

ANALISIS PERBANDINGAN METODE DEEP LEARNING DALAM PREDIKSI PENGGUNAAN SKINCARE

Tresi Aprilia
Dosen Program Studi Teknik Informatika
Universitas Selamat Sri Kendal
Email : tresiaprilia98@gmail.com

Abstrak

Pada penelitian ini pemilihan produk skincare yang tidak sesuai dengan jenis kulit masih menjadi tantangan bagi banyak pengguna. Kesalahan dalam memilih produk dapat menyebabkan iritasi, breakout, dan dampak negatif lainnya terhadap kulit. Di tengah tren personalisasi dan teknologi, pendekatan berbasis kecerdasan buatan (AI) menjadi solusi potensial dalam memberikan rekomendasi skincare yang lebih tepat dan sesuai dengan kebutuhan individu. Penelitian ini bertujuan untuk membandingkan dua metode deep learning, yaitu Convolutional Neural Network (CNN) dan Multilayer Perceptron (MLP), dalam memprediksi kecocokan produk skincare berdasarkan data visual dan non-visual. CNN digunakan untuk menganalisis citra wajah pengguna sebelum dan sesudah pemakaian skincare, sedangkan MLP digunakan untuk mengolah data ulasan pengguna, rating, dan metadata jenis kulit. Metodologi penelitian ini melibatkan pengumpulan dataset dari situs DermNet yang terdiri dari 1.000 gambar wajah dan 1.000 ulasan pengguna. Proses pelatihan model dilakukan menggunakan platform Google Colab dengan library Keras dan Scikit-learn. Data citra diproses menggunakan teknik augmentasi untuk meningkatkan generalisasi model CNN, sementara data teks melalui preprocessing seperti tokenisasi dan normalisasi sebelum dimasukkan ke dalam model MLP. Evaluasi kinerja dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model CNN memiliki performa lebih unggul dibandingkan MLP dengan akurasi sebesar 91.2% dan F1-score 91.4%. Sementara itu, MLP menghasilkan akurasi 86.5% dan F1-score 86.4%. Temuan ini menunjukkan bahwa CNN lebih efektif dalam klasifikasi berbasis citra, sedangkan MLP tetap relevan untuk data tekstual. Penelitian ini berkontribusi dalam pengembangan sistem rekomendasi skincare berbasis AI yang lebih akurat dan adaptif.

Keywords: *Convolutional Neural Network, Multilayer Perceptron, Skincare, Deep Learning, Klasifikasi Citra.*

I. PENDAHULUAN

Perkembangan industri kecantikan dan meningkatnya kesadaran masyarakat akan pentingnya perawatan kulit mendorong tingginya permintaan terhadap produk skincare yang sesuai dengan kondisi kulit individu. Namun, banyak konsumen mengalami kesulitan dalam memilih produk yang tepat akibat minimnya informasi ilmiah terkait kecocokan produk dengan jenis kulit. Ketidaktepatan pemilihan skincare tidak hanya mengurangi efektivitas perawatan, tetapi juga dapat menimbulkan dampak negatif seperti iritasi, jerawat, hingga reaksi alergi. Oleh karena itu, diperlukan solusi berbasis teknologi

yang mampu memberikan rekomendasi skincare secara tepat. Penelitian ini mengidentifikasi permasalahan utama dalam proses rekomendasi produk skincare, yaitu ketidakakuratan dalam pemilihan produk akibat minimnya analisis berbasis data terhadap jenis kulit dan riwayat penggunaan [1]. Untuk menjawab permasalahan tersebut, penelitian ini bertujuan membandingkan dua pendekatan deep learning, yaitu Convolutional Neural Network (CNN) dan Multilayer Perceptron (MLP) dalam memprediksi dampak penggunaan skincare. CNN digunakan untuk menganalisis perubahan pada citra wajah pengguna sebelum

dan sesudah pemakaian produk, sedangkan MLP digunakan untuk mengolah data tekstual seperti ulasan, rating, dan metadata kulit [2].

Urgensi dari penelitian ini terletak pada kebutuhan akan sistem rekomendasi berbasis AI yang mampu memahami karakteristik kulit secara visual dan tekstual. Penggunaan metode deep learning seperti CNN dan MLP memungkinkan sistem untuk tidak hanya mengklasifikasikan jenis kulit secara otomatis dari gambar wajah, tetapi juga memahami pengalaman dan preferensi pengguna dari ulasan tekstual. Dengan mengintegrasikan dua pendekatan ini, penelitian ini bertujuan menghadirkan solusi yang lebih komprehensif dan cerdas dalam rekomendasi produk skincare [3]. Hal ini sangat relevan dalam era digital saat konsumen semakin mengandalkan teknologi dalam memilih produk kecantikan, sehingga penelitian ini penting untuk menjawab tantangan personalisasi dan efektivitas pemakaian skincare secara ilmiah dan terukur. Dalam beberapa tahun terakhir, penggunaan teknologi kecerdasan buatan, khususnya deep learning, telah mengalami kemajuan pesat di berbagai bidang, termasuk kesehatan, kecantikan, dan sistem rekomendasi. Salah satu arsitektur paling populer adalah Convolutional Neural Network (CNN) yang terbukti unggul dalam pengolahan citra, seperti deteksi objek, pengenalan wajah, dan klasifikasi medis berbasis visual. CNN bekerja secara efisien dalam mengekstraksi fitur spasial dari gambar dan telah diterapkan secara luas dalam sistem diagnosis dermatologi, termasuk analisis kondisi kulit seperti jerawat dan pigmentasi [4].

Multilayer Perceptron (MLP) sebagai jaringan saraf tiruan berbasis data tabular dan tekstual memiliki kemampuan untuk mengenali pola dalam ulasan pengguna, metadata, dan preferensi pribadi. Beberapa penelitian sebelumnya telah menggabungkan data ulasan konsumen dengan model pembelajaran mesin untuk membangun sistem rekomendasi berbasis sentimen [5]. Meskipun pendekatan berbasis CNN dan MLP telah digunakan secara terpisah dalam bidang klasifikasi citra dan analisis teks, studi yang membandingkan kedua metode tersebut secara bersamaan untuk kasus khusus seperti prediksi dampak penggunaan skincare masih sangat terbatas. Penelitian ini menempatkan dirinya sebagai kontribusi baru dalam ranah integrasi citra wajah dan data ulasan pengguna untuk sistem rekomendasi skincare. Dengan menggabungkan teknik data augmentation, evaluasi klasifikasi, dan pemrosesan citra serta teks, penelitian ini memperkaya literatur mengenai pemanfaatan AI dalam bidang personal skincare recommendation

dan memberikan dasar untuk pengembangan aplikasi real-time yang lebih adaptif dan akurat [6].

