

**KLASIFIKASI GAMBAR FASE PERTUMBUHAN PADI
MENGUNAKAN ALGORITMA *EXTREME GRADIENT BOOSTING***



**MAYA IFRA SHOBIA
NIM. 21337078**

**PROGRAM STUDI SARJANA STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2025**

**KLASIFIKASI GAMBAR FASE PERTUMBUHAN PADI
MENGUNAKAN ALGORITMA *EXTREME GRADIENT BOOSTING***

SKRIPSI

*Diajukan sebagai salah satu persyaratan guna memperoleh gelar
Sarjana Statistika*



Oleh
MAYA IFRA SHOBIA
NIM. 21337078

**PROGRAM STUDI SARJANA STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2025**

PERSETUJUAN SKRIPSI

KLASIFIKASI GAMBAR FASE PERTUMBUHAN PADI MENGUNAKAN ALGORITMA *EXTREME GRADIENT BOOSTING*

Nama : Maya Ifra Shobia
NIM : 21337078
Program Studi : S1 Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 8 Agustus 2025

Mengetahui:
Kepala Departemen Statistika



Dr. Yenni Kurniawati, S.Si., M.Si
NIP. 198402232010122005

Disetujui Oleh:
Pembimbing



Dr. Yenni Kurniawati, S.Si., M.Si
NIP. 198402232010122005

PENGESAHAN LULUS UJIAN SKRIPSI

Nama : Maya Ifra Shobia
NIM : 21337078
Program Studi : S1 Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

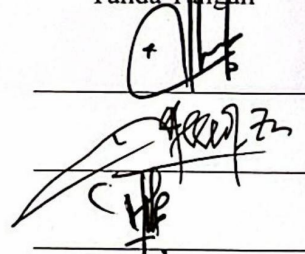
KLASIFIKASI GAMBAR FASE PERTUMBUHAN PADI MENGUNAKAN ALGORITMA *EXTREME GRADIENT BOOSTING*

Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Skripsi
Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Padang

Padang, 8 Agustus 2025

Tim Penguji	Nama
Ketua	: Dr. Yenni Kurniawati, S.Si., M.Si
Anggota	: Dr. Syafriandi, M.Si
Anggota	: Dina Fitria, M.Si

Tanda Tangan



SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan dibawah ini :

Nama : Maya Ifra Shobia
NIM : 21337078
Program Studi : S1 Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan, bahwa skripsi saya dengan judul “Klasifikasi Gambar Fase Pertumbuhan Padi Menggunakan Algoritma *Extreme Gradient Boosting*” adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dan etika yang berlaku dalam tradisi keilmuan. Apabila suatu saat terbukti saya melakukan plagiat maka saya bersedia diproses dan menerima sanksi akademis maupun hukum sesuai dengan hukum dan ketentuan yang berlaku, baik di institusi UNP maupun masyarakat dan negara.

Demikian pernyataan ini saya buat dengan penuh kesadaran dan rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Padang, 17 November 2025

Diketahui Oleh:

Kepala Departemen Statistika



Dr. Yenni Kurniawati, S.Si., M.Si.
NIP. 198402232010122005

Saya yang menyatakan,



Maya Ifra Shobia
NIM. 21337078

Klasifikasi Gambar Fase Pertumbuhan Padi Menggunakan Algoritma *Extreme Gradient Boosting*

Maya Ifra Shobia

ABSTRAK

Klasifikasi gambar merupakan salah satu penerapan penting dari *machine learning* yang banyak digunakan di berbagai bidang, termasuk sektor pertanian. Di Indonesia, proses identifikasi fase pertumbuhan padi pada kegiatan Kerangka Sampel Area (KSA) masih dilakukan secara manual, sehingga memerlukan waktu yang cukup lama dan masih menimbulkan kesalahan dalam pengamatan dikarenakan proses klasifikasi yang dilakukan secara manual. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model otomatis dalam mengklasifikasikan fase pertumbuhan tanaman padi meliputi fase vegetatif awal, vegetatif akhir, generatif, dan panen menggunakan algoritma *Extreme Gradient Boosting* (XGBoost) yang dipadukan dengan ekstraksi fitur VGG16, guna meningkatkan akurasi dan efisiensi proses identifikasi fase pertumbuhan.

