

**PENERAPAN DALAM MEMBEDAKAN GAMBAR BUATAN
ARTIFICIAL INTELLIGENCE DAN GAMBAR ASLI
DENGAN EKSTRAKSI FITUR VISUAL
MENGUNAKAN ALGORITMA
ENSEMBLE LEARNING**

SKRIPSI

DISUSUN OLEH

MOH HAFIZ NAUFAL

2209020194



UMSU

Unggul | Cerdas | Terpercaya

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

2026

**PENERAPAN DALAM MEMBEDAKAN GAMBAR BUATAN
ARTIFICIAL INTELLIGENCE DAN GAMBAR ASLI
DENGAN EKSTRAKSI FITUR VISUAL
MENGUNAKAN ALGORITMA
ENSEMBLE LEARNING**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana
Komputer (S.Kom) Dalam Program Studi Teknologi Informasi pada
Fakultas Ilmu Komputer dan Teknologi Informasi , Univeristas
Muhammadiyah Sumatera Utara**

**MOH HAFIZ NAUFAL
NPM.2209020194**

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

2026

LEMBAR PENGESAHAN

Judul Skripsi : Penerapan dalam membedakan gambar buatan *artificial intelligence* dan gambar asli dengan ekstraksi fitur visual menggunakan algoritma *ensemble learning*

Nama Mahasiswa : Moh Hafiz Naufal

NPM : 2209020194

Program Studi : Teknologi Informasi

Menyetujui
Komisi Pembimbing



(Dr. Al-Khowarizmi, M.kom.)

NIDN.0127099201

Ketua Program Studi



(Fatma Sari Hutagalung S.Kom.,
M.Kom.)

NIDN. 0117019301

Dekan



(Dr. Al-Khowarizmi, M.kom.)

NIDN. 0127099201

PERNYATAAN ORISINALITAS

PENERAPAN DALAM MEMBEDAKAN GAMBAR BUATAN ARTIFICIAL INTELLIGENCE DAN GAMBAR ASLI DENGAN EKSTRAKSI FITUR VISUAL MENGUNAKAN ALGORITMA ENSEMBLE LEARNING

SKRIPSI

Dengan ini penulis menyatakan bahwa seluruh isi karya tulis ini merupakan hasil pemikiran dan pekerjaan penulis sendiri, kecuali pada bagian tertentu yang mengacu pada sumber lain dan telah di cantumkan secara jelas.

Medan, 18 April 2026

Yang membuat pernyataan



Moh Hafiz Naufal
NPM.2209020194

**PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN
AKADEMIS**

Sebagai bagian dari sivitas akademika Universitas Muhammadiyah Sumatera

Utara, saya yang bertanda tangan dibawah ini:

Nama : Moh Hafiz Naufal

NPM : 2209020194

Program Studi : Teknologi Informasi

Karya Ilmiah : Skripsi

Dalam rangka mendukung pengembangan ilmu penguatan, saya memberikan persetujuan kepada Universitas Muhammadiyah Sumatera Utara untuk memanfaatkan karya ilmiah ini dalam bentuk hak bebas Royalti Non-Eksklusif (Non-Eksekutif Royalty-Free Right) atas penelitian saya.

Dengan hak tersebut, pihak Universitas berwenang untuk menyimpan, mengelola mengubah format, serta mendistribusikan karya ini melalui berbagai media, termasuk dalam bentuk database, tanpa perlu meminta izin kembali nama penulis sebagai pemilik karya.

Pernyataan ini dibuat dengan sebenar-benarnya untuk dapat digunakan sebagaimana mestinya.

Medan, 18 April 2026

Yang membuat pernyataan



Moh Hafiz Naufal
NPM. 2209020194

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Moh Hafiz Naufal
Tempat dan Tanggal Lahir : Ujungbatu, 21 Desember 2003
Alamat Rumah : Jl. Lingkar Pematang Tebih
Telepon/HP : 0822 6456 8606
E-mail : mh3604211@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD : SDIT INAYAH UJUNGBATU
SMP : SMPIT INAYAH UJUNGBATU
SMK : SMK ISLAM INAYAH UJUNGBATU

KATA PENGANTAR

Assalamualaikum Warahmatullahi Wabarakatuh

Puji dan Syukur penulis panjatkan ke hadirat Allah SWT atas segala Rahmat, karunia, serta nikmat Kesehatan dan kesempatan yang telah diberikan, sehingga penulis dapat menyelesaikan skripsi ini dengan judul “Penerapan dalam Membedakan Gambar Buatan *Artificial Intelligence* dan Gambar Asli dengan Ekstraksi Fitur Visual Menggunakan Algoritma *Ensemble Learning*”.

Penyusunan skripsi ini dilakukan sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi Teknologi Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara.

Penulis menyadari bahwa proses penyusunan skripsi ini tidak terlepas dari dukungan, bantuan, serta bimbingan dari berbagai pihak. Oleh karena itu penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Agussani, M.AP., Rektor Universitas Muhammadiyah Sumatera Utara (UMSU).
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom. Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU, sekaligus Dosen Pembimbing penulis.
3. Ibu Fatma Sari Hutagalung, S.Kom., M.Kom., selaku Ketua Program Studi Teknologi Informasi sekaligus Dosen Pembimbing Akademik penulis.
4. Bapak Okvi Nugroho, S.Kom., M.Kom selaku Sekretaris Program Studi Teknologi Informasi

5. Seluruh dosen dan staf akademik di lingkungan Fakultas Ilmu Komputer dan Teknologi Informasi, khususnya Program Studi Teknologi Informasi, yang telah memberikan ilmu dan pengalaman berharga selama masa studi.
6. Kedua orang tua tercinta, bapak Bambang Edison dan Ibu Masruroh, yang selalu memberikan doa, semangat, cinta, serta dukungan moral dan material tanpa henti.
7. Semua pihak yang tidak dapat disebutkan satu per satu yang telah membantu dan memberikan dukungan secara langsung maupun tidak langsung dalam penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki berbagai kekurangan dan belum sepenuhnya sempurna. Oleh sebab itu, penulis sangat mengharapkan kritik dan saran yang bersifat membangun guna penyempurnaan di masa yang akan datang. Penulis berharap skripsi ini dapat memberikan manfaat bagi para pembaca serta dapat dijadikan sebagai referensi, khususnya dalam penelitian yang berkaitan dengan klasifikasi atau pembedaan antara gambar hasil *artificial intelligence* dan gambar asli dengan memanfaatkan algoritma *ensemble learning*.

Medan, 18 April 2026
Penulis

Moh Hafiz Naufal

PENERAPAN DALAM MEMBEDAKAN GAMBAR BUATAN ARTIFICIAL INTELLIGENCE DAN GAMBAR ASLI DENGAN EKSTRAKSI FITUR VISUAL MENGGUNAKAN ALGORITMA ENSEMBLE LEARNING

ABSTRAK

Perkembangan teknologi Artificial Intelligence (AI) generatif memungkinkan sistem komputer menghasilkan gambar yang sangat realistis dan menyerupai foto asli. Kondisi ini menimbulkan tantangan dalam membedakan gambar buatan AI dan gambar asli secara visual. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi citra yang mampu membedakan gambar buatan AI dan gambar asli dengan memanfaatkan ekstraksi fitur visual serta algoritma ensemble learning. Metode penelitian yang digunakan meliputi tahap preprocessing citra dengan penyeragaman ukuran menjadi 256×256 piksel, ekstraksi fitur visual yang terdiri dari histogram warna RGB, distribusi intensitas grayscale, fitur tekstur menggunakan Gray Level Co-occurrence Matrix (GLCM), serta fitur tepi menggunakan metode Canny Edge Detection. Fitur yang diperoleh kemudian digunakan sebagai input pada beberapa algoritma klasifikasi yaitu Support Vector Machine (SVM), K-Nearest Neighbor (KNN), dan Random Forest. Selanjutnya dilakukan penggabungan model menggunakan metode ensemble learning dengan teknik hard voting. Hasil pengujian menunjukkan bahwa model Random Forest memperoleh nilai akurasi sebesar 65,71%, sedangkan metode ensemble learning menghasilkan nilai akurasi sebesar 65,00% dengan nilai F1-score sebesar 0,6918. Sistem yang dikembangkan juga diimplementasikan dalam bentuk aplikasi berbasis web menggunakan framework Streamlit sehingga pengguna dapat mengunggah gambar dan memperoleh hasil prediksi secara langsung. Penelitian ini menunjukkan bahwa pendekatan ekstraksi fitur visual dan ensemble learning dapat digunakan sebagai metode untuk membantu mengidentifikasi gambar buatan AI dan gambar asli.

Kata Kunci: Artificial Intelligence, ekstraksi fitur visual, ensemble learning.

APPLICATION IN DISTINGUISHING ARTIFICIAL INTELLIGENCE- MAKED IMAGES AND ORIGINAL IMAGES WITH VISUAL FEATURE EXTRACTION USING ENSEMBLE LEARNING ALGORITHM

ABSTRACT

The development of generative Artificial Intelligence (AI) technology enables computer systems to produce highly realistic images that closely resemble real photographs. This condition creates challenges in distinguishing AI-generated images from real images visually. This research aims to develop an image classification system capable of distinguishing AI-generated images and real images using visual feature extraction and ensemble learning algorithms.

The research method consists of several stages including image preprocessing by resizing images to 256×256 pixels, visual feature extraction including RGB color histogram, grayscale intensity distribution, texture features using Gray Level Co-occurrence Matrix (GLCM), and edge features using the Canny Edge Detection method. The extracted features are then used as input for several classification algorithms such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest. Furthermore, model combination is performed using an ensemble learning method with a hard voting technique. The experimental results show that the Random Forest model achieved an accuracy of 65.71%, while the ensemble learning method achieved an accuracy of 65.00% with an F1-score of 0.6918. The developed system is also implemented as a web-based application using the Streamlit framework, allowing users to upload images and obtain prediction results directly. The results indicate that the combination of visual feature extraction and ensemble learning can be used as an approach to help identify AI-generated images and real images.

Keywords: Artificial Intelligence, visual feature extraction, ensemble learning.

DAFTAR ISI

LEMBAR PENGESAHAN	Error! Bookmark not defined.
PERNYATAAN ORISINALITAS.....	Error! Bookmark not defined.
PERNYATAAN PERSETUJUAN PUBLIKASI ...	Error! Bookmark not defined.
RIWAYAT HIDUP	v
KATA PENGANTAR.....	vi
ABSTRAK	viii
ABTRACT	ix
DAFTAR ISI	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian	4
BAB II LANDASAN TEORI	5
2.1 Artificial Intelligence (AI).....	5
2.2 AI Generatif (Generatif AI).....	5
2.3 Citra Digital.....	5
2.4 Pengolahan Citra Digital.....	6
2.5 Ekstraksi Fitur Visual.....	6
2.6 Machine Learning (Pembelajaran Mesin).....	7
2.7 klasifikasi Citra	7
2.8 Ensemble Learning.....	7
2.9 Tools dan Implementasi	8
2.9.1 Bahasa Pemograman Python.....	8
2.9.2 OpenCV.....	8
2.9.3 Numpy.....	8
2.9.4 Scikit-Learn.....	9
2.9.5 Visual Studio Code.....	9
2.9.6 Pandas	9
2.10 Penelitian Terkait.....	10
2.11 Analisis Gap	12
BAB III METODOLOGI PENELITIAN.....	14
3.1 Pendekatan Penelitian	14

3.2 Sumber Data.....	15
3.2.1 Rentang dan Format Data.....	16
3.2.2 Alat Pengambilan Data.....	17
3.2.3 Karakteristik Data	17
3.3 Teknik Pengumpulan Data	18
3.3.1 Tahapan Pengumpulan Data.....	18
3.3.2 Penjelasan Pengumpulan Data	18
3.4 Teknik Analisis Data	19
3.4.1 <i>Preprocessing</i> Data Citra	20
3.4.2 Ekstraksi Fitur Visual	20
3.4.3 Pembentukan Dataset Fitur	21
3.4.4 Pelatihan Model Menggunakan <i>Ensemble Learning</i>	21
3.4.5 Pengujian dan Evaluasi Hasil.....	22
3.5 Prosedur Pengujian dan Validasi.....	26
3.5.1 Tahapan Pengujian Sistem	26
3.5.2 Validasi Hasil Klasifikasi	26
3.5.3 Kriteria Keberhasilan Sistem	27
BAB IV HASIL DAN PEMBAHASAN.....	31
4.1 Deskripsi Data.....	31
4.1.1 Sumber Data Gambar AI.....	32
4.1.2 Sumber Data Gambar Asli	32
4.2 Hasil <i>Preprocessing</i> Data	33
4.3 Hasil Ekstraksi Fitur Visual	34
4.3.1 Ekstraksi Fitur Warna.....	35
4.3.2 Fitur Tekstur	36
4.3.3 Fitur Tepi.....	37
4.3.4 Distribusi Intesnsitas Piksel	38
4.4 Hasil Pelatihan Model.....	39
4.4.1 Pelatihan Model Support Vector Machine (SVM)	40
4.4.2 Pelatihan Model K-Nearest Neighbor (KNN).....	41
4.4.3 Pelatihan Random Forest	42
4.4.4 Pelatihan <i>Ensemble Learning (Hard Voting)</i>	43
4.5 Evaluasi Model Menggunakan Confusion Matrix	44
4.6 Analisis Peforma Model.....	45
4.7 Implementasi Sistem	47
4.6.1 Tampilan Halaman Utama Sistem.....	48

4.6.2 Tampilan Proses Analisis Citra.....	48
4.6.3 Tampilan Hasil Prediksi Sistem	49
BAB V PENUTUP.....	51
5.1 Kesimpulan	51
5.2 Saran.....	52
DAFTAR PUSTAKA	53

DAFTAR TABEL

Tabel 2. 1 Ringkasan Penelitian Terdahulu	10
Tabel 3. 1 Tabel gambar buatan AI dan gambar asli	15
Tabel 3. 2 simbol ensemble learning.....	22
Tabel 4. 1 Perbandingan Peforma Model Klasifikasi.....	47

DAFTAR GAMBAR

Gambar 4. 1 Contoh pengambilan data gambar AI di platform Grok.....	32
Gambar 4. 2 contoh pengambilan data gambar asli	33
Gambar 4. 3 Hasil Penyeragaman ukuran citra.....	34
Gambar 4. 4 Hasil Ekstraksi Fitur Visual.....	35
Gambar 4. 5 Histogram distribusi warna RGB	36
Gambar 4. 6 Hasil ekstraksi fitur tekstur mengunakan GLCM.....	37
Gambar 4. 7 Hasil deteksi tepi menggunakan metode canny	38
Gambar 4. 8 Hasil Histogram Distribusi Intensitas Grayscale	39
Gambar 4. 9 Hasil Evaluasi SVM.....	40
Gambar 4. 10 Hasil Evaluasi Model KNN.....	41
Gambar 4. 11 Hasil evaluasi Random Forest	42
Gambar 4. 12 Hasil evaluasi ensemble learning	43
Gambar 4. 13 Hasil Confusion Matrix Model Ensemble Learning	45
Gambar 4. 14 Tampilan Halaman Utama Sistem.....	48
Gambar 4. 15 Tampilan Proses Analisis Citra.....	49
Gambar 4. 16 Hasil Prediksi Sistem	50

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan teknologi *Artificial Intelligence* (AI) mengalami kemajuan yang sangat pesat, khususnya pada bidang pengolahan citra digital. Model AI generatif seperti *Generative Adversarial Network* (GAN) dan model AI generatif lainnya mampu menghasilkan gambar dengan kualitas visual yang sangat realistis dan menyerupai gambar asli tangkapan kamera. Kemampuan ini banyak dimanfaatkan berbagai bidang, seperti industri kreatif, desain grafis, periklanan, dan media digital. (A. Alajaji, et al., 2024)

Namun, berkembangnya teknologi *Artificial Intelligence*, gambar buatan AI semakin sulit dibedakan dari gambar asli secara visual. Kondisi ini berpotensi menimbulkan kesalahpahaman di masyarakat, khususnya bagi pengguna awam seperti orang tua yang cenderung mempercayai konten visual digital tanpa verifikasi. Sebagian besar sistem deteksi yang tersedia saat ini menggunakan pendekatan berbasis deep learning yang kompleks dan bersifat black box, sehingga proses pengambilan keputusan model sulit dipahami. Oleh karena itu, diperlukan pendekatan ilmiah yang tidak hanya menghasilkan klasifikasi, tetapi juga mampu menjelaskan karakteristik visual yang menjadi dasar pembeda antara gambar buatan AI dan gambar asli secara lebih transparan.