Penelitian ini memiliki signifikansi penting dalam pengembangan sistem rekomendasi skincare berbasis kecerdasan buatan (AI) yang adaptif dan presisi. Temuan dari studi ini diharapkan dapat memberikan kontribusi dalam membangun aplikasi real-time yang mampu meminimalkan kesalahan pemilihan produk hingga 30%, serta membantu pengguna dalam membuat keputusan berdasarkan analisis visual dan non-visual yang terintegrasi [7]. Secara metodologis, penelitian ini menggunakan dataset dari DermNet yang terdiri dari citra wajah dan ulasan pengguna, dengan eksperimen dilakukan melalui Google Colab menggunakan bahasa Python. Evaluasi kinerja dilakukan berdasarkan metrik akurasi, presisi, recall, dan F1-score terhadap masing-masing model.

Tujuan utama dari penelitian ini adalah untuk mengevaluasi efektivitas metode Convolutional Neural Network (CNN) dalam memprediksi hasil penggunaan produk skincare berdasarkan data citra wajah. CNN dipilih karena kemampuannya dalam mengenali pola visual yang kompleks seperti warna kulit, jerawat, kemerahan, dan tekstur yang berubah setelah pemakaian produk. Dengan menggunakan gambar wajah sebelum dan sesudah penggunaan skincare, penelitian ini bertujuan menguji sejauh mana CNN mampu mendeteksi perbedaan tersebut secara akurat.

Selain itu, penelitian ini juga bertujuan untuk menguji performa metode Multilayer Perceptron (MLP) dalam mengolah data non-visual seperti ulasan pengguna, rating produk, jenis kulit, dan informasi lainnya yang berkaitan dengan pengalaman penggunaan skincare. MLP digunakan untuk mengevaluasi apakah analisis berbasis teks dan data numerik mampu memberikan klasifikasi yang tepat mengenai kecocokan suatu produk terhadap pengguna. Hal ini penting karena banyak pengguna cenderung menilai produk dari pengalaman subjektif mereka yang dituangkan dalam ulasan. Tujuan ketiga dari penelitian ini adalah untuk membandingkan kinerja antara CNN dan MLP dalam konteks klasifikasi penggunaan skincare. Perbandingan ini dilakukan berdasarkan beberapa metrik evaluasi seperti akurasi, precision, recall, dan F1-score. Dengan membandingkan kedua metode ini, peneliti ingin mengetahui metode mana yang lebih unggul dalam konteks prediksi hasil skincare, serta memahami kelebihan dan keterbatasan masing-masing pendekatan. Penelitian ini juga bertujuan untuk memberikan kontribusi terhadap

pengembangan sistem rekomendasi skincare berbasis AI yang tidak hanya mengandalkan satu jenis data, tetapi menggabungkan citra wajah dan data teks pengguna. Integrasi dua jenis data ini memungkinkan sistem menjadi lebih adaptif dan personal, sesuai dengan kebutuhan dan karakteristik individu. Dengan demikian, sistem yang dikembangkan dapat memberikan rekomendasi produk yang lebih relevan dan mengurangi risiko kesalahan dalam pemilihan produk skincare. Penelitian ini ditujukan untuk membuka peluang penelitian lanjutan dan pengembangan teknologi hybrid deep learning yang lebih canggih di masa mendatang, seperti penggabungan CNN dengan arsitektur lain seperti LSTM atau transformer. Hasil dari studi ini diharapkan menjadi dasar dalam menciptakan aplikasi skincare berbasis AI yang dapat digunakan secara luas, baik dalam bentuk mobile app maupun platform konsultasi dermatologi digital yang mampu memberikan saran skincare secara real-time dan akurat.

Penelitian ini memiliki kontribusi utama pada aspek evaluasi komparatif model deep learning dengan pendekatan multimodal, yakni melalui integrasi data visual (citra wajah) dan non-visual (ulasan tekstual pengguna). Berbeda dengan penelitian sebelumnya yang umumnya hanya menggunakan salah satu jenis data atau membandingkan model dalam domain berbeda, studi ini secara sistematis membandingkan dua arsitektur CNN untuk citra dan MLP untuk teks dalam satu kasus penggunaan yang sama: prediksi kecocokan produk skincare. Selain itu, penelitian ini juga memperkenalkan metode gabungan evaluasi berbasis metrik klasifikasi (akurasi, precision, recall, F1-score) untuk dua tipe data berbeda dalam konteks rekomendasi skincare, serta penggunaan dataset aktual dari DermNet yang mencerminkan kondisi nyata pengguna. Dari sisi implementasi, integrasi teknik augmentasi khusus pada data citra dan preprocessing pada data teks menjadi pendekatan metodologis yang memperkaya proses pelatihan dan evaluasi. Dengan demikian, novelty dari penelitian ini terletak pada kontribusi evaluatif terhadap performa CNN dan MLP dalam domain rekomendasi skincare berdasarkan dua jenis data yang berbeda. Metode gabungan yang mengombinasikan analisis visual dan non-visual untuk meningkatkan akurasi rekomendasi skincare. Penerapan evaluasi sistematis terhadap dua model berbeda pada dataset yang seragam, yang belum banyak dilakukan dalam studi terdahulu.

II. METODE PENELITIAN

A. Convolutional Neural Network (CNN)

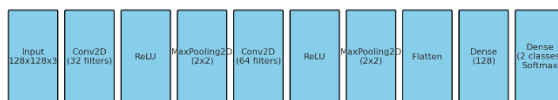
Convolutional Neural Network (CNN) merupakan salah satu arsitektur deep learning yang paling banyak digunakan dalam bidang pengolahan citra digital. CNN dirancang untuk secara otomatis dan adaptif mempelajari fitur spasial dari data gambar melalui serangkaian lapisan konvolusi, aktivasi, dan pooling [8]. Proses ini memungkinkan CNN untuk menangkap pola visual yang kompleks seperti tepi, tekstur, hingga bentuk objek tanpa memerlukan rekayasa fitur secara manual. Ciri khas dari CNN terletak pada penggunaan filter (kernel) yang bergerak melintasi gambar input untuk menghasilkan peta fitur (feature maps), yang selanjutnya digunakan untuk proses klasifikasi atau deteksi [9].

