KLASIFIKASI TINGKAT KEMATANGAN BUAH KELAPA SAWIT BERBASIS EKSTRAKSI FITUR CITRA DIGITAL MENGGUNAKAN ALGORITMA CATBOOST

DISUSUN OLEH

SETYA EKA ARDHINI

2209020271

PROGRAM STUDI TEKNOLOGI INFORMASI

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

MEDAN

2026

KLASIFIKASI TINGKAT KEMATANGAN BUAH KELAPA SAWIT BERBASIS EKSTRAKSI FITUR CITRA DIGITAL MENGGUNAKAN ALGORITMA CATBOOST

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer (S.Kom) dalam Program Studi Tenologi Infromasi. pada Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara

SETYA EKAARDHINI

NPM. 22092020271

PROGRAM STUDI TEKNOLOGI INFORMASI

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

MEDAN

2026

LEMBARAN PENGESAHAN

Judul Skripsi	: Klasifikasi Tingkat Kematangan Buah Kelapa Sawit Berbasis Ekstraksi Fitur Citra Digital Menggunakan Algoritma Catboost
Nama Mahasiswa	: SETYA EKA ARDHINI
NPM	2209020271
Program Studi	: TEKNOLOGI INFORMASI

Menyetujui

Komisi Pembimbing

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

NIDN. 0117019301

Ketua Program Studi **Dekan**

(Fatma Sari Hutagalung, S.Kom, M.Kom) **(Dr. Al-Khowarizmi, S.Kom., M.Kom.)**
NIDN. 0117019301 **NIDN. 0127099201**

PERNYATAAN ORISINALITAS

KLASIFIKASI TINGKAT KEMATANGAN BUAH KELAPA SAWIT BERBASIS EKSTRAKSI FITUR CITRA DIGITAL MENGGUNAKAN ALGORITMA CATBOOST

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, Maret 2026

Yang membuat pernyataan

Materai
10000

Setya Eka Ardhini

NPM. 2209020271

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama	: Setya Eka Ardhini
NPM	2209020271
Program Studi	: Teknologi Informasi
Karya Ilmiah	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bedas Royalti Non-Eksekutif *(Non-Exclusive Royalty free Right)* atas penelitian skripsi saya yang berjudul:

KLASIFIKASI TINGKAT KEMATANGAN BUAH KELAPA SAWIT BERBASIS EKSTRAKSI FITUR CITRA DIGITAL MENGGUNAKAN ALGORITMA CATBOOST.

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta. Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, Maret 2026

Yang membuat pernyataan

Setya Eka Ardhini

NPM. 2209020271

RIWAYAT HIDUP

DATAPRIBADI

Nama Lengkap	: Setya Eka Ardhini
Tempat dan Tanggal Lahir	: Laut Tador, 08 Desember 2003
Alamat Rumah	: JL Arjoe Utomo Dusun VI Cendana, Kec Laut Tador, Kab Batu Bara, Sumatera Utara
Telepon/Faks/HP	082162916448
E-mail	: setyaekaardhini512@gmail.com
Instansi Tempat Kerja	: -
Alamat Kantor	: -

DATAPENDIDIKAN

SD	: SDN 014713 Laut Tador	TAMAT: 2015
SMP	: SMPN 2 Sei Suka	TAMAT: 2018
SMA	: SMA Swasta Mitra Inalum Tanjung Gading	TAMAT: 2021

KATA PENGANTAR



Assalamualaikum Warahmatullahi Wabarakatuh

Puji Syukur penulis panjatkan ke hadirat Allah SWT karena atas rahmat dan karunia-Nya, penulis dapat menyelesaikan skripsi ini yang berjudul " Klasifikasi Tingkat Kematangan Buah Kelapa Sawit Berbasis Ekstraksi Fitur Citra Digital Menggunakan Algoritma Catboost". Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan program Sarjana Strata satu (S1) program studi Teknologi Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI), Universitas Muhammadiyah Sumatera Utara.

Dalam menyelesaikan skripsi ini penulis banyak mendapatkan bimbingan dan motivasi dari berbagai pihak yang tidak ternilai harganya. Untuk itu dalam kesempatan ini dengan ketulusan hati penulis ingin mengucapkan banyak terima kasih dan penghargaan sebesar-besarnya kepada semua pihak yang telah bersedia membantu, memotivasi, membimbing, dan mengarahkan selama proses skripsi ini. Penulis ingin mengucapkan banyak terima kasih sebesar-besarnya terutama kepada:

1. Allah SWT yang telah memberikan rahmat, hidayah, serta kekuatan kepada penulis sehingga skripsi ini dapat disusun dengan baik
2. Bapak Prof. Dr. Agussani, M.AP., selaku Rektor Universitas Muhammadiyah Sumatera Utara.
3. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom. selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.
4. Ibu Dr. Firahmi Rizky, S.Kom., M.Kom selaku Wakil Dekan 1 Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.

5. Bapak Mhd Basri, S.Si., M.Kom selaku Wakil Dekan III Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.
6. Ibu Fatma Sari Hutagalung, S.Kom., M.Kom selaku Ketua Program Studi Teknologi Informasi Fakultas Ilmu Komputer Dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.
7. Bapak Okvi Nugroho, S.Kom., M.Kom selaku Sekretaris Program Studi Teknologi Informasi Fakultas Ilmu Komputer Dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.
8. Ibu Fatma Sari Hutagalung, S.Kom., M.Kom selaku dosen pembimbing yang telah meluangkan waktu, memberikan masukan , serta memberikan semangat kepada penulis selama mengerjakan skripsi ini.
9. Ucapan terimakasih yang sebesar-besarnya penulis sampaikan kepada yang teristimewa yaitu Ayahanda Jemingan dan Ibunda Miswati atas segala doa yang tak pernah putus, kasih sayang yang tulus, serta dukungan moril dan materil tiada henti yang diberikan kepada penulis. Terimakasih atas setiap pengorbanan, kesabaran, dan cinta yang selalu menguatkan penulis dalam menyelesaikan skripsi ini. Penulis berharap pencapaian ini dapat menjadi bentuk kecil dari rasa terima kasih dan kebanggaan untuk Ayah dan Mama.
10. Penulis mengucapkan terima kasih yang sebesar-besarnya kepada seluruh keluarga tercinta atas segala doa, dukungan , perhatian dan kasih sayang yang tiada henti diberikan selama proses penyusunan skripsi ini. Kehadiran dan dukungan keluarga menjadi kekuatan utama bagi penulis dalam menghadapi setiap tantangan selama proses penyusunan skripsi.
11. Penulis mengucapkan terima kasih yang tulus kepada adik Perempuan tercinta, Almirah Dwi Antari, atas doa , dukungan, serta semangat yang selalu diberikan kepada penulis selama proses penyusunan skripsi ini. Terimakasih atas kehadiran , keceriaan, dan perhatian yang menjadi penyemangat bagi penulis. Penulis berharap dapat menjadi role model yang baik bagi adik Perempuan tercinta, dalam meraih cita-cita dan menghadapi masa depan.

12. Penulis juga mengucapkan terima kasih yang sebesar-besarnya kepada seluruh teman dan sahabat yang tidak dapat penulis sebutkan satu persatu, atas segala dukungan, bantuan, dan kebersamaan yang telah diberikan selama proses penyusunan skripsi ini. Terima kasih atas setiap semangat, motivasi, serta kenangan yang telah menjadi bagian penting dalam perjalanan penulis hingga mampu menyelesaikan skripsi ini.
13. Kepada pemilik NPM 2209020180, penulis mengucapkan terima kasih yang tulus atas segala dukungan, waktu , doa , perhatian, serta kesabaran yang telah diberikan selama proses penyusunan skripsi ini. Terimakasih karena selalu hadir memberikan semangat di setiap kondisi, menjadi tempat bertukar pikiran,dan menjadi salah satu sumber kekuatan bagi penulis dalam menghadapi berbagai tantangan.
14. Dan yang terakhir penulis ucapkan terima kasih kepada Diri Sendiri, SETYA EKA ARDHINI. Terima kasih atas segala perjuangan, ketekunan, dan kesabaran dalam menyelesaikan skripsi ini. Perjalanan ini tidak mudah, penuh dengan rasa lelah dan keraguan, namun penulis mampu bertahan dan terus melangkah hingga sampai pada titik ini.

Terima kasih karena tidak menyerah di saat keadaan terasa berat, dan terus percaya bahwa setiap proses memiliki makna dan tujuan. Setiap usaha, waktu, dan pengorbanan yang telah dilalui menjadi bagian penting dalam membentuk pribadi yang lebih kuat dan dewasa. Penulis menyadari bahwa pencapaian ini bukanlah akhir , melainkan awal dari perjalanan yang lebih panjang kedepan. Semoga apa yang telah diraih dapat menjadi pengingat bahwa kerja keras, doa , dan keteguhan hati akan selalu menemukan jalan nya. Penulis menyadari bahwa penulisan skripsi ini masih jauh dari kata sempurna , untuk itu dengan kerendahan hati penulis sangat mengharapkan kritik dan saran yang membangun guna menyempurnakan skripsi ini dari semua pihak. Wassalamualaikum Warahmatullahi Wabarakatuh.

KLASIFIKASI TINGKAT KEMATANGAN BUAH KELAPA SAWIT BERBASIS EKSTRAKSI FITUR CITRA DIGITAL MENGGUNAKAN ALGORITMA CATBOOST.

ABSTRAK

Penentuan tingkat kematangan buah kelapa sawit merupakan faktor penting dalam meningkatkan kualitas hasil produksi. Proses penentuan kematangan secara manual seringkali bersifat subjektif dan kurang konsisten. Oleh karena itu, diperlukan suatu sistem yang dapat membantu dalam mengklasifikasikan tingkat kematangan buah kelapa sawit secara otomatis dan objektif. Penelitian ini beujuan untuk membangun sistem klasifikasi tingkat kematangan buah kelapa sawit berbasis citra digital menggunakan algoritma CatBoost. Dataset yang digunakan berjumlah 1000 citra yang terbagi ke dalam empat kelas, yaitu mentah, mengkal, matang, dan terlalu matang, masing-masing sebanyak 250 citra. Tahapan yang dilakukan meliputi preprocessing citra berupa resize, normalisasi, dan konversi warna, kemudian dilanjutkan dengan ekstraksi fitur warna (RGB) dan fitur tekstur menggunakan metode Gray Level Co-occurrence Matrix (GLCM). Hasil ekstraksi fitur digunakan sebagai input dalam proses klasifikasi menggunakan algoritma CatBoost. Sistem yang dibangun berbasis website menggunakan framework Flask, sehingga memudahkan pengguna dalam melakukan klasifikasi secara langsung. Berdasarkan hasil pengujian, sistem mampu mengklasifikasikan tingkat kematangan buah kelapa sawit dengan cukup baik sesuai dengan karakteristik visual citra yang diinputkan. Dengan demikian, sistem ini dapat digunakan sebagai alat bantu dalam menentukan tingkat kematangan buah kelapa sawit secara otomatis, sehingga diharapkan dapat meningkatkan efisiensi dan objektivitas dalam proses penilaian kematangan buah.

Kata Kunci: Klasifikasi Citra, Kelapa Sawit, CatBoost, Ekstraksi Fitur, GLCM, Website

CLASSIFICATION OF RIPENESS LEVEL OF COCONUT FRUIT PALM BASED EXTRACTION DIGITAL IMAGE FEATURES USING THE CATBOOST ALGORITHM.

ABSTRACT

Determining the maturity level of oil palm fruit is an important factor in improving production quality. Manual assessment of fruit maturity is often subjective and inconsistent. Therefore, a system is needed to assist in automatically and objectively classifying the maturity level of oil palm fruit. This study aims to develop an image-based classification system for determining the maturity level of oil palm fruit using the CatBoost algorithm. The dataset used consists of 1000 images divided into four classes, namely unripe, under-ripe, ripe, and overripe, with 250 images in each class. The research process includes image preprocessing such as resizing, normalization, and color space conversion, followed by feature extraction of color (RGB) and texture using the Gray Level Co-occurrence Matrix (GLCM) method. The extracted features are then used as input for the classification process using the CatBoost algorithm. The system is developed as a web-based application using the Flask framework, allowing users to perform classification easily. Based on the testing results, the system is able to classify the maturity level of oil palm fruit effectively according to the visual characteristics of the input images. Therefore, this system can be used as a tool to assist in determining the maturity level of oil palm fruit automatically, which is expected to improve efficiency and objectivity in the assessment process.

