

p-ISSN: 2723-567X

e-ISSN: 2723-5661

Jurnal Computer Science and Information Technology (CoSciTech)

http://ejurnal.umri.ac.id/index.php/coscitech/index



IMPLEMENTASI MODEL VISION TRANSFORMER PADA KLASIFIKASI JENIS KULIT WAJAH BERBASIS WEBSITE

Dila Aura Futri*1, Ivana Lucia Kharisma2, Somantri3

Email: ¹dila.aura ti21@nusaputra.ac.id, ²ivana.lucia@nusaputra.ac.id, ³somantri@nusaputra.ac.id

1,2,3,4Program Studi Teknik Informatika, Teknik Komputer dan Desain, Universitas Nusa Putra

Diterima: 16 Agustus 2025 | Direvisi: - | Disetujui: 31 Agustus 2025 © 2020 Program Studi Teknik Informatika Fakultas Ilmu Komputer,

Universitas Muhammadiyah Riau, Indonesia

Abstrak

Permasalahan dalam mengenali jenis kulit wajah seringkali menyebabkan kesalahan dalam memilih produk skincare yang dapat berdampak buruk pada kesehatan kulit. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi jenis kulit wajah secara otomatis berbasis web dengan memanfaatkan arsitektur Vision Transformer (ViT). Model yang digunakan adalah ViT Base Patch 16 yang telah dipra-latih pada dataset ImageNet, kemudian dilakukan fine-tuning dengan dataset citra wajah sebanyak 10.000 gambar yang terbagi rata dalam empat kelas: normal, kering, berminyak, dan kombinasi. Data diproses dengan teknik augmentasi dan normalisasi sebelum dilakukan pelatihan. Hasil pelatihan menunjukkan akurasi sebesar 78% pada data uji, dengan performa terbaik pada kelas combination (F1-score 0.86) dan performa terendah pada kelas normal (F1-score 0.72). Model kemudian diintegrasikan ke dalam sistem berbasis Flask yang memungkinkan pengguna melakukan klasifikasi melalui unggah gambar atau tangkapan kamera secara langsung. Pengujian sistem dilakukan dengan metode functional testing dan API testing menggunakan Postman. Hasil pengujian menunjukkan bahwa seluruh fitur utama sistem berjalan dengan baik, dan API berhasil memberikan respons klasifikasi sesuai format JSON. Dengan demikian, sistem ini dapat membantu pengguna dalam mengenali jenis kulit wajah serta menjadi acuan dalam memilih kandungan skincare yang sesuai.

Kata kunci: Deep Learning, Klasifikasi, Kulit Wajah, Transformer, Vision Transformer.

Abstract

Skin type misidentification often leads to inappropriate skincare product selection, which can negatively affect skin health. This study aims to develop a web-based automatic facial skin type classification system using the Vision Transformer (ViT) architecture. The model implemented is ViT Base Patch 16, pre-trained on the ImageNet dataset and fine-tuned using 10,000 facial images evenly distributed across four classes: normal, dry, oily, and combination. The dataset underwent augmentation and normalization during preprocessing. The training results showed an accuracy of 78% on the test data, with the best performance in the combination skin class (F1-score of 0.86) and the lowest in the normal skin class (F1-score of 0.72). The model was integrated into a Flask-based system that enables users to classify their skin type by either uploading an image or capturing it via camera. System testing was conducted using functional testing and API testing via Postman. The results demonstrated that all key features of the system functioned properly, and the API successfully returned classification responses in JSON format. This system can assist users in identifying their skin type and serve as a reference for selecting appropriate skincare ingredients.

Keywords: Classification, Deep Learning, Skin Type, Transformer, Vision Transformer.

1. PENDAHULUAN

Industri perawatan kulit atau skincare menjadi salah satu sektor yang paling berkembang pesat di pasar global dan lokal. Menurut laporan Allied Market Research, pasar dermatologis global diproyeksikan mencapai \$59,309 miliar pada tahun 2030, dengan tingkat pertumbuhan tahunan gabungan (CAGR) sebesar 11,5% dari tahun 2021 hingga 2030 [1]. Hal ini mengindikasikan pergeseran perawatan kulit dari sekadar kebutuhan estetika menjadi bagian integral gaya hidup sehat. Sejalan dengan tren ini, brand kecantikan lokal terus berinovasi dengan menghadirkan produk berbasis bahan alami dan calming ingredients, yang dirancang untuk memenuhi kebutuhan konsumen yang semakin kritis terhadap kandungan produk yang digunakan [2].

Di Indonesia, minat masyarakat terhadap produk skincare, baik lokal maupun internasional terus meningkat. Banyaknya pilihan produk skincare di pasaran kerap membuat konsumen merasa bingung dan kesulitan memilih produk yang sesuai dengan jenis dan kebutuhan kulit mereka [3]. Ketidakmampuan mengenali kondisi kulit atau memilih kandungan yang sesuai sering kali berujung pada kesalahan pembelian seperti penggunaan produk yang tidak tepat yang berisiko menimbulkan efek samping seperti iritasi atau breakout [4] yang tidak hanya memperburuk kondisi kulit tetapi juga meningkatkan biaya untuk konsultasi tambahan atau pembelian produk yang kurang optimal dalam memberikan hasil yang diharapkan.

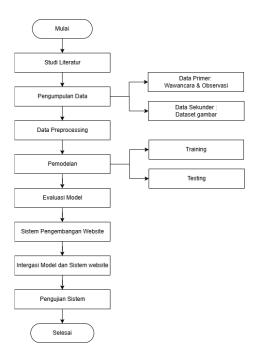
Perkembangan teknologi kecerdasan buatan, khususnya Vision Transformer (ViT) menawarkan peluang baru dalam mengatasi permasalahan klasifikasi citra kulit. Meskipun hingga saat ini belum banyak penelitian yang menerapkan teknologi ViT untuk klasifikasi jenis kulit wajah secara interaktif. Model ini telah menunjukkan hasil yang unggul dalam klasifikasi kanker kulit [5], [6]. Vision Transformer mampu mengolah data dengan kompleksitas tinggi melalui mekanisme self-attention, memberikan akurasi yang lebih baik dibandingkan model tradisional seperti CNN dalam beberapa kasus seperti klasifikasi citra medis dan perawatan kesehatan [7].

