncul dalam penelitian ini, yaitu ketika kelas *Bottoms* salah diidentifikasi sebagai kelas *Shoes*.



Gambar 5. Visualisasi Kesalahan Klasifikasi (Bottoms diprediksi sebagai Shoes)

Pada Gambar 5 terlihat citra pakaian bawah (*Bottoms*) yang salah diprediksi sebagai sepatu (*Shoes*). Kesalahan ini mungkin disebabkan oleh pencahayaan yang tidak memadai kurang optimal, warna kain yang gelap, serta sudut pengambilan gambar yang membuat siluet celana terlihat serupa dengan bentuk sepatu tertutup. Temuan ini sejalan dengan hasil analisis kuantitatif yang terdapat pada Tabel 2, yang menunjukkan bahwa kesamaan fitur visual antara kelas *Bottoms* dan *Shoes* menjadi salah satu faktor utama yang menyebabkan penurunan akurasi model pada kedua kelas tersebut.

#### 4.5. Analisis Kinerja Berdasarkan Matrik Klasifikasi

Berdasarkan analisis yang dilakukan, model *CNN* yang menggunakan *MobileNetV2* memperoleh akurasi pelatihan sebesar 98.05% dan akurasi validasi sebesar 91.75%. Nilai *macro-F1 score* mencapai 0.91 dan *weighted-F1 score* sebesar 0.92, yang menunjukkan performa yang baik di semua kelas. Detail metrik untuk setiap kelas dapat dilihat pada Tabel 2 di bawah ini.

Tabel 2. Metrik Evaluasi Per Kelas

Kelas	Precision	Recall	F1-Score
Tops	0.95	0.97	0.96
Bottoms	0.87	0.85	0.86
Outerwear	1.00	1.00	1.00
Shoes	0.89	0.88	0.88

Dari tabel yang ada, dapat dilihat bahwa kelas *Outerwear* menunjukkan hasil yang sangat baik dengan nilai *F1-Score* 1.00, sedangkan kelas *Bottoms* merupakan kelas dengan kesalahan klasifikasi tertinggi. Hal ini menunjukkan bahwa model lebih mudah mengenali pakaian yang mempunyai ciri khas visual yang jelas, seperti jaket atau pakaian luaran, namun mengalami kesulitan dalam membedakan antara *bottoms* dan *shoes* karena adanya kesamaan dalam bentuk atau tekstur.

#### 4.6. Pembahasan

Secara keseluruhan, hasil penelitian ini dengan *validation accuracy* 91.75% dan *Macro-F1* 0.91 membuktikan bahwa *MobileNetV2* adalah arsitektur yang sangat kapabel untuk tugas ini.



Nilai *F1-score* sebesar 0.91 yang diperoleh dalam penelitian ini lebih unggul jika dibandingkan dengan hasil dari CNN konvensional pada dataset *Fashion-MNIST* (*F1* 0.88) yang dihasilkan oleh Kayed et al. (2020), hal ini menunjukkan bahwa *MobileNetV2* lebih efisien dalam menganalisis fitur citra pakaian meskipun dengan jumlah data yang terbatas [6]. Tren peningkatan efisiensi model yang serupa juga terlihat pada beberapa penelitian terbaru, seperti *L-Net* (2024) dan *LitE-SNN* (2024), yang fokus pada peningkatan kinerja di perangkat dengan keterbatasan sumber daya melalui pendekatan arsitektur yang ringan [7][10].

Hasil ini sejalan dengan penelitian yang dilakukan oleh Sandler et al. (2018) bahwa arsitektur *inverted residuals* efektif menangkap fitur kompleks dengan parameter minim. Namun, isu *overfitting* yang terdeteksi melalui grafik *loss* menyarankan bahwa pelatihan selama 250 *epoch* mungkin terlalu lama untuk dataset sekecil ini. Penerapan mekanisme *Early Stopping* di masa depan sangat disarankan untuk menghentikan pelatihan saat *validation loss* mulai meningkat, sehingga efisiensi waktu pelatihan dapat ditingkatkan dan risiko *overfitting* dapat diminimalisir.