Keywords: Image Classification, Oil Palm, CatBoost, Feature Extraction, GLCM, Web-Based System

DAFTAR ISI

LEMBAR PENGESAHAN ..... **i**
PERNYATAAN ORISINALITAS ..... **ii**
PERNYATAAN PERSETUJUAN PUBLIKASI ..... **iii**
RIWAYAT HIDUP ..... **iv**
KATA PENGANTAR ..... **v**
ABSTRAK ..... **viii**
ABSTRACT ..... **ix**
DAFTAR ISI ..... **x**
DAFTAR TABEL ..... **xv**
DAFTAR GAMBAR ..... **xvi**
BAB 1. PENDAHULUAN **1**
1.1. Latar Belakang Masalah 1
1.2. Rumusan Masalah 4
1.3. Batasan Masalah 4
1.4. Tujuan Penelitian 5
1.5. Manfaat Penelitian 6
BAB II. LANDASAN TEORI **7**
2.1. Kelapa Sawit dan Tingkat Kematangan Buah 7
2.2. Citra Digital 9
2.3. Pengolahan Citra Digital 10
2.4. Pra-Pengolahan (Pre – Processing) Citra 10
2.5. Ekstraksi Fitur Citra Digital 12
2.5.1 Fitur Warna 13
2.5.2 Fitur Tekstur 13
2.6. Pembelajaran Mesin Untuk Klasifikasi 13
2.6.1. Pembelajaran Mesin Tak Terawasi (Unsupervised Learning) 14
2.6.2. Pembelajaran Mesin Terawasi (Supervised Learning) 15
2.7. Algoritma CatBoost 15
2.7.1. Konsep Dasar Gradient Boosting 16
2.7.2. Prinsip Kerja Algoritma CatBoost 17
2.7.3. Penerapan CatBoost pada Klasifikasi Citra Digital 19
2.7.4. Parameter Utama Pada Algoritma CatBoost 20
2.7.5. Kelebihan Algoritma CatBoost 22
2.8. Evaluasi Kinerja Model CatBoost 24
2.9. Ringkasan Penelitian Terdahulu 25
2.10. Analisis GAP 28
BAB III METODOLOGI PENELITIAN **31**
3.1. Jenis dan Pendekatan Penelitian 31
3.2. Lokasi dan Waktu Penelitian 31
3.3. Tahapan Penelitian 31
3.4. Data dan Sumber Data 33
3.4.1. Data Primer 33
3.4.2. Data Sekunder 33

3.5. Teknik Pengumpulan Data 33

3.5.1. Observasi Lapangan 33
3.5.2. Dokumentasi 34
3.6. Alur PenPelitian 34
3.7. Perancangan Sistem 36
3.7.1. Use Case Diagram 36
3.7.2. Activity Diagram 37
3.7.3. Class Diagram 38
3.7.4. Sequence Diagram 38
3.8. Sketsa Perancangan Website 39
3.8.1. Tampilan Menu Utama Website 39
3.8.2. Tampilan Menu Klasifikasi Website 39
3.9. Pra-Pemrosesan (Pre-Processing) Citra 40
3.9.1. Resize Citra 40
3.9.2. Normalisasi Nilai Piksel 40
3.9.3. Konversi Ruang Warna 41
3.10. Ekstraksi Fitur 41
3.10.1. Ekstraksi Fitur Warna 41
3.10.2. Ekstraksi Fitur Tekstur 42
3.10.3. Output Ekstraksi Fitur 43
3.11. Pembagian Dataset 43
3.12. Metode Klasifikasi menggunakan Algoritma CatBoost 43
3.13. Perangkat dan Alat Penelitian 44
BAB IV HASIL DAN PEMBAHASAN 46
4.1. Hasil Pengumpulan Data 46
4.1.1. Data Pada Kelas Mentah 46
4.1.2. Data Pada Kelas Mengkal 47
4.1.3. Data Pada Kelas Matang 48
4.1.4. Data Pada Kelas Terlalu Matang 48
4.2. Hasil Implementasi Sistem 49
4.2.1. Tampilan Halaman Login 49
4.2.2. Tampilan Dasboard 50
4.2.3. Tampilan Upload Citra 50
4.2.4. Hasil Preprocessing 51
4.2.5. Hasil Ekstraksi Fitur 51
4.2.6. Hasil Klasifikasi 52
4.3. Hasil Pengujian Sistem 53
4.3.1. Pengujian Sistem 53
4.3.2. Analisis Hasil 53
4.3.3. Kinerja Sistem 54
4.4. Implementasi Sistem (coding) 54
4.4.1. Inisialisasi Sistem 54
4.4.2. Upload dan Pembacaan Citra 54
4.4.3. Ekstraksi Fitur Warna 55
4.4.4. Ekstraksi Fitur Tekstur 55
4.4.5. Klasifikasi CatBoost 56
4.5. Pembahasan 56

BAB V PENUTUP 57
5.1. Kesimpulan 57
5.2. Saran 58
DAFTAR PUSTAKA 60

DAFTAR TABEL

Tabel 2.9. Ringkasan Penelitian Terdahulu 25
Tabel 3.13. Perangkat dan Alat Penelitian.....44

DAFTAR GAMBAR

Gambar 3.2.Lokasi Penelitian 31
Gambar 3.2.Area Sortasi Pabrik Kelapa Sawit..... 32
Gambar 3.6.Alur Penelitian..... 35
Gambar 3.7.1. Use Case Diagram 36
Gambar 3.7.2. Activity Diagram 37
Gambar 3.7.3. Class Diagram..... 38
Gambar 3.7.4. Sequence Diagram..... 38
Gambar 3.8.1. Tampilan Menu Utama Website 39
Gambar 3.8.2. Tampilan Menu Klasifikasi Website..... 40
Gambar 4.1.1. Kelas Mentah..... 47
Gambar 4.1.2. Kelas Mengkal..... 47
Gambar 4.1.3. Kelas Matang..... 48
Gambar 4.1.4. Kelas Terlalu Matang..... 49
Gambar 4.2.1. Halaman Login 49
Gambar 4.2.2. Tampilan Dasboard..... 50
Gambar 4.2.3. Upload Citra 50
Gambar 4.2.4. Preprosesing Citra 51
Gambar 4.2.5. Ekstraksi Fitur 51
Gambar 4.2.6. Hasil Klasifikasi 52
Gambar 4.3.1. Hasil Pengujian Sistem..... 53
Gambar 4.4.1. Inisialisasi Sistem..... 54
Gambar 4.4.2. Upload dan Pembacaan Citra..... 54
Gambar 4.4.3. Ekstraksi Fitur Warna..... 55
Gambar 4.4.4. Ekstraksi Fitur Tekstur 55
Gambar 4.4.5. Klasifikasi Catboost..... 56

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Kelapa sawit merupakan komoditas perkebunan unggulan di Indonesia yang berperan penting dalam mendukung perekonomian nasional. Mutu minyak kelapa sawit sangat dipengaruhi oleh tingkat kematangan buah pada saat panen. Buah yang dipanen pada kondisi kematangan optimal mampu menghasilkan rendemen minyak yang lebih tinggi serta kualitas minyak yang lebih baik. Sebaliknya, pemanenan buah dalam keadaan belum matang atau terlalu matang dapat menyebabkan penurunan kualitas dan efisiensi proses produksi (Siregar et al., 2020).

Dalam praktik lapangan, penentuan tingkat kematangan buah kelapa sawit umumnya masih dilakukan secara manual melalui pengamatan visual, seperti warna permukaan buah dan jumlah brondolan yang terlepas. Pendekatan ini sangat bergantung pada pengalaman serta penilaian subjektif pekerja. Selain itu, hasil pengamatan mudah dipengaruhi oleh faktor lingkungan, termasuk variasi pencahayaan dan sudut pengamatan. Kondisi tersebut dapat menyebabkan ketidakkonsistenan dalam penentuan tingkat kematangan buah, yang berpotensi menimbulkan kesalahan panen serta kerugian bagi perusahaan perkebunan (Rahmawati et al., 2021).

Perkembangan teknologi mendorong pemanfaatan pengolahan citra digital untuk mengonversi informasi visual suatu objek menjadi fitur numerik. Pada citra buah kelapa sawit, variasi tingkat kematangan dapat direpresentasikan melalui perbedaan karakteristik warna dan tekstur pada setiap kelas kematangan. Sejumlah

penelitian menunjukkan bahwa proses ekstraksi fitur citra digital mampu. membedakan tingkat kematangan buah kelapa sawit secara efektif (Siregar et al., 2020; Putra & Nugroho, 2022). Temuan tersebut selaras dengan rumusan masalah penelitian yang menelaah pemanfaatan ekstraksi fitur citra digital dalam mengidentifikasi tingkat kematangan buah kelapa sawit. Selain proses ekstraksi fitur, pemilihan algoritma klasifikasi merupakan aspek krusial yang memengaruhi kinerja sistem. Berbagai algoritma pembelajaran mesin, seperti *K-Nearest Neighbor*, *Support Vector Machine*, dan *Convolutional Neural Network*, telah diterapkan dalam klasifikasi citra, termasuk untuk menentukan tingkat kematangan buah kelapa sawit (Ridla, 2022; Rahmawati et al., 2021). Meskipun demikian, sejumlah algoritma tersebut masih memiliki keterbatasan dalam menangani keragaman data serta berisiko mengalami *overfitting*, khususnya pada dataset dengan tingkat kompleksitas yang tinggi.

Algoritma CatBoost merupakan salah satu metode *gradient boosting* yang memiliki keunggulan dalam menangani data dengan tingkat kompleksitas tinggi. Algoritma ini mampu meningkatkan akurasi prediksi sekaligus menekan risiko *overfitting* (Prokhorenkova et al., 2018). Sejumlah penelitian terkini melaporkan bahwa CatBoost menunjukkan kinerja yang unggul pada berbagai permasalahan klasifikasi dibandingkan dengan algoritma *boosting* lainnya (Fitriani et al., 2023; Hidayat & Prasetyo, 2024). Meskipun demikian, pemanfaatan CatBoost untuk klasifikasi tingkat kematangan buah kelapa sawit berbasis citra digital masih tergolong terbatas. Kondisi tersebut membuka peluang untuk dilakukan penelitian lebih lanjut dan sejalan dengan rumusan masalah terkait kemampuan algoritma CatBoost dalam mengklasifikasikan tingkat kematangan buah kelapa sawit.

Berdasarkan permasalahan yang telah diuraikan, penelitian ini bertujuan untuk melakukan ekstraksi fitur citra digital serta membangun model klasifikasi tingkat kematangan buah kelapa sawit menggunakan algoritma CatBoost. Dalam penelitian klasifikasi tingkat kematangan buah kelapa sawit, evaluasi kinerja model merupakan tahapan krusial untuk menilai kemampuan model dalam mengelompokkan citra buah ke dalam empat kelas, yaitu mentah, mengkal, matang, dan terlalu matang. Penilaian kinerja dilakukan menggunakan beberapa metrik, meliputi akurasi, presisi, recall, dan F1-score. Akurasi digunakan untuk menggambarkan tingkat ketepatan model secara keseluruhan terhadap seluruh data citra. Sementara itu, presisi, recall, dan F1-score dihitung pada setiap kelas kematangan untuk mengukur ketepatan dan konsistensi model dalam mengklasifikasikan masing-masing kategori. Penerapan keempat metrik tersebut diharapkan menghasilkan model yang tidak hanya memiliki kinerja global yang baik, tetapi juga mampu mengenali setiap tingkat kematangan buah kelapa sawit secara optimal. Untuk menjaga fokus dan keterarahan penelitian, ruang lingkup kajian dibatasi pada penggunaan citra buah kelapa sawit dengan kelas kematangan tertentu. Variabel input yang digunakan berupa fitur warna dan tekstur, serta pengujian model dilakukan pada variasi pencahayaan dan sudut pengambilan gambar yang telah ditetapkan. Dengan pendekatan tersebut, penelitian ini diharapkan mampu menghasilkan solusi yang objektif, akurat, dan aplikatif dalam mendukung proses penentuan tingkat kematangan buah kelapa sawit di sektor perkebunan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam

penelitian ini adalah bagaimana melakukan proses pengolahan citra digital buah kelapa sawit mulai dari tahap preprocessing hingga ekstraksi fitur warna dan tekstur, bagaimana menerapkan algoritma CatBoost untuk mengklasifikasikan tingkat kematangan buah kelapa sawit ke dalam empat kategori yaitu mentah, mengkal, matang, dan terlalu matang, bagaimana mengevaluasi kinerja model klasifikasi yang dihasilkan menggunakan metrik akurasi, presisi, recall, dan F1-score, serta bagaimana mengimplementasikan hasil klasifikasi tersebut ke dalam sebuah sistem berbasis website yang dapat digunakan secara praktis.

1.3 Batasan Masalah

Adapun Batasan masalah penelitian ini sebagai berikut:

1. Penelitian hanya difokuskan pada buah kelapa sawit (TBS) dengan beberapa kelas tingkat kematangan yaitu mentah, mengkal, matang, terlalu matang sesuai kriteria yang ditetapkan oleh peneliti.
2. Penelitian ini hanya membahas identifikasi tingkat kematangan buah kelapa sawit berdasarkan citra digital, tanpa mempertimbangkan faktor fisik lain seperti berat, tekstur manual, atau kandungan kimia buah.
3. Citra buah sawit yang digunakan diperoleh melalui kamera digital / smartphone dengan resolusi tertentu dan tidak membahas perbedaan kualitas kamera secara mendalam.
4. Algoritma klasifikasi yang digunakan dalam penelitian ini dibatasi hanya pada algoritma CatBoost, tanpa melakukan perbandingan dengan algoritma machine learning lainnya.
5. Output dari penelitian ini berupa sistem aplikasi berbasis website yang digunakan untuk menampilkan hasil klasifikasi tingkat kematangan buah

kelapa sawit berdasarkan citra digital.