Seiring dengan berkembangnya permasalahan tersebut, berbagai penelitian telah dilakukan untuk membedakan gambar buatan *Artificial Intelligence* dan gambar asli menggunakan Teknik pengolahan citra dan pembelajaran mesin. Salah satu pendekatan yang umum digunakan adalah ekstraksi visual, yaitu proses

pengambilan karakteristik penting dari citra seperti tekstur, warna, pola tepi, dan distribusi piksel. Fitur-fitur visual tersebut terbukti mampu merepresentasikan perbedaan karakteristik antara gambar buatan AI dan gambar asli, meskipun secara visual sulit dibedakan oleh mata manusia. (Shofia & Aulia, 2023).

Meskipun demikian, penggunaan satu algoritma klasifikasi saja sering kali belum memberikan hasil yang optimal, terutama ketika dihadapkan pada variasi data yang semakin kompleks. Oleh karena itu, diperlukan metode yang mampu meningkatkan akurasi dan stabilitas klasifikasi. Algoritma *ensemble learning* merupakan pendekatan yang mengombinasikan beberapa model pembelajaran untuk menghasilkan Keputusan yang lebih akurat dibandingkan model Tunggal (Alotaibi, 2025).

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan penerapan sistem untuk membedakan gambar buatan *Artificial Intelligence* dan gambar asli dengan ekstraksi fitur visual menggunakan algoritma *Ensemble Learning*. Penelitian ini diharapkan dapat membantu proses identifikasi keaslian gambar secara otomatis serta menjadi salah satu solusi dalam meminimalkan penyebaran hoaks berbasis gambar dan meningkatkan literasi digital di Masyarakat.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini Adalah bagaimana menerapkan sistem untuk membedakan gambar buatan *artificial intelligence* dan gambar asli dengan memanfaatkan ekstraksi visual. Selain itu penelitian ini merumuskan bagaimana penerapan algoritma *Ensemble Learning* dapat meningkatkan ketepatan dan kestabilan klasifikasi dalam membedakan gambar buatan AI dan gambar asli.

1.3 Batasan Masalah

Agar penelitian ini lebih terarah dan tidak meluas, maka Batasan masalah dalam penelitian ini Adalah sebagai berikut:

1. Penelitian ini hanya berfokus pada proses klasifikasi untuk membedakan gambar buatan *artificial intelligence* dan gambar asli.
2. Data yang digunakan berupa dataset gambar yang dikumpulkan dari sumber yang relevan (gambar AI dan gambar asli)
3. Metode yang digunakan untuk membedakan gambar Adalah ekstraksi visual.
4. Algoritma klasifikasi yang digunakan Adalah *ensemble learning* sebagai metode utama dalam proses pengambilan Keputusan.
5. Penelitian ini tidak mencakup pendeteksi gambar buatan AI yg bersifat ilustratif atau seni digital, seperti anime art, digital painting dan gambar kartun, melainkan hanya berfokuskan pada gambar yang menyurapai foto nyata.
6. Sistem dibangun hanya bertujuan sebagai prototype /penelitian akademik, Bukan sistem komersial atau sistem deteksi tingkat industri.

1.4 Tujuan Penelitian

Tujuan penelitian ini Adalah sebagai berikut:

1. Menerapkan sistem untuk membedakan gambar buatan *artificial intelligence* (AI) dan gambar asli menggunakan ekstraksi fitur visual.
2. Menentukan dan menganalisis fitur visual yang dapat digunakan untuk mempresentasikan karakteristik gambar sebagai dasar proses klasifikasi .

3. Menerapkan algoritma *ensemble learning* dalam proses klasifikasi untuk meningkatkan ketepatan hasil dalam membedakan gambar buatan AI dan gambar asli.
4. Menghasilkan prototype sistem deteksi yang dapat digunakan sebagai referensi pengembangan penelitian selanjutnya terkait identifikasi gambar buatan AI.

1.5 Manfaat Penelitian

Manfaat penelitian ini diharapkan dapat memberikan kontribusi sebagai berikut:

1. Mendukung upaya peningkatan literasi digital Masyarakat, terutama dalam mengurangi penyebaran informasi palsu (hoaks) berbasis gambar melalui penggunaan sistem deteksi keaslian gambar.
2. Membantu meningkatkan pemahaman mengenai karakteristik visual yang membedakan gambar buatan AI dan gambar asli.
3. Memberikan kontribusi akademik dalam bidang pengolahan citra digital dan pembelajaran mesin, khususnya terkait klasifikasi untuk membedakan gambar buatan AI dan gambar asli.
4. Penelitian ini diharapkan dapat memberikan pemahaman yang lebih transparan mengenai proses klasifikasi gambar buatan AI melalui analisis fitur visual, sehingga hasil penelitian dapat dijadikan referensi akademik dan edukatif bagi masyarakat.

BAB II

LANDASAN TEORI

2.1 Artificial Intelligence (AI)

Artificial intelligence merupakan cabang ilmu computer yang berfokus pada pengembangan sistem super yang mampu meniru kemampuan manusia seperti belajar, memahami pola dan pengambil Keputusan. AI banyak diterapkan dalam berbagai bidang, salah satunya pada pengolahan citra digital. Dalam konteks penelitian ini, AI menjadi perhatian utama karena perkembangannya teknologi mampu menciptakan gambar secara otomatis sehingga menyerupai foto nyata. (Citra Kurniawan, 2015)

2.2 AI Generatif (Generatif AI)

AI Generative Adalah teknologi AI yang mampu menghasilkan konten baru, seperti teks, gambar, audio, dan video berdasarkan data yang telah dipelajari sebelumnya. Pada bidang citra, AI generative mampu membuat gambar realistis menggunakan model tertentu, seperti *Generative Adversarial Network (GAN)*. Hasil gambar AI generative seringkali tampak nyata dan sulit dibedakan dari gambar asli, sehingga menimbulkan tantangan dalam autentifikasi visual. (Anggarwal, 2021)

2.3 Citra Digital

Citra digital adalah representasi suatu objek dalam bentuk matriks piksel yang tersusun dari nilai intensitas tertentu. Citra digital umumnya memiliki format warna seperti grayscale atau RGB (Red, Green, Blue). Setiap citra mengandung informasi visual yang dapat dianalisis lebih lanjut untuk tujuan pengenalan pola, klasifikasi, maupun deteksi keaslian. Pada penelitian ini, citra digital menjadi

objek utama yang akan dianalisis untuk membedakan citra buatan AI dan citra asli. (Melo Milarez & Pontes , 2014)

2.4 Pengolahan Citra Digital

Pengolahan citra digital merupakan proses pengolahan data berupa gambar dengan tujuan meningkatkan kualitas gambar atau mengambil informasi penting dari gambar tersebut. Tahapan pengolahan citra dapat meliputi preprocessing seperti resizing, normalisasi, noise reduction, serta proses ekstraksi fitur. Pengolahan citra pada penelitian ini digunakan sebagai dasar sebelum dilakukan proses klasifikasi antara gambar buatan AI dan gambar asli. (Duhan & Bhatia, 2026)

2.5 Ekstraksi Fitur Visual

Ekstraksi fitur visual adalah tahapan untuk mengambil karakteristik penting dari citra yang dapat digunakan sebagai representasi data dalam proses klasifikasi. Fitur visual yang umum digunakan dalam pengolahan citra antara lain:

1. Tekstur, yaitu pola permukaan gambar yang menunjukkan detail atau keragaman intensitas.
2. Warna, yaitu informasi spektrum warna pada gambar yang dapat digunakan untuk membedakan karakteristik tertentu.
3. Pola tepi (edge), yaitu perubahan kontras intensitas yang menunjukkan batas-batas objek.

Ekstraksi fitur visual penting karena meskipun gambar AI terlihat menyerupai gambar asli, keduanya dapat memiliki pola statistik atau karakteristik visual tertentu yang berbeda. (Hallur & Gavade, 2025)

2.6 Machine Learning (Pembelajaran Mesin)

Machine learning merupakan metode yang memungkinkan komputer belajar dari data tanpa diprogram secara eksplisit. Dalam *machine learning*, model mempelajari pola dari data latih dan digunakan untuk memprediksi atau mengklasifikasi data baru. Pada penelitian ini, *machine learning* digunakan untuk mengklasifikasikan apakah suatu gambar termasuk kategori buatan AI atau gambar asli berdasarkan fitur yang diambil. (Baruggada & Kalluri, 2025)

2.7 klasifikasi Citra

Klasifikasi citra adalah proses pengelompokan citra ke dalam kelas tertentu berdasarkan ciri atau fitur yang dimiliki. Dalam penelitian ini, klasifikasi citra bertujuan untuk mengelompokkan citra ke dalam dua kelas utama, yaitu gambar AI dan gambar asli, Klasifikasi dilakukan setelah citra melalui proses ekstraksi fitur visual sehingga model dapat mengenali perbedaan karakteristik antar kedua jenis gambar. (Maulana & Rochmawati, 2019)

2.8 Ensemble Learning

Ensemble Learning merupakan pendekatan pembelajaran mesin yang menggabungkan beberapa model untuk meningkatkan performa prediksi. Metode ini bekerja dengan cara menyatukan beberapa model klasifikasi sehingga hasil keputusan lebih stabil dan akurat dibandingkan hanya menggunakan satu model. *Ensemble Learning* digunakan karena mampu mengurangi kesalahan prediksi serta meningkatkan ketahanan (robustness) terhadap variasi data yang kompleks. (Indahyanti, Azizah, & Setiawan, 2022)

2.9 Tools dan Implementasi

Pada penelitian ini, proses implementasi sistem dilakukan menggunakan perangkat lunak dan pustaka pemrograman yang mendukung pengolahan citra digital serta pembelajaran mesin. Pemilihan tools dan pustaka dilakukan agar proses pengolahan dataset, ekstraksi fitur visual, pelatihan model klasifikasi, dan evaluasi hasil dapat dilakukan secara terstruktur dan terprogram.

2.9.1 Bahasa Pemrograman Python

Bahasa pemrograman yang digunakan dalam penelitian ini adalah Python. Python dipilih karena memiliki sintaks yang sederhana serta didukung oleh banyak pustaka yang umum digunakan untuk pengolahan citra dan *machine learning*. Selain itu, Python juga sering digunakan dalam penelitian akademik karena fleksibel untuk proses eksperimen dan pengembangan sistem. (Fondation, 2026)

2.9.2 OpenCV

OpenCV (Open Source Computer Vision Library) digunakan pada penelitian ini untuk membantu proses pengolahan citra digital, khususnya pada tahap preprocessing. OpenCV digunakan untuk membaca citra dari dataset, mengubah ukuran citra (*resize*), konversi ruang warna (misalnya RGB ke grayscale), serta operasi pengolahan citra lainnya yang dibutuhkan sebelum ekstraksi fitur dilakukan. (Team, 2025/2026)

2.9.3 Numpy

NumPy digunakan untuk melakukan pengolahan data numerik dalam bentuk array atau matriks. Dalam konteks penelitian ini, citra digital direpresentasikan sebagai matriks piksel sehingga NumPy berperan penting dalam

proses manipulasi data citra, perhitungan fitur, serta penyusunan vektor fitur yang akan digunakan dalam proses klasifikasi. (NumPy, 2024)

2.9.4 Scikit-Learn

scikit-learn digunakan untuk membangun model klasifikasi dan proses pembelajaran mesin, termasuk implementasi algoritma *ensemble learning*. Pada penelitian ini, scikit-learn dimanfaatkan untuk membentuk model Voting Classifier yang menggabungkan beberapa algoritma klasifikasi untuk menghasilkan prediksi akhir. Selain itu, pustaka ini juga menyediakan fitur evaluasi seperti confusion matrix dan penghitungan akurasi, precision, recall, serta F1-score. (Developers, 2024)

2.9.5 Visual Studio Code

Visual Studio Code digunakan sebagai lingkungan pengembangan dalam penelitian ini untuk menjalankan program Python secara terstruktur. VS Code memudahkan proses implementasi sistem, pengelolaan folder dataset, penulisan kode, serta proses debugging. Selain itu, penggunaan VS Code mendukung integrasi dengan pustaka Python yang digunakan dalam pengolahan citra dan pembelajaran mesin. (Corporation, 2024)

2.9.6 Pandas

pandas digunakan untuk mengelola data hasil ekstraksi fitur dalam bentuk tabel (dataframe). Pada penelitian ini pandas digunakan untuk membantu proses penyusunan dataset fitur, pengelompokan label kelas, serta penyimpanan data fitur ke dalam format CSV sehingga memudahkan proses pelatihan model dan analisis hasil. (team, 2024)

2.10 Penelitian Terkait

Penelitian terkait diperlukan untuk mengetahui perkembangan penelitian sebelumnya dan menentukan posisi penelitian yang dilakukan. Pada penelitian terkait, dibahas beberapa penelitian yang relevan dengan topik klasifikasi citra, pendeteksian gambar buatan AI, ekstraksi fitur visual, serta penggunaan *ensemble learning* untuk meningkatkan performa klasifikasi.