Penelitian terdahulu telah menunjukkan efektivitas ViT dalam berbagai konteks. Dyah Ayu Purboningtyas [8] menggunakan model ViT-B/16 untuk klasifikasi tiga kelas jenis kulit wajah dan mengintegrasikannya ke website berbasis Streamlit dengan akurasi mencapai 90,63%. Penelitian ini menegaskan kemampuan ViT dalam mengidentifikasi label positif dengan presisi tinggi. Sementara itu, Penelitian lain oleh Carolin Flosdorf dkk. [9] juga mendukung superioritas ViT, walaupun dengan tantangan pada aspek recall melanoma. Di sisi lain, penelitian seperti yang dilakukan oleh Muhamad Ath-Thariq dkk. [10] masih menggunakan arsitektur CNN untuk klasifikasi kulit wajah dan deteksi penyakit kulit, dengan keterbatasan pada ketidakseimbangan data dan keterbatasan platform berbasis mobile. Dari hasil penelitian-penelitian tersebut dapat disimpulkan bahwa Vision Transformer memiliki potensi kuat dalam mengklasifikasikan jenis kulit wajah karena kemampuannya memahami konteks visual secara global melalui mekanisme self-attention.

Penelitian ini menghadirkan novelty dalam bentuk penggabungan model ViT-Base-patch16-224-in21k dengan sistem berbasis web interaktif menggunakan Flask yang tidak hanya melakukan klasifikasi ke dalam empat kelas kulit (berminyak, kering, normal, dan kombinasi), tetapi juga memberikan rekomendasi kandungan skincare yang sesuai dengan tujuan untuk menghadirkan solusi berbasis teknologi dalam mengatasi permasalahan identifikasi jenis kulit wajah yang sering menjadi kendala bagi konsumen dalam memilih kandungan skincare yang sesuai. Hasil dari penelitian ini dapat dimanfaatkan oleh industri skincare untuk mengemecbangkan produk yang lebih tepat sasaran dan sesuai dengan kebutuhan konsumen di era digital.

2. METODE PENELITIAN

Penelitian ini menggunakan metodologi penelitian kualitatif yang bertujuan untuk memahami dan menggali secara mendalam persepsi dan pengetahuan ahli mengenai klasifikasi jenis kulit wajah serta kandungan skincare yang sesuai. Metodologi ini memungkinkan peneliti untuk memperoleh informasi secara eksploratif melalui pendekatan yang bersifat deskriptif dan kontekstual. Tahapan metode penelitian yang dilakukan dapat dillihat pada gambar di bawah.



Gambar 1. Tahapan Penelitian

2.1. Studi Literatur

Studi literatur merujuk pada penelusuran dan pengkajian sumber-sumber tertulis yang relevan, seperti buku, jurnal nasional, artikel ilmiah, dan skripsi terdahulu yang memiliki keterkaitan tema dengan penelitian [11]. Literatur dilakukan untuk memahami teori dan konsep dari topik penelitian untuk menunjang penguatan argumen dari hipotesis yang diteliti. Pada penulisan ini, dilakukan dengan cara mengunjungi situs website, penelitian terdahulu seperti jurnal, buku dan thesis baik skala nasional maupun internasional. Hasil dari studi literatur digunakan untuk menyusun dasar teori atau kerangka konseptual yang menjadi fondasi penelitian.

2.2. Pengumpulan Data

2.2.2 Data Primer

Data primer adalah informasi utama yang diperoleh langsung dari sumber pertama oleh peneliti, seperti individu atau kelompok yang berkaitan dengan topik penelitian [12]. Pada penelitian ini digunakan metode wawancara secara semi-terstruktur yang dilaksanakan secara langsung bersama seorang dokter yang memiliki kompetensi di bidang dermatologi atau perawatan kulit bertempat di Klinik Arora, Kota Sukabumi. Metode ini digunakan untuk memperoleh pemahaman yang mendalam mengenai suatu fenomena atau permasalahan tertentu terkait jenis kulit [13]. Metode selanjutnya adalah observasi yaitu mengamati objek yang sedang diteliti seperti kondisi kulit wajah orang di sekitar [14].

2.2.3 Data Sekunder

Data sekunder adalah data yang diperoleh secara tidak langsung, biasanya melalui media atau sumber yang telah ada sebelumnya [12]. Pada penelitian ini, dataset yang dikumpulkan di dapat dari beberapa platform seperti kaggle sebanyak 3.152 gambar [15], roboflow sebanyak 3.321 gambar [16], pinterest sebanyak 1.106 gambar dan pengambilan gambar secara langsung sebanyak 12 gamabr. Total dataset yang dikumpulkan berjumlah 7.591 yang terdiri dari 4 kelas diantaranya normal, berminyak, kering dan kelas kombinasi.

2.3. Data Preprocessing

Preprocessing merupakan tahapan awal penyesuaian dataset yang dilakukan dengan membersihkan data mentah yang telah diperoleh seperti data yang tidak lengkap dan tidak akurat. Dalam penelitian ini preprocessing dilakukan dengan transformasi gambar, augmentasi dan menyeimbangkan dataset.

- 2.3.1 Resize Gambar merupakan proses mengubah ukuran gambar ke dimensi yang diperlukan untuk input model sehingga jaringan tidak perlu mengatasi perbedaan ukuran yang diinputkan [17]. Dalam penelitian ini ukuran gambar yang diperlukan adalah 224x224 pixel meggunakan transforms.Resize.
- 2.3.2 Augmentasi adalah teknik yang digunakan untuk meningkatkan variasi data pelatihan tanpa perlu menambah data baru secara manual. Metode ini bekerja dengan memodifikasi data yang ada, seperti memutar, memotong, atau mengubah warna

gambar, sehingga menghasilkan variasi yang lebih beragam [17]. Beberapa teknik augmentasi yang digunakan adalah RandomHorizontalFlip(), RandomRotation(15), ColorJitter(0.2, 0.2, 0.2), RandomResizedCrop(224, scale=(0.8, 1.0)).