6. Penelitian ini dilaksanakan di PT Sumber Sawit Makmur yang berlokasi di Laut Tador.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka tujuan dari penelitian ini Adalah:

1. Mengembangkan metode klasifikasi tingkat kematangan buah kelapa sawit berbasis ekstraksi fitur citra digital yang terstruktur dan dapat diimplementasikan secara konsisten pada dataset penelitian.
2. Menguji dan mengevaluasi kinerja algoritma CatBoost dalam mengklasifikasikan tingkat kematangan buah kelapa sawit pada variasi kondisi pencahayaan dan sudut pengambilan gambar.
3. Mengimplementasikan sistem klasifikasi ke dalam aplikasi berbasis website untuk mempermudah proses identifikasi tingkat kematangan buah kelapa sawit secara otomatis.

1.5 Manfaat Penelitian

Adapun manfaat yang didapat dari penelitian ini adalah:

1. Manfaat Teoritis
    a. Memberikan gambaran tentang stabilitas performa CatBoost pada variasi pencahayaan dan sudut pengambilan gambar, yang berguna untuk pengembangan model klasifikasi citra pada bidang lain.
    b. Memberikan kontribusi terhadap pengembangan ilmu di bidang pengolahan citra digital dan machine learning, khususnya pada penerapan algoritma CatBoost untuk klasifikkasi objek berbasis fitur

citra

2. Manfaat Praktis

   a. Membantu petani atau pihak perkebunan dalam mendapatkan alat bantu penilaian tingkat kematangan buah sawit yang lebih objektif dan konsisten, dibandingkan metode pengamatan manual.

   b. Mendukung proses penentuan waktu panen yang lebih tepat, sehingga dapat meningkatkan kualitas hasil panen dan efisiensi produksi.

BAB II

LANDASAN TEORI

2.1. Kelapa Sawit dan Tingkat Kematangan Buah

Kelapa sawit (*Elaeis guineensis Jacq*.) merupakan tanaman industri penghasil minyak nabati yang memiliki peranan strategis dalam perekonomian wilayah tropis, khususnya di Indonesia dan Malaysia. Tingkat produktivitas kelapa sawit dipengaruhi oleh faktor genetik, varietas tanaman, serta sistem pengelolaan budidaya. Upaya peningkatan produktivitas tersebut terus dilakukan melalui proses seleksi dan perbaikan mutu bahan tanam (Wahyuni Eka Sari, Muslimin, Annafi Franz, & Putu Sugiartawan, 2022). Dalam praktik perkebunan, tingkat kematangan umumnya dikategorikan menjadi empat kondisi, yaitu:

1. Label 0 (Mentah / Unripe)

   Buah kelapa sawit dengan label 0 merupakan buah yang belum matang. Ciri utama buah mentah adalah warna kulit yang masih hijau atau keunguan, tekstur buah yang keras, serta belum adanya brondolan yang terlepas dari tandan. Pada kondisi ini, kandungan minyak dalam buah masih rendah sehingga belum layak untuk dipanen. Menurut (Siregar et al. 2020), buah kelapa sawit yang belum matang memiliki karakteristik warna dominan hijau dan belum menunjukkan perubahan warna menuju kematangan sehingga tidak optimal untuk produksi minyak.

2. Label 1 (Mengkal / Under-ripe)

   Label 1 menunjukkan kondisi buah mengkal atau setengah matang. Pada

tahap ini, warna buah mulai berubah dari hijau menuju kemerahan atau oranye, namun perubahan tersebut belum merata. Jumlah brondolan yang terlepas masih sedikit atau bahkan belum ada. (Rahmawati et al. 2021) menyatakan bahwa buah mengkal menunjukkan transisi warna yang belum sempurna, sehingga kualitas minyak yang dihasilkan masih belum maksimal dibandingkan buah matang.

3 Label 2 (Matang / Ripe)

Label 2 merupakan kondisi buah yang telah matang optimal. Ciri utama buah matang adalah warna merah atau oranye yang dominan dan merata pada permukaan buah, serta mulai terlepasnya brondolan secara alami dalam jumlah tertentu. Kondisi ini merupakan waktu panen terbaik karena kandungan minyak berada pada tingkat maksimum. Menurut penelitian (Putra & Nugroho 2022), buah kelapa sawit matang memiliki karakteristik warna yang jelas dan konsisten, serta menghasilkan rendemen minyak tertinggi dibandingkan tingkat kematangan lainnya.

4. Label 3 (Terlalu Matang / Over-ripe)

Label 3 menunjukkan kondisi buah yang terlalu matang. Pada tahap ini, jumlah brondolan yang terlepas sangat banyak, warna buah menjadi lebih gelap, dan kualitas minyak mulai menurun akibat proses oksidasi. (Siregar et al. 2020) menjelaskan bahwa buah yang terlalu matang mengalami penurunan kualitas karena kandungan asam lemak bebas meningkat, sehingga tidak ideal untuk proses produksi minyak kelapa sawit. Meskipun demikian, penilaian tingkat kematangan buah di lapangan umumnya masih dilakukan secara manual melalui pengamatan visual oleh tenaga panen.

Penggunaan label numerik (0–3) dalam penelitian ini bertujuan untuk mempermudah proses pelatihan model klasifikasi berbasis machine learning. Setiap label merepresentasikan kelas kematangan tertentu yang akan dipelajari oleh model berdasarkan fitur citra seperti warna dan tekstur. Seperti yang dijelaskan oleh Nanda & Prabowo (2022), dalam pembelajaran mesin terawasi, data berlabel digunakan untuk membangun model yang mampu mengenali pola dan mengklasifikasikan data baru secara otomatis berdasarkan karakteristik yang telah dipelajari. Pendekatan ini bersifat subjektif dan sangat bergantung pada pengalaman pekerja. Selain itu, hasil penilaian dipengaruhi oleh kondisi pencahayaan dan faktor lingkungan lainnya. Kondisi tersebut berpotensi menyebabkan kesalahan klasifikasi serta ketidakkonsistenan hasil panen (Nur Rali Rahma Wati, Asrul Abdullah, & Sucipto, 2025)

2.2 Citra Digital

Citra digital merupakan representasi visual dari suatu objek atau pemandangan nyata yang telah dikonversi ke dalam bentuk digital sehingga dapat diproses oleh komputer. Menurut Gonzalez dan Woods, citra digital adalah fungsi dua dimensi yang telah didiskretkan baik pada koordinat spasial maupun nilai intensitasnya, sehingga citra tersebut dapat direpresentasikan dalam bentuk matriks berhingga yang elemen-elemennya disebut piksel (Gonzalez & Woods). Setiap piksel memiliki posisi tertentu dan menyimpan nilai numerik yang merepresentasikan tingkat keabuan atau warna pada titik tersebut.

2.3 Pengolahan Citra Digital

Pengolahan citra digital merupakan cabang ilmu yang mempelajari teknik pemrosesan citra dua dimensi berbasis komputer dengan tujuan meningkatkan kualitas citra agar dapat dianalisis secara lebih efektif oleh manusia maupun sistem komputasi. Citra digital didefinisikan sebagai representasi numerik dari citra analog yang telah melalui proses digitalisasi. Dalam bentuk ini, setiap elemen citra direpresentasikan sebagai piksel yang memiliki nilai intensitas tertentu (Putri Aisyiyah Rakhma Devi & Harunur Rosyid, 2022).

Analisis warna merupakan salah satu komponen penting dalam pengolahan citra digital. Model warna yang paling banyak digunakan adalah RGB (*Red, Green,Blue*) dan HSV (*Hue, Saturation, Value*). Model RGB menggambarkan warna melalui kombinasi intensitas komponen merah, hijau, dan biru. Sementara itu, model HSV memisahkan informasi warna, tingkat kejenuhan, dan kecerahan, sehingga lebih mendekati cara manusia dalam mempersepsikan warna (Juandri & Anwar, 2023). Dalam pengolahan citra, pemilihan model warna yang tepat berperan penting dalam meningkatkan keberhasilan proses ekstraksi fitur dan klasifikasi.

2.4 Pra-Pengolahan Citra

Preprocessing citra merupakan tahap awal yang krusial dalam pengolahan citra digital sebelum dilakukan proses segmentasi, ekstraksi fitur, dan klasifikasi. Tahapan ini bertujuan untuk meningkatkan kualitas citra serta meminimalkan gangguan yang dapat memengaruhi hasil analisis pada tahap berikutnya. Menurut Salma Ashillah, Steven Adventino Gulo, Yunita Rahmi, dan Debi Yandra Niska (2025), preprocessing citra berfungsi untuk memperbaiki kualitas visual dengan

cara mereduksi noise, meningkatkan ketajaman, dan menormalkan intensitas citra agar informasi yang terkandung dapat diproses secara optimal. Secara umum, tahap ini meliputi beberapa teknik utama, seperti konversi warna, pengurangan *noise*, peningkatan kontras, dan normalisasi citra. Penerapan teknik tersebut bertujuan menghasilkan citra yang lebih stabil dan representatif terhadap objek sebenarnya.

Konversi warna merupakan proses transformasi citra dari satu ruang warna ke ruang warna lain, misalnya dari RGB ke grayscale. Proses ini bertujuan untuk menyederhanakan perhitungan komputasi dengan tetap mempertahankan informasi intensitas utama. Menurut (Ratna 2020), citra grayscale sering digunakan pada tahap awal preprocessing karena dapat menurunkan kompleksitas data tanpa menghilangkan struktur visual penting pada citra. *Noise* merupakan gangguan acak yang dapat muncul akibat proses akuisisi maupun transmisi citra. Salah satu teknik yang umum digunakan untuk mereduksi noise adalah median filter. Penelitian yang dilakukan oleh (Ashillah et al.2025) menunjukkan bahwa median filter efektif dalam mengurangi noise jenis salt and pepper tanpa menghilangkan detail penting serta tepi objek pada citra. Metode ini bekerja dengan mengganti nilai suatu piksel dengan nilai median dari piksel-piksel di sekitarnya. Dengan mekanisme tersebut, struktur citra dapat dipertahankan dengan lebih baik dibandingkan penggunaan filter rata-rata.

Peningkatan kontras bertujuan untuk mempertegas perbedaan antara objek dan latar belakang pada citra. Teknik yang umum digunakan dalam tahap ini antara lain histogram equalization dan contrast stretching, yang berfungsi memperbaiki distribusi intensitas piksel. Menurut (Muzahardin, Fauzi, dan Nurhayati 2021), penerapan peningkatan kontras pada tahap preprocessing dapat meningkatkan

kejelasan detail objek. Dengan demikian, proses segmentasi dan analisis citra pada tahap selanjutnya dapat dilakukan dengan lebih optimal. Normalisasi citra bertujuan untuk menyeragamkan nilai intensitas piksel agar berada dalam rentang tertentu, sedangkan penyesuaian ukuran (resizing) dilakukan untuk menyamakan dimensi citra. Menurut (Fajarsari et al.2022), penerapan normalisasi dan resizing merupakan langkah penting dalam menjaga konsistensi data, khususnya ketika citra digunakan sebagai masukan pada algoritma pembelajaran mesin.

2.5 Ekstraksi Fitur Citra Digital

Ekstraksi fitur citra digital merupakan tahapan penting dalam pengolahan citra yang bertujuan untuk memperoleh karakteristik utama suatu citra agar objek dapat direpresentasikan secara numerik dan informatif. Menurut (Yayah Miyantih, Yana Cahyana, dan Santi Arum Puspita Lestari 2025), proses ekstraksi fitur berperan dalam mengonversi informasi visual pada citra menjadi data kuantitatif. Data tersebut selanjutnya digunakan dalam proses klasifikasi dan pengenalan pola secara otomatis. Menurut (Indah Purnama Sari, Al-Khowarizmi dan, Ismail Hanif Batubara) Dalam proses analisis data, penggunaan variabel atau fitur dalam bentuk numerik menjadi faktor penting untuk menentukan hasil pengelompokan. Setiap data direpresentasikan oleh sejumlah atribut yang mencerminkan karakteristik objek yang dianalisis, sehingga memungkinkan sistem untuk mengidentifikasi pola dan hubungan antar data secara lebih akurat

2.5.1 Fitur Warna

Fitur warna merupakan salah satu karakteristik yang paling umum digunakan dalam analisis citra buah. Model warna RGB (*Red, Green, Blue*) dan HSV (*Hue,*

Saturation, Value) banyak dimanfaatkan karena mampu merepresentasikan informasi warna dalam bentuk numerik. Menurut (Miyantih et al.2025), ekstraksi fitur warna menggunakan model HSV lebih mendekati persepsi visual manusia dibandingkan dengan RGB, karena informasi warna, tingkat kejenuhan, dan kecerahan dipisahkan secara jelas.

2.5.2 Fitur Tekstur

Fitur tekstur merupakan salah satu karakteristik penting dalam ekstraksi fitur citra digital yang digunakan untuk merepresentasikan pola keteraturan serta hubungan spasial antar piksel pada suatu citra. Salah satu metode yang paling banyak diterapkan untuk mengekstraksi fitur tekstur adalah Gray Level Co-occurrence Matrix (GLCM). Menurut (Yayah Miyantih, Yana Cahyana, dan Santi Arum Puspita Lestari 2025), GLCM merupakan matriks yang menggambarkan frekuensi kemunculan pasangan nilai intensitas piksel tertentu pada jarak dan arah yang telah ditentukan dalam suatu citra. Penelitian yang dilakukan oleh (Miyantih et al.2025) menunjukkan bahwa fitur tekstur berbasis Gray Level Co-occurrence Matrix (GLCM) mampu menyediakan informasi tambahan yang tidak dapat diperoleh hanya dari fitur warna. Pada objek dengan perbedaan visual yang relatif halus, seperti tingkat kematangan buah, GLCM berperan dalam membedakan kelas objek berdasarkan pola permukaan dan distribusi intensitas piksel. Dengan demikian, pemanfaatan fitur tekstur berbasis GLCM dapat meningkatkan kinerja sistem klasifikasi citra.