Tabel 2. 1 Ringkasan Penelitian Terdahulu

No	Judul dan penelitian	Pembahasan	Metode	Kelebihan dan kekurangan
1	Artifact feature purification for cross-domain detection of AI-generated images (Meng at al., 2024)	Menangani masalah penurunan performa deteksi ketika pindah domain/generator AI berbeda, dengan pemurnian fitur artifact	Artifact purification network	(+) kuat cross-domain (-) membutuhkan training lebih rumit
2	Extreme learning machine approach on heart abnormalities identification in ECG images (Nababan, et al., 2024)	Penelitian ini mengklasifikasikan citra ECG untuk mendeteksi kelainan jantung menggunakan teknik pengolahan citra dan machine learning. Fitur diekstraksi menggunakan invariant moments, kemudian diklasifikasikan dengan ELM.	Preprocessing, segmentasi, ekstraksi fitur, klasifikasi: Extreme Machine Learning	(+) Alurasi Tinggi, proses klasifikasi cepat (-) Fokus pada citra ECG
3	A minimalist approach to robust GAN-generated face detection (Ghost at al.,2024)	Mengusulkan pendekatan deteksi gambar wajah GAN yang ringan dengan fitur chrominance space	Lightweight detector (chrominance feutures)	(+) ringan dan efisien, (-) focus pada wajah GAN
4	Dual-path AI-generated image detection:	Pendeteksian gambar AI menggunakan strategi jalur ganda dan fusion fitur	Dual-path feature fusion network	(+) akurat (-) implementasi cukup berat

	Leveraging texture-rich and texture-poor patches with global semantic features (Lin at al., 2025)	untuk meningkatkan generalisasi dan robustness		
5	Lightweight AI-Generated image detection based on enhanced common artifact features(Li at al., 2025)	Mengembangkan model deteksi gambar AI dengan fokus ekstraksi & penguatan artifact features, lebih ringan dari model besar	Gradient-aware detection network	(+) ringan & efisien, (-) tantangan generalisasi
6	Towards universal AI-generated image detection (Lin at al., 2025)	CLIP untuk deteksi gambar AI namun ditingkatkan agar sensitif pada artifact low-level melalui mekanisme attention dan fusion	CLIP-based detection + fusion	(+) memanfaatkan model besar, (-) butuh resource tinggi
7	Harnessing attention for cropping and fusion in CLIP-based AIGC detection (Xu et al., (Xu et al., 2025)	CLIP untuk deteksi gambar AI namun ditingkatkan agar sensitif pada artifact low-level melalui mekanisme attention dan fusion	CLIP-based detection + fusion	(+) memanfaatkan model besar, (-) butuh resource tinggi
8	GCS-Net: A universal AI-generated visual content detection method based on CLIP (Xu at al., 2025)	Deteksi konten visual AI secara universal dengan fine-tuning dan tambahan fitur dari ruang warna	CLIP-based generalized detection	(+) generalisasi kuat, (-) perlu tuning parameter
9	A comprehensive review on ensemble deep learning (Mohammed et al., 2023)	Review komprehensif metode ensemble pada deep learning termasuk faktor-faktor yang mempengaruhi keberhasilan ensemble	Review ensemble deep learning	(+) sangat kuat untuk teori, (-) tidak spesifik gambar AI

10	Ensemble deep learning: Approaches in health care : A review (aziz Alotaibi., 2025)	Review strategi ensemble: bagging, boosting, stacking, decision fusion, dll untuk meningkatkan akurasi dan stabilitas	Review ensemble learning	(+) cocok untuk latar belakang ensemble, (-) tidak spesifik gambar AI
----	-------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------	--------------------------	-----------------------------------------------------------------------

2.11 Analisis Gap

Analisis gap dilakukan untuk mengidentifikasi kesenjangan antara penelitian terdahulu dengan penelitian yang akan dilakukan, sehingga dapat diketahui kontribusi dari penelitian ini. Berdasarkan hasil kajian penelitian terkait pada Tabel 2.1, dapat dilihat bahwa penelitian mengenai deteksi gambar buatan *Artificial Intelligence* telah banyak berkembang, terutama dengan pendekatan berbasis deep learning yang memanfaatkan jejak fitur (deep trace representation), pemurnian fitur artefak (artifact feature purification), serta penggabungan fitur (feature fusion) untuk meningkatkan kemampuan pendeteksian gambar buatan AI (Xu et al., 2024; Meng et al., 2024). Namun, sebagian besar pendekatan tersebut memiliki kompleksitas model yang cukup tinggi serta membutuhkan sumber daya komputasi yang besar.

Selain itu, beberapa penelitian juga telah mengarah pada pengembangan metode pendeteksian yang bersifat universal untuk berbagai jenis gambar buatan AI, baik yang berasal dari GAN maupun diffusion model. Meskipun menunjukkan peningkatan performa, penelitian tersebut masih menghadapi tantangan dalam hal generalisasi model, terutama ketika gambar uji berasal dari generator AI yang berbeda atau mengalami modifikasi tertentu (Lin et al., 2025). Kondisi ini menunjukkan bahwa proses deteksi gambar AI masih membutuhkan metode yang lebih stabil, sederhana, dan efektif.

Di sisi lain, penelitian terkait juga menunjukkan bahwa penggunaan satu algoritma klasifikasi saja sering kali belum cukup optimal untuk menghadapi variasi data yang kompleks. Karena itu, pendekatan *Ensemble Learning* dipandang sebagai salah satu solusi untuk meningkatkan akurasi dan kestabilan klasifikasi melalui penggabungan beberapa model pembelajaran (Ganaie et al., 2022; Mohammed et al., 2023). Akan tetapi, penerapan *ensemble learning* pada kasus pendeteksian gambar buatan AI masih memerlukan kajian lebih lanjut, terutama jika dipadukan dengan proses ekstraksi fitur visual yang sesuai dengan karakteristik citra.

Berdasarkan kesenjangan tersebut, penelitian ini diusulkan untuk mengisi celah penelitian dengan merancang sistem klasifikasi untuk membedakan gambar buatan *Artificial Intelligence* dan gambar asli menggunakan ekstraksi fitur visual serta penerapan metode *Ensemble Learning*. Berbeda dengan pendekatan deep learning yang bersifat kompleks dan cenderung black box, penelitian ini menekankan pendekatan yang lebih sederhana dan transparan, sehingga proses klasifikasi dapat dijelaskan melalui fitur visual yang digunakan. Dengan demikian, penelitian ini tidak hanya berfokus pada peningkatan akurasi, tetapi juga pada aspek interpretabilitas model sebagai kontribusi akademik dalam pengembangan metode deteksi gambar buatan AI.

BAB III

METODOLOGI PENELITIAN

3.1 Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komputasional pada data citra digital. Pendekatan kuantitatif dipilih karena penelitian ini berfokus pada pengukuran performa sistem klasifikasi secara numerik melalui hasil pengujian dan evaluasi model pada dataset citra. Model yang dibangun akan melakukan klasifikasi terhadap dua kelas, yaitu gambar buatan AI dan gambar asli.

Secara kerangka kerja, penelitian ini termasuk dalam kategori supervised learning, di mana sistem dilatih menggunakan dataset berlabel yang telah ditentukan kategorinya. Dataset yang digunakan terdiri dari citra buatan AI sebagai kelas pertama dan citra asli sebagai kelas kedua. Sistem akan mempelajari perbedaan karakteristik visual pada data latih dan kemudian melakukan prediksi pada data uji.

Selain itu, penelitian ini menggunakan pendekatan rancang bangun sistem (research and development) yang menghasilkan prototype sistem deteksi. Prototype tersebut disusun melalui tahapan: pengumpulan data, preprocessing, ekstraksi fitur, pelatihan model, evaluasi performa, dan validasi hasil. Pendekatan ini penting agar penelitian tidak hanya menghasilkan teori, tetapi juga menghasilkan implementasi sistem yang dapat diuji dan dikembangkan.

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersifat kuantitatif dalam bentuk dataset citra digital. Data sekunder merupakan

data yang diperoleh dari sumber yang sudah tersedia dan telah dipublikasikan oleh pihak lain sehingga dapat digunakan kembali untuk kebutuhan penelitian.





Dalam konteks penelitian ini, data berupa citra digunakan sebagai bahan utama untuk membangun sistem klasifikasi. Data dibagi menjadi dua kategori, yaitu gambar buatan AI (AI-generated images) dan gambar asli (real images). Data citra yang digunakan akan diproses secara digital dan dianalisis menggunakan metode ekstraksi fitur visual.

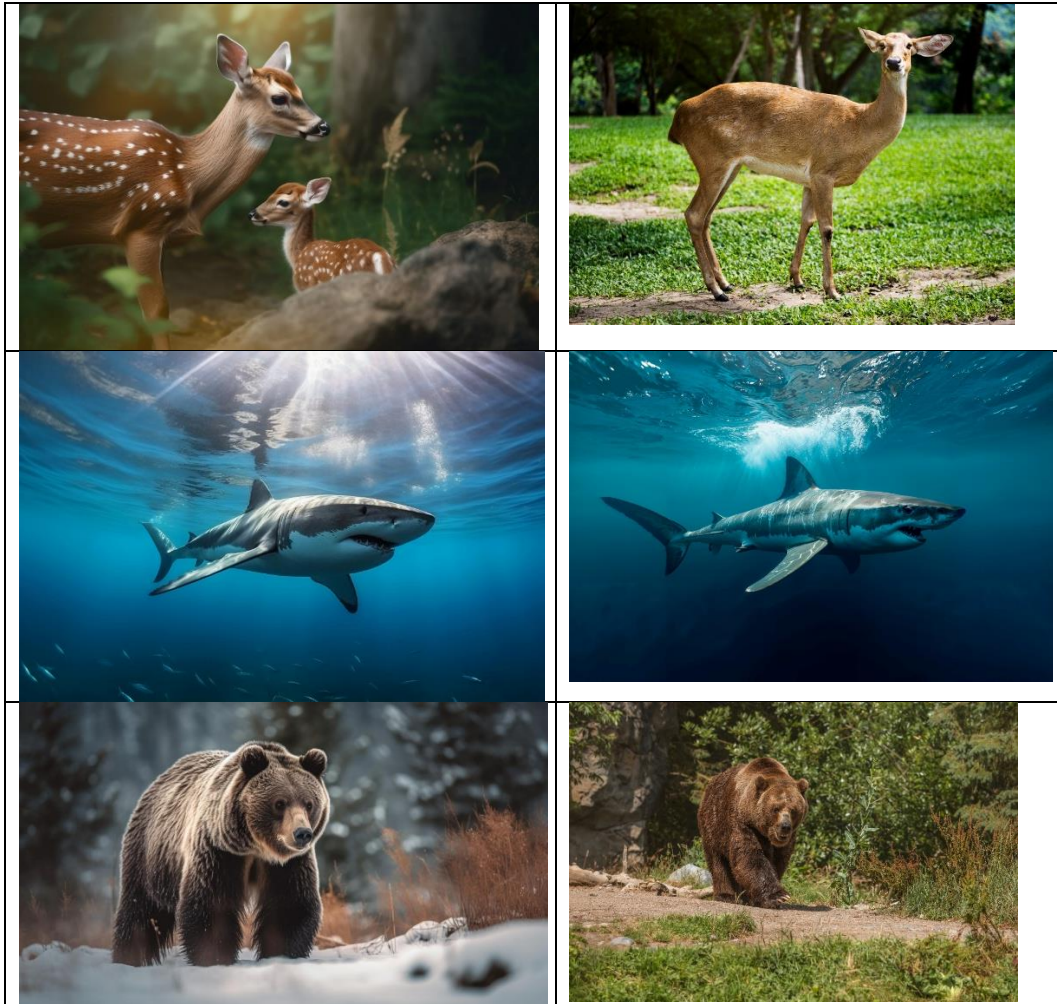
3.2 Sumber Data

Sumber data penelitian ini diperoleh dari data citra yang dikumpulkan secara mandiri dan/atau dari sumber terbuka. Data citra dibagi menjadi dua kategori, yaitu gambar buatan AI dan gambar asli. Seperti contoh berikut:

Tabel 3. 1 Tabel gambar buatan AI dan gambar asli

Sumber : Freepik (<https://www.Freepik.com>), 2025

Gambar Buatan AI	Gambar Asli
	
	



3.2.1 Rentang dan Format Data

Format data yang digunakan dalam penelitian ini adalah citra digital dengan ekstensi umum seperti .jpg, .jpeg, atau .png. Dataset terdiri dari dua kelas, sehingga struktur dataset disusun dalam dua folder utama: folder citra AI dan folder citra asli.

Selanjutnya, dataset dipisahkan menjadi dua bagian utama: data latih (training) dan data uji (testing). Pembagian ini dilakukan agar model dapat mempelajari pola pada data latih serta diuji kemampuannya pada data uji yang belum pernah dilihat sebelumnya.

3.2.2 Alat Pengambilan Data

Pengambilan data dilakukan dengan metode unduh dataset serta pengorganisasian file citra ke dalam struktur folder sesuai kebutuhan pelatihan model. Selanjutnya, data diproses menggunakan bahasa pemrograman Python melalui beberapa pustaka pendukung seperti OpenCV untuk mendukung proses pengolahan citra secara terprogram, sehingga seluruh tahapan penelitian dapat dilakukan secara sistematis dan konsisten (Nguyen et al., 2025).

3.2.3 Karakteristik Data

Karakteristik data yang digunakan dalam penelitian ini Adalah sebagai berikut:

1. Data yang digunakan berupa citra digital.
2. Dataset terdiri dari dua kelas, yaitu gambar buatan AI dan gambar asli.
3. Format citra yang digunakan Adalah JPG dan PNG.
4. Citra memiliki variasi ukuran/resolusi, sehingga diperlukan proses preprocessing seperti resize agar ukuran seragam.
5. Dataset disusun dalam bentuk folder berdasarkan label kelas untuk memudahkan proses pelatihan model.
6. Dataset dibagi menjadi data latih (training) dan uji (testing) untuk memastikan evaluasi model dilakukan secara objektif.

Selanjutnya, data citra yang digunakan memiliki karakteristik visual yang beragam seperti perbedaan tekstur, pola tepi, intensitas piksel, dan variasi warna. Oleh karena itu, diperlukan preprocessing untuk menyesuaikan ukuran citra serta memastikan data siap digunakan pada proses ekstraksi fitur visual. Selain itu, penyusunan dataset dalam folder berlabel dan pembagian data

menjadi training dan testing merupakan praktik umum dalam penelitian berbasis pengolahan citra dan pembelajaran mesin, karena dapat membantu model mempelajari pola dari data latih dan diuji pada data yang belum pernah dilihat sebelumnya (Vásquez-Salazar et al., 2024).

3.3 Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini dilakukan untuk memperoleh dataset citra yang akan digunakan dalam proses pelatihan dan pengujian sistem klasifikasi. Data yang dikumpulkan terdiri dari dua kategori, yaitu gambar buatan AI dan gambar asli, sehingga sistem dapat mempelajari perbedaan karakteristik visual di antara keduanya.