- 2.3.4 ToTensor adalah transformasi yang digunakan untuk mengonversi gambar dari format PIL atau NumPy array menjadi tensor dalam PyTorch [18]. Transformasi ini juga secara otomatis menyusun ulang dimensi gambar dan mengubah skala nilai piksel dari 0–255 menjadi 0.0–1.0. Jika gambar berasal dari format yang didukung seperti RGB, maka hasil konversi berupa FloatTensor yang siap diproses oleh model. Konversi ini akan mengubah urutan dimensi dari (tinggi, lebar, saluran) menjadi (saluran, tinggi, lebar).
- 2.3.5 Normalisasi adalah proses yang dilakukan setelah gambar diubah menjadi tensor melalui ToTensor(). Tujuan dari normalisasi adalah untuk menyetarakan distribusi nilai piksel berdasarkan nilai rata-rata (mean) dan standar deviasi (std) dari dataset yang dijadikan acuan, seperti ImageNet. Proses ini membantu mempercepat konvergensi selama pelatihan dan menstabilkan perubahan nilai gradien [19]. Nilai piksel gambar yang semula berada dalam rentang [0.0, 1.0] akan disesuaikan menjadi distribusi yang lebih terstandar, misalnya mendekati distribusi dengan rata-rata 0 dan deviasi 1.

2.4. Pemodelan

Pemodelan sistem pada penelitian ini menggunakan arsitektur Vision Transformer (ViT) dengan model dasar google/vit-base-patch16-224-in21k yang telah dilatih sebelumnya (pretrained). Pemodelan terdiri dari dua tahap yaitu training dan validation. Model training atau pelatihan model merupakan tahapan penting yang bertujuan mengajarkan model cara mengenali pola dari data. Proses ini dilakukan berulang kali, di mana model terus memperbaiki prediksinya dengan membandingkan hasil dengan label yang benar [20]. Pada penelitian ini proses pelatihan model dilakukan dalam 30 epoch menggunakan data latih. Optimasi model menggunakan fungsi CrossEntropyLoss sebagai loss function dan algoritma Adam sebagai optimizer dengan laju pembelajaran awal (learning rate) sebesar 1e-4 dan weight_decay sebesar 1e-5. Scheduler ReduceLROnPlateau digunakan untuk menyesuaikan learning rate secara otomatis ketika validasi stagnan.

Proses validasi dilakukan setiap epoch menggunakan data validasi. Validasi bertujuan untuk mengevaluasi performa model terhadap data yang tidak dilatih guna mengetahui sejauh mana model mampu melakukan generalisasi. Model terbaik disimpan berdasarkan nilai loss validasi terendah. Jika dalam beberapa epoch tidak ada peningkatan performa, maka pelatihan dihentikan lebih awal melalui mekanisme early stopping.

2.5. Evaluasi Model

Evaluasi model merupakan langkah menilai kinerja model klasifikasi setelah proses pelatihan selesai. Hal ini bertujuan untuk menemukan model terbaik yang sesuai dengan data dan dapat berkinerja baik [20]. Langkah pertama dalam evaluasi adalah penyajian confusion matrix, yang menggambarkan hasil klasifikasi model ke dalam empat kategori. True Positive (TP) adalah saat model berhasil memprediksi data sebagai positif (mirip) dan memang benar positif. True Negative (TN) saat model berhasil memprediksi data sebagai negatif (tidak mirip) dan memang benar negatif. False Positive (FP) yaitu kesalahan model memprediksi data sebagai positif padahal sebenarnya negatif. Sedangkan False Negative (FN) yaitu kesalahan model memprediksi data sebagai negatif padahal sebenarnya positif. Struktur confusion matrix ditunjukkan pada tabel berikut:

Tabel 1. Confusion Matrix

| | Prediksi Mirip | Prediksi Tidak Mirip |
|------------------------|---------------------|----------------------|
| Sebenarnya Mirip | True Positive (TP) | False Negative (FN) |
| Sebenarnya Tidak Mirip | False Positive (FP) | True Negative (TN) |

Setelah nilai TP, TN, FP, dan FN diperoleh, barulah dilakukan penghitungan metrik evaluasi yang memberikan gambaran menyeluruh mengenai akurasi dan efektivitas model dalam mengenali pola dan membuat prediksi. Dalam penelitian ini, empat parameter digunakan, yaitu akurasi, presisi, recall, dan skor F1. Rumus masing-masing metrik adalah sebagai berikut:

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

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

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

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$
 (4)

2.6. Sistem Pengembangan Website

Metode yang digunakan dalam mengembangkan website ini adalah waterfall. Metode Waterfall merupakan pendekatan dalam pengembangan perangkat lunak yang dilakukan secara bertahap dan terurut, di mana setiap tahap harus diselesaikan terlebih dahulu sebelum melanjutkan ke tahap berikutnya [21].

2.6.1 Requirement

Requirement atau kebutuhan sistem merupakan tahapan awal dalam pengembangan perangkat lunak yang berfungsi untuk mengidentifikasi, menganalisis, dan merinci seluruh kebutuhan yang diperlukan dalam membangun sistem, baik dari sisi perangkat lunak maupun perangkat keras yang digunakan [21].

| Perangkat Lunak | Fungsi | |
|--------------------|---|--|
| Visual Studio Code | Editor kode sumber untuk pengembangan antarmuka pengguna dan layanan backend. | |
| HTML5 | Merancang struktur dasar halaman web pada sisi frontend. | |
| CSS3 | Mengatur tampilan visual dan tata letak halaman agar responsif dan menarik. | |
| JavaScript | Menangani interaktivitas pada antarmuka pengguna serta integrasi ke API backend. | |
| Python | Bahasa pemrograman utama untuk proses backend dan integrasi model AI | |
| Flask | Framework backend untuk membangun REST API yang memproses dan mengembalikan output. | |

Tabel 2. Kebutuhan Perangkat Lunak

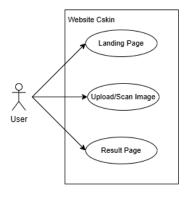
Selain kebutuhan perangkat lunak, sistem ini juga memerlukan dukungan perangkat keras yang memadai untuk menjamin proses pengembangan dan eksekusi sistem berjalan dengan lancar. Berikut spesifikasi perangkat keras yang dibutuhkan:

| Perangkat Keras | Spesifikasi | | |
|---|--|--|--|
| Monitor | 14" | | |
| Keyboard | Merancang struktur dasar halaman web pada sisi frontend. | | |
| Memory | Minimal 8GB DDR4 | | |
| Processor 12th Gen Intel(R) Core(TM) i5-1235U @ 1.30GHz up to 4.40GHz | | | |
| Storage | SSD 512 GB NVMe | | |

Tabel 3. Kebutuhan Perangkat Keras

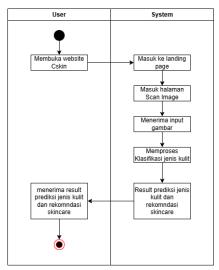
2.6.2 Design

Setelah dilakukan kebutuhan perangkat lunak, selanjutnya adalah perancangan design untuk menggambarkan secara visual bagaimana sistem akan berjalan, baik dari sisi alur fungsionalitas, interaksi pengguna, hingga tampilan antarmuka.