2.6 Pembelajaran Mesin Untuk Klasifikasi

Pembelajaran mesin (*machine learning*) merupakan cabang dari kecerdasan buatan yang memungkinkan sistem komputer untuk belajar dari data dan membuat keputusan atau prediksi tanpa diprogram secara eksplisit. Dalam konteks pengolahan citra digital, pembelajaran mesin banyak digunakan untuk klasifikasi, yaitu proses pengelompokan data citra ke dalam kelas-kelas tertentu berdasarkan fitur yang telah diekstraksi**.** Menurut (R. E. Nanda dan Y. D. Prabowo 2022), klasifikasi berbasis pembelajaran mesin bertujuan untuk mengidentifikasi pola tertentu pada data citra sehingga sistem dapat mengelompokkan citra secara otomatis ke dalam kategori yang telah ditetapkan. Proses klasifikasi tersebut umumnya mencakup beberapa tahapan utama, yaitu pengumpulan data, preprocessing, ekstraksi fitur, pelatihan model, serta pengujian kinerja model klasifikasi.

2.6.1 Pembelajaran Mesin Tak Terawasi

Salah satu pendekatan pembelajaran mesin yang diterapkan dalam klasifikasi citra adalah *unsupervised learning*, yaitu metode pembelajaran yang tidak menggunakan label kelas yang telah ditentukan. (R. E. Nanda dan Y. D. Prabowo 2022) mengembangkan model klasifikasi citra lukisan dengan memanfaatkan Self-Organizing Map (SOM) menggunakan library Minisom. SOM merupakan jaringan saraf tiruan yang berfungsi memetakan data berdimensi tinggi ke dalam ruang berdimensi lebih rendah dengan tetap mempertahankan struktur topologi data. Hasil penelitian tersebut menunjukkan bahwa Self-Organizing Map (SOM) mampu mengelompokkan citra lukisan berdasarkan tingkat kemiripan fitur visual yang dimilikinya. Namun, karena menggunakan pendekatan tidak terawasi, hasil

klasifikasi sangat bergantung pada kualitas fitur yang diekstraksi serta proses interpretasi terhadap klaster yang terbentuk (Nanda & Prabowo, 2022).

2.6.2 Pembelajaran Mesin Terawasi (Supervised Learning)

Selain pendekatan tidak terawasi, *supervised learning* merupakan metode yang paling umum digunakan dalam klasifikasi citra. Pada metode ini, model dilatih menggunakan data berlabel sehingga mampu mempelajari hubungan antara fitur dan kelas target. Penelitian oleh (M. A. Ridla 2022) membandingkan beberapa algoritma pembelajaran mesin untuk klasifikasi warna kulit berdasarkan nilai piksel citra. (Ridla 2022) menyatakan bahwa algoritma supervised learning mampu mengklasifikasikan warna kulit secara efektif berdasarkan fitur warna piksel. Namun, kinerja setiap algoritma dapat berbeda, bergantung pada karakteristik data serta metode ekstraksi fitur yang diterapkan. Temuan tersebut menegaskan bahwa pemilihan algoritma pembelajaran mesin berpengaruh signifikan terhadap tingkat akurasi dan kestabilan hasil klasifikasi.

2.7 Algoritma CatBoost

CatBoost (*Categorical Boosting*) merupakan salah satu algoritma gradient boosting yang dikembangkan oleh perusahaan Yandex dan dirancang khusus untuk menangani fitur kategorikal serta mengatasi bias akibat target leakage pada proses pembelajaran model. Prinsip utama algoritma ini adalah mengoptimalkan hasil prediksi dengan cara menggabungkan beberapa model pohon keputusan lemah (weak learners) secara bertahap untuk memperkecil error residual di setiap iterasi (F. Madani & A.H. Lubis, 2025).CatBoost merupakan algoritma gradient boosting

yang telah banyak diterapkan dalam tugas klasifikasi pada berbagai domain. (Jasman dkk.2022) membandingkan CatBoost dengan beberapa metode boosting lain dalam klasifikasi kualitas air dan menemukan performa CatBoost yang unggul. Selain itu, penelitian oleh (Aditya dkk. 2025) menunjukkan efektivitas CatBoost dalam klasifikasi indeks pencemaran udara di Jakarta dengan teknik balancing data. Studi lain oleh peneliti internasional juga menegaskan bahwa CatBoost dapat memberikan akurasi tinggi pada berbagai dataset bila dibandingkan dengan algoritma serupa seperti XGBoost.

2.7.1 Konsep Dasar Gradient Boosting

Gradient Boosting merupakan metode ensemble yang membangun model secara iteratif dengan meminimalkan kesalahan residual pada setiap tahap pembelajaran (Jasman et al., 2022). CatBoost adalah implementasi *open-source* dari Gradient Boosted Decision Tree yang menggunakan struktur pohon simetris dalam pembentukan model. Algoritma ini dirancang untuk mengendalikan bias terhadap fitur kategorikal (Rizkallah, 2025; Madani & Lubis, 2025). Sejumlah penelitian juga menunjukkan bahwa CatBoost memiliki kinerja yang lebih unggul dibandingkan dengan algoritma boosting lainnya (Murtiningsih et al., 2025).

2.7.2 Prinsip Kerja Algoritma CatBoost

Menurut (M. T. Syamkalla, S. Khomsah, dan Y. S. R. Nur 2024), CatBoost menerapkan pendekatan khusus yang dikenal sebagai ordered boosting. Pada pendekatan ini, data pelatihan diacak berdasarkan urutan tertentu (permutation) sehingga model tidak memperoleh akses terhadap nilai target aktual selama proses

pembelajaran. Teknik tersebut dirancang untuk mengatasi permasalahan target leakage, yaitu kondisi ketika model mempelajari informasi yang seharusnya belum tersedia pada saat proses prediksi. (Aditya, F. R. Umbara, dan P. N. Sabrina 2025) menyatakan bahwa CatBoost memiliki keunggulan dalam menangani data kategorikal secara langsung tanpa memerlukan proses one-hot encoding atau label encoding secara manual. Kemampuan tersebut dimungkinkan melalui penerapan statistical target encoding, yaitu proses transformasi fitur kategorikal menjadi nilai numerik berdasarkan distribusi statistik target yang dilakukan secara terkontrol dan berurutan (ordered statistics). Pendekatan ini berperan dalam menurunkan variansi model serta meningkatkan kestabilan hasil prediksi.

Prinsip kerja CatBoost juga dicirikan oleh penggunaan oblivious tree structure, yaitu struktur pohon keputusan yang bersifat simetris, di mana setiap tingkat menerapkan kriteria pemisahan yang sama (S. Utomo, S. Sigit, & N. Sulistyowati, 2025). Struktur ini memberikan keuntungan berupa efisiensi komputasi yang lebih tinggi. Hal tersebut disebabkan oleh bentuk pohon yang seragam, sehingga proses evaluasi dapat dilakukan lebih cepat dibandingkan dengan pohon keputusan konvensional yang tidak seimbang. Studi yang dilakukan oleh (D. A. Murtiningsih dan B. W. Sari 2025) menunjukkan bahwa CatBoost dilengkapi dengan komponen regularization berbasis gradient descent yang berperan dalam menjaga kestabilan proses pembelajaran serta menekan risiko overfitting, khususnya pada dataset yang tidak seimbang. Proses pelatihan model tersebut dioptimalkan melalui fungsi *loss* yang diminimalkan secara bertahap dengan memanfaatkan gradien negatif dari kesalahan prediksi.

CatBoost juga sering dimanfaatkan dalam teknik ensembling bersama

algoritma lain, seperti XGBoost dan LightGBM. Menurut (A. S. Alfath dan A. K. Wardhana 2025), penerapan CatBoost dalam pendekatan stacking ensemble mampu meningkatkan akurasi prediksi risiko hipertensi hingga mencapai 99%. Temuan tersebut menunjukkan bahwa integrasi berbagai model *boosting* dapat memperkuat kemampuan generalisasi model.

Dari perspektif aplikasi, CatBoost terbukti efektif dalam berbagai bidang seperti kesehatan, keuangan, dan lingkungan**. (**N. Kadir 2024) menunjukkan bahwa CatBoost dapat memodelkan kualitas udara dengan stabil karena mampu memproses fitur numerik dan kategorikal secara simultan. Sementara itu, (M.R.P. Putra, S. Juwariyah, dan M. Ridwan 2025) menemukan bahwa prinsip kerja CatBoost semakin optimal ketika dikombinasikan dengan teknik resampling seperti SMOTE-ENN, yang membantu menyeimbangkan data dan meningkatkan akurasi prediksi kelayakan pinjaman hingga 91,67%. Secara keseluruhan, prinsip kerja CatBoost dapat diringkas menjadi lima aspek utama:

a. Ordered Boosting untuk mencegah target leakage dan bias data.
b. Oblivious Tree Structure yang menjamin efisiensi dan kestabilan model.
c. Statistical Target Encoding untuk pengolahan fitur kategorikal tanpa *manual preprocessing*.
d. Gradient Descent Optimization dalam pembaruan bobot model di setiap iterasi.
e. Regularization dan Ensemble Integration untuk meningkatkan generalisasi model dan mencegah *overfitting*.

Dengan memadukan berbagai mekanisme tersebut, CatBoost dikenal sebagai salah satu algoritma *gradient boosting* yang efisien, stabil, dan memiliki tingkat

akurasi tinggi dalam pemodelan data tabular. Keunggulan ini terutama terlihat pada dataset yang mengandung fitur kategorikal dalam jumlah besar (I. Maulana, A. M. Siregar, & S. A. P. Lestari, 2024).

2.7.3 Penerapan CatBoost pada Klasifikasi Citra Digital

CatBoost telah banyak digunakan dalam berbagai permasalahan klasifikasi yang memanfaatkan fitur numerik. (Madani dan Lubis 2025) menyatakan bahwa algoritma ini mampu mempelajari hubungan nonlinier antarfitur tanpa memerlukan proses normalisasi data yang rumit. Selain itu, (Aditya et al. 2025) menunjukkan bahwa CatBoost memiliki kemampuan generalisasi yang baik pada data dengan variasi kondisi lingkungan. Dalam konteks citra digital, fitur warna dan tekstur terlebih dahulu direpresentasikan sebagai vektor numerik sebelum proses klasifikasi dilakukan (Rahma Wati et al., 2025).

2.7.4 Parameter Utama pada Algoritma CatBoost

Algoritma CatBoost memiliki berbagai parameter yang berpengaruh signifikan terhadap kinerja serta kemampuan generalisasi model. Parameter tersebut berfungsi mengatur proses pembelajaran, tingkat kompleksitas struktur pohon keputusan, dan metode evaluasi performa. Oleh karena itu, penentuan nilai parameter yang tepat menjadi hal penting agar model mampu menghasilkan klasifikasi yang akurat dan konsisten, terutama pada data yang memiliki variasi kondisi lingkungan dan pencahayaan (Jasman et al., 2022; Ptr et al., 2024; Madani & Lubis, 2025).

a. Iterations

Parameter *iterations* merepresentasikan jumlah pohon keputusan yang dibangun secara bertahap selama proses pelatihan model. Peningkatan nilai

iterations memungkinkan model membentuk lebih banyak pohon sehingga pola pada data dapat dipelajari secara lebih mendalam. Namun demikian, jumlah iterasi yang terlalu tinggi berpotensi menimbulkan overfitting serta meningkatkan beban komputasi (Jasman et al., 2022).

b. Learning Rate

Learning rate menentukan besarnya kontribusi setiap pohon keputusan dalam memperbarui model. Nilai learning rate yang rendah mendukung proses pembelajaran yang lebih stabil, namun memerlukan jumlah iterasi yang lebih besar. Sebaliknya, nilai yang terlalu tinggi dapat mempercepat proses pelatihan, tetapi berisiko menurunkan kemampuan generalisasi model (Madani & Lubis, 2025*)*.

c. Depth

Parameter *depth* menetapkan batas kedalaman maksimum pohon keputusan yang dibangun dalam model. Nilai *depth* memiliki pengaruh langsung terhadap tingkat kompleksitas model serta kemampuannya dalam menangkap hubungan nonlinier antar fitur. Penentuan nilai *depth* yang sesuai dapat meningkatkan akurasi klasifikasi tanpa memperbesar risiko *overfitting* (Aditya et al., 2025).

d. Loss Function

Pada algoritma *gradient boosting* seperti CatBoost, mekanisme pembelajaran difokuskan pada upaya meminimalkan *loss function* yang digunakan, misalnya *log-loss* pada tugas klasifikasi. Pendekatan ini memungkinkan model menurunkan kesalahan prediksi secara bertahap. (Ptr et al. 2024) menunjukkan bahwa proses optimasi terhadap *loss function* berperan penting dalam menentukan kinerja akhir model pada pengklasifikasian data dunia nyata.

e. Evaluation Metric

Evaluation metric digunakan untuk menilai kinerja model selama tahap pelatihan dan pengujian. Metrik yang lazim diterapkan antara lain accuracy, precision, recall, dan F1-score. Pemilihan metrik yang tepat perlu disesuaikan dengan tujuan penelitian serta karakteristik dataset yang digunakan (Murtiningsih et al., 2025).