3.3.1 Tahapan Pengumpulan Data

Tahapan pengumpulan data pada penelitian ini dilakukan sebagai berikut:

1. Menentukan kebutuhan data citra sesuai tujuan penelitian (gambar AI dan gambar asli)
2. Mengumpulkan gambar gambar buatan AI melalui proses generasi menggunakan model AI atau dari website
3. Mengumpulkan gambar asli melalui pengambilan foto secara langsung menggunakan kamera/handphone atau melalui sumber terbuka yang kredibel
4. Menyusun struktur folder dataset berdasarkan kelas (AI dan asli)

3.3.2 Penjelasan Pengumpulan Data

Pengumpulan data dilakukan dengan cara mengorganisasikan citra ke dalam dua kategori utama, yaitu citra buatan AI dan citra asli. Data gambar AI

diperoleh melalui proses generasi menggunakan model *Artificial Intelligence* seperti ChatGPT dan juga mengambil gambar dari website seperti Freepik, sedangkan gambar asli diperoleh dari hasil pengambilan foto maupun dari sumber terbuka yang valid. Selanjutnya, seluruh file citra diseleksi agar memiliki kualitas yang layak untuk dilakukan pengolahan citra.

Setelah data terkumpul, file citra disusun dalam struktur folder berdasarkan label kelas dan dilakukan pembagian dataset menjadi data latih serta data uji. Penyusunan dataset secara terstruktur dan pembagian training/testing merupakan praktik umum dalam penelitian berbasis computer vision dan pembelajaran mesin, karena membantu memastikan proses pelatihan model berjalan sistematis dan hasil evaluasi lebih objektif (Vásquez-Salazar et al., 2024).

3.4 Teknik Analisis Data

Teknik analisis data dalam penelitian ini dilakukan untuk mengolah dataset citra menjadi data yang siap digunakan dalam proses klasifikasi. Sistem yang dibangun tidak bertujuan sebagai produk komersial, melainkan sebagai prototype penelitian untuk menganalisis peran ekstraksi fitur visual dan *ensemble learning* dalam membedakan gambar buatan AI dan gambar asli secara transparan. Pendekatan ini menekankan pada pemahaman proses pembentukan fitur dan mekanisme pengambilan keputusan model, bukan hanya pada hasil prediksi akhir. Adapun tahapan analisis data pada penelitian ini adalah sebagai berikut:

1. Preprocessing data citra
2. Ekstraksi fitur visual
3. Pembentukan label kelas
4. Pembagian data training dan testing

5. Pelatihan model menggunakan *ensemble learning*
6. Pengujian model dan evaluasi

3.4.1 Preprocessing Data Citra

Preprocessing dilakukan untuk menyamakan karakteristik citra agar proses ekstraksi fitur dapat berjalan lebih konsisten. Tahapan preprocessing yang dilakukan meliputi:

1. Resize Citra ke ukuran tertentu agar seragam
2. Normalisasi nilai pixel agar berada pada rentang tertentu
3. Konversi warna (misalnya RGB ke grayscale) bila diperlukan

Tahap preprocessing penting karena data citra biasanya memiliki perbedaan resolusi, komposisi warna, dan kualitas gambar yang bervariasi. Dengan preprocessing, data menjadi lebih terstruktur sehingga model klasifikasi dapat mempelajari pola secara lebih stabil dan tidak bias terhadap perbedaan ukuran gambar (Vásquez-Salazar et al., 2024).

3.4.2 Ekstraksi Fitur Visual

Setelah processing, citra diproses untuk mengambil karakteristik penting yang disebut sebagai fitur visual. Fitur visual digunakan sebagai pembeda antara gambar buatan AI dan gambar asli. Pada penelitian ini fitur yang digunakan meliputi:

1. Fitur warna
2. Fitur tekstur
3. Fitur tepi
4. Distribusi intensitas piksel

Fitur-fitur tersebut dipilih karena karakteristik visual pada gambar buatan AI dan gambar asli dapat memiliki perbedaan tertentu, baik dalam pola tekstur maupun distribusi intensitas warna dan piksel. Pemanfaatan representasi fitur dan teknik penggabungan fitur juga sering digunakan untuk meningkatkan kemampuan sistem dalam mendeteksi gambar buatan AI (Xu et al., 2024).

3.4.3 Pembentukan Dataset Fitur

Hasil ekstraksi fitur dari setiap citra akan menghasilkan data berbentuk vector numerik. Vector ini mempresentasikan karakteristik visual gambar. Selanjutnya setiap vector fitur diberikan label kelas, yaitu:

1. Label (0) untuk gambar asli
2. Label (1) untuk gambar buatan AI

Dataset fitur yang sudah terbentuk kemudian digunakan sebagai input pelatihan model klasifikasi.

3.4.4 Pelatihan Model Menggunakan *Ensemble Learning*

Pada tahap ini dilakukan pelatihan model klasifikasi untuk membedakan gambar buatan *Artificial Intelligence* (AI) dan gambar asli berdasarkan fitur visual yang telah diekstraksi pada tahap sebelumnya. Metode yang digunakan dalam proses klasifikasi adalah *Ensemble Learning*, yaitu pendekatan pembelajaran mesin yang menggabungkan beberapa model (base learner) untuk menghasilkan keputusan akhir yang lebih stabil dan akurat dibandingkan model tunggal. Konsep penggabungan beberapa model untuk meningkatkan performa prediksi banyak digunakan karena dapat mengurangi error dan meningkatkan generalisasi model pada data yang kompleks (Ganaie et al., 2022).

Pendekatan *Ensemble Learning* dipilih karena data citra pada penelitian ini memiliki variasi yang tinggi, baik dari sisi tekstur, pola tepi, maupun distribusi intensitas piksel. Dengan menggabungkan beberapa model klasifikasi, sistem diharapkan mampu meningkatkan kemampuan generalisasi serta mengurangi kesalahan prediksi ketika dihadapkan pada data citra baru.

Secara umum, prinsip *ensemble* Adalah menggabungkan hasil prediksi dari sejumlah model M untuk menghasilkan prediksi akhir. Dalam penelitian ini, proses penggabungan Keputusan dapat dilakukan menggunakan metode voting (pemungutan suara), baik hard voting maupun soft voting (Ganaie et al., 2022).

1. Hard Voting

Hard voting adalah metode *ensemble* yang menentukan hasil klasifikasi berdasarkan kelas yang paling banyak dipilih oleh model-model penyusun *ensemble*. Rumus hard voting dapat dituliskan sebagai berikut:

$$\{y\} = \text{mode}(h_{1(x)}, h_{2(x)}, \dots, h_{M(x)})$$

Tabel 3. 2 simbol *ensemble learning*

Simbol	Keterangan
$\{y\}$	hasil Prediksi akhir dari <i>ensemble</i>
$\{\text{mode}\}(\cdot)$	nilai yang paling sering muncul
$h_{i(x)}$	hasil prediksi model ke-i terhadap data x
m	Jumlah model dalam <i>ensemble</i>

3.4.5 Pengujian dan Evaluasi Hasil

Setelah proses pelatihan model selesai dilakukan pada tahap sebelumnya, langkah berikutnya adalah melakukan pengujian dan evaluasi untuk mengetahui

performa model dalam membedakan gambar buatan AI dan gambar asli. Pengujian dilakukan menggunakan data uji (testing data) yang tidak digunakan pada proses pelatihan sehingga hasil evaluasi dapat menggambarkan kemampuan model dalam melakukan generalisasi terhadap data baru.

Pada penelitian ini, model yang telah dibangun menggunakan Voting Classifier akan melakukan prediksi terhadap data uji. Selanjutnya, hasil prediksi dibandingkan dengan label sebenarnya (ground truth) untuk menghitung nilai evaluasi berdasarkan beberapa metrik klasifikasi. Evaluasi ini diperlukan untuk mengukur tingkat keberhasilan model dan menilai kualitas sistem klasifikasi yang dibangun.

1. Confusion Matrix

Evaluasi model dilakukan dengan menggunakan confusion matrix, yaitu tabel yang menggambarkan jumlah prediksi yang benar dan salah untuk masing-masing kelas. Confusion matrix terdiri dari empat komponen utama yaitu:

- a. True Positive (TP): data gambar AI yang diprediksi sebagai gambar AI.
- b. True Negative (TN): data gambar asli yang diprediksi sebagai asli.
- c. False Positive (FP): data gambar asli tetapi diprediksi sebagai AI.
- d. False Negative (FN): data gambar AI tetapi diprediksi sebagai asli.

Confusion matrix membantu dalam melihat kesalahan klasifikasi secara detail, terutama untuk mengetahui jenis kesalahan prediksi yang sering terjadi.

2. Akurasi (*Accuracy*)

Akurasi merupakan presentase prediksi yang benar terhadap seluruh data uji. Rumus akurasi Adalah sebagai berikut:

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

Keterangan:

- a. TP: True Positive
- b. TN: True Negative
- c. FP: False Positive
- d. FN: False Negative

3. Presisi (*Precision*)

Precision menunjukkan tingkat ketepatan prediksi model pada kelas tertentu, yaitu perbandingan antara jumlah prediksi positif yang benar dengan seluruh prediksi positif. Rumus precision Adalah:

$$Precision = \frac{TP}{TP + FP}$$

Keterangan:

- a. TP: True Positive
- b. FP: False Positive

4. Recall (Sensitivity / True Positive Rate)

Recall menunjukkan kemampuan model dalam menemukan seluruh data yang benar-bener termasuk kelas positif. Rumus recall Adalah:

$$Recall = \frac{TP}{TP + FN}$$

Keterangan:

- a. TP: True Positive
- b. FN: False Negative

5. F1-Score

F1-score merupakan nilai rata-rata harmonis antara precision dan recall.

F1-score digunakan untuk menilai performa model secara seimbang, terutama ketika jumlah data antar kelas tidak seimbang. Rumus F1-score

Adalah:

$$F1 - score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

Keterangan:

- a. Precision: nilai presisi
- b. Recall: nilai recall

6. Interpretasi Hasil

Nilai evaluasi yang diperoleh dari akurasi, precision, recall, dan F1-score akan dianalisis untuk mengetahui apakah model Voting Classifier mampu membedakan gambar buatan AI dan gambar asli dengan baik. Jika hasil evaluasi menunjukkan nilai performa yang tinggi, maka dapat disimpulkan bahwa ekstraksi fitur visual dan metode *ensemble* voting classifier mampu memberikan klasifikasi yang efektif. Sebaliknya, apabila performa masih rendah, maka dilakukan analisis terhadap penyebab kesalahan prediksi, seperti variasi citra, noise pada gambar, atau keterbatasan fitur yang digunakan.

3.5 Prosedur Pengujian dan Validasi

Prosedur pengujian dan validasi dilakukan untuk memastikan bahwa sistem klasifikasi yang dibangun mampu membedakan gambar buatan AI dan gambar dan gambar asli dengan baik. Pengujian dilakukan terhadap dataset uji yang digunakan dalam proses pelatihan, sehingga hasil evaluasi dapat menggambarkan kemampuan model dalam menggenerasikan data baru.

3.5.1 Tahapan Pengujian Sistem

Tahapan pengujian sistem pada penelitian ini dilakukan sebagai berikut:

- a. Menentukan pembagian data menjadi data latih dan data uji.
- b. Melakukan processing citra pada data latih dan data uji.
- c. Melakukan ekstraksi fitur visual pada seluruh data citra.
- d. Melatih model klasifikasi menggunakan voting classifier dengan data latih.
- e. Menguji model menggunakan data uji untuk menghasilkan prediksi kelas.
- f. Membandingkan hasil prediksi dengan label sebenarnya (*ground truth*).
- g. Menghitung nilai evaluasi menggunakan confusion matrix dan metrik evaluasi (akurasi, precision, recall, dan F1-score).

3.5.2 Validasi Hasil Klasifikasi

Validasi dilakukan untuk memastikan hasil klasifikasi dapat dipercaya. Proses validasi dilakukan dengan cara membandingkan prediksi model terhadap data uji dengan label asli yang telah ditentukan. Jika hasil evaluasi menunjukkan nilai yang baik, maka sistem dianggap mampu mengklasifikasikan gambar buatan AI dan gambar asli sesuai tujuan penelitian.

Selain itu, validasi juga bertujuan untuk menganalisis kesalahan klasifikasi yang terjadi. Kesalahan dapat disebabkan oleh beberapa faktor, seperti kualitas citra yang tidak konsisten, adanya noise pada gambar, serta karakteristik gambar AI yang semakin mirip dengan gambar asli. Analisis tersebut digunakan sebagai bahan evaluasi untuk perbaikan metode, misalnya dengan menambah fitur visual atau penyesuaian parameter model.

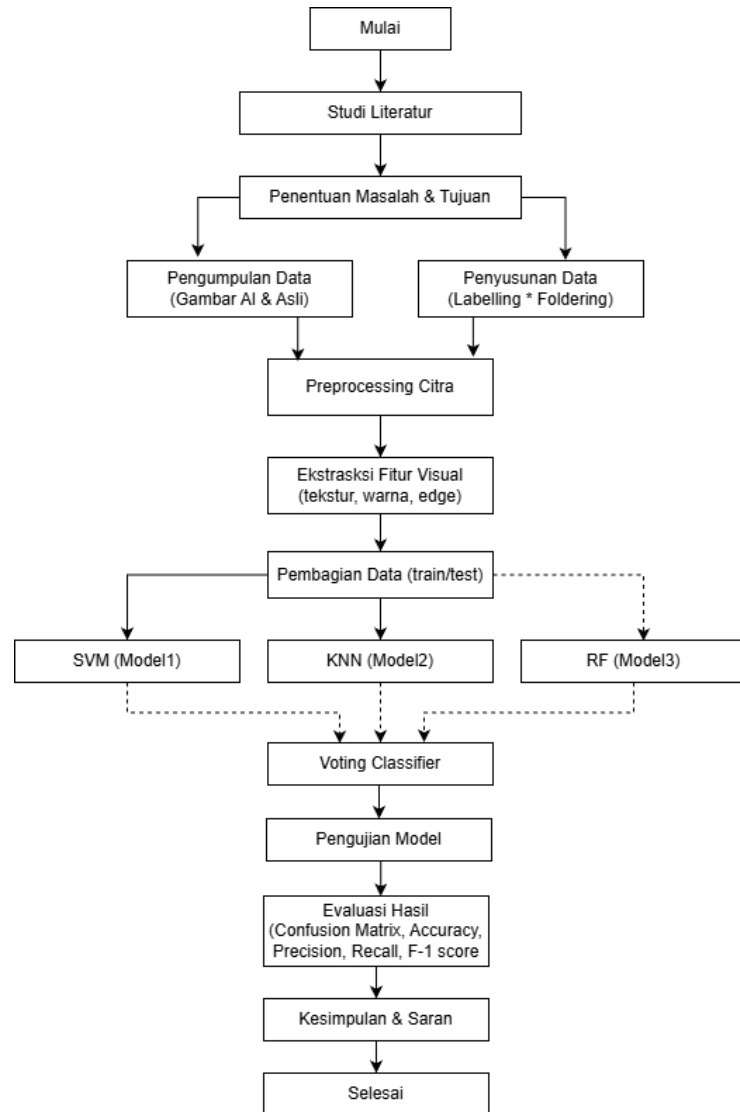
3.5.3 Kriteria Keberhasilan Sistem

Kriteria keberhasilan sistem dalam penelitian ini ditentukan berdasarkan hasil evaluasi performa model. Sistem dikatakan berjalan dengan baik apabila model mampu menghasilkan Tingkat akurasi dan nilai evaluasi lainnya yang tinggi serta memiliki kesalahan klasifikasi yang rendah pada data uji.