Gambar 2. Use Case Diagram

Diagram di atas menunjukkan alur sederhana dari struktur navigasi pada website Cskin yang terdiri dari tiga halaman utama yaitu Landing Page, Upload/Scan Image dan Result Page. Landing Page berfungsi sebagai titik masuk utama dan memperkenalkan layanan yang ditawarkan oleh Cskin. Dari halaman ini, pengguna dapat diarahkan ke halaman berikutnya untuk memulai proses pemindaian. Halaman Upload/Scan Image memungkinkan pengguna untuk mengunggah gambar kulit atau melakukan pemindaian langsung menggunakan kamera perangkat. Result Page menampilkan hasil probablitas jenis kulit yang di analisis berdasarkan gambar yang diberikan dan rekomendasi kandungan yang sesuai dengan kondisi kulit pengguna.



Gambar 3. Activity Diagram

Gambar activity diagram ini menjelaskan alur aktivitas antara pengguna dan sistem pada website Cskin. Proses dimulai saat pengguna membuka website, kemudian sistem menampilkan halaman awal (landing page). Selanjutnya, pengguna diarahkan ke halaman untuk mengunggah gambar wajah. Setelah gambar diunggah, sistem akan menerima dan memproses gambar tersebut untuk melakukan klasifikasi jenis kulit. Setelah proses analisis selesai, sistem memberikan hasil prediksi jenis kulit beserta rekomendasi skincare yang sesuai. Hasil ini kemudian diterima oleh pengguna sebagai output dari layanan yang disediakan oleh website.

2.6.3 Implementation

Website ini dibangun dengan pendekatan arsitektur client—server, di mana bagian frontend dirancang menggunakan kombinasi HTML, CSS, dan JavaScript untuk menghadirkan antarmuka yang interaktif bagi pengguna. Sementara itu, bagian backend menggunakan bahasa pemrograman Python dengan framework Flask yang berfungsi memproses permintaan pengguna dan menghubungkannya dengan model klasifikasi. Model Vision Transformer (ViT) yang telah melalui tahap pelatihan sebelumnya disimpan dalam format .pth dan dimuat ke dalam backend menggunakan library PyTorch agar dapat diakses secara langsung melalui API.



Gambar 4. Arsitektur Sistem

Alur kerja sistem ditunjukkan pada Gambar 4 Pengguna mengunggah foto wajah melalui antarmuka web, lalu gambar dikirim ke backend melalui permintaan HTTP POST. Backend melakukan praproses seperti resize dan normalisasi sebelum gambar diteruskan ke model ViT. Model kemudian menganalisis citra wajah dan menghasilkan output berupa prediksi tipe kulit (normal, oily, atau dry). Hasil klasifikasi tersebut dikirim kembali ke backend, diteruskan melalui API, dan akhirnya ditampilkan kepada pengguna dalam bentuk informasi jenis kulit beserta rekomendasi kandungan skincare yang sesuai. Dengan demikian, sistem ini mengintegrasikan model klasifikasi kulit berbasis ViT dengan website secara menyeluruh sehingga mampu memberikan hasil analisis secara real-time dan mudah diakses oleh pengguna.

2.6.4 Testing

Secara umum, terdapat tiga jenis pengujian software yang paling umum digunakan, yaitu black box testing, white box testing, dan usability testing [22]. Pada penelitian ini dilakukan pengujian blackbox testing atau yang juga dikenal sebagai behavioral testing. Black box testing menekankan pada pengujian fungsional sistem dari sisi pengguna, dengan fokus utama pada kesesuaian antara input dan output, tanpa melibatkan pemeriksaan struktur internal atau kode program [22].

2.6.5 Operations

Website Cskin yang terhubung dengan model klasifikasi Vision Transformer diimplementasikan di backend menggunakan Flask dan dioperasikan melalui server lokal selama fase pengembangan. Sistem selanjutnya diuji untuk memastikan semua fungsionalitas beroperasi dengan baik. Setelah dinyatakan siap, sistem bisa di-deploy ke platform hosting supaya dapat diakses secara daring oleh pengguna secara luas.

3. HASIL DAN PEMBAHASAN

3.1 Preprocessing Data

Berdasarkan hasil wawancara, diketahui bahwa jenis kulit wajah secara umum terbagi menjadi empat, yaitu kulit normal, berminyak, kering, dan kombinasi. Sedangkan jenis kulit menurut Baumann Skin Type System dibagi menjadi 16 jenis kulit. Cara membedakan jenis kulit dapat dilakukan melalui pengamatan langsung berdasarkan tampilan pori-pori, kilap pada wajah, serta tingkat kelembapan permukaan kulit. Selain itu, dijelaskan pula bahwa jenis kulit dipengaruhi oleh beberapa faktor, seperti genetik, pola makan, pola tidur, kondisi hormonal, lingkungan, dan penggunaan produk skincare. Kemudian untuk kulit kombinasi, narasumber menjelaskan bahwa setiap orang pasti mempunyai kulit kombinasi dengan oily di area T zone karena daerah tersebut memiliki kelenjar minyak paling banyak.

Penelitian ini juga menggunakan dataset sekunder yang di dapat dari beberapa platform yang kemudian dilakukan penyortiran sehingga total dataset yang diambil berjumlah 10.000 gambar. Dataset selanjutnya dibagi ke dalam data latih, data validasi dan data uji dengan perbandingan 70:15:15. Berikut akumulasi dataset yang digunakan.

| Kelas | Jumlah Dataset | | | | |
|-----------|----------------|---------------|----------|--|--|
| | Data Latih | Data Validasi | Data Uji | | |
| Berminyak | 2.000 | 250 | 250 | | |
| Kering | 2.000 | 250 | 250 | | |
| Kombinasi | 2.000 | 250 | 250 | | |
| Normal | 2.000 | 250 | 250 | | |
| Total | | 10.000 | | | |

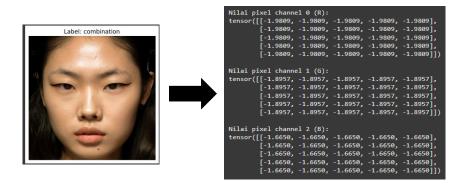
Tabel 4. Jumlah Dataset

Dataset yang telah dikumpulkan selanjutnya di proses dengan teknik resize gambar tujuannya untuk mengubah dimensi gambar dari yang bervarisi menjadi sama yaitu 224x224 piksel. Selanjutnya gambar akan di augmentasi dengan 4 teknik yaitu RandomHorizontalFlip(), RandomRotation(15), ColorJitter(0.2, 0.2, 0.2) dan RandomResizedCrop(224, scale=(0.8, 1.0)). Berikut hasil augmentasi sesuai dengan teknik yang diterapkan.