Secara umum, pemilihan serta pengaturan parameter pada algoritma CatBoost memiliki peran penting dalam meningkatkan akurasi dan kemampuan generalisasi model untuk mengklasifikasikan tingkat kematangan buah kelapa sawit. Oleh sebab itu, proses penentuan parameter harus dilakukan secara teliti agar model yang dihasilkan dapat bekerja secara optimal pada berbagai kondisi citra (Ptr et al., 2024).

2.7.5 Kelebihan Algoritma CatBoost

Adapun kelebihan Algoritma CatBoost Adalah sebagai berikut:

a. Kemampuan mengolah data kategorikal secara otomatis tanpa encoding manual

Keunggulan utama CatBoost terletak pada kemampuannya dalam menangani fitur kategorikal secara langsung tanpa memerlukan Teknik pengodean manual, seperti one-hot encoding atau label encoding. Mekanisme ini memungkinkan representasi data tetap terjaga secara alami serta meminimalkan risiko terjadinya data leakage yang kerao dijumpai pada algoritma lain, seperti XGBoost(Madani & Lubis, 2025).

b. Struktur pohon simetris (oblivious trees) yang efisien dan stabil.

Selain itu, (Utomo, Sigit, dan Sulistyowati 2025) menyatakan bahwa CatBoost menerapkan struktur pohon keputusan simetris yang dikenal

sebagai oblivious trees. Struktur ini meningkatkan efisiensi proses pelatihan sekaligus menjaga kestabilan hasil prediksi.Pada setiap pohon, seluruh node menggunakan kriteris pemisahan yang sama, sehingga waktu komputasi dapat dipercepat tanpa menurunkan Tingkat akurasi model.

c. Performa tinggi pada data tidak seimbang dan non-linear.

(Royana, Via, dan Putra 2025) juga mengemukakan keunggulan CatBoost dalam penelitian mengenai prediksi customer churn pada layanan musik digital. Hasil penelitian menunjukkan bahwa CatBoost lebih efektif dalam menangani data tidak seimbang serta fitur kategorikal yang kompleks dibandingkan LightGBM. Keunggulan tersebut tercermin dari peningkatan akurasi dalam mengklasifikasikan perilaku pelanggan.

d. Kecepatan pelatihan berkat optimasi *ordered boosting*.

(Marwan dan Setiawan 2024) mengungkapkan bahwa penggunaan *oblivious tree* pada algoritma CatBoost memberikan keunggulan dalam efisiensi proses pelatihan serta penghematan memori komputasi. Selain itu, struktur tersebut mampu menjaga kestabilan hasil prediksi, khususnya pada pengolahan dataset berskala besar, seperti sistem rekomendasi anime berbasis MyAnimeList.

e. Kemampuan generalisasi kuat yang mencegah *overfitting*.

(Alfath dan Wardhana 2025), menyatakan bahwa CatBoost memiliki tingkat stabilitas model yang lebih baik dalam mengatasi overfitting. Pada penelitian mengenai prediksi risiko hipertensi, CatBoost mampu mencapai akurasi yang tinggi pada data medis. Selain itu, algoritma ini menunjukkan kemampuan generalisasi yang lebih unggul dibandingkan dengan XGBoost

dan LightGBM.

f. Fleksibilitas tinggi dalam integrasi dengan model *ensemble* lain seperti XGBoost dan LightGBM. **(**Nganro, 2024) mendukung temuan sebelumnya dengan menunjukkan bahwa penerapan CatBoost dapat meningkatkan kestabilan model *ensemble* ketika dikombinasikan melalui pendekatan *stacking* dan *voting* bersama XGBoost dan LightGBM. Dalam skema tersebut, CatBoost berkontribusi menjaga konsistensi hasil prediksi dengan tingkat variansi yang rendah pada dataset lintas domain.

2.8 Evaluasi Kinerja Model CatBoost

Evaluasi kinerja pada model klasifikasi bertujuan untuk menilai tingkat ketepatan model dalam menghasilkan prediksi terhadap data yang diuji. Pada penerapan algoritma CatBoost, proses evaluasi digunakan untuk mengukur efektivitas model dalam melakukan klasifikasi sekaligus mengidentifikasi variabel yang memiliki pengaruh dominan terhadap hasil prediksi. Penilaian kinerja dilakukan menggunakan metrik utama, meliputi akurasi, presisi, *recall*, dan *F1-score*, yang secara komprehensif mencerminkan kemampuan model dalam mengklasifikasikan data dengan benar (Madani & Lubis, 2025).

Algoritma CatBoost memiliki performa yang unggul dalam berbagai domain klasifikasi karena:

a. Menghasilkan nilai akurasi dan F1-score yang tinggi, yang mencerminkan keseimbangan optimal antara ketepatan prediksi dan kemampuan model dalam mengenali seluruh kelas secara menyeluruh.
b. Menunjukkan nilai presisi dan recall yang konsisten , sehingga mengindikasikan kemampuan generalisasi model yang baik terhadap data

yang belum pernah dilatih.

c. Mampu menangani data tidak seimbang secara efektif melalui penerapan mekanisme ordered boosting dan teknik penyeimbangan data.
d. Memiliki Tingkat efisiensi komputasi yang tinggi dengan variansi prediksi yang rendah, sehingga sesuai untuk diterapkan pada system klasifikasi dengan skala data yang besar.

Dengan demikian, CatBoost terbukti menjadi algoritma boosting yang memiliki kinerja evaluasi yang kuat, stabil, dan adaptif untuk berbagai jenis dataset dan domain penelitian.

2.9 Ringkasan Peneliti Terdahulu

Penelitian sebelumnya terkait klasifikasi berbasis citra digital telah banyak dikembangkan, terutama pada sektor pertanian dan perkebunan. Sejumlah studi memanfaatkan fitur warna dan tekstur untuk mengidentifikasi tingkat kematangan buah dengan menerapkan berbagai algoritma pembelajaran mesin. Ringkasan penelitian terdahulu yang relevan dengan penelitian ini disajikan dalam Tabel berikut.

Tabel 2.9. Ringkasan Penelitian Terdahulu

No	Judul dan Peneliti	Pembahasan	Metode	Kelebihan dan Kekurangan
1	Perbandingan Algoritma Pembelajaran Mesin untuk Klasifikasi Warna	Penelitian ini membahas klasifikasi warna kulit manusia berbasis citra	Pengolahan Citra Digital, Ekstraksi fitur warna RGB, Klasifikasi menggunakan	**Kelebihan:** Fitur Warna RGB untuk membedakan kelas warna **Kekurangan:**

	Kulit Berdasarkan Warna Piksel Citra Ridla (2022).	digital dengan membandingkan kinerja algoritma KNN dan SVM menggunakan fitur warna piksel.	KNN dan SVM	Tidak menggunakan fitur tekstur dan bentuk dan sensitif terhadap variasi pencahayaan.
2	Klasifikasi Tingkat Kematangan Buah Kelapa Sawit Menggunakan Fitur Warna dan Tekstur (Siregar et.al 2020)	Penelitian ini mengkaji klasifikasi tingkat kematangan buah kelapa sawit berbasis citra digital dengan memanfaatkan perubahan warna dan tekstur sebagai indikator utama.	Ekstraksi Fitur warna dan tekstur, menggunakan Algoritma SVM.	**Kelebihan:** Metode cukup stabil pada dataset terbatas **Kekurangan:** Hanya menggunakan satu algoritma klasifikasi.
3	Klasifikasi Tingkat Kematangan Buah Kelapa Sawit	Penelitian ini menerapkan CNN untuk mengklasifikasikan Tingkat	Menggunakan CNN.	**Kelebihan:** Akurasi tinggi, Ekstraksi fitur otomatis **Kekurangan:**

	Menggunakan Convolutional Neutral Network (Rahma Wati et.al 2025)	kematangan buah kelapa sawit secara otomatis. CNN mampu mengekstraksi fitur warna dan tekstur secara langsung dari citra tanpa proses manual.		Membutuhkan dataset besar.
4	Ekstraksi Fitur Citra Digital untuk Identifikasi Tingkat Kematangan Buah Berbasis Computer Vision (Claudia Delvia Pangati et al. 2025)	Penelitian ini berfokus pada proses ekstraksi fitur citra digital untuk mengidentifikasi Tingkat kematangan buah. Penelitian menekankan pentingnya representasi fitur visual yang kuat untuk klasifikasi	Ekstraksi fitur warna, tekstur, dan bentuk dan Klasifikasi berbasis vektor fitur.	**Kelebihan:** Mampu menangkap karakteristik visual secara detail dan cocok untuk berbagai jenis buah **Kekurangan:** Tidak membahas algoritma klasifikasi secara mendalam dan belum menggunakan algoritma ensemble

		kematangan.		modern.
5	Identifikasi Gejala Penyakit Tanaman Jeruk melalui Preprocessing, FFT & LBP Maura Widyaningsih & Agus Harjoko (2021)	Penelitian ini membahas identifikasi penyakit tanaman jeruk berdasarkan citra daun. Tekstur dan pola frekuensi citra digunakan sebagai indikator utama gejala penyakit.	Preprocessing citra (resize, clipping, sharpening, grayscale), Ekstraksi fitur FFT dan LBP dan Klasifikasi menggunakan KNN.	**Kelebihan:** Preprosesing meningkatkan kualitas citra dan Fitur dapat efektif membedakan kelas **Kekurangan:**Tidak menggunakan fitur warna dan Algoritma KNN kurang optimal untuk data besar.

2.10 Analisis Gap

Berdasarkan kajian terhadap beberapa penelitian terdahulu, dapat diidentifikasi beberapa celah penelitian (*research gap*) yang menjadi dasar dilakukannya penelitian ini. Penelitian yang dilakukan oleh (Ridla 2022) dan (Siregar et al. 2020) menunjukkan bahwa fitur warna dan tekstur citra digital efektif digunakan dalam klasifikasi objek visual. Namun, algoritma yang diterapkan masih didominasi oleh metode konvensional, seperti KNN dan SVM. Metode tersebut memiliki keterbatasan dalam memodelkan hubungan nonlinier yang kompleks antar fitur, terutama ketika jumlah fitur meningkat dan data dipengaruhi oleh variasi pencahayaan serta kondisi lingkungan.

Penelitian lain oleh (Rahma Wati et al. 2025) menerapkan Convolutional

Neural Network (CNN) untuk mengklasifikasikan tingkat kematangan buah kelapa sawit dan memperoleh tingkat akurasi yang tinggi. Meskipun demikian, pendekatan berbasis CNN memerlukan dataset berukuran besar, sumber daya komputasi yang tinggi, serta proses pelatihan yang kompleks. Kondisi ini menjadikannya kurang efisien untuk sistem klasifikasi berbasis fitur numerik dengan keterbatasan data dan perangkat. Sementara itu, studi yang dilakukan oleh (Claudia Delvia Pangati et al. 2025) serta (Widyaningsih dan Harjoko, 2021) membuktikan bahwa kombinasi fitur warna, tekstur, dan bentuk yang diekstraksi menggunakan teknik computer vision mampu merepresentasikan karakteristik visual objek secara efektif. Namun, penelitian tersebut belum mengeksplorasi pemanfaatan algoritma boosting modern yang memiliki kemampuan generalisasi lebih baik dan lebih stabil terhadap variasi data citra. Selain itu, belum banyak penelitian yang secara khusus mengkaji penerapan algoritma CatBoost untuk klasifikasi tingkat kematangan buah kelapa sawit berbasis ekstraksi fitur citra digital, terutama dalam konteks pengaruh parameter CatBoost terhadap kinerja model. Padahal, CatBoost memiliki keunggulan dalam menangani fitur numerik, mengurangi *overfitting,* serta mempelajari hubungan nonlinier antar fitur tanpa memerlukan proses normalisasi data yang kompleks. Penelitian ini mengisi celah penelitian dengan mengusulkan pendekatan klasifikasi tingkat kematangan buah kelapa sawit berbasis ekstraksi fitur warna, tekstur, dan bentuk dari citra digital, yang ke6mudian diklasifikasikan menggunakan algoritma CatBoost. Pendekatan ini diharapkan mampu menghasilkan model klasifikasi yang akurat, stabil, dan efisien, serta menjadi alternatif yang lebih optimal dibandingkan metode konvensional dan deep learning pada kondisi dataset terbatas.

BAB III

METODOLOGI PENELITIAN

3.1. Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan metode kuantitatif dengan pendekatan eksperimental Pendekatan tersebut diterapkan untuk mengevaluasi kinerja algoritma CatBoost dalam mengklasifikasikan tingkat kematangan buah kelapa sawit berdasarkan fitur yang diekstraksi dari citra digital. Proses penelitian mencakup tahapan pengolahan citra dan penerapan pembelajaran mesin guna menghasilkan model klasifikasi yang optimal.