3.6 Diagram Alir Penelitian

Diagram alir penelitian digunakan untuk menggambarkan alur proses penelitian secara sistematis dan terstruktur, mulai dari tahap awal perencanaan hingga tahap akhir penyusunan kesimpulan. Diagram ini membantu peneliti dalam memahami setiap tahapan yang harus dilalui, mulai dari pengumpulan data, pengolahan data, hingga evaluasi hasil, sehingga proses penelitian dapat berjalan secara terarah, jelas, dan mudah dipahami.

Penyajian dalam bentuk diagram ini membantu memperjelas hubungan antar proses serta memudahk dalam melakukan analisis apabila terjadi kendala pada salah satu tahapan, dengan demikian keseluruhan proses penelitian dapat berjalan secara lebih terarah, efisien, dan mudah dipahami oleh peneliti maupun pembaca.



Gambar 3. 1 Alur Penelitian

Diagram alir penelitian pada Gambar 3.1 menggambarkan tahapan-tahapan yang dilakukan dalam penelitian ini secara sistematis, mulai dari tahap awal hingga tahap akhir. Diagram tersebut digunakan untuk mempermudah pemahaman terhadap alur proses penelitian dalam membedakan citra buatan *Artificial Intelligence* (AI) dan citra asli.

Tahap pertama dimulai dengan studi literatur, yaitu pengumpulan referensi dari jurnal dan penelitian terdahulu yang berkaitan dengan pengolahan citra

digital, ekstraksi fitur visual, dan metode klasifikasi. Tahap ini bertujuan untuk memperoleh landasan teori yang mendukung penelitian.

Selanjutnya dilakukan pengumpulan data berupa gambar buatan AI yang dihasilkan menggunakan teknologi generatif serta gambar asli yang diperoleh dari kamera atau sumber terpercaya. Data yang terkumpul kemudian disusun dalam bentuk dataset dengan struktur folder yang rapi dan penamaan file yang sesuai.

Pada tahap berikutnya, dilakukan preprocessing data citra yang meliputi proses resize, normalisasi, dan konversi ruang warna jika diperlukan. Tahap ini bertujuan untuk menyeragamkan ukuran dan kualitas citra agar siap diproses lebih lanjut.

Setelah preprocessing, dilakukan ekstraksi fitur visual untuk memperoleh karakteristik penting dari citra, seperti fitur warna, tekstur, tepi, dan distribusi piksel. Fitur-fitur tersebut digunakan sebagai representasi data dalam proses klasifikasi. Dengan adanya ekstraksi fitur, informasi yang semula berbentuk citra mentah dapat diolah menjadi data yang lebih terstruktur dan mudah dianalisis oleh model.

Selanjutnya dilakukan pembentukan vektor fitur dan pemberian label kelas, yaitu kelas citra AI dan citra asli. Data yang telah dilabeli kemudian dibagi menjadi data training dan data testing. Pembagian ini bertujuan untuk melatih model sekaligus menguji kemampuan tersebut.

Tahap selanjutnya adalah pelatihan model menggunakan metode *ensemble learning* berdasarkan data training yang telah disiapkan. Model yang telah dilatih kemudian digunakan untuk melakukan pengujian terhadap data testing.

Hasil pengujian dievaluasi menggunakan confusion matrix untuk memperoleh nilai akurasi, presisi, recall, dan F1-score. Evaluasi ini bertujuan untuk mengetahui kinerja model dalam membedakan citra buatan AI dan citra asli.

Tahap terakhir adalah penarikan kesimpulan dan pemberian saran berdasarkan hasil evaluasi yang telah diperoleh. Kesimpulan digunakan untuk menjawab tujuan penelitian, sedangkan saran diberikan sebagai bahan pengembangan penelitian selanjutnya, saran tersebut dapat berupa peningkatan jumlah dataset, penggunaan metode ekstraksi fitur yang lebih optimal, maupun penerapan algoritma klasifikasi lain yang berpotensi memberikan hasil yang lebih baik.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Deskripsi Data

Dataset yang digunakan dalam penelitian ini terdiri dari citra digital yang terbagi ke dalam dua kategori, yaitu gambar hasil kecerdasan buatan (AI-generated image) dan gambar asli (real image). Total dataset yang digunakan berjumlah 560 gambar, dengan distribusi yang seimbang pada masing-masing kelas.

Rincian dataset adalah sebagai berikut:

- a. Gambar AI Sebanyak 280 citra
- b. Gambar asli sebanyak 280 citra

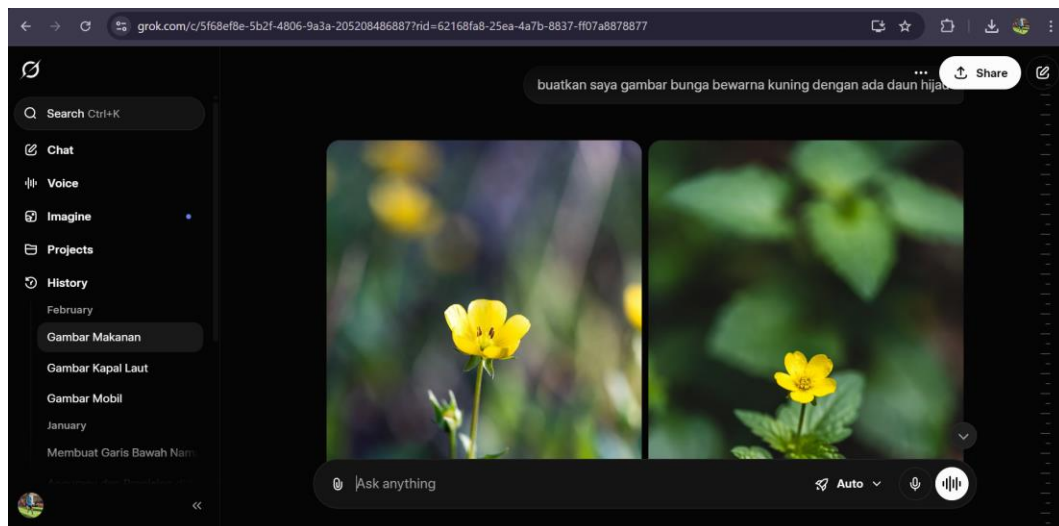
Dataset kemudian dibagi menjadi dua bagian, yaitu data latih (training data) dan data uji (testing data) dengan pembagian terpisah sejak awal proses pengumpulan data. Data latih terdiri dari 210 gambar hasil *artificial intelligence* dan 210 gambar asli, data uji berjumlah 140 gambar, yang terdiri dari 70 gambar AI dan 70 gambar asli.

Pembagian dataset dilakukan secara seimbang antara kedua kelas dengan tujuan untuk menghindari terjadinya bias pada model terhadap salah satu kategori. Dengan komposisi data yang seimbang tersebut, diharapkan model klasifikasi dapat bekerja secara optimal dalam membedakan antara gambar buatan AI dan gambar asli.

4.1.1 Sumber Data Gambar AI

Data gambar AI dalam penelitian ini diperoleh melalui proses generasi citra menggunakan platform kecerdasan buatan Grok yang dikembangkan oleh xAI. Proses pengambilan data dilakukan dengan memasukkan prompt tertentu untuk menghasilkan variasi gambar sesuai kebutuhan penelitian.

Tangkapan layar proses generasi gambar AI ditunjukkan pada Gambar 4.1 sebagai bukti proses pengumpulan data.



Gambar 4. 1 Contoh pengambilan data gambar AI di platform Grok

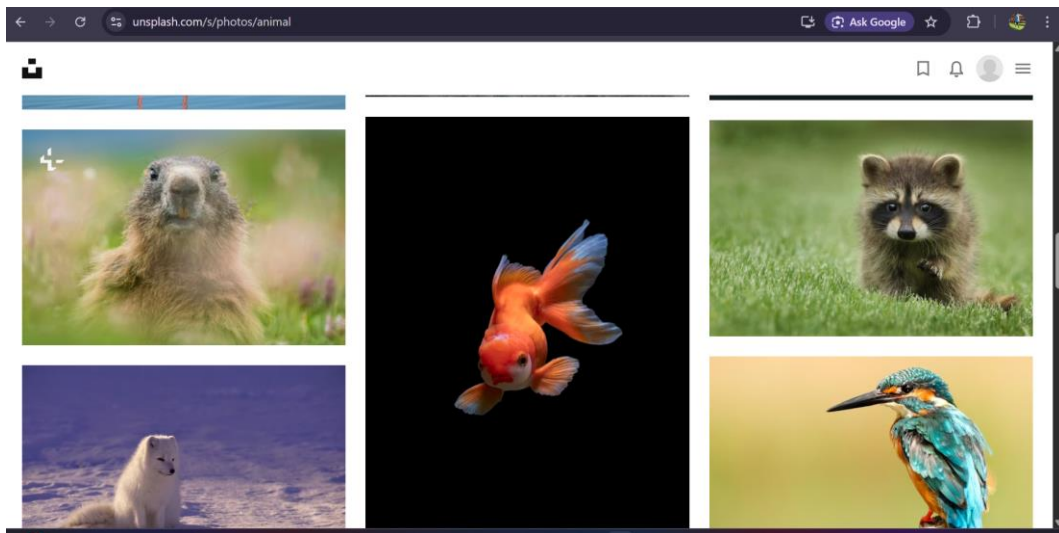
4.1.2 Sumber Data Gambar Asli

Data gambar asli (real image) dalam penelitian ini diperoleh dari platform penyedia foto bebas lisensi Unsplash. Platform ini menyediakan citra fotografi berkualitas tinggi yang diunggah oleh fotografer profesional maupun komunitas global.

Pengambilan data dilakukan dengan memilih gambar yang memiliki karakteristik visual alami, seperti pencahayaan realistis, variasi tekstur, dan komposisi fotografi

umum. Pemilihan dilakukan secara selektif untuk memastikan bahwa gambar yang digunakan benar-benar merupakan hasil fotografi nyata dan bukan hasil generasi kecerdasan buatan.

Tangkapan layar proses pengambilan gambar dari Unsplash ditunjukkan pada Gambar 4.2 sebagai dokumentasi sumber data.



Gambar 4. 2 contoh pengambilan data gambar asli

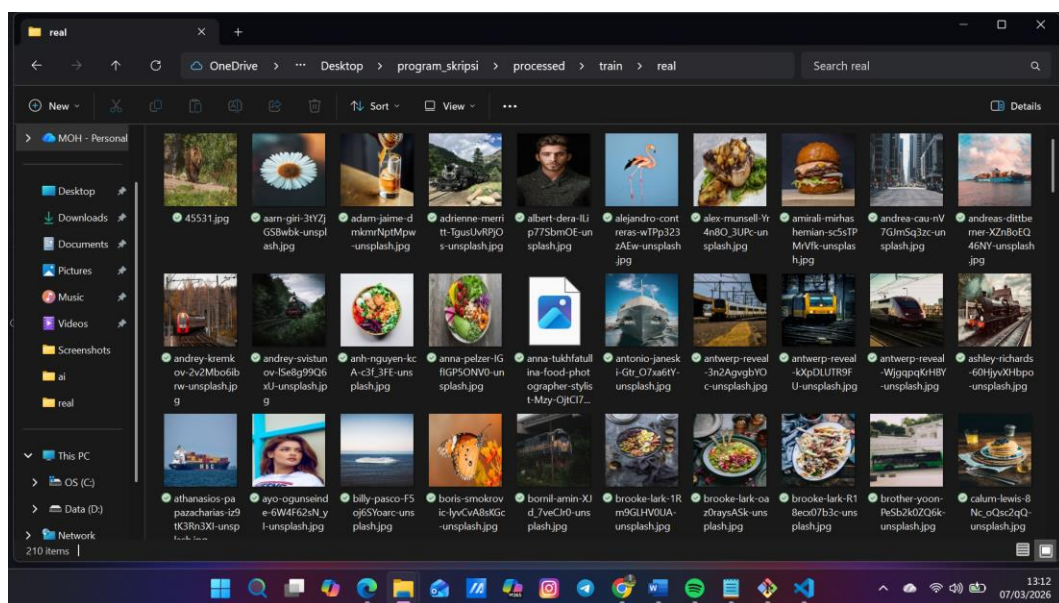
4.2 Hasil Preprocessing Data

Tahap preprocessing dilakukan untuk memastikan seluruh citra memiliki karakteristik yang seragam sebelum dilakukan proses ekstraksi fitur. Dataset yang digunakan dalam penelitian ini memiliki variasi ukuran dan resolusi yang berbeda, sehingga diperlukan proses penyesuaian agar data dapat diproses secara konsisten oleh sistem klasifikasi.

Preprocessing merupakan tahapan penting dalam pengolahan citra digital karena dapat meningkatkan kualitas representasi data serta membantu model dalam mempelajari pola visual secara lebih stabil (Renu, Neelam, & Bhatia, 2026). Salah satu langkah yang umum dilakukan pada tahap ini adalah penyeragaman ukuran

citra agar setiap gambar memiliki dimensi yang sama sehingga dapat diproses secara lebih konsisten oleh algoritma pembelajaran mesin.

Pada penelitian ini, seluruh citra pada dataset diubah menjadi ukuran 256 x 256 piksel menggunakan pustaka OpenCV pada bahasa pemrograman Python. Proses resize bertujuan untuk menyeragamkan dimensi citra sehingga perbedaan resolusi antar gambar tidak mempengaruhi proses ekstraksi fitur dan klasifikasi.