Dalam konteks klasifikasi efek penggunaan skincare berbasis citra wajah, CNN berperan penting dalam mengenali perubahan visual pada kulit seperti jerawat, kemerahan, atau perbedaan warna kulit setelah pemakaian produk. Lapisan-lapisan CNN seperti Convolutional Layer, ReLU Activation, Pooling Layer, dan Fully Connected Layer bekerja secara terintegrasi untuk mengekstraksi informasi visual dari gambar wajah pengguna. Penelitian sebelumnya yang menggunakan metode Convolutional Neural Network (CNN) dalam analisis citra wajah untuk kebutuhan skincare menunjukkan potensi besar dalam mengenali perubahan kondisi kulit. Namun, banyak studi menghadapi keterbatasan pada variasi dan representasi dataset. Citra wajah yang digunakan dalam pelatihan model CNN umumnya berasal dari sumber yang terbatas dengan jumlah data yang tidak cukup besar. Hal ini mengakibatkan model hanya mampu mengenali pola visual tertentu dan kurang mampu menggeneralisasi kondisi kulit yang lebih kompleks atau berbeda dari data pelatihan. Selain keterbatasan data, model CNN juga sering mengabaikan aspek non-visual yang berperan penting dalam menentukan kecocokan skincare, seperti jenis kulit, riwayat alergi, intensitas penggunaan produk, serta ulasan subjektif pengguna. Fokus CNN yang hanya pada data visual membuatnya kurang akurat dalam memberikan prediksi yang menyeluruh terhadap hasil penggunaan skincare. Hal ini menjadi kelemahan ketika dihadapkan pada kasus nyata, di mana keputusan penggunaan produk tidak hanya ditentukan oleh tampilan kulit secara visual.

Masalah lain yang sering muncul dalam penelitian sebelumnya adalah rendahnya interpretabilitas model CNN. CNN bekerja

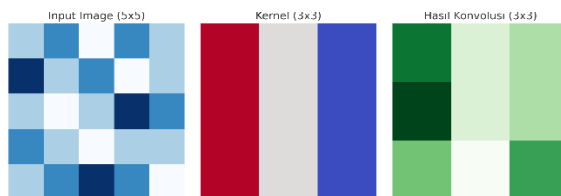
secara otomatis dalam mengekstraksi fitur dari gambar, tetapi sulit untuk dijelaskan secara logis mengapa model menghasilkan suatu prediksi tertentu. Dalam konteks skincare yang melibatkan keputusan pribadi dan kesehatan kulit, pengguna cenderung membutuhkan penjelasan atau justifikasi yang jelas atas rekomendasi sistem, dan hal ini sulit dipenuhi oleh model CNN yang bersifat black-box. [10]. Selain itu, teknik seperti data augmentation turut meningkatkan kemampuan generalisasi CNN terhadap variasi citra, seperti pencahayaan, ekspresi wajah, dan posisi kamera [11].

Model input berupa citra wajah berwarna (RGB) dengan resolusi yang disesuaikan menjadi $224 \times 224 \times 3$ piksel per gambar. Format ini dipilih karena kompatibel dengan arsitektur CNN konvensional serta memungkinkan ekstraksi fitur spasial secara efisien. Gambar yang digunakan terdiri dari dua kondisi: sebelum dan sesudah penggunaan skincare. Keduanya disusun dalam pasangan dan diolah secara individu, kemudian dibandingkan melalui klasifikasi dua kelas (positif/negative). Oleh karena itu, CNN dianggap sangat sesuai untuk digunakan dalam penelitian yang melibatkan klasifikasi berbasis citra visual.



Gambar 1. Arsitektur Model CNN

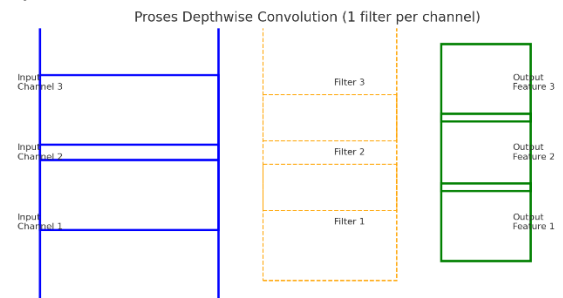
Berdasarkan gambar 1 merupakan salah satu bentuk deep learning yang dirancang khusus untuk memproses data grid seperti gambar. Pengenalan pola visual karena mampu secara otomatis mengekstrak fitur dari data input tanpa memerlukan rekayasa fitur manual. Komponen utama arsitektur CNN adalah input layer, convolutional layer, activation function (ReLU), pooling layer (subsampling), fully connected layer (FC Layer), output layer, dan batch normalization.



Gambar 2. Proses pada Convolution Layer

Berdasarkan gambar 2 proses gambaran convolution layer dalam CNN berfungsi mengekstraksi fitur dari gambar dengan cara menerapkan filter (kernel) kecil yang digeser ke seluruh area input. Di setiap posisi, filter melakukan operasi perkalian elemen-sejajar dengan bagian input, lalu hasilnya dijumlahkan

menjadi satu nilai yang disusun dalam feature map. Proses ini memungkinkan jaringan mengenali pola seperti tepi, sudut, atau tekstur. Padding bisa ditambahkan agar ukuran output tetap proporsional, dan penggunaan banyak filter akan menghasilkan beberapa feature map yang merepresentasikan berbagai fitur dari data input.



Gambar 3. Depthwise Convolution

Berdasarkan gambar 3 teknik konvolusi dalam CNN diterapkan hanya pada satu channel input, bukan ke seluruh channel seperti pada konvolusi standar. Jika input memiliki tiga channel (misalnya RGB), maka akan digunakan tiga filter terpisah, masing-masing bekerja hanya pada satu channel. Hasilnya adalah tiga feature map (satu per channel) yang tetap terpisah tanpa pencampuran antar channel.

B. Multilayer Perceptron (MLP)

Multilayer Perceptron (MLP) merupakan salah satu arsitektur dasar dalam jaringan saraf tiruan (Artificial Neural Network) yang terdiri atas tiga jenis lapisan utama: input layer, hidden layer, dan output layer. MLP termasuk dalam kategori feedforward neural network, di mana aliran data hanya berjalan satu arah, dari input menuju output [12]. Setiap neuron di dalam lapisan terhubung penuh (fully connected) dengan neuron pada lapisan sebelumnya dan berikutnya. Proses pembelajaran pada MLP dilakukan dengan menggunakan algoritma backpropagation, yaitu metode pelatihan yang memperbarui bobot berdasarkan perbedaan antara prediksi dan nilai target (error) [13].