3.2 Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan di area pabrik kelapa sawit tepatnya di PT Sumber Sawit Makmur PKS Laut Tador, Dusun PKS, Desa Laut Tador, Kec Laut Tador, Kab Batubara.



Gambar 3.2 Lokasi Penelitian

tepatnya pada bagian sortasi, yaitu tempat pengumpulan dan pemilahan buah kelapa sawit sebelum dilakukan proses pengolahan. Pemilihan lokasi penelitian ini didasarkan pada ketersediaan buah kelapa sawit dalam jumlah yang cukup serta variasi tingkat kematangan yang beragam, sehingga sesuai dengan kebutuhan penelitian.



Gambar 3.2 Area Sortasi Pabrik Kelapa Sawit

Waktu pelaksanaan penelitian, khususnya pada tahap pengambilan data di lapangan, dilakukan pada tanggal 4 Maret 2026 sampai dengan 8 Maret 2026.

3.3 Tahapan Penelitian

Penelitian diawali dengan pengumpulan data berupa citra digital buah kelapa sawit yang diperoleh melalui observasi dan dokumentasi di lapangan. Citra yang terkumpul kemudian melalui tahap pra-pemrosesan untuk menyeragamkan dan meningkatkan kualitas citra. Selanjutnya dilakukan ekstraksi fitur warna dan tekstur sebagai informasi utama dalam proses klasifikasi. Data hasil ekstraksi kemudian

dibagi menjadi data latih dan data uji. Data latih digunakan untuk melatih model menggunakan algoritma CatBoost, sedangkan data uji digunakan untuk menguji performa model. Tahap akhir adalah evaluasi kinerja model menggunakan metrik akurasi, presisi, recall, dan F1-score untuk menilai kemampuan model dalam mengklasifikasikan tingkat kematangan buah kelapa sawit.

3.4 Data dan Sumber Data

3.4.1 Data Primer

Data primer diperoleh melalui pengambilan gambar buah kelapa sawit secara langsung di lapangan menggunakan kamera. Namun, dalam proses pengumpulan data tersebut terdapat beberapa keterbatasan, seperti kondisi lingkungan dan ketersediaan objek, sehingga jumlah data yang terkumpul belum mencukupi untuk kebutuhan penelitian.

3.4.2 Data Sekunder

Untuk mengatasi keterbatasan tersebut, peneliti menambahkan data sekunder yang diperoleh dari platform Kaggle. Data ii dimanfaatkan sebagai pelengkap untuk meningkatkan jumlah serta variasi data, sehingga dapat mendukung peningkatan kualitas dataset dalam proses pelatihan model klasifikasi.

3.5 Teknik Pengumpulan Data

3.5.1 Observasi Lapangan

Observasi lapangan dilakukan untuk memperoleh gambaran langsung mengenai kondisi buah kelapa sawit dan lingkungan pengambilan data. Kegiatan ini dilakukan di area kebun atau tempat pengumpulan hasil panen pada perusahaan kelapa sawit. Observasi difokuskan pada pengamatan karakteristik buah kelapa

sawit pada berbagai tingkat kematangan serta kondisi pencahayaan saat pengambilan citra.

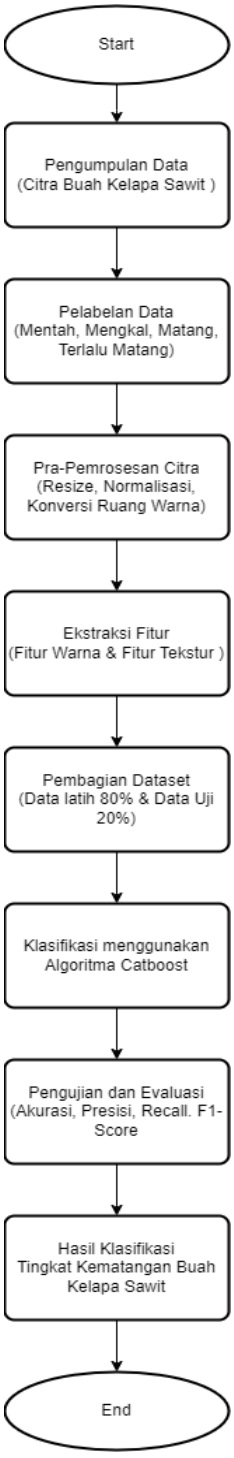
Hasil observasi digunakan sebagai dasar dalam menentukan prosedur pengambilan citra dan pelabelan data, sehingga data yang diperoleh sesuai dengan kondisi nyata di lapangan dan dapat digunakan dalam proses klasifikasi tingkat kematangan buah kelapa sawit.

3.5.2. Dokumentasi

Dokumentasi digunakan sebagai teknik pengumpulan data untuk memperoleh citra digital buah kelapa sawit. Proses ini dilakukan melalui pengambilan foto secara langsung di lapangan menggunakan kamera digital atau kamera pada perangkat smartphone. Seluruh citra yang dihasilkan disimpan dalam format digital dan diklasifikasikan berdasarkan tingkat kematangan buah. Data dokumentasi berupa citra digital tersebut digunakan sebagai dataset penelitian yang selanjutnya diproses melalui tahap pra-pemrosesan, ekstraksi fitur, dan klasifikasi menggunakan algoritma CatBoost. Selain pengambilan data secara langsung di lapangan, penelitian ini juga memanfaatkan dataset dari Kaggle sebagai sumber data tambahan untuk melengkapi keterbatasan data.

3.6 Alur Penelitian

Berikut ini merupakan tahapan yang dilakukan pada penelitian ini dan digambarkan melalui diagram berikut ini :



Gambar 3.6 Alur Penelitian

Alur penelitian menggambarkan tahapan penelitian yang dimulai dari pengumpulan data citra buah kelapa sawit, pra-pemrosesan citra, ekstraksi fitur, pembagian dataset, klasifikasi menggunakan algoritma CatBoost, hingga tahap

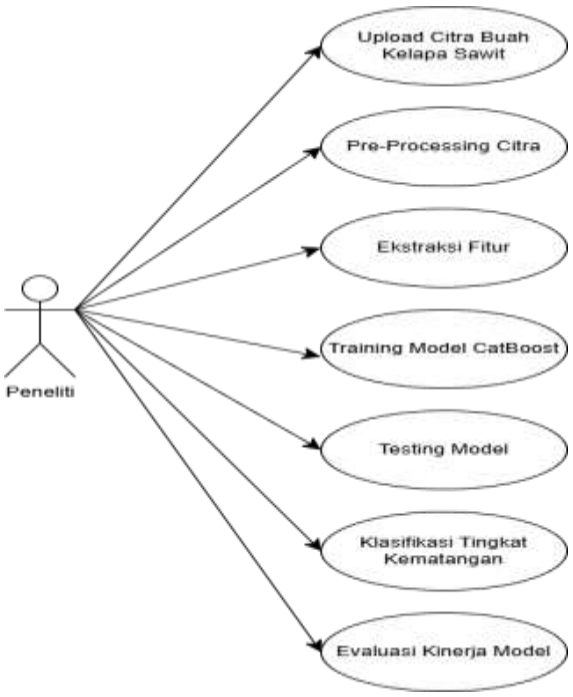
pengujian dan evaluasi untuk memperoleh hasil klasifikasi tingkat kematangan buah.

3.7 Perancangan Sistem

Untuk memodelkan perancangan sistem secara visual, penelitian ini menggunakan Unified Modeling Language (UML) sebagai alat bantu pemodelan sistem.

3.7.1 Use Case Diagram

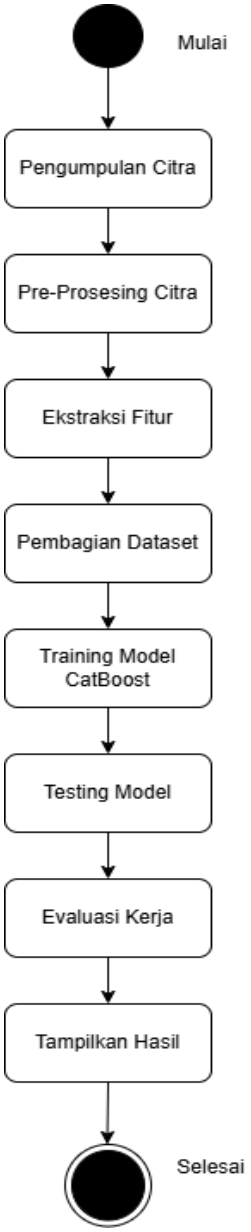
Use Case Diagram digunakan untuk menggambarkan interaksi antara pengguna dengan sistem klasifikasi tingkat kematangan buah kelapa sawit berbasis website. Diagram ini menunjukkan fungsi-fungsi utama yang dapat dilakukan oleh pengguna terhadap system.



Gambar 3.7.1. Use Case Diagram

3.7.2 Activity Diagram

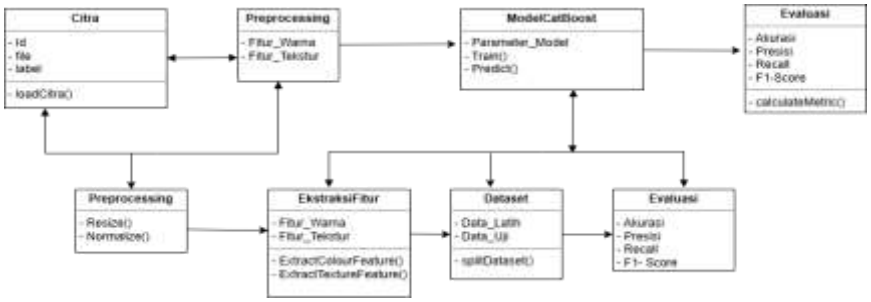
Activity Diagram digunakan untuk menggambarkan alur aktivitas sistem dalam proses klasifikasi tingkat kematangan buah kelapa sawit, mulai dari pengguna mengunggah citra hingga sistem menampilkan hasil klasifikasi.



Gambar 3.7.2. Activity Diagram

3.7.3 Class Diagram

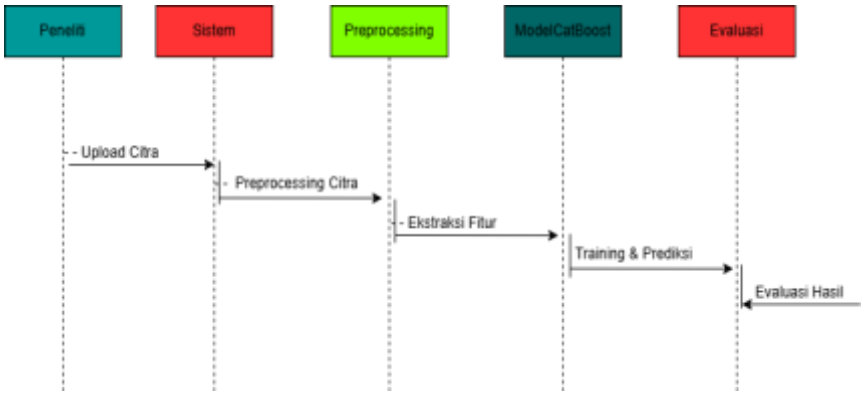
Class Diagram digunakan untuk menggambarkan struktur sistem klasifikasi tingkat kematangan buah kelapa sawit, termasuk kelas, atribut, metode, serta hubungan antar kelas.



Gambar 3.7.3. Class Diagram

3.7.4 Sequence Diagram

Sequence Diagram digunakan untuk menggambarkan urutan interaksi antara pengguna dan sistem dalam proses klasifikasi tingkat kematangan buah kelapa sawit berbasis website.



Gambar 3.7.4. Sequence Diagram

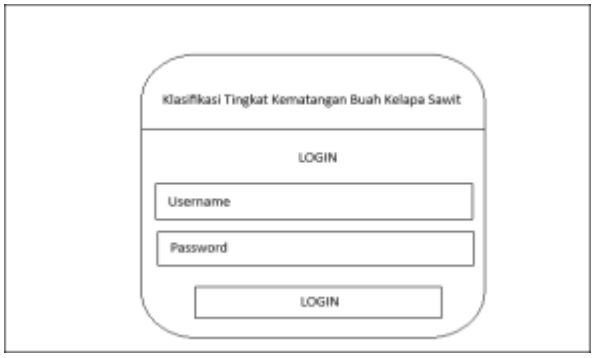
3.8 Sketsa Perancangan Website

Sketsa perancangan website merupakan representasi awal dari desain antarmuka sistem yang disusun sebelum tahap implementasi. Sketsa ini umumnya disajikan dalam bentuk wireframe sederhana yang menunjukkan tata letak

komponen utama pada setiap halaman website, seperti menu dashboard, pengunggahan citra, kolom input, dan area penampilan hasil.

3.8.1 Tampilan Menu Utama Website

Berikut merupakan tampilan utama website klasifikasi Tingkat kematangan buah kelapa sawit.



Gambar 3.8.1. Tampilan Menu Utama Website

3.8.2 Tampilan Menu Klasifikasi Website

Berikut merupakan tampilan klasifikasi website klasifikasi Tingkat kematangan buah kelapa sawit.