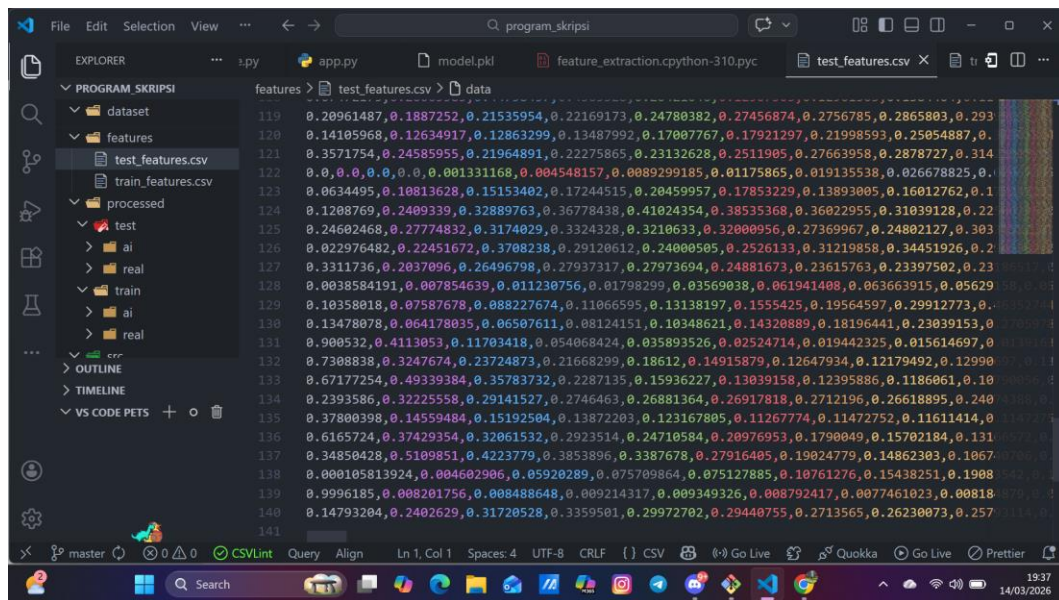
Gambar 4.3 Hasil Penyeragaman ukuran citra

Dengan dilakukannya proses resize, seluruh citra pada dataset memiliki ukuran yang seragam sehingga dapat diproses secara lebih efektif pada tahap ekstraksi fitur visual. Proses ini membantu sistem dalam menganalisis karakteristik citra secara lebih konsisten tanpa dipengaruhi oleh variasi ukuran gambar.

4.3 Hasil Ekstraksi Fitur Visual

Pada tahap ini dilakukan proses ekstraksi fitur visual dari citra yang telah melalui tahap preprocessing. Ekstraksi fitur bertujuan untuk mengubah citra digital

menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi. Fitur visual berperan penting dalam sistem pengenalan citra karena mampu merepresentasikan karakteristik visual dari suatu objek dalam bentuk data numerik yang dapat dianalisis oleh model *machine learning* (Rekhan, Rishi, & Aditya, 2025)



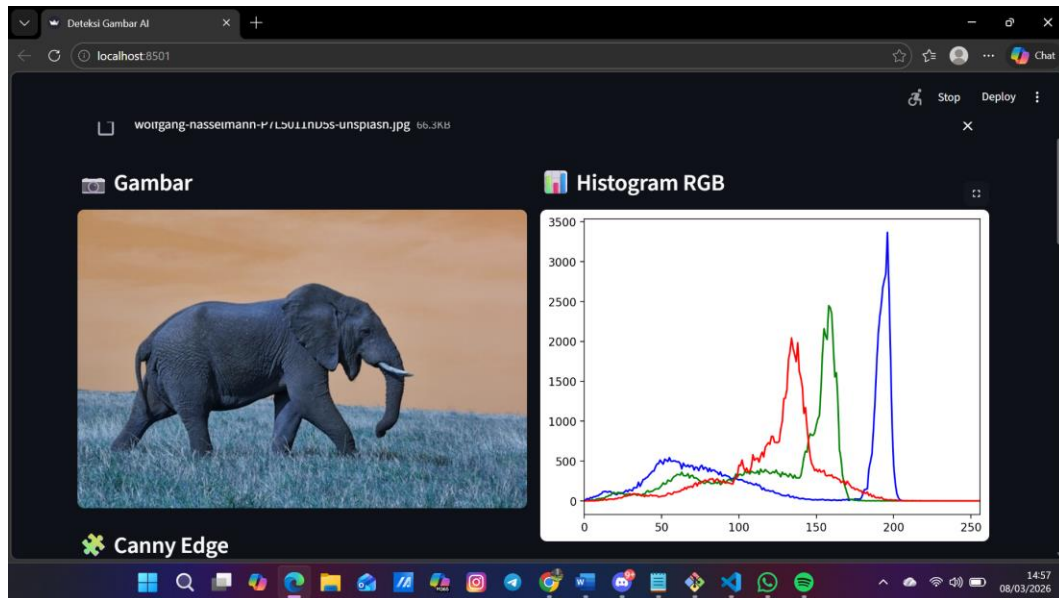
Gambar 4. 4 Hasil Ekstraksi Fitur Visual

Dalam penelitian ini, fitur visual yang digunakan meliputi fitur warna, fitur tekstur, fitur tepi, serta distribusi intensitas piksel. Proses ekstraksi fitur dilakukan menggunakan bahasa pemrograman Python dengan bantuan pustaka OpenCV dan Scikit-image. Setiap citra yang telah distandarisasi ukurannya menjadi 256×256 piksel kemudian diproses untuk menghasilkan vektor fitur yang merepresentasikan karakteristik visual citra tersebut.

4.3.1 Ekstraksi Fitur Warna

Fitur warna diperoleh menggunakan histogram warna pada ruang warna RGB. Histogram warna digunakan untuk menggambarkan distribusi intensitas

warna merah (Red), hijau (Green), dan biru (Blue) pada citra. Pendekatan ini banyak digunakan dalam analisis citra karena mampu mempresentasikan karakteristik warna secara efektif dalam bentuk distribusi statistic (Jokin, Pizarro, Olga, Juan, & Ricard, 2023).




Gambar 4. 5 Histogram distribusi warna RGB

Pada proses ini, setiap kanal warna dianalisis untuk menghitung distribusi intensitas piksel. Nilai histogram yang dihasilkan kemudian dinormalisasi sehingga menghasilkan representasi fitur yang lebih stabil dan dapat digunakan sebagai input dalam proses klasifikasi.

4.3.2 Fitur Tekstur

Fitur tekstur pada penelitian ini diperoleh menggunakan metode Gray Level Co-occurrence Matrix (GLCM). Metode ini digunakan untuk menganalisis hubungan spasial antar piksel pada citra grayscale sehingga dapat menggambarkan pola tekstur yang terdapat pada citra.

 Nilai Fitur yang Digunakan		
Contrast	Energy	Edge Ratio
308.31	0.0693	0.097
Correlation	Homogeneity	
0.891	0.418	

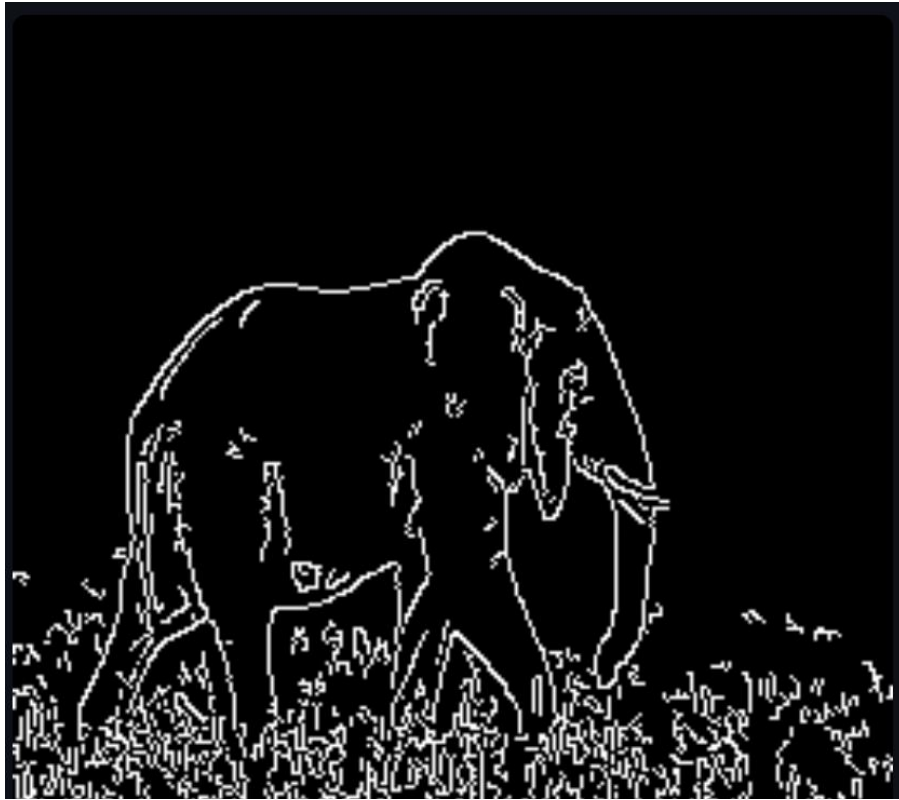
Gambar 4. 6 Hasil ekstraksi fitur tekstur menggunakan GLCM

Dari matriks GLCM tersebut dihitung beberapa parameter statistik yang digunakan sebagai fitur tekstur, yaitu contrast, correlation, energy, dan homogeneity. Parameter-parameter tersebut digunakan untuk merepresentasikan kompleksitas tekstur serta hubungan intensitas antar piksel dalam citra.

Nilai-nilai fitur tekstur tersebut kemudian digunakan sebagai bagian dari vektor fitur yang akan diproses pada tahap klasifikasi.

4.3.3 Fitur Tepi

Fitur tepi diperoleh menggunakan metode Canny Edge Detection. Metode ini digunakan untuk mendeteksi perubahan intensitas piksel yang signifikan sehingga dapat mengidentifikasi batas atau kontur objek pada citra. Informasi tepi dapat membantu sistem dalam mengenali struktur objek yang terdapat dalam gambar (Afolabi I., Omolegho A., Idama , Jones U., & Michael O. , 2025).

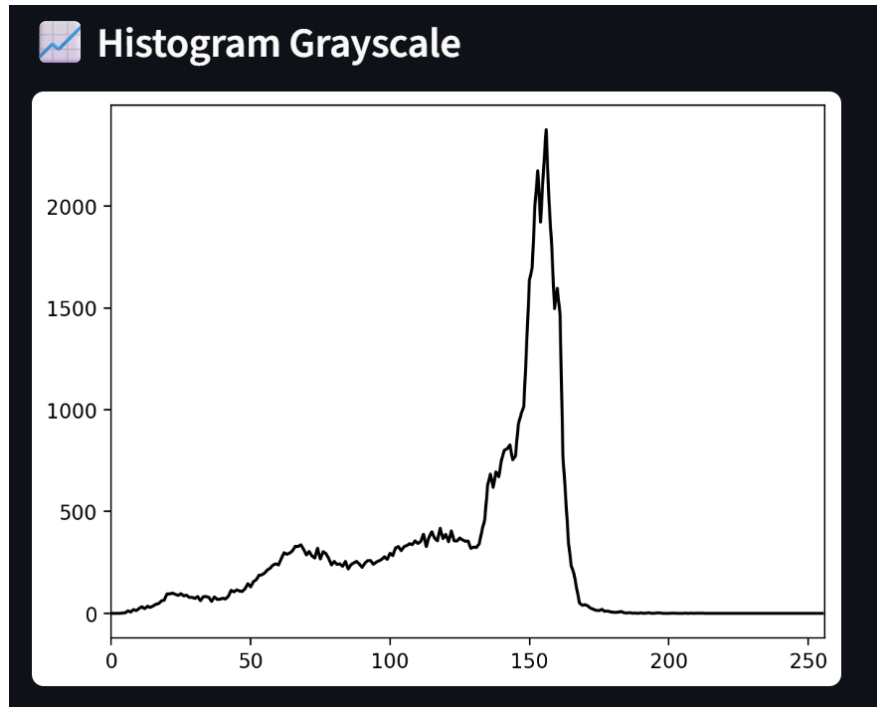


Gambar 4. 7 Hasil deteksi tepi menggunakan metode canny

Pada penelitian ini, selain menghasilkan citra tepi, juga dihitung rasio jumlah piksel tepi terhadap total piksel pada citra yang disebut sebagai edge ratio. Nilai ini digunakan sebagai salah satu fitur untuk menggambarkan kompleksitas struktur objek dalam citra.

4.3.4 Distribusi Intesnsitas Piksel

Distribusi intensitas piksel diperoleh melalui histogram grayscale. Histogram ini menunjukkan penyebaran nilai intensitas piksel pada citra setelah dikonversi ke grayscale. Distribusi intensitas ini dapat menggambarkan karakteristik terang dan gelap pada citra yang menjadi salah satu indikator perbedaan antara gambar buatan AI dan gambar asli.



Gambar 4. 8 Hasil Histogram Distribusi Intensitas Grayscale

Histogram grayscale kemudian dinormalisasikan sehingga dapat digunakan sebagai bagian dari representasi fitur dalam proses klasifikasi.

4.4 Hasil Pelatihan Model

Setelah proses ekstraksi fitur visual selesai dilakukan, tahap selanjutnya adalah pelatihan model klasifikasi menggunakan fitur yang telah diperoleh dari citra. Pada penelitian ini digunakan beberapa algoritma *machine learning*, yaitu Support Vector Machine (SVM), K-Nearest Neighbor (KNN), dan Random Forest. Selain itu, dilakukan juga penggabungan beberapa model menggunakan metode *Ensemble Learning* dengan teknik Hard Voting untuk meningkatkan performa klasifikasi.

Pelatihan model dilakukan menggunakan dataset fitur yang telah dipisahkan menjadi data training dan data testing. Data training digunakan untuk melatih model agar mampu mempelajari pola dari fitur visual citra, sedangkan

data testing digunakan untuk menguji kemampuan model dalam melakukan klasifikasi terhadap data yang belum pernah dilihat sebelumnya. Penggunaan data training dan testing bertujuan untuk mengukur kemampuan generalisasi model dalam melakukan klasifikasi citra. (Yang, Xueyi, Zelin, & Fang, 2023).

4.4.1 Pelatihan Model Support Vector Machine (SVM)

Support Vector Machine merupakan algoritma klasifikasi yang bekerja dengan mencari hyperplane optimal yang dapat memisahkan dua kelas data secara maksimal dalam ruang fitur. Metode ini banyak digunakan dalam pengolahan citra dan pengenalan pola karena mampu menangani data berdimensi tinggi dengan baik (Siyuan, Qianfei, Mengyang, Weisong, & Jianying, 2025).