Gambar 5. Hasil Augmentasi

Pada gambar 5 dapat dilihat bahwa terdapat perbedaan antara gambar original yaitu gambar paling kiri dengan gambar selanjutnya. Berikut keterangan dari masing-masing teknik augmentasi yang digunakan. Original merupakan gambar asli dari dataset yang belum di augmentasi. RandomHorizontalFlip() digunakan untuk membalik gambar secara horizontal secara acak. RandomRotation(15) yaitu memutar gambar dalam -15° hingga +15° derajat secara acak. ColorJitter(0.6, 0.6, 0.6) digunakan mengubah kecerahan, kontras, saturasi, dan hue gambar secara acar ±60%. RandomResizedCrop(224, scale=(0.8, 1.0)) berfungsi untuk memotong acak bagian gambar (antara 80–100% ukuran gambar), lalu resize ke 224x224. Teknik ini akan membuat posisi wajah menjadi sedikit menaik, menurun, ke kanan, ke kiri dan tidak di tengah sempurna. Gambar yang sudah di augmentasi kemudian dikonversi dari format PIL (Python Imaging Library) atau NumPy array ke bentuk tensor agar dapat diolah oleh model deep learning dalam bentuk PyTorch. Konversi ini akan mengubah urutan dimensi dari (tinggi, lebar, saluran) menjadi (saluran, tinggi, lebar).



Gambar 6. Hasil Totensor

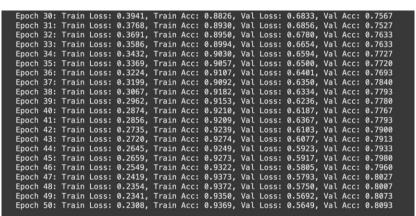
Setelah gambar dikonversi menjadi tensor, langkah berikutnya adalah melakukan normalisasi piksel agar nilai tiap channel (R, G, B) berada dalam rentang tertentu. Proses normalisasi ini dilakukan berdasarkan rata-rata (mean) dan standar deviasi (std) dari dataset ImageNet, yaitu mean = [0.485, 0.456, 0.406] dan std = [0.229, 0.224, 0.225]. Dengan normalisasi tersebut, nilai piksel pada gambar berubah menjadi angka negatif maupun positif seperti ditunjukkan pada Gambar 3.x, di mana channel merah (R), hijau (G), dan biru (B) masing-masing memiliki distribusi nilai yang berbeda. Tujuan dari normalisasi ini adalah untuk menyamakan skala data dengan distribusi data pelatihan awal model sehingga proses training menjadi lebih stabil dan konvergensi dapat dicapai lebih cepat.

3.2 Pemodelan

Model klasifikasi jenis kulit dibangun menggunakan arsitektur Vision Transformer (ViT) tepatnya ViT Base Patch 16 yang telah dipra-latih (pretrained) pada dataset ImageNet. Parameter yang digunakan dalam pelatihan adalah sebagai berikut:

Parameter Nilai Arsitektur Model Vision Transformer (ViT Base Patch 16) Optimizer Adam Loss Function Cross Entropy Loss Metric Accuracy, Loss Epoch 30 Batch 32 Learning Rate 0,0001 Early Stopping Patience = 3

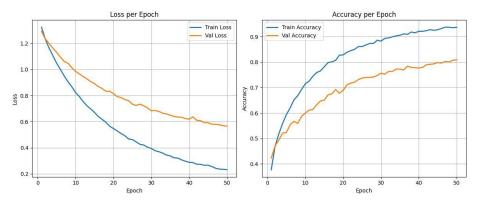
Tabel 5. Parameter Pelatihan



Gambar 7. Proses Pelatihan

Berdasarkan hasil pelatihan pada epoch 21 hingga 50, Train Accuracy meningkat konsisten dari 83.87% menjadi 93.69%, disertai penurunan Train Loss dari 0.5284 menjadi 0.2308. Di sisi lain, Validation Accuracy juga mengalami peningkatan bertahap dari 71.07% menjadi 80.93% dan Validation Loss menurun dari 0.7914 menjadi 0.5649.

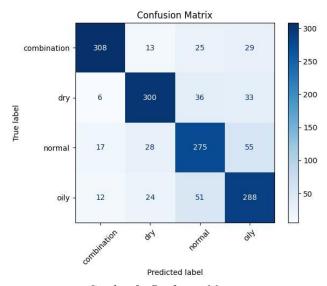
Namun, terdapat gap yang cukup besar antara akurasi pelatihan dan validasi pada epoch akhir misalnya, pada epoch ke-50, Train Accuracy mencapai 93.69% sedangkan Validation Accuracy berada di 80.93%. Hal ini mengindikasikan bahwa model cenderung lebih optimal dalam menghafal data pelatihan dibandingkan menggeneralisasi ke data yang belum pernah dilihat seperti dataset yang ada pada validasi. Dari pola hasil pelatihan diketahui bahwa akurasi pelatihan meningkat secara konsisten sementara peningkatan akurasi validasi cenderung melambat atau stagnan sehingga hal ini mengindikasikan adanya overfitting [23].



Gambar 8. Grafik Loss dan Accuracy

Train Loss mengalami penurunan yang konsisten dari awal hingga akhir pelatihan, yaitu dari sekitar 1.3 menjadi 0.25. Seiring dengan itu, Train Accuracy meningkat signifikan dari 0.4 hingga lebih dari 90% pada akhir epoch, yang menunjukkan bahwa model mampu mempelajari pola data latih dengan baik dan melakukan klasifikasi dengan akurasi tinggi. Pada sisi lain, Validation Loss juga menurun dari sekitar 1.2 menjadi 0.55, sementara Validation Accuracy meningkat bertahap dari 0.4 hingga mencapai lebih dari 80%, dengan tren stabil tanpa penurunan tajam. Meskipun hasil validasi menunjukkan perbaikan, adanya perbedaan yang cukup besar antara akurasi pelatihan dan validasi menandakan bahwa model mengalami overfitting. Kondisi ini menggambarkan bahwa model terlalu menyesuaikan diri dengan data latih sehingga performanya pada data baru tidak sebaik pada data pelatihan. Dengan demikian, meskipun model terlihat optimal pada data latih, kemampuan generalisasinya masih terbatas.