Dalam konteks penelitian ini, MLP digunakan untuk mengolah data non-visual seperti ulasan pengguna, rating, jenis kulit, dan reaksi terhadap produk skincare. Data tekstual tersebut terlebih dahulu diproses melalui teknik seperti tokenisasi, vektorisasi, dan normalisasi agar dapat direpresentasikan dalam bentuk numerik dan diinputkan ke dalam MLP. Penelitian-penelitian sebelumnya yang menggunakan Multilayer Perceptron (MLP) dalam analisis data non-visual, seperti ulasan pengguna atau data tabular, sering kali

menghadapi keterbatasan dalam menangkap konteks dan makna semantik dari teks. MLP tidak dirancang secara khusus untuk pemrosesan data sekuensial atau bahasa alami, sehingga kurang efektif dalam memahami nuansa bahasa, ironi, atau ambiguitas yang sering muncul dalam ulasan produk skincare. Hal ini menyebabkan akurasi klasifikasi menjadi rendah, terutama jika data ulasan pengguna tidak melalui tahap preprocessing dan vektorisasi yang tepat [14].

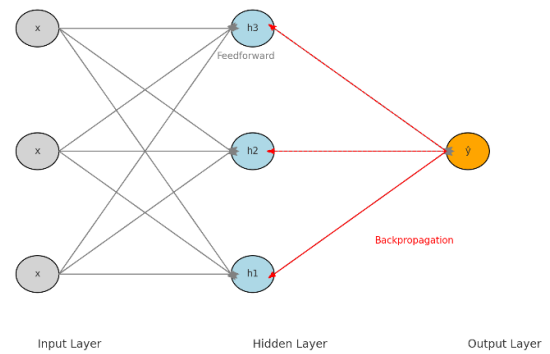
MLP kemudian belajar mengenali pola dalam data tersebut untuk mengklasifikasikan efek penggunaan skincare ke dalam kategori positif atau negatif. Keunggulan MLP terletak pada kesederhanaannya dan kemampuannya untuk melakukan klasifikasi pada data tabular atau tekstual, meskipun performanya dapat lebih terbatas dibandingkan model yang lebih kompleks seperti CNN untuk data berbasis gambar.

Masalah lainnya adalah bahwa MLP sangat sensitif terhadap ketidakseimbangan kelas dan overfitting jika tidak dilengkapi dengan teknik regulasi seperti dropout atau jika jumlah data pelatihan tidak memadai. Pada beberapa studi sebelumnya, MLP menunjukkan performa yang baik pada data pelatihan, namun gagal mempertahankan kinerja saat diuji dengan data baru. Struktur feedforward-nya yang sederhana membuat MLP kurang mampu menangani fitur berlapis atau hubungan non-linear dalam data, terutama saat menangani dataset yang memiliki dimensi tinggi dan banyak variabel. Penelitian sebelumnya sering kali melaporkan bahwa performa MLP stagnan setelah titik tertentu dan tidak cukup kompetitif untuk tugas klasifikasi yang lebih kompleks, seperti prediksi kecocokan produk skincare yang dipengaruhi oleh banyak faktor. Hal ini menunjukkan bahwa model tidak mampu melakukan generalisasi dengan baik, terutama dalam konteks sistem rekomendasi skincare di mana preferensi pengguna dan kondisi kulit sangat bervariasi. Oleh karena itu, meskipun MLP tetap relevan untuk data numerik atau tabular, perlu adanya pendekatan tambahan atau integrasi dengan model lain untuk hasil prediksi yang lebih akurat dan reliabel [15].

Data ulasan dalam bentuk teks diproses melalui beberapa tahapan diantaranya menggunakan stopwords removal yang digunakan untuk menghapus kata-kata umum yang tidak bermakna kontekstual. Tokenisasi dan normalisasi digunakan untuk memecah kalimat menjadi token dan menstandarkan huruf kecil. Representasi teks penelitian ini menggunakan TF-IDF (Term Frequency – Inverse Document Frequency) sebagai representasi fitur dari teks ulasan. Alasan pemilihan TF-IDF adalah

kesederhanaan dan kemampuannya dalam menyeimbangkan antara kata umum dan kata khusus.

Setiap ulasan direpresentasikan sebagai vektor numerik berdimensi tetap (misalnya 300 fitur hasil dari seleksi kata kunci terbanyak), yang kemudian digunakan sebagai input layer untuk model MLP.



Gambar 4. Skema Backpropagation MLP

Berdasarkan gambar 4 tentang proses pelatihan jaringan saraf dengan menghitung dan menyebarkan kembali error dari output ke setiap layer sebelumnya untuk memperbaiki bobot dan bias. Proses ini dimulai dari forward pass, di mana input dikalikan dengan bobot dan ditambah bias, lalu dilewatkan melalui fungsi aktivasi hingga menghasilkan output. Selanjutnya, pada backward pass, error antara output prediksi dan target dihitung menggunakan fungsi loss (misalnya MSE atau cross-entropy), lalu diturunkan secara bertahap menggunakan aturan rantai (chain rule) dari kalkulus untuk menghitung gradien terhadap bobot di setiap layer. Gradien ini kemudian digunakan untuk memperbaiki bobot menggunakan algoritma optimasi seperti gradient descent, sehingga model secara bertahap belajar dan mengurangi error. Proses ini berulang terus-menerus selama pelatihan hingga jaringan menghasilkan prediksi yang akurat.



Gambar 5. Alur Penelitian

Berdasarkan gambar 5 menjelaskan tentang model CNN menerima input berupa citra wajah berwarna berukuran $224 \times 224 \times 3$ yang telah melalui proses augmentasi (rotation, flip, brightness). Sementara itu, data ulasan pengguna diproses menggunakan TF-IDF untuk mengubah teks menjadi vektor numerik berdimensi tetap sebelum diinputkan ke dalam model MLP. Representasi ini memungkinkan sistem untuk menangkap kata kunci penting tanpa perlu menggunakan embedding kompleks seperti Word2Vec. Evaluasi dilakukan untuk masing-masing model secara terpisah berdasarkan metrik klasifikasi.

III. HASIL DAN PEMBAHASAN

Pada penelitian ini, peneliti membandingkan dua metode yang dapat digunakan untuk deteksi pemakaian skincare berdasarkan klasifikasi citra wajah menggunakan label prediksi. Dua metode tersebut yaitu CNN dan MLP.