Hasil pelatihan model Support Vector Machine (SVM) pada penelitian ini ditampilkan pada Gambar 4.8 Hasil tersebut menunjukkan nilai evaluasi model yang meliputi accuracy, precision, recall, dan F1-score yang diperoleh dari proses pengujian menggunakan data testing.

```
=== SVM ===  
Accuracy : 0.5786  
Precision: 0.5470  
Recall    : 0.9143  
F1-Score  : 0.6845
```

Gambar 4. 9 Hasil Evaluasi SVM

Berdasarkan hasil evaluasi pada Gambar 4.8, model Support Vector Machine menghasilkan nilai accuracy sebesar 0.5786, precision sebesar 0.5470, recall sebesar 0.9143, dan F1-score sebesar 0.6845. Nilai recall yang cukup tinggi menunjukkan bahwa model mampu mengenali sebagian besar citra pada kelas

target, namun nilai precision yang relatif lebih rendah menunjukkan bahwa masih terdapat beberapa kesalahan klasifikasi pada proses prediksi. Hasil ini menunjukkan bahwa model SVM memiliki kemampuan yang cukup baik dalam mendeteksi citra,

4.4.2 Pelatihan Model K-Nearest Neighbor (KNN)

Algoritma K-Nearest Neighbor merupakan metode klasifikasi berbasis jarak yang menentukan kelas suatu data berdasarkan sejumlah tetangga terdekat dalam ruang fitur. Metode ini sering digunakan dalam pengenalan pola karena konsepnya yang sederhana namun efektif dalam mengklasifikasikan data berdasarkan kemiripan fitur (Lenz, Henri , & Chris , 2025)

Pada penelitian ini, model KNN dilatih menggunakan fitur visual yang telah diekstraksi dari citra. Proses pelatihan dilakukan menggunakan data training, sedangkan evaluasi model dilakukan menggunakan data testing untuk mengetahui performa model dalam melakukan klasifikasi citra AI dan citra asli.

```
=== KNN ===  
Accuracy : 0.6000  
Precision: 0.5833  
Recall    : 0.7000  
F1-Score  : 0.6364
```

Gambar 4. 10 Hasil Evaluasi Model KNN

Berdasarkan hasil evaluasi pada Gambar 4.9, model K-Nearest Neighbor menghasilkan nilai accuracy sebesar 0.6000, precision sebesar 0.5833, recall sebesar 0.7000, dan F1-score sebesar 0.6364. Nilai recall yang cukup baik menunjukkan bahwa model mampu mengenali sebagian besar citra pada kelas target, sedangkan nilai precision menunjukkan tingkat ketepatan model dalam

melakukan prediksi. Hasil ini menunjukkan bahwa model KNN memiliki performa yang cukup stabil dalam melakukan klasifikasi citra dibandingkan dengan model lainnya.

4.4.3 Pelatihan Random Forest

Random Forest merupakan algoritma klasifikasi berbasis *ensemble decision tree* yang bekerja dengan membangun sejumlah pohon keputusan secara acak. Setiap pohon keputusan melakukan proses klasifikasi secara independen, kemudian hasil prediksi dari seluruh pohon digabungkan untuk menghasilkan keputusan akhir. Metode ini dikenal memiliki kemampuan yang baik dalam menangani data dengan dimensi fitur yang cukup banyak serta mampu mengurangi risiko overfitting pada proses pelatihan model (Markus & Andre , 2023).

Pada penelitian ini, model Random Forest dilatih menggunakan fitur visual yang telah diperoleh dari proses ekstraksi fitur citra. Model kemudian diuji menggunakan data testing untuk mengetahui performa klasifikasi dalam membedakan citra yang dihasilkan oleh AI dan citra asli.

```
=== Random Forest ===  
Accuracy : 0.6571  
Precision: 0.7115  
Recall    : 0.5286  
F1-Score  : 0.6066
```

Gambar 4. 11 Hasil evaluasi Random Forest

Berdasarkan hasil evaluasi pada Gambar 4.10, model Random Forest menghasilkan nilai accuracy sebesar 0.6571, precision sebesar 0.7115, recall

sebesar 0.5286, dan F1-score sebesar 0.6066. Nilai precision yang cukup tinggi menunjukkan bahwa model memiliki tingkat ketepatan yang baik dalam memprediksi kelas citra. Selain itu, nilai accuracy yang lebih tinggi dibandingkan model lainnya menunjukkan bahwa Random Forest memiliki performa yang lebih baik dalam melakukan klasifikasi pada dataset yang digunakan dalam penelitian ini.

4.4.4 Pelatihan *Ensemble Learning (Hard Voting)*

Ensemble learning merupakan metode dalam *machine learning* yang menggabungkan beberapa model klasifikasi untuk menghasilkan prediksi yang lebih stabil dan akurat. Pendekatan ini bertujuan untuk memanfaatkan kelebihan dari masing-masing model sehingga mampu meningkatkan performa sistem secara keseluruhan. Salah satu teknik *ensemble* yang sering digunakan adalah Hard Voting, yaitu metode yang menentukan hasil klasifikasi berdasarkan suara terbanyak dari beberapa model yang digunakan (Sivagurunathan & Sridharan, 2024)

```
=== Ensemble (Hard Voting) ===  
Accuracy : 0.6500  
Precision: 0.6180  
Recall   : 0.7857  
F1-Score : 0.6918
```

Gambar 4. 12 Hasil evaluasi *ensemble learning*

Berdasarkan hasil evaluasi pada Gambar 4.11, model *ensemble learning* menghasilkan nilai accuracy sebesar 0.6500, precision sebesar 0.6180, recall sebesar 0.7857, dan F1-score sebesar 0.6918. Nilai recall yang cukup tinggi

menunjukkan bahwa metode *ensemble* mampu mengenali sebagian besar citra pada kelas target. Selain itu, nilai F1-score yang lebih tinggi dibandingkan beberapa model individual menunjukkan bahwa metode *ensemble* dapat memberikan keseimbangan yang lebih baik antara precision dan recall dalam proses klasifikasi citra.

4.5 Evaluasi Model Menggunakan Confusion Matrix

Setelah proses pelatihan model selesai dilakukan, tahap selanjutnya adalah melakukan evaluasi terhadap performa model klasifikasi. Evaluasi model bertujuan untuk mengetahui sejauh mana kemampuan model dalam mengklasifikasikan citra secara benar berdasarkan data uji yang digunakan. Salah satu metode evaluasi yang umum digunakan dalam sistem klasifikasi adalah Confusion Matrix.

Confusion Matrix merupakan metode evaluasi yang digunakan untuk menganalisis performa model klasifikasi dengan membandingkan hasil prediksi model dengan label sebenarnya dari data uji sehingga dapat diketahui jumlah prediksi yang benar maupun yang salah (Joanne, Veronique, Bertrand, Edith , & Come, 2026).

Dalam confusion matrix terdapat empat komponen utama yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). True Positive menunjukkan jumlah data yang diprediksi sebagai kelas positif dan memang benar termasuk dalam kelas tersebut. True Negative menunjukkan jumlah data yang diprediksi sebagai kelas negatif dan memang benar termasuk dalam kelas tersebut. False Positive merupakan data yang sebenarnya termasuk kelas negatif namun

diprediksi sebagai kelas positif oleh model, sedangkan False Negative merupakan data yang sebenarnya termasuk kelas positif tetapi diprediksi sebagai kelas negatif (Yujia & Jun, 2025)

```
=== Confusion Matrix (Ensemble) ===  
[[36 34]  
 [15 55]]
```

Gambar 4. 13 Hasil Confusion Matrix Model *Ensemble Learning*

Berdasarkan confusion matrix pada Gambar 4.12, model berhasil mengklasifikasikan 36 citra real dengan benar (True Negative) dan 55 citra AI dengan benar (True Positive). Namun masih terdapat kesalahan klasifikasi yaitu 34 citra real yang diprediksi sebagai AI (False Positive) dan 15 citra AI yang diprediksi sebagai real (False Negative).

Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengenali citra yang dihasilkan oleh AI, yang terlihat dari jumlah True Positive yang cukup tinggi. Namun masih terdapat kesalahan klasifikasi pada citra real yang diprediksi sebagai citra AI. Hal ini dapat disebabkan oleh kemiripan karakteristik visual antara citra real dan citra yang dihasilkan oleh m AI pada beberapa kondisi tertentu.

4.6 Analisis Peforma Model

Berdasarkan hasil pelatihan dan evaluasi model yang telah dilakukan, dapat dilakukan analisis terhadap performa masing-masing algoritma klasifikasi yang digunakan dalam penelitian ini. Model yang digunakan meliputi Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest, serta metode *Ensemble Learning* dengan teknik Hard Voting.

Hasil pengujian menunjukkan bahwa setiap model memiliki tingkat performa yang berbeda dalam melakukan klasifikasi citra. Model Support Vector Machine (SVM) menghasilkan nilai accuracy sebesar 0.5786, precision sebesar 0.5470, recall sebesar 0.9143, serta F1-score sebesar 0.6845. Nilai recall yang cukup tinggi menunjukkan bahwa model mampu mengenali sebagian besar citra yang termasuk dalam kelas target, meskipun masih terdapat kesalahan klasifikasi yang menyebabkan nilai precision relatif lebih rendah.

Model K-Nearest Neighbor (KNN) menghasilkan nilai accuracy sebesar 0.6000, precision sebesar 0.5833, recall sebesar 0.7000, serta F1-score sebesar 0.6364. Hasil ini menunjukkan bahwa model KNN memiliki performa yang cukup stabil dalam melakukan klasifikasi citra berdasarkan kedekatan fitur antar data.

Sementara itu, model Random Forest menunjukkan performa yang lebih baik dibandingkan dua model sebelumnya dengan nilai accuracy sebesar 0.6571, precision sebesar 0.7115, recall sebesar 0.5286, serta F1-score sebesar 0.6066. Nilai precision yang cukup tinggi menunjukkan bahwa model memiliki tingkat ketepatan yang baik dalam memprediksi kelas citra.

Selain itu, metode Ensemble Learning dengan teknik Hard Voting yang menggabungkan model SVM, KNN, dan Random Forest menghasilkan nilai accuracy sebesar 0.6500, precision sebesar 0.6180, recall sebesar 0.7857, serta F1-score sebesar 0.6918. Nilai F1-score yang lebih tinggi dibandingkan model lainnya menunjukkan bahwa metode *ensemble* mampu memberikan keseimbangan yang lebih baik antara precision dan recall dalam proses klasifikasi citra.

Secara keseluruhan, model Random Forest dan *Ensemble Learning* menunjukkan performa yang lebih baik dibandingkan model lainnya dalam melakukan klasifikasi citra. Hal ini menunjukkan bahwa metode berbasis *ensemble* mampu meningkatkan kemampuan sistem dalam mengenali pola fitur visual pada citra.

Tabel 4. 1 Perbandingan Peforma Model Klasifikasi

Model	Accuracy	Precision	Recall	F1-Score
SVM	0.5786	0.5470	0.9143	0.6845
KNN	0.6000	0.5833	0.7000	0.6364
Random Forest	0.6571	0.7115	0.5286	0.6066
Ensemble	0.6500	0.6180	0.7857	0.6918

Berdasarkan Tabel 4.1, model Random Forest memiliki nilai accuracy tertinggi yaitu sebesar 0.6571, sedangkan metode *Ensemble Learning* menghasilkan nilai F1-score tertinggi yaitu sebesar 0.6918. Hal ini menunjukkan bahwa metode *ensemble* mampu memberikan keseimbangan performa yang lebih baik dalam proses klasifikasi citra.

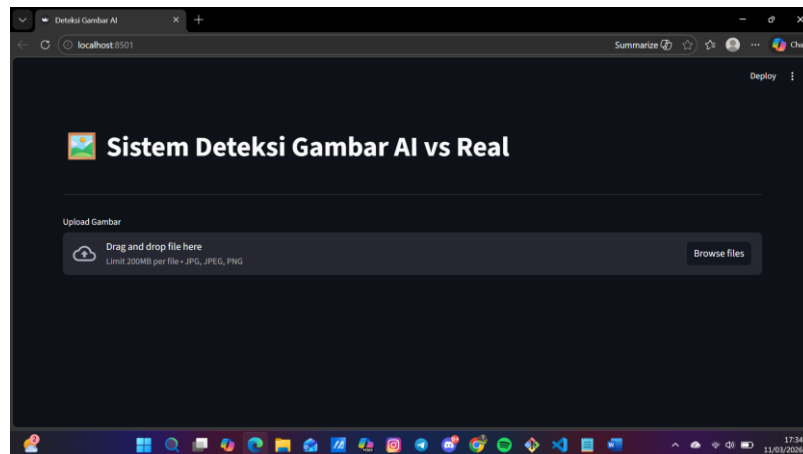
4.7 Implementasi Sistem

Pada tahap implementasi sistem, model klasifikasi yang telah dilatih diintegrasikan ke dalam sebuah aplikasi berbasis web menggunakan framework Streamlit. Implementasi sistem ini bertujuan untuk mempermudah pengguna dalam melakukan proses deteksi terhadap citra yang diunggah, sehingga sistem dapat secara otomatis memprediksi apakah citra tersebut merupakan gambar yang dihasilkan oleh AI atau gambar asli.

Aplikasi ini dikembangkan menggunakan bahasa pemrograman Python dengan memanfaatkan beberapa pustaka seperti OpenCV untuk pengolahan citra, Scikit-learn untuk proses klasifikasi, serta Streamlit sebagai antarmuka pengguna.

4.6.1 Tampilan Halaman Utama Sistem

Halaman utama sistem merupakan tampilan awal aplikasi yang digunakan oleh pengguna untuk mengunggah citra yang akan dianalisis oleh sistem. Pada halaman ini tersedia fitur upload gambar yang memungkinkan pengguna memilih citra dari perangkat yang digunakan



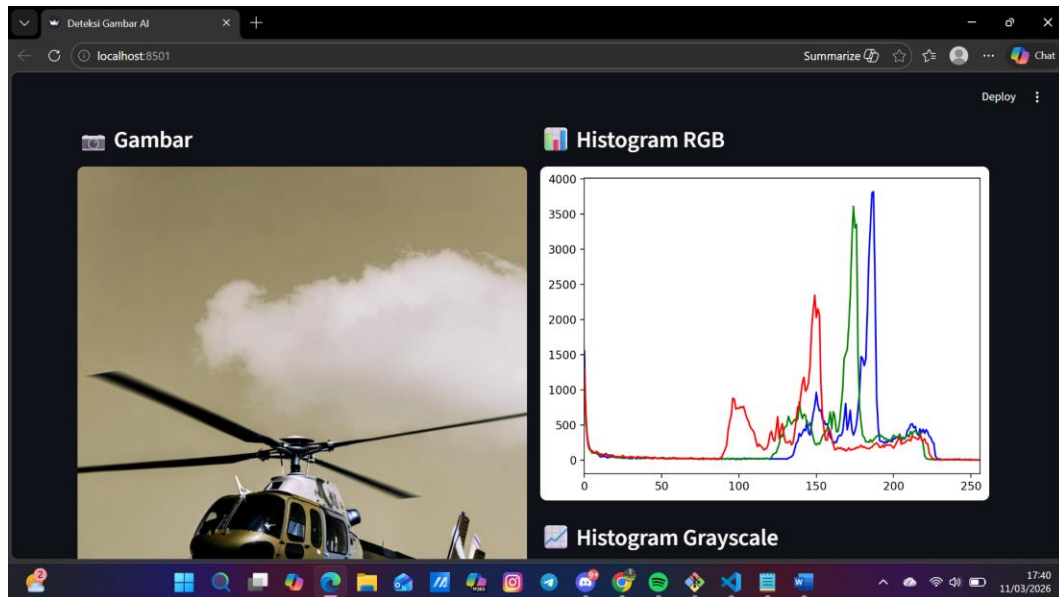
Gambar 4. 14 Tampilan Halaman Utama Sistem

Gambar menunjukkan tampilan awal aplikasi deteksi citra berbasis web yang dikembangkan menggunakan Streamlit. Pada halaman ini pengguna dapat mengunggah citra yang akan dianalisis oleh sistem.