3.3 Evaluasi Model



Gambar 9. Confusion Matrix

Berdasarkan hasil Confussion Matrix di atas dapat dinyatakan bahwa model paling akurat mengenali kulit combination karena kelas combination memiliki prediksi benar terbanyak yaitu 308 dari 375 dataset, hal ini menunjukkan model mampu membedakan ciri khas kulit kombinasi dengan baik. Kulit normal paling sering salah dikenali yaitu hanya sekitar 275 gambar kulit normal yang diprediksi dengan benar. Sisanya sering keliru diklasifikasikan sebagai oily sebanyak 55 gambar. Kesalahan umum terjadi antar kelas yang tampak mirip yaitu normal dan oily. Hal ini mungkin disebabkan oleh kemiripan visual antar jenis kulit, seperti tekstur atau pencahayaan gambar yang menyerupai, sehingga membuat model sulit membedakannya secara akurat.

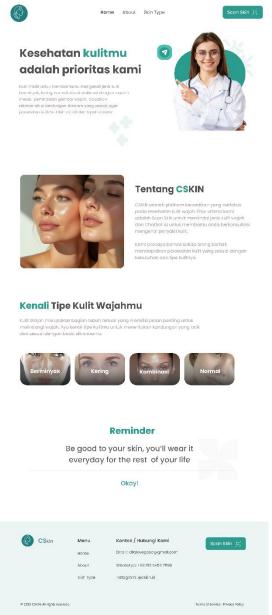
Tabel 6. Classification Report

| Kelas | Precision | Recall | F1-Score | Support |
|-------------|-----------|--------|----------|---------|
| Combination | 0.90 | 0.82 | 0.86 | 375 |
| Dry | 0.82 | 0.80 | 0.81 | 375 |
| Normal | 0.71 | 0.73 | 0.72 | 375 |
| Oily | 0.71 | 0.77 | 0.74 | 375 |
| | | 0.78 | | |

Berdasarkan Tabel 6, model mencapai akurasi keseluruhan sebesar 78%. Kelas Combination menunjukkan performa terbaik dengan precision 0.90 dan F1-score 0.86, menandakan model sangat akurat dalam mengenali kulit kombinasi. Kelas Dry juga cukup baik dengan precision 0.82 dan recall 0.80. Sementara itu, kelas Normal memiliki performa lebih rendah (precision 0.71, F1-score 0.72) sehingga masih sering terjadi kesalahan prediksi. Pada kelas Oily, recall tergolong tinggi (0.77) namun precision rendah (0.71), yang menunjukkan adanya false positive dalam klasifikasi kulit berminyak.

3.4 Implementasi Sistem Website

3.4.1 Landing Page



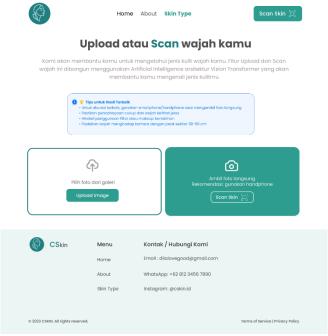
Gambar 10. Landing page

Halaman awal website menampilkan pengenalan sistem klasifikasi kulit wajah dengan beberapa section utama: Home, About, Skin Type, dan Skin Scan. Pada Home, pengguna disambut dengan tujuan platform CSKIN sebagai alat bantu mengenali jenis kulit dan memberikan rekomendasi skincare, lengkap dengan tombol "Konsultasi Sekarang". Bagian About menjelaskan bahwa CSKIN menggabungkan teknologi AI dengan kecantikan untuk mendukung kesehatan kulit. Pada Skin Type, ditampilkan penjelasan singkat mengenai tipe kulit (berminyak, kering, kombinasi, dan normal) beserta gambar pendukung. Fitur utama ada pada Skin Scan, yang memungkinkan pengguna memindai wajah menggunakan model Vision Transformer (ViT) untuk mendeteksi jenis kulit secara otomatis, dengan tombol "Scan Skin" sebagai akses cepat.

3.4.2 Halaman Input Gambar

Jurnal Computer Science and Information Technology (CoSciTech) Vol. 6, No. 2, Agustus 2025, hal. 214-229

Tampilan halaman dirancang minimalis dan intuitif dengan dua kotak interaktif besar yang mudah dikenali dan diakses. Pada opsi upload image memungkinkan pengguna untuk mengunggah gambar wajah dari perangkat mereka. Fitur ini memberikan fleksibilitas bagi pengguna yang telah memiliki foto wajah sebelumnya. Sedangkan pada opsi Scan Skin akan mengaktifkan kamera perangkat untuk melakukan pemindaian wajah secara langsung



Gambar 11. Halaman Input Gambar

3.4.3 Halaman Hasil

Halaman ini ditampilkan setelah proses klasifikasi jenis kulit berhasil dilakukan. Di bagian atas halaman terdapat hasil klasifikasi berupa jenis kulit pengguna, contohnya: "Tipe Kulit Anda Berminyak", disertai dengan foto wajah yang telah di-scan atau di-upload. Berikut tampilan halaman hasil klasifikasi jenis kulit wajah tertera pada gambar 12. Di bawah.



Gambar 12. Halaman Hasil

Pada Gambar 12 ditunjukkan bahwa sistem memberikan rekomendasi kandungan skincare yang disesuaikan dengan jenis kulit yang terdeteksi. Rekomendasi ini dibagi ke dalam tiga kategori dasar perawatan kulit, yaitu cleanser, moisturizer, dan protector. Pada kategori cleanser, sistem menyarankan penggunaan bahan aktif seperti Niacinamide, Salicylic Acid, dan Glycolic Acid yang efektif untuk mengatasi kulit berminyak. Pada kategori moisturizer, kandungan yang direkomendasikan antara lain Ceramide, Hyaluronic Acid, dan Niacinamide dengan tujuan menjaga kelembapan kulit serta memperkuat skin barrier tanpa menyumbat pori-pori. Sementara itu, pada kategori protector, sistem menganjurkan bahan seperti Zinc Oxide dan Titanium Dioxide untuk memberikan perlindungan dari paparan sinar matahari sekaligus membantu mencegah timbulnya jerawat.