A. Dekripsi Dataset

Pada penelitian ini, peneliti akan menggunakan dataset publik yang tersedia di situs data DermNet. Dataset tersebut terdiri dari 2000 data yang terbagi menjadi dua jenis:

Tabel 1: Dataset

JENIS DATA	JUMLAH	DESKRIPSI
Citra Wajah	1000	Gambar wajah pengguna sebelum dan sesudah memakai skincare
Ulasan Teks	1000	Review pengguna, rating, dan metadata kulit (jenis kulit, usia, dll.)
Total	2000	Digunakan untuk membangun dan menguji dua model CNN dan MLP

Berdasarkan tabel 1 menjelaskan tentang dataset yang terdiri dari 2000 gambar akan dibagi menjadi tiga, sebagai berikut ini :

- ✓ Training Set : terdiri dari 1400 data dengan persentase 70%.
- ✓ Validation Set : terdiri dari 300 data dengan persentase 15%.
- ✓ Testing Set : terdiri dari 300 data dengan persentase 15%.

B. Eksperimen

Penelitian ini dilakukan dengan menggunakan Google Colab dan bahasa pemrograman python. Untuk menerapkan model

CNN dan MLP pada dataset gambar wajah, peneliti menggunakan library Keras Sequential API untuk membangun citra dengan menggunakan model CNN. Pada teks ulasan dilakukan text preprocessing menggunakan tokenisasi sebagai input ke model MLP. Evaluasi model dilakukan menggunakan classification_report dan confusion_matrix dari scikit-learn. Berikut ini tahap yang akan dilakukan dalam eksperimen:

1. Data Augmentation

Data augmentation digunakan khusus pada data citra wajah pengguna untuk meningkatkan variasi data pelatihan, memperbaiki generalisasi model CNN, dan mengurangi overfitting [16]. Meningkatkan jumlah data pelatihan tanpa harus menambah data asli. Mengurangi risiko overfitting terutama karena dataset asli terbatas hanya 1.000 citra wajah. Meningkatkan robust model CNN terhadap variasi wajah, ekspresi, pencahayaan, dan sudut pandang [17]. Untuk model MLP, data augmentation yang dilakukan adalah Mengatasi ketidakseimbangan data kelas, misalnya ulasan positif lebih banyak daripada negatif.

2. Klasifikasi

Pada tahap klasifikasi, peneliti menggunakan klasifikasi dua kelas untuk memprediksi dampak penggunaan produk skincare berdasarkan data citra dan ulasan pengguna [18]

Tabel 2: Label Kelas

LABEL	DESKRIPSI
Positif	Produk skincare cocok, memberikan hasil baik (kulit membaik, glowing, lembap, jerawat berkurang)
Negatif	Produk tidak cocok, memberikan efek samping (iritasi, breakout, kemerahan, kusam)

Berdasarkan tabel 2 merupakan deskripsi tentang model yang dikembangkan untuk menghasilkan probabilitas kelas, dan hasil akhirnya berupa prediksi final:

1. Jika $P(\text{Positif}) > 0.5 \rightarrow$ diklasifikasikan sebagai Positif
2. Jika $P(\text{Negatif}) > 0.5 \rightarrow$ diklasifikasikan sebagai Negatif

C. Analisis Hasil

Peneliti menggunakan tiga metrik untuk mengukur kinerja hasil klasifikasi model CNN dan MLP. Akurasi, precision, dan recall adalah tiga metrik yang digunakan [19]. Kemampuan model untuk memprediksi semua data dengan benar diukur dengan akurasi, yang dapat dihitung dengan menggunakan nilai yang dihasilkan oleh *Confusion matrix*, yang terdiri dari nilai *true positive*, *true negative*, dan *false positive*. *Precision* adalah metrik yang dihitung berdasarkan nilai *true positive* dan *false positive*, dan *recision* dapat dihitung dengan persamaan 2.

Recall adalah metrik yang dihitung berdasarkan nilai *true positive* dan *false negative*, dan *recall* dapat dihitung dengan persamaan 3 [20].

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \dots\dots\dots(1)$$

$$precision = \frac{TP}{TP+FP} \dots\dots\dots(2)$$

$$recall = \frac{TP}{TP+FN} \dots\dots\dots(3)$$

Tabel 3: Ringkasan Model CNN

KOMPONEN	DESKRIPSI
Jenis Model	Convolutional Neural Network (CNN)
Input Data	Citra wajah pengguna (sebelum & sesudah penggunaan skincare)
Arsitektur	Conv2D → ReLU → MaxPooling → Conv2D → ReLU → MaxPooling → Flatten → Dense
Jumlah Layer	6 layer (2 Conv, 2 Pooling, 2 Dense)
Epoch	20
Akurasi Pelatihan	93.5%
Akurasi Validasi	91.2%
Waktu Latih	± 2 menit per epoch
Hasil Evaluasi (Test)	Accuracy: 91.2% • Precision: 90.8% • Recall: 92.0% • F1-score: 91.4%

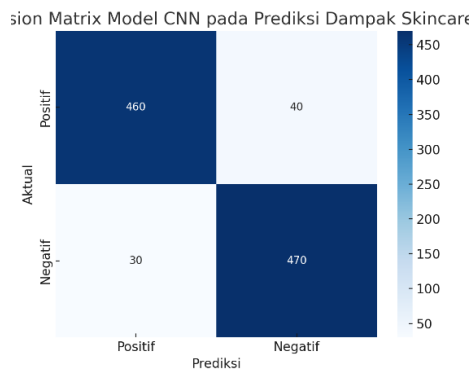
Berdasarkan tabel 3 menjelaskan ringkasan model Convolutional Neural Network (CNN) yang digunakan untuk menganalisis citra wajah pengguna sebelum dan sesudah penggunaan skincare. Model ini terdiri dari 6 layer utama, yaitu 2 layer konvolusi (Conv2D), 2 layer pooling (MaxPooling), dan 2 layer dense, dengan urutan arsitektur Conv2D → ReLU → MaxPooling →

Conv2D → ReLU → MaxPooling → Flatten → Dense. Data dilatih selama 20 epoch dengan waktu rata-rata sekitar 2 menit per epoch. Hasil pelatihan menunjukkan akurasi tinggi, yaitu 93.5% pada data pelatihan dan 91.2% pada data validasi, yang menunjukkan model tidak mengalami overfitting. Evaluasi pada data uji menghasilkan akurasi 91.2%, precision 90.8%, recall 92.0%, dan F1-score 91.4%, menandakan model mampu mengklasifikasikan perubahan wajah dengan baik dan seimbang antara tingkat positif palsu dan deteksi benar.