4.6.2 Tampilan Proses Analisis Citra

Setelah citra berhasil diunggah, sistem akan memproses citra tersebut dengan melakukan ekstraksi fitur visual. Pada tahap ini sistem menampilkan

beberapa visualisasi fitur citra seperti histogram warna serta hasil deteksi tepi sebagai bentuk transparansi proses analisis citra.

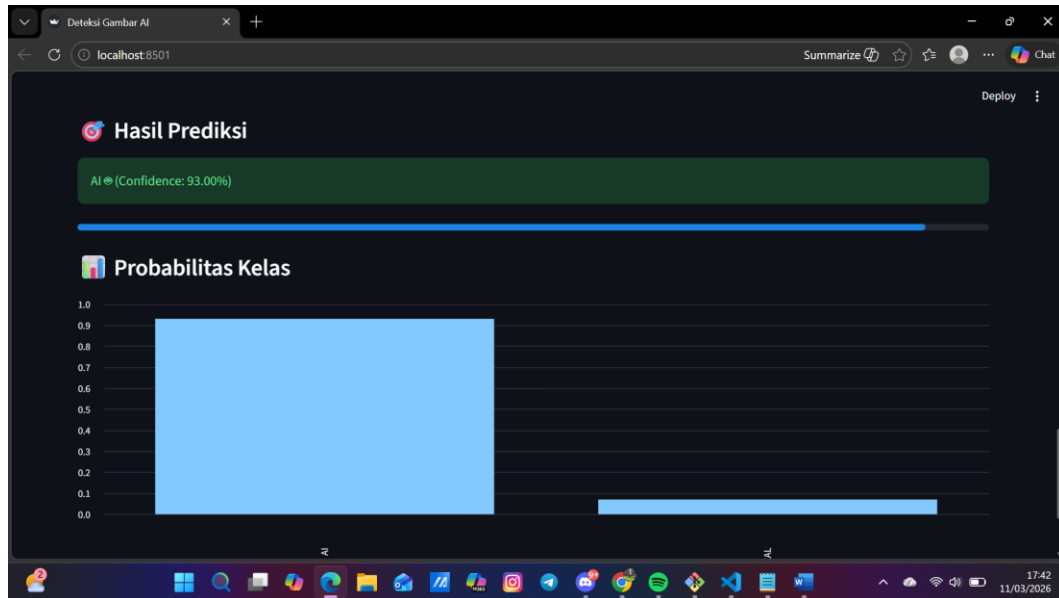


Gambar 4. 15 Tampilan Proses Analisis Citra

Gambar menunjukkan proses analisis citra yang dilakukan oleh sistem setelah citra diunggah oleh pengguna. Sistem menampilkan visualisasi fitur citra seperti histogram warna dan deteksi tepi sebagai bagian dari proses ekstraksi fitur.

4.6.3 Tampilan Hasil Prediksi Sistem

Setelah proses analisis citra selesai dilakukan, sistem akan menampilkan hasil prediksi berupa kategori citra serta tingkat kepercayaan model terhadap hasil prediksi tersebut. Kategori tersebut menunjukkan apakah citra termasuk dalam kelompok gambar yang dihasilkan oleh AI atau merupakan gambar asli. Selain itu sistem juga menyajikan nilai Tingkat kepercayaan(*confidence*) dari model klasifikasi terhadap hasil prediksi yang diberikan.



Gambar 4. 16 Hasil Prediksi Sistem

Gambar menunjukkan hasil prediksi sistem yang menampilkan kategori citra apakah termasuk gambar yang dihasilkan oleh AI atau gambar asli, serta nilai tingkat kepercayaan (confidence) dari model klasifikasi. Nilai confidence ini menunjukkan seberapa yakin model dalam menentukan kelas dari citra yang diuji. Semakin tinggi nilai confidence, maka semakin besar Tingkat keyakinan model terhadap hasil prediksi tersebut. Sebaliknya, nilai confidence yang rendah mengindikasikan bahwa model masih memiliki keraguan dalam menentukan ketegori citra

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian mengenai sistem deteksi citra berbasis *machine learning* untuk membedakan gambar yang dihasilkan oleh kecerdasan buatan (AI) dan gambar asli, maka dapat diperoleh beberapa kesimpulan sebagai berikut.

1. Penelitian ini berhasil mengembangkan sistem deteksi gambar AI menggunakan pendekatan *machine learning* dengan memanfaatkan fitur visual dari citra digital. Proses penelitian dimulai dari tahap preprocessing berupa penyeragaman ukuran citra menjadi 256 x 256 piksel, kemudian dilanjutkan dengan proses ekstraksi fitur visual yang meliputi histogram warna RGB, distribusi intensitas grayscale, fitur tekstur menggunakan metode Gray Level Co-occurrence Matrix (GLCM), serta deteksi tepi menggunakan metode Canny Edge Detection.
2. Hasil ekstraksi fitur tersebut kemudian digunakan sebagai input pada beberapa algoritma klasifikasi yaitu Support Vector Machine (SVM), K-Nearest Neighbor (KNN), dan Random Forest. Selain itu, penelitian ini juga menerapkan metode *ensemble learning* dengan teknik hard voting untuk menggabungkan hasil prediksi dari beberapa model klasifikasi guna meningkatkan performa sistem.
3. Berdasarkan hasil pengujian menggunakan data uji, model klasifikasi menunjukkan performa yang berbeda-beda. Model Random Forest memperoleh nilai akurasi sebesar 65,71%, sedangkan metode *ensemble*

learning (hard voting) memperoleh nilai akurasi sebesar 65,00% dengan nilai F1-score sebesar 0,6918. Hasil ini menunjukkan bahwa metode *ensemble* mampu memberikan performa yang cukup baik dalam proses klasifikasi citra AI dan citra asli.

4. Berdasarkan hasil penelitian yang dilakukan, pendekatan ekstraksi fitur visual dan klasifikasi *machine learning* dapat digunakan sebagai metode untuk membantu mengidentifikasi apakah suatu citra merupakan gambar hasil generasi AI atau gambar asli.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya:

1. Penelitian selanjutnya dapat menggunakan dataset dengan jumlah yang lebih besar dan lebih beragam agar model klasifikasi dapat mempelajari pola visual secara lebih baik dan menghasilkan tingkat akurasi yang lebih tinggi.
2. Penelitian selanjutnya juga dapat menambahkan fitur visual lainnya, seperti fitur bentuk atau fitur frekuensi citra, untuk meningkatkan kemampuan model dalam membedakan citra AI dan citra asli.
3. Sistem yang telah dikembangkan dapat dikembangkan lebih lanjut dengan menambahkan fitur seperti analisis batch image, sehingga sistem dapat memproses beberapa gambar sekaligus dalam satu waktu.

DAFTAR PUSTAKA

- A. A., O. I., I. O., J. E., & M. I. (2025). Effective preprocessing techniques for improved facial recognition under variable conditions. *Franklin Open*, 4.
- A. Alajaji, S., H. Khoury, Z., Elgharib, M., Saeed, M., Ahmed, A. R., & B.Khan, M. (2024). Generative Adversarial Networks in Digital Histopathology: Current Applications, Limitations, Ethical Considerations, and Future Directions. *Modern Phatology*, 37.
- Alotaibi, A. (2025). Ensemble Deep Learning Approaches in Health Care: A Review. *computers, materials and continua*, 3771.
- Anggarwal, A. (2021). Generative adversarial network: An overview of theory and applications. *International Journal of Information Management data Insights*, 9.
- Baruggada, P., & Kalluri, H. K. (2025). Intelligence, Machine Learning, Deep Learning and Computer Vision in Food Science. *Journal of Future Foods*, 51.
- Citra Kurniawan, S. M. (2015). IMPLEMENTASI ARTIFICIAL INTELLIGENCE DALAM PENYELESAIAN MASALAH. *Jurnal Teknik*, 15.
- Corporation, M. (2024). *Visual studio code documentation*. Retrieved from visual studio code: <https://code.visualstudio.com/>
- Developers, S.-L. (2024). *Scikit-Learn: Machine Learning in Python*. Retrieved from Scikit-Learn: <https://scikit-learn.org/stable/>

- Duhan, N., & Bhatia, K. K. (2026). Exploring Image Processing Techniques for Skin Lesion Detection focusing on Segmentation and Classification: A Systematic Review. *Biomedical Signal Processing and Control*, 17.
- Fondation, P. S. (2026, 1). *Python Documentation*. Retrieved from Python.org: <https://www.python.org/>
- Hallur, S., & Gavade, A. (2025). Image feature extraction techniques: A comprehensive review. *Franklin Open*, 19.
- Indahyanti, U., Azizah, N. L., & Setiawan, H. (2022). Pendekatan Ensemble Learning Untuk Meningkatkan Akurasi. *jurnal sains dan informatika*, 10.
- J. B., V. E., B. K., E. L., & Come, J. M. (2026). Soil image classification and segmentation: A survey from deep learning, multimodal data and hybrid models. *Geodata and AI*, 16.
- J. E., Pizarro, A. G., O. B., Juan, A. d., & R. B. (2023). Analysing olive ripening with digital image RGB histogram. *Analytica Chimica Acta*, 8.
- Lenz, O. U., H. B., & C. C. (2025). A unified weighting framework for evaluating nearest neighbour. *Fuzzy Sets and Systems*, 21.
- M. O., & A. H. (2023). Classification of FIB/SEM-tomography images for highly porous multiphase material using random forest classifiers. *Journal of Power Sources*, 10.
- Maulana, F. F., & Rochmawati, N. (2019). Klasifikasi Citra Buah Menggunakan Convolutional Neural. *Journal of Informatics and Computer Science*, 5.
- Melo Milanez, K. D., & Coelho Pontes, M. J. (2014). Classification of edible vegetable oil using digital image and pattern. *Microchemical Journal*, 7.

- Melo Milarez, K. D., & Pontes, M. J. (2014). Classification of edible vegetable oil using digital image and pattern recognition techniques. *Microchemical Journal*, 16.
- Nababan, A. M., Nasution, U. R., Pandiangan, T. D., Nadi, F., Al-Khowarizmi, Budiarto, R., & Rahmat, R. F. (2024). Extreme learning machine approach on heart. *INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS*, 8.
- NumPy. (2024). *NumPy*. Retrieved from NumPy: <https://numpy.org/>
- R. B., N. D., & Bhatia, K. K. (2026). Exploring Image Processing Techniques for Skin Lesion Detection focusing on Segmentation and Classification: A Systematic Review. *Biomedical signal processing and control*, 17.
- R. S., R. G., & A. S. (2025). From image processing to artificial intelligence-driven tools: A comprehensive survey on the evolution of feature extraction methods in. *Engineering Application of Artificial Intelligence*, 14.
- S. T., V. V., M. V., S. G., Awotunde, J. B., Adeniyi, A. E., & Aroba, O. J. (2026). An enhanced diabetic retinopathy blindness detection using deep learning. *Array*, 13.
- S. V., & Sridharan, N. V. (2024). Brake fault diagnosis using a voting ensemble of machine learning classifier. *Results in Engineering*, 12.
- S. Z., Q. L., M. F., W. M., & J. F. (2025). Multi-view least squares support vector classifiers with the principles of. *Neurocomputing*, 21.
- Shofia, R., & Aulia, W. (2023). Perancangan Perangkat Makan Bersama Bertolak. *Jurnal Sains dan Teknologi*, 13.

Team, O. (2025/2026). *OpenCV Documentation*. Retrieved from OpenCV.org:

<https://opencv.org/>

team, p. d. (2024). *pandas documentation*. Retrieved from pandas:

<https://pandas.pydata.org/>

Y. L., X. W., Z. Z., & F. D. (2023). Deep learning based data augmentation for

large-scale mineral image recognition and classification. *Minerals*

Engineering , 12.

Y. w., & J. w. (2025). A survey of text classification based on pre-trained

language model. *Neurocomputing*, 16.

LAMPIRAN

Lampiran 1. SK-1 Penetapan Dosen Pembimbing



UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
 UMSU Terakreditasi Topik 1B (Sangat Baik) dan 1C (Sangat Baik) oleh Badan Penyelenggara Akreditasi Universitas
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20236 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003
 Email: umsu@umsu.ac.id umsu@umsu.ac.id umsu@umsu.ac.id umsu@umsu.ac.id

Lampiran Dosen Pembimbing Prodi Teknologi Informasi
 Nomor : 87/KEP/IL.3.AU/UMSU-09/F/2026
 Tanggal : 19 Rajab 1447 H /08 Januari 2026 M

**PENGANGKATAN DOSEN PEMBIMBING SKRIPSI
 PRODI TEKNOLOGI INFORMASI
 FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
 UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

NO	NAMA	NPM	JUDUL	DOSEN
1	Moh Hafiz Naufal	2209020194	Penerapan Dalam Membedakan Gambar Buatan Artificial Intelligence Dan Gambar Asli Dengan Ekstraksi Fitur Visual Menggunakan Algoritma Ensemble Learning	Dr. Al-Khowarizmi, M.Kom.
2	Sulthan Muhadzib Melfan	2209020149	Pengembangan Prototype Sepatu Lari Pintar Terintegrasi Aplikasi Mobile Untuk Monitoring Pace Dan Jarak Tempuh	Dr. Firahmi Rizky, M.Kom
3	Silabudiprasetya	2209020294	Sistem Akuisisi Dan Analisis Data Lingkungan Berbasis Internet Of Things	Mhd. Basri, S.Si. M.Kom
4	Alwi Azhari Hasibuan	2209020133	Perancangan Sistem Monitoring Kesehatan Kelapa Sawit Berbasis Iot Dengan Deteksi Penyakit Menggunakan Sensor Multispectral Dan Convolutional Neural Network(CNN)	Dr. Al-Khowarizmi, M.Kom.
5	Abdillah Riyanda	2209020238	Pengembangan Smart Medicine Box Menggunakan Teknologi Iot Terintegrasi Aplikasi Android	Farid Akbar Siregar S.Kom., M.Kom.
6	Yayang Ramadani	2209020112	Sistem IOT Berbasis Mikrokontroler Untuk Pengendalian Fotoperiodisme Tanaman Buah Naga Melalui Suplementasi Cahaya Malam Guna Produksi Di Luar Musim	Fatma Sari Hutagalung, S.Kom., M.Kom
7	Khairul Amri Annazri	2209020130	Perancangan Sistem Pemantauan Kualitas Air Kolam Ikan Menggunakan Teknologi IOT Berbasis Android	Farid Akbar Siregar S.Kom., M.Kom.





Dr. Al-Khowarizmi, M.Kom
NIDN : 0127099201