3.5 Pengujian Sistem

Pengujian sistem dibagi menjadi dua pengujian yaitu pengujian API dan pengujian functional testing. Pengujian Application Programming Interface (API) menggunakan Postman untuk memastikan bahwa endpoint dapat merespon permintaan dengan benar. Pengujian dilakukan dengan metode HTTP POST di mana pengguna mengirimkan file gambar melalui form-data dengan key image. Berikut hasil dari API testing dengan Postman.

Tabel 7. Hasil API Testing

| TTP Verb | Endpoint | Action | Deskripsi Pengujian | Hasil yang Diharapkan | Hasil yang Diperoleh | Status |
|-------------|----------|--------|--|--|--|--------|
| POST | /predict | Create | Mengirim gambar kulit wajah melalui form-data menggunakan key image | Server merespons status 200 OK dan mengembalikan JSON berisi hasil klasifikasi | Respon JSON berisi key: predicted_class, confidence, probabilities, dan success dengan nilai prediksi seperti "normal", confidence score, dan probabilitas tiap kelas. | > |

Hasil pengujian menunjukkan bahwa server merespons dengan kode status 200 OK, serta menampilkan hasil klasifikasi dalam format JSON yang mencakup nama file, kelas prediksi, confidence score dan persentase probabilitas untuk masing-masing kelas kulit. Respon yang diterima sesuai dengan ekspektasi yang menunjukkan bahwa API dapat memproses input gambar dan menghasilkan keluaran klasifikasi secara fungsional dan akurat. Hal ini membuktikan bahwa API berjalan dengan baik dan telah terintegrasi secara optimal pada sistem.

Selanjutnya pengujian sistem dilakukan dengan cara functional testing dimana fitur dari masing-masing halaman dan section akan di uji berdasarkan skenario pengujian. Berikut hasil functional testing:

Tabel 8. Functional Testing

| HALAMAN UTAMA | | | | |
|---|---|---|------------|--|
| Skenario Pengujian | Hasil yang diharapkan | Hasil yang didapatkan | Keterangan | |
| Menampilkan navbar di semua halaman | Navbar muncul dan dapat diakses di setiap halaman | Navbar muncul dengan tombol "Home", "Scan", "About", dan "Contact" | √ | |
| Klik tombol "Home" | Mengarahkan user kembali ke halaman utama | Halaman utama berhasil ditampilkan | ✓ | |
| Klik tombol "about" | Mengarahkan user ke section about di halaman utama | Section about berhasil ditampilkan | √ | |
| Klik tombol "Skin type" | Mengarahkan user ke section <i>Skin type</i> di halaman utama | Halaman <i>Skin type</i> di halaman utama berhasil ditampilkan | ✓ | |
| Klik card di section "Skin type" pada setiap jenis kulit | Menampilkan modal berisi informasi karakteristik dari tipe kulit yang dipilih | Modal berhasil ditampilkan dan menampilkan informasi yang sesuai dengan jenis kulit yang dipilih | √ | |
| | HALAM | AN SCAN SKIN | 1 | |
| Skenario Pengujian | Hasil yang diharapkan | Hasil yang didapatkan | Keterangan | |
| Klik tombol "scan skin" | Pindah halaman ke <i>page</i> scan skin | Berhasil pindah page | ✓ | |
| Klik tombol "Upload Image" | Website menampilkan hasil klasifikasi dengan persentase probabilitas berdasarkan gambar yang diunggah. | Hasil klasifikasi berhasil ditampilkan setelah gambar diunggah | √ | |
| Klik tombol "Scan Skin" untuk Capture gambar | Website menampilkan hasil klasifikasi dengan persentase probabilitas berdasarkan hasil tangkapan kamera. | Hasil klasifikasi ditampilkan setelah user melakukan scan wajah secara langsung (Contoh hasil: 53.4%.) berminyak. | ✓ | |
| Responsif di perangkat mobile | Tampilan menyesuaikan dengan ukuran layar kecil | Tampilan tetap proporsional dan navigasi tetap dapat digunakan | ✓ | |

Berdasarkan hasil pengujian menggunakan metode Functional Testing pada sistem klasifikasi tipe kulit berbasis web, dapat disimpulkan bahwa seluruh fitur utama berjalan dengan baik dan sesuai dengan fungsinya. Semua skenario pengujian pada halaman utama maupun halaman scan telah berhasil dilakukan tanpa ditemukan kesalahan fungsional (functional errors).

4. KESIMPULAN

Hasil penelitian menunjukkan bahwa penggunaan arsitektur Vision Transformer versi ViT Base Patch 16 dalam klasifikasi jenis kulit mampu memberikan performa yang cukup baik dengan akurasi sebesar 78% pada data uji. Performa terbaik diperoleh pada kelas kulit kombinasi dengan nilai precision 0.90 dan F1-score 0.86 yang menunjukkan akurasi dan keseimbangan klasifikasi yang tinggi. Kelas kulit kering juga menunjukkan performa yang konsisten dengan precision 0.82 dan recall 0.80. Sementara itu, kelas kulit normal dan berminyak memiliki performa yang lebih rendah khususnya pada precision 0.71, meskipun recall pada kelas berminyak cukup tinggi yaitu 0.77. Hal ini menunjukkan bahwa model masih mengalami kesulitan dalam membedakan ciri visual antar jenis kulit tertentu terutama antara kulit normal dan berminyak.

Sistem klasifikasi berhasil diintegrasikan ke dalam platform website berbasis Flask yang memungkinkan pengguna melakukan klasifikasi kulit melalui dua metode input yaitu unggah gambar dan pemindaian kamera secara langsung. Pengujian API dilakukan menggunakan Postman dengan metode HTTP POST dan hasilnya menunjukkan bahwa server mampu memberikan respons klasifikasi yang tepat dalam format JSON. Dari sisi pengujian sistem, metode functional testing menunjukkan bahwa semua fitur utama berjalan sesuai fungsinya mulai dari navigasi halaman, pengolahan input gambar hingga penampilan hasil klasifikasi disertai persentase probabilitas tiap kelas.