Gambar 3.8.2. Tampilan Menu Klasifikasi Website

3.9 Pra-Pemrosesan (Pre-Processing) Citra

Pra-pemrosesan citra merupakan tahap awal dalam pengolahan citra digital

yang bertujuan untuk meningkatkan kualitas citra serta menyeragamkan data sebelum dilakukan ekstraksi fitur. Tahapan pra-pemrosesan yang dilakukan dalam penelitian ini adalah sebagai berikut:

3.9.1 Resize Citra

Setiap citra buah kelapa sawit diubah ukurannya (resize) ke dimensi yang sama untuk menyeragamkan ukuran citra dan mengurangi kompleksitas komputasi. Proses resize dilakukan tanpa mengubah rasio aspek secara signifikan agar karakteristik visual objek tetap terjaga.

3.9.2 Normalisasi Nilai Piksel

Normalisasi dilakukan untuk menyeragamkan rentang nilai piksel citra. Proses ini bertujuan agar perbedaan intensitas warna akibat pencahayaan dapat diminimalkan sehingga fitur yang diekstraksi menjadi lebih stabil.

3.9.3 Konversi Ruang Warna

Citra buah kelapa sawit dikonversi dari ruang warna RGB ke ruang warna lain yang lebih representatif, seperti HSV atau Grayscale, sesuai kebutuhan ekstraksi fitur. Konversi ini bertujuan untuk memudahkan pemisahan karakteristik warna dan intensitas cahaya.

3.10 Ekstraksi Fitur

Ekstraksi fitur merupakan tahap untuk mengambil informasi penting dari citra buah kelapa sawit yang digunakan sebagai dasar dalam proses klasifikasi tingkat kematangan. Pada penelitian ini, fitur yang digunakan meliputi fitur warna dan fitur tekstur, karena kedua fitur tersebut mampu merepresentasikan perubahan visual buah kelapa sawit seiring tingkat kematangannya.

3.10.1 Ekstraksi Fitur Warna

Fitur warna diekstraksi untuk merepresentasikan karakteristik warna kulit buah kelapa sawit yang berubah dari mentah hingga matang. Proses ini dilakukan dengan menghitung nilai statistik dari setiap kanal warna pada citra, seperti mean (rata-rata) dan standar deviasi.

Rumus nilai rata-rata (mean) warna :

$$\mu = \frac{1}{N}\sum_{i=1}^{N} \mathcal{P}i \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (1)$$

Keterangan:

a) μ = nilai rata-rata warna

b) pi = nilai piksel ke-i

c) N = jumlah total piksel

Rumus Standar Deviasi:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(pi - \mu)^2} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (2)$$

3.10.2 Ekstraksi Fitur Tekstur

Fitur tekstur diekstraksi untuk menggambarkan pola permukaan buah kelapa sawit. Metode ekstraksi tekstur yang digunakan mampu merepresentasikan tingkat kekasaran dan pola permukaan buah.

a. Contrast = Mengukur perbedaan intensitas piksel dalam citra

$$\sum_{i,j}(i-j)^2\ p(i,j) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (1)$$

b. Homogenity = Mengukur keseragaman distribusi piksel.

$$\sum_{i,j}(i-j)^2\ p(i,j) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (2)$$

c. Energi = Mengukur Tingkat keseragaman tekstur.

$$\sum_{i,j} P(i,j)^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

d. Correlation = Mengukur hubungan linier antar piksel.

$$\sum_{ij} \frac{(i-\mu i)(j-\mu i)P\ (i,j)}{\sigma i \sigma i} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Keterangan:

a) P(i,j) = probabilitas kemunculan pasangan piksel

b) μ = nilai rata-rata

c) σ = standar deviasi

3.10.3 Output Ekstraksi Fitur

Hasil ekstraksi fitur warna dan tekstur kemudian digabungkan menjadi vektor fitur numerik, yang selanjutnya digunakan sebagai data masukan (variabel X) dalam proses klasifikasi menggunakan algoritma CatBoost.

3.11 Pembagian Dataset

Dataset citra buah kelapa sawit yang telah melalui tahap pra-pemrosesan dan ekstraksi fitur dibagi menjadi data latih dan data uji untuk keperluan pelatihan dan pengujian model klasifikasi menggunakan algoritma CatBoost. Pembagian dataset dilakukan dengan perbandingan 80% data latih dan 20% data uji secara acak untuk menghindari bias dan menjaga distribusi data pada setiap kelas kematangan. Data latih digunakan untuk melatih model, sedangkan data uji digunakan untuk mengevaluasi kinerja model dalam mengklasifikasikan tingkat kematangan buah kelapa sawit.

3.12 Metode Klasifikasi Menggunakan Algoritma CatBoost

Penelitian ini menggunakan algoritma CatBoost sebagai metode klasifikasi

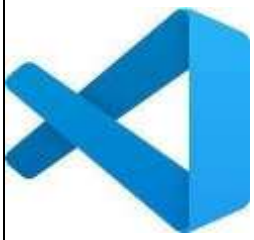
untuk menentukan tingkat kematangan buah kelapa sawit berdasarkan fitur citra yang telah diekstraksi. Model CatBoost dilatih menggunakan data latih dan diuji menggunakan data uji untuk mengukur kemampuannya dalam mengklasifikasikan data baru. Pemilihan algoritma CatBoost didasarkan pada kemampuannya dalam menangani data numerik, mengurangi overfitting, serta menghasilkan performa klasifikasi yang baik. Hasil klasifikasi selanjutnya digunakan sebagai dasar dalam evaluasi kinerja model.

3.13 Perangkat dan Alat Penelitian

Perangkat dan tools penelitian digunakan untuk mendukung seluruh proses penelitian, mulai dari pengumpulan data hingga evaluasi kinerja model klasifikasi. Perangkat yang digunakan dalam penelitian ini terdiri dari perangkat keras (hardware) dan perangkat lunak (software).

Tabel 3.13 Perangkat dan Alat Penelitian

No	Alat	Kegunaan
1	Smartphone / HP 	Digunakan untuk mengambil citra buah kelapa sawit secara langsung di lapangan sebagai data utama penelitian
2	Laptop/Komputer 	Digunakan untuk melakukan proses pengolahan citra, ekstraksi fitur, pelatihan model, pengujian model, dan evaluasi hasil klasifikasi

3	Vs Code 	penulisan kode (syntax highlighting), debugging, integrasi Git (GitHub), serta kustomisasi melalui ribuan ekstensi untuk mempercepat alur kerja developer.
4	Python 	Digunakan sebagaiBahasa pemrograman utama dalam pengolahan citra digital dan implementasi algoritma CatBoost.
5	CatBoost 	Digunakan sebagai algoritma klasifikasi Tingkat kematangan buah kelapa sawit.
6	Draw.io 	Digunakan untuk menggambarkan alur dan tahapan penelitian secara sistematis.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Implementasi Sistem

Data yang digunakan dalam penelitian ini diperoleh dari dua sumber, yaitu data primer dan data sekunder. Data primer diperoleh melalui pengambilan citra secara langsung di area sortasi pabrik kelapa sawit, yaitu tempat pengumpulan dan pemilahan buah sebelum proses pengolahan. Pada area sortasi tersebut tersedia buah kelapa sawit dengan berbagai tingkat kematangan, seperti mentah, mengkal, matang, dan terlalu matang. Pengambilan citra dilakukan menggunakan kamera dengan tujuan memperoleh data yang sesuai dengan kondisi nyata di lapangan. Selain itu, untuk melengkapi keterbatasan jumlah data, peneliti menggunakan data sekunder yang diperoleh dari platform Kaggle. Data dari kedua sumber tersebut kemudian digabungkan untuk meningkatkan jumlah dan variasi dataset.

4.1.1. Data Pada Kelas Mentah

Pada kelas mentah berjumlah 250 data atau foto buah kelapa sawit.



Gambar 4.1 Kelas Mentah

Penjelasan:Citra pada kelas mentah umumnya menunjukkan dominasi warna hijau hingga keunguan dengan tekstur buah yang masih keras. Pada tahap ini, buah belum memperlihatkan indikator kematangan, sehingga belum memenuhi kriteria untuk dilakukan pemanenan.

4.1.2 Data Pada Kelas Mengkal

Pada kelas mengkal terdapat sebanyak 250 citra buah kelapa sawit.



Gambar 4.2 Kelas Mengkal

Keterangan gambar 4.2 Citra pada kelas mengkal memperlihatkan adanya perubahan warna dari hijau menuju kemerahan yang belum merata. Kondisi ini menunjukkan bahwa buah sedang berada dalam tahap awal proses pematangan.

4.1.3 Data Pada kelas Matang

Pada kelas matang terdapat sebanyak 250 citra buah kelapa sawit.



Gambar 4.1.3. Kelas Matang

Penjelasan: Citra pada kelas matang umumnya didominasi oleh warna merah hingga oranye yang tersebar merata. Kondisi ini mengindikasikan bahwa buah telah mencapai tingkat kematangan optimal sehingga layak untuk dilakukan pemanenan

4.1.4 Data Pada Kelas Terlalu Matang

Pada kelas terlalu matang terdapat sebanyak 250 citra buah kelapa sawit.



Gambar 4.1.4 Kelas Terlalu Matang

Penjelasan: Citra pada kelas terlalu matang umumnya memperlihatkan warna yang lebih gelap disertai dengan jumlah brondolan yang terlepas dalam jumlah signifikan. Kondisi ini menunjukkan bahwa buah telah melampaui tingkat

kematangan optimal.

4.2 Hasil Implementasi Sistem

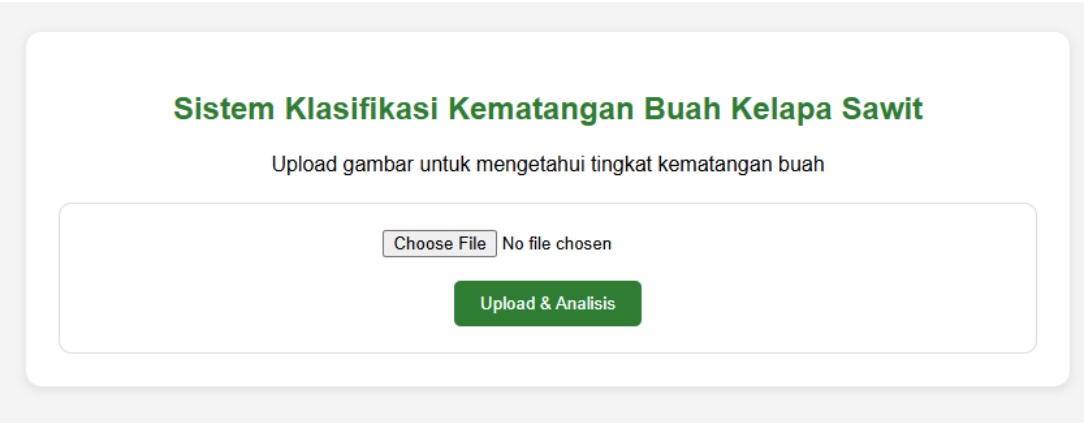
4.2.1 Tampilan Halaman Login



Gambar 4.2.1 Halaman Login

Penjelasan:Halaman login berfungsi sebagai mekanisme pembatasan akses pengguna ke dalam sistem. Pengguna diwajibkan memasukkan username dan password yang valid agar dapat mengakses berbagai fitur yang tersedia pada sistem.

4.2.2 Tampilan Dasboard

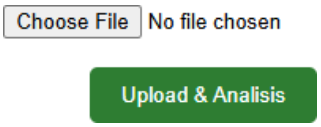


Gambar 4.2.2 Tampilan Dasboard

Penjelasan: Dashboard merupakan halaman utama yang ditampilkan setelah

pengguna berhasil melakukan proses login. Pada halaman ini tersedia menu navigasi yang digunakan untuk mengakses fitur klasifikasi serta riwayat data.

4.2.3 Tampilan Upload Citra



Gambar 4.2.3 Upload Citra

Penjelasan: Halaman ini berfungsi untuk mengunggah citra buah kelapa sawit yang akan diklasifikasikan. Sistem akan melakukan proses pengolahan citra setelah pengguna menekan tombol unggah.

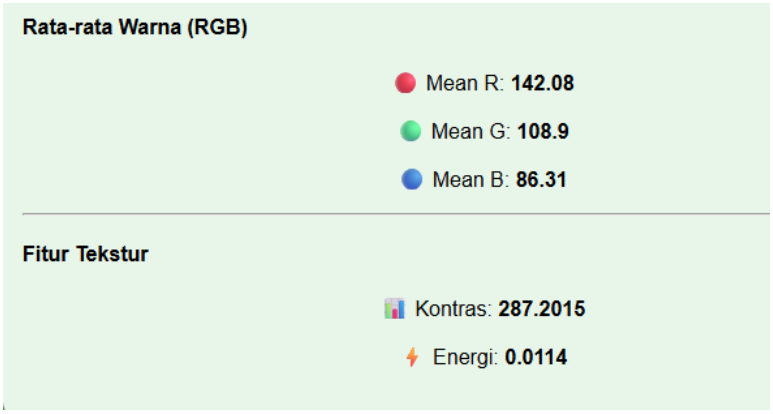
4.2.4 Hasil Preprocessing



Gambar 4.2.4 Preprocessing Citra

Penjelasan: Pada tahap praproses (preprocessing), citra diseragamkan ukurannya, dinormalisasi, serta dikonversi ke dalam format tertentu. Tahap ini bertujuan untuk meningkatkan kualitas citra sehingga proses ekstraksi fitur dapat dilakukan secara lebih optimal.