Dataset yang digunakan terdiri atas 1.417 gambar tanaman padi yang dikumpulkan di Kabupaten Padang Pariaman selama periode November 2023 hingga Februari 2024. Tahapan praproses mencakup penyesuaian ukuran dimensi gambar menjadi 224×224 piksel dan normalisasi dalam rentang [0,1], kemudian dilakukan augmentasi data untuk meningkatkan variasi gambar dan mengurangi risiko *overfitting*. Dataset dibagi menjadi data pelatihan (80%) dan data pengujian (20%). Ekstraksi fitur dilakukan menggunakan VGG16, dan hasilnya digunakan sebagai masukan (*input*) pada algoritma XGBoost. Proses penyetelan parameter (*grid search*) menghasilkan parameter optimal, yaitu $n_estimators = 500$, $max_depth = 5$, $learning_rate = 0,01$, $subsample = 0,5$, dan $colsample_bytree = 0,9$.

Model yang dikembangkan berhasil mencapai akurasi sebesar 80,63%, yang termasuk dalam kategori *good classification*. Hasil ini menunjukkan bahwa kombinasi VGG16 dan XGBoost mampu mengenali serta membedakan setiap fase pertumbuhan padi secara efektif. Pendekatan ini berpotensi mendukung proses pemantauan pertanian, serta memberikan kontribusi terhadap peningkatan ketahanan pangan nasional melalui penerapan kecerdasan buatan.

Kata Kunci: Fase Pertumbuhan Padi, Kerangka Sampel Area, Klasifikasi Gambar, VGG16, XGBoost.

Image Classification of Rice Growth Phase Using the Extreme Gradient Boosting Algorithm

Maya Ifra Shobia

ABSTRACT

Image classification is one of the most important applications of machine learning, widely used across various fields, including agriculture. In Indonesia, the identification of rice growth phases within the Area Sampling Frame (KSA) program is still performed manually, which requires considerable time and is prone to observation errors. Therefore, this study aims to develop an automated model capable of classifying rice growth phases namely early vegetative, late vegetative, generative, and harvest stages using the Extreme Gradient Boosting (XGBoost) algorithm combined with VGG16 feature extraction to improve the accuracy and efficiency of the identification process.

The dataset used in this study consists of 1,417 rice plant images collected in Padang Pariaman Regency between November 2023 and February 2024. The preprocessing stage includes resizing the images to 224×224 pixels and normalizing them within the range [0,1], followed by data augmentation to increase image variation and reduce overfitting. The dataset was divided into training data (80%) and testing data (20%). Feature extraction was carried out using VGG16, and the extracted features were used as input for the XGBoost algorithm. Parameter tuning using grid search produced optimal parameters: `n_estimators = 500`, `max_depth = 5`, `learning_rate = 0.01`, `subsample = 0.5`, and `colsample_bytree = 0.9`.

The developed model achieved an accuracy of 80.63%, which falls into the category of good classification. These results demonstrate that the combination of VGG16 and XGBoost can effectively recognize and distinguish each rice growth phase. This approach has the potential to support more efficient and objective agricultural monitoring and contribute to strengthening national food security through the application of artificial intelligence in agriculture.

Keywords: Rice Growth Phases, Area Sampling Frame, Image Classification, VGG16, XGBoost.

KATA PENGANTAR

Alhamdulillah, puji syukur penulis panjatkan kepada Allah SWT yang telah melimpahkan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Klasifikasi Gambar Fase Pertumbuhan Padi Menggunakan Algoritma *Extreme Gradient Boosting*”. Shalawat dan salam senantiasa tercurah kepada Nabi Muhammad SAW, beserta keluarga, sahabat, dan seluruh umatnya.

Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan pendidikan pada Program Studi Sarjana Statistika, Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Padang. Penulisan skripsi ini tidak lepas dari dukungan, bimbingan, dan bantuan berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang mendalam kepada:

1. Ibu Dr.Yenni Kurniawati, M.Si., sebagai Dosen Pembimbing Skripsi dan Kepala Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang. Bimbingan, arahan, dan ilmu yang ibu berikan selama pembimbingan akademik maupun penyusunan skripsi sangat membantu dalam menyelesaikan studi ini dengan baik.
2. Bapak Syafriandi, M.Si., dan Ibu Dina Fitria, M.Si., sebagai Dosen Penguji. Masukan, arahan, koreksi, serta bimbingan yang diberikan selama perkuliahan dan penyusunan skripsi sangat membantu dalam meningkatkan kualitas skripsi ini.

3. Seluruh Dosen dan Staf Pengajar Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang, atas ilmu dan dukungan moral yang selalu diberikan selama masa studi.
4. Kedua orang tua, Bapak Firmansyah dan Ibu Desnawati, atas kasih sayang, doa, serta dukungan moral dan materiil yang selalu diberikan, menjadi fondasi utama dalam perjalanan pendidikan dan kehidupan penulis.
5. Adik tercinta, Gilang Nahdhatul Fikri, yang selalu memberikan motivasi dan semangat dalam setiap kesempatan.
6. Teman-teman seperjuangan Angkatan 2021, atas kebersamaan dan kerja sama selama proses perkuliahan maupun penulisan skripsi ini.
7. Semua pihak yang telah memberikan dukungan, bantuan, dan kontribusi, baik langsung maupun tidak langsung, selama proses penyusunan skripsi ini sehingga dapat terselesaikan dengan baik.

Semoga semua bimbingan, arahan, masukan, bantuan, dan do'a yang telah diberikan dibalas oleh Allah SWT dengan pahala dan amal ibadah yang berlipat. Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, oleh karena itu, saran dan masukan dari berbagai pihak sangat penulis harapkan demi perbaikan dan peningkatan kualitas di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak yang membutuhkan.

Padang, 28 Juli 2025

Penulis

Maya Ifra Shobia

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR.....	iii
DAFTAR ISI.....	v
DAFTAR TABEL	vii
DAFTAR GAMBAR.....	viii
DAFTAR LAMPIRAN	ix
BAB I PENDAHULUAN.....	1
A. Latar Belakang Masalah.....	1
B. Batasan Masalah.....	7
C. Rumusan Masalah	7
D. Tujuan Penelitian	7
E. Manfaat Penelitian	8
BAB II KERANGKA TEORITIS	10
A. Kajian Teori	10
B. Penelitian Relevan.....	36
BAB III METODOLOGI PENELITIAN	38
A. Jenis Penelitian.....	38
B. Jenis dan Sumber Data	38
C. Variabel Penelitian	38
D. Teknik Analisis Data.....	39
BAB IV HASIL DAN PEMBAHASAN	42
A. Deskripsi Data.....	42
B. Hasil Analisis	44
C. Pembahasan.....	52
BAB V PENUTUP	55
A. Kesimpulan	55
B. Saran.....	55
DAFTAR PUSTAKA	57
LAMPIRAN.....	61

DAFTAR TABEL

Tabel	Halaman
1. Fase Amatan KSA.....	13
2. Parameter XGBoost	33
3. <i>Confusion Matrix</i> 4×4	34
4. Tingkatan Akurasi.....	36
5. Proporsi Data Pelatihan dan Pengujian	45
6. Model VGG16.....	48
7. Hasil <i>Tuning</i> Parameter.....	49
8. <i>Confusion Matrix</i> XGBoost	49