## 5. KESIMPULAN

Penelitian ini berhasil merancang dan mensimulasikan model klasifikasi citra pakaian menggunakan CNN berbasis *MobileNetV2* dengan pendekatan *Transfer Learning*. Hasil eksperimen menunjukkan bahwa model mampu mencapai kinerja tinggi dengan *Validation Accuracy* 91.75% dan *Macro-F1 Score* 0.91 pada dataset terbatas (2.000 citra). Arsitektur *MobileNetV2* terbukti efisien dan efektif dalam mengekstraksi fitur visual pakaian, terutama pada kelas *Outerwear* yang terdeteksi dengan sempurna. Meskipun demikian, evaluasi menunjukkan adanya indikasi *overfitting* ringan akibat jumlah *epoch* yang besar, serta kebingungan model dalam membedakan kelas *Bottoms* dan *Shoes*.

Untuk pengembangan selanjutnya, disarankan menerapkan teknik *Early Stopping* guna mencegah pelatihan berlebih dan melakukan augmentasi data yang lebih spesifik pada kelas *Bottoms* dan *Shoes* untuk mempertegas perbedaan fitur visual keduanya. Penelitian ini menyimpulkan bahwa *Transfer Learning* dengan *MobileNetV2* adalah solusi yang layak dan efisien untuk implementasi sistem kecerdasan buatan pada lingkungan *e-commerce* dengan sumber daya komputasi terbatas.

## DAFTAR PUSTAKA

- [1] Z. O. Harris, G. G. Katta, R. Slater, and J. L. Woodall IV, "Deep Learning for Online Fashion: A Novel Solution for the Retail E-Commerce Industry," *SMU Data Sci. Rev.*, vol. 6, no. 2, p. 17, 2022.
- [2] B. Samia, Z. Soraya, and M. Malika, "Fashion images classification using machine learning, deep learning and transfer learning models," in *2022 7th international*

- conference on image and signal processing and their applications (ISPA), IEEE, 2022, pp. 1–5.
- [3] K. Kandoi, R. N. Ravikumar, S. G. Singh, R. Bediya, K. Mishra, and S. K. Singh, “Comparative Analysis of Deep Learning Models for Fashion Recommendation in E-Commerce,” in *2023 IEEE Fifth International Conference on Advances in Electronics, Computers and Communications (ICA ECC)*, IEEE, 2023, pp. 1–6.
- [4] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.
- [5] S. Duseja and J. Jha, “Transfer learning-based fashion image classification using hybrid 2D-CNN and ImageNet neural network,” *Int. J. Res. Appl. Sci. Eng. Technol*, vol. 9, pp. 1537–1545, 2021.
- [6] M. Kayed, A. Anter, and H. Mohamed, “Classification of garments from fashion MNIST dataset using CNN LeNet-5 architecture,” in *2020 international conference on innovative trends in communication and computer engineering (ITCE)*, IEEE, 2020, pp. 238–243.
- [7] H. Shen, Z. Wang, J. Zhang, and M. Zhang, “L-Net: A lightweight convolutional neural network for devices with low computing power,” *Inf. Sci. (Ny)*, vol. 660, p. 120131, 2024.
- [8] A. George, C. Ecabert, H. O. Shahreza, K. Kotwal, and S. Marcel, “Edgeface: Efficient face recognition model for edge devices,” *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 6, no. 2, pp. 158–168, 2024.
- [9] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *J. big data*, vol. 6, no. 1, pp. 1–48, 2019.
- [10] Q. Liu, J. Yan, M. Zhang, G. Pan, and H. Li, “Lite-snn: Designing lightweight and efficient spiking neural network through spatial-temporal compressive network search and joint optimization,” *arXiv Prepr. arXiv2401.14652*, 2024.