4.2.5 Hasil Ekstraksi Fitur



Gambar 4.2.5 Ekstraksi Fitur

Penjelasan: Sistem melakukan ekstraksi fitur warna berupa nilai rata-rata RGB serta fitur tekstur dengan menggunakan metode GLCM. Fitur-fitur tersebut selanjutnya dimanfaatkan sebagai masukan dalam proses klasifikasi.

4.2.6 Hasil Klasifikasi



Gambar 4.2.6 Hasil Klasifikasi

Penjelasan: Sistem menyajikan hasil klasifikasi dalam bentuk kategori tingkat

kematangan buah kelapa sawit, yaitu mentah, mengkal, matang, dan terlalu matang. Selain itu, sistem juga menampilkan nilai probabilitas sebagai representasi tingkat kepercayaan model terhadap hasil prediksi.

Ada 4 pelabelan yaitu:

a. Mentah dilabeli dengan angka 0
b. mengkal dilabeli dengan angka 1
c. matang dilabeli dengan angka 2
d. dan terlalu matang dilabeli dengan angka 3

4.3 Hasil Pengujian Sistem

4.3.1 Pengujian Sistem





Gambar 4.3.1 Hasil Pengujian Sistem

Penjelasan: Pengujian dilakukan dengan mengunggah beberapa citra ke dalam sistem. Sistem mampu memberikan hasil klasifikasi sesuai dengan karakteristik citra yang diinputkan.

4.3.2 Analisis Hasil

Berdasarkan hasil pengujian, sistem mampu mengidentifikasi tingkat kematangan buah kelapa sawit dengan tingkat akurasi yang cukup baik. Fitur warna menjadi faktor dominan dalam proses klasifikasi, mengingat perubahan warna yang signifikan terjadi pada setiap tingkat kematangan.

Namun, terdapat beberapa kesalahan klasifikasi yang disebabkan oleh:

1. Variasi pencahayaan
2. Sudut pengambilan citra
3. Kemiripan warna antar kelas

4.3.3 Kinerja Sistem

Secara keseluruhan, sistem menunjukkan kinerja yang baik dalam mengklasifikasikan citra buah kelapa sawit. Sistem ini berpotensi dimanfaatkan

sebagai alat bantu untuk menentukan tingkat kematangan buah secara otomatis.

4.4 Implementasi Sistem (Coding)

4.4.1 Inisialisasi Sistem

```
app = Flask(__name__)
app.secret_key = 'secret123'
```

Gambar 4.4.1 Inisialisasi Sistem

Penjelasan: Kode ini digunakan untuk membangun aplikasi web menggunakan Flask serta mengatur keamanan sesi pengguna.

4.4.2 Upload dan Pembacaan Citra

```
# ===== BACA GAMBAR =====
img = cv2.imread(filepath)
img = cv2.resize(img, (256, 256))
```

Gambar 4.4.2 Upload dan Pembacaan Citra

Penjelasan: Citra yang diunggah akan dibaca menggunakan OpenCV dan diubah ukurannya agar seragam.

4.4.3 Ekstraksi Fitur Warna

```
# ===== RGB =====
Mean_R = float(np.mean(img[:, :, 2]))
Mean_G = float(np.mean(img[:, :, 1]))
Mean_B = float(np.mean(img[:, :, 0]))
```

Gambar 4.4.3 Ekstraksi Fitur Warna

Penjelasan: Digunakan untuk menghitung nilai rata-rata warna RGB sebagai fitur utama dalam klasifikasi.

4.4.4 Ekstraksi Fitur Tekstur

```
# ===== TEKSTUR (GLCM) =====
gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

glcm = graycomatrix(
    gray,
    distances=[1],
    angles=[0],
    levels=256,
    symmetric=True,
    normed=True
)

kontras = float(graycoprops(glcm, 'contrast')[0
energi = float(graycoprops(glcm, 'energy')[0][0
```

Gambar 4.4.4 Ekstraksi Fitur Tekstur

Penjelasan: Digunakan untuk memperoleh informasi tekstur citra menggunakan metode GLCM.

4.4.5 Klasifikasi CatBoost

```
# ================== LOAD MODEL ==================
model = CatBoostClassifier()
model.load_model("model.cbm")
```

Gambar 4.4.5 Klasifikasi CatBoost

Penjelasan: Model CatBoost digunakan untuk menentukan kategori kematangan buah berdasarkan fitur yang telah diekstraksi.

4.5 Pembahasan

Berdasarkan hasil penelitian, sistem yang dikembangkan mampu mengklasifikasikan tingkat kematangan buah kelapa sawit dengan kinerja yang baik. Kombinasi fitur warna dan tekstur terbukti efektif dalam merepresentasikan karakteristik visual buah sehingga mendukung proses klasifikasi.

Algoritma CatBoost menunjukkan performa yang stabil dalam melakukan proses klasifikasi. Selain itu, tahap praproses (preprocessing) memiliki peran yang signifikan dalam meningkatkan kualitas data, sehingga berkontribusi terhadap peningkatan akurasi hasil klasifikasi.

Sistem berbasis web yang dibangun juga memberikan kemudahan bagi pengguna dalam melakukan proses klasifikasi secara cepat dan praktis. Dengan demikian, sistem ini dapat dimanfaatkan sebagai alat bantu dalam menentukan tingkat kematangan buah kelapa sawit secara otomatis.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan terkait klasifikasi tingkat kematangan buah kelapa sawit berbasis ekstraksi fitur citra digital menggunakan algoritma CatBoost, diperoleh beberapa kesimpulan sebagai berikut:

a. Sistem klasifikasi tingkat kematangan buah kelapa sawit berhasil dikembangkan dalam bentuk aplikasi berbasis web dengan memanfaatkan bahasa pemrograman Python dan framework Flask.

b. Sistem ini mampu menerima masukan barupa citra buah kelapa sawit, yang kemudian diproses melalui tahapan praproses (preprocessing), ekstraksi fitur, hingga menghasilkan keluaran berupa hasil klasifikasi

c. Tahap praproses yang mencakup perubahan ukuran citra (resize), normalisasi, serta konversi warna terbukti mampu meningkatkan kualitas citra, sehingga data menjadi lebih seragam dan siap digunakan pada tahap ekstraksi fitur.

d. Ekstraksi fitur citra yang diterapkan dalam penelitian ini meliputi fitur warna (RGB) dan fitur tekstur (GLCM). Fitur warna berperan sebagai indikator utama dalam membedakan tingkat kematangan buah, sedangkan fitur tekstur berfungsi untuk memperkuat representasi karakteristik permukaan buah.

e. Algoritma CatBoost dapat diimplementasikan secara efektif dalam proses

klasifikasi dan mampu mengelompokkan citra buah kelapa sawit ke dalam empat kategori, yaitu mentah, mengkal, matang, dan terlalu matang.

f. Berdasarkan hasil pengujian sistem, diperoleh bahwa sistem mampu menghasilkan klasifikasi yang sesuai dengan karakteristik visual citra. Hal ini menunjukkan bahwa sistem memiliki potensi untuk digunakan sebagai alat bantu dalam menentukan tingkat kematangan buah kelapa sawit secara otomatis.

5.2 Saran

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

a. Penelitian selanjutnya disarankan untuk menggunakan jumlah dataset yang lebih besar serta lebih bervariasi agar model klasifikasi memiliki kemampuan generalisasi yang lebih optimal.
b. Pengembangan lebih lanjut dapat dilakukan dengan menambahkan jenis fitur citra lainnya, seperti fitur bentuk maupun fitur berbasis *deep learning*, guna meningkatkan tingkat akurasi klasifikasi.
c. Penggunaan algoritma lain, seperti Convolutional Neural Network (CNN), dapat dipertimbangkan sebagai pembanding terhadap algoritma CatBoost untuk mengetahui metode dengan performa terbaik dalam klasifikasi citra.
d. Sistem yang telah dikembangkan juga dapat ditingkatkan menjadi aplikasi berbasis mobile, sehingga lebih mudah diakses dan digunakan oleh petani maupun pihak perkebunan secara langsung di lapangan.
e. Selain itu, diperlukan pengujian lanjutan pada kondisi

pencahayaan serta lingkungan yang lebih beragam untuk meningkatkan ketahanan sistem terhadap variasi data pada kondisi nyata.

DAFTAR PUSTAKA

Aditya, F. R., Umbara, F. R., & Sabrina, P. N. (2025). Penerapan algoritma CatBoost untuk klasifikasi berbasis fitur kategorikal. Jurnal Teknologi Informasi, 12(1), 15–22.

Alfath, A. S., & Wardhana, A. K. (2025). Penerapan CatBoost untuk prediksi medis. Jurnal Sistem Cerdas, 9(2), 44–52.

Ashillah, S., Gulo, S. A., Rahmi, Y., & Niska, D. Y. (2025). Preprocessing citra dan reduksi noise dalam pengolahan citra digital. Jurnal Pengolahan Citra, 8(1), 10–18.

Claudia Delvia Pangati, dkk. (2025). Ekstraksi fitur citra digital untuk identifikasi tingkat kematangan buah berbasis computer vision. Jurnal Vision Informatika, 6(2), 70–78.

Fajarsari, dkk. (2022). Normalisasi dan resizing citra untuk pembelajaran mesin. Jurnal Komputasi Terapan, 5(3), 88–95.

Fitriani, dkk. (2023). Analisis kinerja algoritma CatBoost pada kasus klasifikasi. Jurnal Data Science Indonesia, 7(2), 101–109.

Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson Education Inc. ISBN 978-0-13-335672-4

Hidayat, & Prasetyo. (2024). Performa CatBoost dibanding algoritma boosting lainnya. Jurnal Informatika Terapan, 11(1), 33–41.

Jasman, dkk. (2022). Gradient boosting dan performa CatBoost dalam klasifikasi data. Jurnal Sistem Informasi, 10(2), 55–63.

Juandri, & Anwar. (2023). Model warna RGB dan HSV dalam analisis citra digital. Jurnal Multimedia & Citra, 9(1), 12–20.

Kadir, N. (2024). Pemodelan kualitas udara menggunakan algoritma CatBoost. Jurnal Sains Data, 4(2), 25–32.

Madani, F., & Lubis, A. H. (2025). Prinsip kerja algoritma CatBoost dan keunggulannya. Jurnal Teknologi Komputasi, 13(1), 5–13.

Maulana, I., Siregar, A. M., & Lestari, S. A. P. (2024). Kemampuan generalisasi CatBoost pada data tabular. Jurnal Artificial Intelligence, 8(2), 60–68.

Miyantih, Y., Cahyana, Y., & Lestari, S. A. P. (2025). Ekstraksi fitur warna dan tekstur berbasis GLCM pada citra digital. Jurnal Image Processing, 6(1), 27–35.

Murtiningsih, D. A., & Sari, B. W. (2025). Regularization pada algoritma CatBoost untuk mencegah overfitting. Jurnal Machine Learning, 7(1), 40–48.

Muzahardin, Fauzi, & Nurhayati. (2021). Peningkatan kontras citra menggunakan histogram equalization. Jurnal Sistem Visual, 2(2), 44–52.

Nanda, R. E., & Prabowo, Y. D. (2022). Klasifikasi citra menggunakan self-organizing map (SOM). Jurnal Sistem Komputer, 5(2), 77–85.

Nur Rali Rahma Wati, Abdullah, A., & Sucipto. (2025). Klasifikasi kematangan buah kelapa sawit menggunakan convolutional neural network. Jurnal Publikasi Teknik Informatika, 10(2), 45–52.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. Advances in Neural Information Processing Systems, 31, 6638–6648.

Putra, & Nugroho. (2022). Pemanfaatan citra digital untuk analisis karakteristik visual objek. Jurnal Komputer Grafis, 4(1), 14–21.

Putri Aisyiyah Rakhma Devi, & Rosyid, H. (2022). Konsep dasar pengolahan citra

digital. Jurnal Teknologi Multimedia, 3(2), 50–57.

Rahmawati, dkk. (2021). Subjektivitas penentuan kematangan buah sawit secara visual. Jurnal Pertanian Digital, 2(1), 33–39.

Ratna. (2020). Konversi grayscale pada tahap preprocessing citra digital. Jurnal Pengolahan Sinyal, 3(1), 9–15.

Ridla, M. A. (2022). Perbandingan algoritma KNN dan SVM untuk klasifikasi warna piksel citra. Jurnal Ilmu Komputasi, 6(2), 98–106.

Sari, I. P., Al-Khowarizmi, & Batubara, I. H. (2025). *Cluster analysis using K-Means algorithm and fuzzy C-means clustering for grouping students' abilities in online learning process*. Jurnal Teknologi Informasi, 12(1), 1–10.

Siregar, dkk. (2020). Klasifikasi tingkat kematangan buah kelapa sawit berbasis fitur warna dan tekstur. Jurnal Agroinformatika, 4(2), 72–80.

Utomo, S., Sigit, S., & Sulistyowati, N. (2025). Struktur oblivious tree pada algoritma CatBoost. Jurnal Data Mining, 9(1), 18–26.

Widyaningsih, M., & Harjoko, A. (2021). Identifikasi gejala penyakit tanaman jeruk melalui preprocessing, FFT, dan LBP. Jurnal Teknologi Pertanian, 7(3), 120–12