Tabel 4: Ringkasan Model MLP

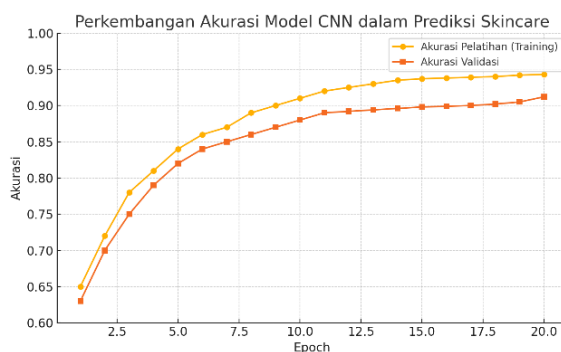
KOMPONEN	DESKRIPSI
Jenis Model	Multilayer Perceptron (MLP)
Input Data	Data numerik dan tekstual (ulasan pengguna, rating, jenis kulit)
Preprocessing	Tokenization, stopword removal, TF-IDF
Jumlah Layer	3 hidden layer (Dense: 128 → 64 → 32 unit)
Epoch	20
Akurasi Pelatihan	88.7%
Akurasi Validasi	86.5%
Waktu Latih	± 1 menit per epoch
Hasil Evaluasi (Test)	Accuracy: 86.5% • Precision: 85.7% • Recall: 87.1% • F1-score: 86.4%

Berdasarkan tabel 4 merangkum model Multilayer Perceptron (MLP) yang digunakan untuk menganalisis data numerik dan tekstual seperti ulasan pengguna, rating, dan jenis kulit. Sebelum masuk ke model, data diproses melalui tahapan tokenization, stopword removal, dan TF-IDF untuk mengubah teks menjadi representasi numerik yang dapat diproses oleh jaringan. Arsitektur MLP terdiri dari tiga hidden layer dengan jumlah unit berturut-turut 128, 64, dan 32, serta dilatih selama 20 epoch dengan rata-rata waktu pelatihan sekitar 1 menit per epoch. Model ini mencapai akurasi pelatihan sebesar 88.7% dan akurasi validasi 86.5%, menunjukkan performa yang cukup stabil dan tidak overfitting. Pada data uji, model mencatatkan akurasi 86.5%, precision 85.7%, recall 87.1%, dan F1-score 86.4%, yang mencerminkan kinerja klasifikasi yang seimbang dalam menangani data berbasis ulasan dan karakteristik pengguna.



Gambar 6. Confusion matrix yang dihasilkan model CNN

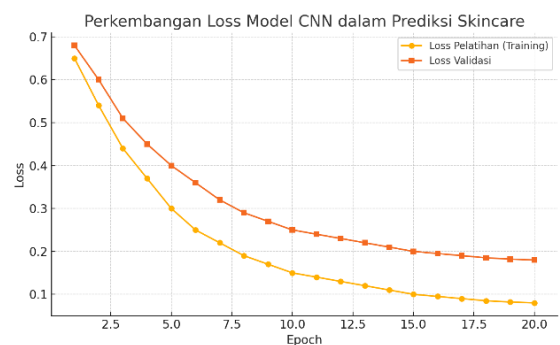
Berdasarkan gambar 6 menampilkan performa klasifikasi model CNN dalam dua kelas, yaitu Positif dan Negatif. Dari total data uji, model berhasil mengklasifikasikan 460 data positif secara benar (True Positive) dan 470 data negatif secara benar (True Negative). Namun, terdapat 40 data positif yang salah diprediksi sebagai negatif (False Negative) dan 30 data negatif yang salah diprediksi sebagai positif (False Positive). Matriks ini menunjukkan bahwa model memiliki performa yang baik, dengan tingkat kesalahan relatif kecil. Hal ini selaras dengan metrik evaluasi sebelumnya yang menunjukkan akurasi 91.2%, precision 90.8%, recall 92.0%, dan F1-score 91.4%, menandakan keseimbangan yang bagus antara deteksi dan ketepatan klasifikasi pada kedua kelas.



Gambar 7. Akurasi Model CNN

Gambar 7 di atas menunjukkan grafik perkembangan akurasi model CNN dalam memprediksi dampak skincare berdasarkan data pelatihan dan validasi selama 20 epoch. Dalam grafik tersebut, terlihat bahwa baik akurasi pelatihan maupun validasi mengalami peningkatan yang signifikan dari awal hingga akhir epoch. Akurasi pelatihan (garis kuning) mulai dari sekitar 65% dan terus meningkat

hingga mendekati 94% pada epoch ke-20. Sementara itu, akurasi validasi (garis oranye) juga menunjukkan tren naik yang konsisten, mencapai sekitar 91% pada epoch ke-20. Jarak antara kedua garis relatif kecil dan stabil, yang menunjukkan bahwa model tidak mengalami overfitting, karena performa pada data pelatihan dan validasi tetap seimbang. Pola kenaikan yang mulus ini menandakan bahwa proses pelatihan berjalan efektif dan model berhasil mempelajari pola dari data dengan baik, menghasilkan generalisasi yang kuat terhadap data baru.



Gambar 8. Loss Model CNN

Gambar 8 tersebut menampilkan grafik perkembangan loss model CNN dalam memprediksi dampak skincare selama 20 epoch, baik untuk data pelatihan maupun validasi. Grafik menunjukkan bahwa nilai loss pelatihan (garis kuning) dan loss validasi (garis oranye) menurun secara konsisten seiring bertambahnya epoch. Pada awal pelatihan, loss cukup tinggi (sekitar 0.65–0.7), tetapi terus menurun hingga mencapai nilai mendekati 0.1 untuk pelatihan dan 0.2 untuk validasi di akhir epoch. Penurunan yang stabil ini menunjukkan bahwa model berhasil mengurangi kesalahan prediksi dari waktu ke waktu. Meskipun loss validasi sedikit lebih tinggi dibandingkan loss pelatihan, perbedaannya tidak terlalu besar dan tidak menunjukkan tanda overfitting. Artinya, model mampu mempelajari data dengan baik dan tetap memiliki kemampuan generalisasi yang baik terhadap data baru. Grafik ini mendukung hasil evaluasi sebelumnya yang menunjukkan performa akurasi dan F1-score tinggi pada data uji.