DAFTAR PUSTAKAS

- [1] O. Sumant, "Dermatologicals Market Expected to Reach \$59.309 Billion by 2030," Allied Market Research. Accessed: Jan. 23, 2025. [Online]. Available: https://www.alliedmarketresearch.com/press-release/dermatological-drugs-market.html
- [2] K. Hidayat, "Industri Kosmetik Diprediksi Tetap Tumbuh di Tahun 2025," Kontan.co.id. Accessed: Jan. 23, 2025. [Online]. Available: https://industri.kontan.co.id/news/industri-kosmetik-diprediksi-tetap-tumbuh-di-tahun-2025
- [3] S. D. Kusumaningrum, "Kajian Pustaka Dalam Penentuan Tipe Dan Permasalahan Kulit Wajah," *Jurnal SNATi*, vol. 1, no. 1, p. 17, 2021, [Online]. Available: https://www.google.com/
- [4] R. Fadil, "Ini yang Dimaksud dengan Breakout Wajah, Penyebab dan Cara Mengobatinya," Halodoc. Accessed: Jan. 28, 2025. [Online]. Available: https://www.halodoc.com/artikel/ini-yang-dimaksud-dengan-breakout-wajah-penyebab-dan-cara-mengobatinya?srsltid=AfmBOoqliIH8dsDRY5VyPv4WRJDExWq9eXOwJCPXPFOhjFI00gwN_ydZ
- [5] M. A. Arshed, S. Mumtaz, M. Ibrahim, S. Ahmed, M. Tahir, and M. Shafi, "Multi-Class Skin Cancer Classification Using Vision Transformer Networks and Convolutional Neural Network-Based Pre-Trained Models," *Information (Switzerland)*, vol. 14, no. 7, Jul. 2023, doi: 10.3390/info14070415.
- [6] C. Flosdorf, J. Engelker, I. Keller, and N. Mohr, "Skin Cancer Detection utilizing Deep Learning: Classification of Skin Lesion Images using a Vision Transformer," 2024. [Online]. Available: www.aaai.org
- [7] G. Boesch, "Vision Transformers (ViT) in Image Recognition 2024 Guide," viso.ai. Accessed: Jan. 28, 2025. [Online]. Available: https://viso.ai/deep-learning/vision-transformer-vit/
- [8] D. A. Purboningtyas, "Klasifikasi Jenis Kulit Wajah Menggunakan Metode Vision Transformers," Universitas Islam Sultan Agung, Semarang, 2024.
- [9] Dian Anisa Agustina, "Klasifikasi Citra Jenis Kulit Wajah Dengan Algoritma Convolutional Neural Network (Cnn) Resnet-50," *Jurnal Riset Sistem Informasi*, vol. 1, no. 3, pp. 01–07, Jul. 2024, doi: 10.69714/13sbby24.
- [10] M. Ath-Thariq and T. N. Suharsono, "Deteksi Penyakit Kulit Serupa Pada Wajah Berbasis Mobile dengan Metode Convolutional Neural Network," *INNOVATIVE: Journal Of Social Science Research*, vol. 3, no. 5, pp. 876–887, Oct. 2023, Accessed: May 28, 2025. [Online]. Available: https://j-innovative.org/index.php/Innovative/article/view/4936
- [11] A. Munib and F. Wulandari, "Studi Literatur: Efektivitas Model Kooperatif Tipe Course Review Horay Dalam Pembelajaran Ipa Di Sekolah Dasar," *Jurnal Pendidikan Dasar Nusantara*, vol. 7, no. 1, pp. 160–172, 2021.
- [12] U. Sulung and M. Muspawi, "Memahami Sumber Data Penelitian: Primer, Sekunder, Dan Tersier," *EDU RESEARCH*, vol. 5, no. 3, pp. 110–116, Sep. 2024.
- [13] Ridwan and N. F. Tungkas, Metode Penelitian, 1st ed. Yogyakarta: Yayasan Sahabat Alam Rafflesia, 2024.
- [14] R. O. Waruwu, K. S. Zai, M. M. Bate'e, and J. B. I. J. Gea, "Pengoperasian Sistem Aplikasi E-Arsip Dalam Memaksimalkan Manajemen Operasi Pelayanan Surat Masuk Dan Surat Keluar Berbasis Digital Di Dinas Komunikasi Dan Informatika Kabupaten Nias Utara," *Jurnal EMBA*, vol. 12, no. 1, pp. 1044–1051, Jan. 2024.
- [15] S. Dissanayake, "Oily, Dry and Normal Skin Types Dataset," 2024. Accessed: Jun. 30, 2025. [Online]. Available: https://www.kaggle.com/datasets/shakyadissanayake/oily-dry-and-normal-skin-types-dataset
- [16] Skin Type, "Skin_Type Computer Vision Project," 2025. Accessed: Jun. 30, 2025. [Online]. Available: https://universe.roboflow.com/skin-type-jtczu/skin_type-vtn2q
- [17] D. M. Pratama, "Implementasi Metode Vision Transformer (Vit) Dalam Klasifikasi Citra Mri Penyakit Alzheimer," UNIVERSITAS LAMPUNG, Lampung, 2024.

Jurnal Computer Science and Information Technology (CoSciTech) Vol. 6, No. 2, Agustus 2025, hal. 214-229

- [18] "Source Code for Torchvision," PyTorch Foundation. Accessed: Jul. 01, 2025. [Online]. Available: https://docs.pytorch.org/vision/main/generated/torchvision.transforms.ToTensor.html
- [19] "Introduction to ViT (Vision Transformers): Everything You Need to Know," Lightly. Accessed: Jul. 01, 2025. [Online]. Available: https://www.lightly.ai/blog/vision-transformers-vit?utm_source=chatgpt.com
- [20] M. D. Junas, "Pemodelan Convolutional Vision Transformer Pada Pemrosesan Gambar Spektogram Untuk Deteksi North Atlantic Right Whales Up-Call," Universitas Islam Negeri Syarif Hidayatullah Jakarta, Jakarta, 2023.
- [21] Y. Anis, A. B. Mukti, and A. N. Rosyid, "Penerapan Model Waterfall Dalam Pengembangan Sistem Informasi Aset Destinasi Wisata Berbasis Website," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 2, pp. 1134–1142, Oct. 2023, doi: 10.30865/klik.v4i2.1287.
- [22] M. T. Abdillah *et al.*, "Implementasi Black box Testing dan Usability Testing pada Website Sekolah MI Miftahul Ulum Warugunung Surabaya," *Jurnal Ilmu Komputer dan Desain Komunikasi Visual*, vol. 8, no. 1, pp. 234–242, Jul. 2023.
- [23] W. A. Firmansyach, U. Hayati, and Y. A. Wijaya, "Analisa Terjadinya Overfitting Dan Underfitting Pada Algoritma Naive Bayes Dan Decision Tree Dengan Teknik Cross Validation," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 1, pp. 262–269, Feb. 2023.