DAFTAR GAMBAR

Gambar	Halaman
1. Tahap Pembangunan Kerangka Sampel.....	10
2. Ilustrasi Pembagian Wilayah dalam Blok dan Segmen	11
3. Amatan KSA	13
4. Fase Pertumbuhan Padi	15
5. <i>Supervised classification</i>	16
6. <i>Unsupervised classification</i>	17
7. <i>Flip</i> (a) Vertikal dan (b) Horizontal	21
8. Rotasi Gambar.....	22
9. Arsitektur VGG16.....	24
10. Diagram Skema Algoritma XGBoost	28
11. Gambar Kategori Amatan KSA	39
12. Diagram Alir Penelitian	39
13. Potret Amatan Lahan Sawah (a) Vegetatif Awal, (b) Vegetatif Akhir, (c) Generatif, (d) Panen.....	42
14. Hasil <i>Preprocessing Data</i> (a) Sebelum Penyesuaian, (b) Sesudah Penyesuaian.....	44
15. Hasil Augmentasi Gambar (a) Asli, (b) Augmentasi 1, (c) Augmentasi 2, (d) Augmentasi 3.....	46
16. Proses Ekstraksi Fitur Gambar	47
17. Hasil Klasifikasi Gambar	52

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Gambar Padi Fase Vegetatif Awal	62
2. Data Gambar Padi Fase Vegetatif Akhir.....	63
3. Data Gambar Padi Fase Generatif.....	64
4. Data Gambar Padi Fase Panen	65
5. <i>Syntax</i> Algoritma XGBoost.....	66

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Klasifikasi adalah proses atau metode untuk mengelompokkan data ke dalam label atau kelas yang ditentukan berdasarkan kesamaan karakteristik atau atributnya (Gordon, 1999). Proses ini bertujuan untuk mempermudah analisis, pengambilan keputusan, dan meningkatkan pemahaman terhadap pola dalam data. Klasifikasi sangat penting karena memungkinkan pengelompokan data yang kompleks menjadi lebih terstruktur dan mudah dipahami. Dengan metode metode klasifikasi, data yang beragam dapat diorganisir secara sistematis, sehingga memungkinkan penarikan kesimpulan yang lebih akurat.

Klasifikasi tidak hanya mencakup data berupa teks, tetapi juga dapat diterapkan pada berbagai jenis data, seperti suara, video, data numerik, dan gambar. Salah satu bentuk klasifikasi yang mengalami perkembangan pesat adalah klasifikasi gambar, yang menjadi aspek penting dalam pengolahan gambar, visi komputer dan pembelajaran mesin (Mohamed dkk., 2022). Kemampuan klasifikasi gambar untuk memberikan analisis yang lebih mendalam terhadap data visual menjadikannya sangat relevan dalam berbagai aplikasi.

Menurut Sarraf dkk. (2021), klasifikasi gambar adalah prosedur pembelajaran mesin dimana gambar dapat dikelompokkan ke dalam kelas yang berbeda sesuai dengan fitur dan karakteristik visual. Klasifikasi gambar merupakan bagian dalam pemrosesan gambar yang berfokus pada analisis dan interpretasi data visual. Tujuan dari klasifikasi gambar adalah untuk

mengidentifikasi semua piksel dalam gambar digital, mengklasifikasikan, dan menetapkan ke dalam kategori yang disebut dengan kelas.

Penerapan klasifikasi gambar sangat luas, mencakup berbagai bidang, seperti kesehatan, otomotif, dan keamanan. Dalam bidang kesehatan, klasifikasi gambar digunakan untuk mendiagnosis penyakit melalui analisis gambar medis, seperti *Magnetic Resonance Imaging* (MRI) dan *Computed Tomography* (CT) scan, dimana algoritma pembelajaran mesin membantu mendeteksi kelainan dengan akurasi tinggi (Litjens dkk., 2017). Di industri otomotif, klasifikasi gambar berperan penting dalam pengembangan sistem kendaraan otonom untuk mengenali objek di sekitarnya, seperti pejalan kaki dan rambu lalu lintas (Bojarski dkk., 2016). Sementara itu, dalam sektor keamanan, klasifikasi gambar diterapkan dalam sistem pengawasan untuk mendeteksi perilaku mencurigakan dan mengenali wajah, yang membantu dalam meningkatkan keamanan publik (Kumar dkk., 2020).