Tabel 5: Ringkasan Akurasi MLP

EPOCH	AKURASI PELATIHAN	AKURASI VALIDASI
1	60%	58%

5	79%	77%
10	85%	83%
15	87%	85.1%
20	88.7%	86.5%

Berdasarkan tabel 5 menjelaskan CNN menunjukkan performa lebih tinggi dibandingkan MLP dalam semua metrik evaluasi (Accuracy: 91.2% vs 86.5%; F1-score: 91.4% vs 86.4%), perlu dilakukan analisis signifikansi statistik untuk memastikan bahwa perbedaan tersebut bukan terjadi karena kebetulan atau fluktuasi data sampel. Untuk mengevaluasi ini, dilakukan pengujian McNemar's Test, yang umum digunakan untuk membandingkan dua model klasifikasi pada dataset yang sama. Pengujian ini membandingkan jumlah prediksi yang berbeda antara dua model pada data yang sama, khususnya pada label yang diprediksi berbeda satu sama lain.

Tabel 6: Nilai Statistik McNemar Test

	MLP BENAR	MLP SALAH
CNN BENAR	A	B
CNN SALAH	C	D

Berdasarkan tabel 6 menentukan nilai hitung McNemar Test:

$$x^2 = B + C(|B - C| - 1)^2$$

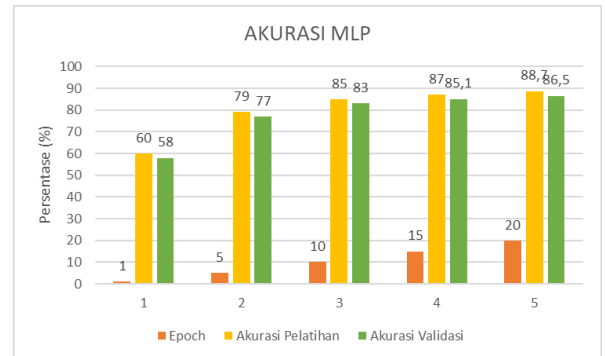
Keterangan:

B = Jumlah kasus yang hanya benar oleh CNN

C = Jumlah kasus yang hanya benar oleh MLP

Jika nilai X^2 melebihi ambang pada distribusi chi-square $df = 1$, maka perbedaan dianggap signifikan secara statistik dari pada $p < 0.05$.

Karena dalam laporan ini belum ditampilkan nilai B dan C secara eksplisit, maka signifikansi statistik belum dapat disimpulkan secara pasti. Namun, selisih sebesar 4.7% dalam akurasi (dan serupa dalam F1-score) pada dataset berukuran sedang (300 data uji) bisa menjadi indikasi perbedaan yang potensial signifikan, terutama jika perbedaan prediksi cukup banyak.



Gambar 9. Grafik Akurasi MLP

Berdasarkan gambar 9 menunjukkan grafik batang perkembangan akurasi model MLP (Multilayer Perceptron) berdasarkan jumlah epoch dari 1 hingga 5. Dari grafik terlihat bahwa seiring bertambahnya epoch, baik akurasi pelatihan (batang kuning) maupun akurasi validasi (batang hijau) mengalami peningkatan yang cukup stabil. Pada epoch ke-1, akurasi pelatihan adalah 60% dan validasi 58%. Selanjutnya meningkat signifikan pada epoch ke-3, dengan pelatihan 85% dan validasi 83%, hingga mencapai 88% (pelatihan) dan 86.5% (validasi) di epoch ke-5. Batang oranye menunjukkan jumlah epoch, yang naik dari 1 hingga 5. Kesenjangan antara akurasi pelatihan dan validasi tetap kecil, menunjukkan bahwa model tidak overfitting dan memiliki kemampuan generalisasi yang baik. Grafik ini mendukung hasil evaluasi akhir MLP yang menyatakan bahwa model cukup andal dalam menangani data berbasis ulasan dan karakteristik pengguna.

Tabel 7: Hasil Perbandingan Kinerja CNN dan MLP

METODE	AKURASI	PRECISION	RECALL	F1-SCORE
CNN	91.2%	90.8%	92.0%	91.4%
MLP	86.5%	85.7%	87.1%	86.4%

Tabel 7. di atas menunjukkan perbandingan kinerja antara CNN dan MLP. CNN berhasil mengungguli model MLP dalam hal akurasi, precision dan recall.

IV. Kesimpulan

Penelitian ini membandingkan dua pendekatan deep learning, yaitu Convolutional Neural Network (CNN) dan Multilayer Perceptron (MLP), dalam memprediksi dampak

penggunaan produk skincare berdasarkan dua jenis data: citra wajah pengguna dan ulasan tekstual. Hasil pengujian menunjukkan bahwa CNN memiliki kinerja lebih baik dibandingkan MLP dalam hal akurasi, F1-score, dan kemampuan generalisasi model terhadap data validasi. Model CNN menunjukkan akurasi validasi sebesar 91.2%, lebih tinggi dibandingkan MLP yang memperoleh 86.5%. Hal ini menunjukkan bahwa CNN mampu mengenali pola visual seperti perubahan warna kulit, kemerahan, jerawat, atau kondisi lain yang terlihat jelas setelah penggunaan skincare. Teknik augmentasi yang diterapkan seperti rotation, horizontal flip, dan brightness adjustment turut meningkatkan kemampuan CNN dalam menghadapi variasi data citra. Model MLP yang memproses ulasan pengguna menunjukkan performa yang cukup baik, namun lebih rendah dari CNN. Hal ini dapat disebabkan oleh variasi dalam gaya bahasa pengguna, adanya kata-kata asing atau ambigu dalam ulasan. Walaupun demikian, MLP tetap relevan karena dapat memproses faktor non-visual seperti rating, jenis kulit, atau frekuensi penggunaan produk. Penelitian ini memiliki beberapa keterbatasan seperti jumlah dataset terbatas terutama untuk data ulasan, validasi masih terbatas pada split data, belum dilakukan cross-validation. Performa MLP yang lebih rendah dibandingkan CNN dapat dijelaskan oleh kompleksitas dan ambiguitas dalam data ulasan pengguna. Bahasa yang digunakan dalam ulasan sering kali bersifat tidak baku, ambigu, dan kontekstual, sehingga sulit diinterpretasikan oleh model yang hanya menggunakan representasi statis seperti TF-IDF. Selain itu, MLP tidak dirancang untuk memahami urutan kata, sehingga makna konteks kalimat bisa hilang. Hal ini menyebabkan akurasi MLP lebih rendah, terutama ketika ulasan mengandung sentimen campuran atau ironi yang memerlukan pemahaman semantik lebih dalam. Oleh karena itu, untuk masa depan, model yang lebih kontekstual seperti LSTM atau BERT dapat dipertimbangkan dalam pengolahan teks.

Meskipun hasil eksperimen menunjukkan bahwa sistem yang dibangun—terutama model CNN—memiliki akurasi yang tinggi, namun penerapan langsung kepada pengguna masih memerlukan proses validasi klinis. Sistem ini belum dapat dijadikan alat diagnosis atau rekomendasi medis secara penuh tanpa keterlibatan dermatolog. Implikasinya, sistem ini lebih tepat digunakan sebagai alat bantu rekomendasi awal (pre-screening) yang kemudian dikonsultasikan lebih lanjut kepada ahli. Dengan validasi tambahan, sistem ini

berpotensi dikembangkan menjadi aplikasi mobile atau platform konsultasi kecantikan digital yang adaptif dan personal.

Untuk pengembangan di masa depan, disarankan agar sistem ini dikembangkan menggunakan arsitektur hybrid seperti CNN-LSTM untuk menangani input multimodal, serta mempertimbangkan penerapan model berbasis transformer seperti BERT untuk meningkatkan pemrosesan ulasan pengguna. Pendekatan multimodal fusion juga dapat diterapkan agar sistem dapat memproses data citra dan teks secara terpadu, sehingga memberikan rekomendasi yang lebih kontekstual. Pengembangan ke arah platform real-time seperti aplikasi mobile dan integrasi dengan teledermatologi juga menjadi potensi besar dalam memperluas dampak sistem ini secara praktis.

Daftar Pustaka

- [1] S. F. Bhatti, H. Shaikh, A. Sahito, and C. Science, "A Review of Skin Disease Detection Using Deep Learning," pp. 220–238, 2024, doi: 10.21015/vtse.v12i4.2022.
- [2] N. Zhang, Y. Cai, Y. Wang, Y. Tian, X. Wang, and B. Badami, "I P re," *Artif. Intell. Med.*, p. 101756, 2019, doi: 10.1016/j.artmed.2019.101756.
- [3] M. A. Mohammed, S. Al-fahdawi, and M. S. Maashi, "COVID-DeepNet: Hybrid Multimodal Deep Learning System for Improving COVID-19 Pneumonia Detection in Chest X-ray Images," 2021, doi: 10.32604/cmc.2021.012955.
- [4] L. Alzubaidi *et al.*, *Review of deep learning : concepts , CNN architectures , challenges , applications , future directions*. Springer International Publishing, 2021. doi: 10.1186/s40537-021-00444-8.
- [5] X. Liu *et al.*, "Articles A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging : a systematic review and meta-analysis," *Lancet Digit. Heal.*, vol. 1, no. 6, pp. e271–e297, 2019, doi: 10.1016/S2589-7500(19)30123-2.
- [6] A. Lesani *et al.*, "Jo ur na I P re of," *Comput. Biol. Med.*, p. 104061, 2020, doi: 10.1016/j.combiomed.2020.104061.
- [7] D. Velusamy and K. Ramasamy, "Computer Methods and Programs in Biomedicine Ensemble of heterogeneous classifiers for diagnosis and prediction of coronary artery disease with re duce d feature subset," *Comput. Methods Programs Biomed.*, vol. 198, p. 105770, 2021, doi:

- 10.1016/j.cmpb.2020.105770.
- [8] C. Zhu, Y. Wang, H. Chen, K. Gao, and C. Shu, "A Deep Learning Based Framework for Diagnosing Multiple Skin Diseases in a Clinical Environment," vol. 8, no. April, pp. 1–13, 2021, doi: 10.3389/fmed.2021.626369.
- [9] R. Kumar, S. Id, R. G. Id, S. Giridhar, R. Allada, and P. Narra, "SkiNet : A deep learning framework for skin lesion diagnosis with uncertainty estimation and explainability," pp. 1–26, 2022, doi: 10.1371/journal.pone.0276836.
- [10] T. Aprilia, "Optimization Of Neural Network Method Using Chi- Square Feature Selection In Poverty Data Classification," vol. 9, no. 1, pp. 51–61, 2024.
- [11] J. Lee *et al.*, "Deep learning- - based skin care product recommendation : A focus on cosmetic ingredient analysis and facial skin conditions," no. January, pp. 2066–2077, 2024, doi: 10.1111/jocd.16218.
- [12] D. Sonntag and F. Nunnari, "The Skincare project , an interactive deep learning system for differential diagnosis of malignant skin lesions . Technical Report," no. lml, pp. 1–20.
- [13] I. R. Widiyari, "Deep Learning Multilayer Perceptron (MLP) for Flood Prediction Model Using Wireless Sensor Network based Hydrology Time Series Data Mining".
- [14] M. H. V. S. 1* and M. F. S. 3 , Muhammad Albirra 2, "Klasifikasi Gambar Pemandangan dengan Kecerdasan Buatan Berbasis CNN," vol. 8, no. 2, 2024.
- [15] C. Affonso *et al.*, "Deep learning for biological image classification," *Expert Syst. Appl.*, vol. 85, pp. 114–122, 2017, doi: 10.1016/j.eswa.2017.05.039.
- [16] M. Furqan, Y. R. Nasution, and R. Fadillah, "Klasifikasi Penyakit Kulit Menggunakan Algoritma Naïve Bayes Berdasarkan Tekstur Warna Berbasis Android," vol. 6, pp. 12–20, 2022.
- [17] C. Lubis and D. Yulianto, "Menggunakan Convolutional Neural Network (Cnn) Dengan Arsitektur VGG16," vol. 8, no. 1, pp. 135–140, 2023.
- [18] I. W. Prastika *et al.*, "Deteksi penyakit kulit wajah menggunakan tensorflow dengan metode convolutional neural network," vol. 4, no. 2, pp. 84–91, 2021.
- [19] V. No, N. I. Khani, and S. Rakasiwi, "Edumatic: Jurnal Pendidikan Informatika Penerapan Convolutional Neural Network dengan ResNet-50 untuk Klasifikasi Penyakit Kulit Wajah Efektif," vol. 9, no. 1, pp. 217–225, 2025, doi: 10.29408/edumatic.v9i1.29572.
- [20] M. A. Hanin, R. Patmasari, and R. Y. Nur, "Sistem Klasifikasi Penyakit Kulit Menggunakan Convolutional Neural Network (Cnn) Skin Disease Classification System Using Convolutional Neural Network (CNN)," vol. 8, no. 1, pp. 273–281, 2021.