Klasifikasi gambar juga memberikan kontribusi penting di bidang pertanian dalam membantu memantau kesehatan tanaman dan mendeteksi hama atau penyakit melalui analisis gambar satelit. Melalui cara ini, petani dapat mengambil tindakan yang tepat untuk meningkatkan hasil panen (Zhang dkk., 2019). Selain itu, klasifikasi gambar juga berperan dalam pengelolaan sumber daya pertanian. Salah satu contohnya adalah melalui kegiatan Survei Kerangka Sampel Area (KSA), yang dilakukan untuk mengumpulkan data akurat mengenai luas panen tanaman pangan.

Survei KSA merupakan metode survei yang dilakukan oleh Badan Pusat Statistik (BPS) untuk mengumpulkan data akurat mengenai luas panen tanaman

pangan, seperti padi dan jagung. Metode ini melibatkan pengamatan langsung terhadap sampel segmen (Khasanah dkk., 2020). Selain itu, KSA memanfaatkan teknologi gambar untuk mengumpulkan data luas dan kondisi lahan pertanian (BPS, 2024).

KSA memungkinkan pengumpulan informasi akurat mengenai fase pertumbuhan tanaman padi setiap bulan. Tanaman padi menjadi fokus utama dalam survei ini karena perannya yang penting sebagai komoditas pangan strategis di Indonesia (Mubekti & Sumargana, 2016). Padi tidak hanya menjadi sumber karbohidrat utama bagi sebagian besar penduduk Indonesia, tetapi juga memiliki nilai ekonomi yang tinggi. Selain itu, padi merupakan bagian integral dari budaya dan tradisi masyarakat Indonesia, yang menjadikannya lebih dari sekadar komoditas.

Di Indonesia, khususnya di Provinsi Sumatera Barat, terdapat dukungan yang signifikan terhadap produksi padi skala besar. Pada tahun 2024, provinsi Sumatera Barat diperkirakan akan menghasilkan 1.352.048,77 ton padi, menjadikannya salah satu penghasil padi terbesar di Indonesia (BPS, 2024). Pada tingkat regional, Kabupaten Padang Pariaman memegang peranan penting dalam mendukung ketahanan pangan Sumatera Barat. Tahun 2024, Padang Pariaman menghasilkan 132 ribu ton padi, memberikan kontribusi signifikan bagi total produksi provinsi (BPS, 2024). Selain bertujuan memenuhi kebutuhan lokal, produksi padi di Padang Pariaman juga berperan krusial dalam menjaga stabilitas pasokan beras, yang merupakan komponen vital untuk ketahanan pangan berkelanjutan di wilayah Sumatera Barat.

Pelaksanaan KSA melibatkan pengamatan pada setiap titik pusat subsegmen dalam segmen terpilih untuk mengidentifikasi fase pertumbuhan padi (BPS, 2018). Informasi yang dikumpulkan dari pengamatan langsung kemudian dikirimkan ke pusat pengolahan data secara *online* untuk divalidasi guna memastikan akurasi dan mengurangi subjektivitas dalam klasifikasi (Dzakwan & Buana, 2020). Menurut Ruslan (2019) data fase pertumbuhan padi yang dihasilkan tidak hanya digunakan untuk memprediksi potensi luas panen untuk tiga bulan ke depan, tetapi juga menjadi dasar penting dalam pengambilan kebijakan pemerintah berkaitan dengan pangan nasional.

Survei KSA merupakan metode yang diharapkan dapat meningkatkan akurasi estimasi luas panen padi di Indonesia. Namun, KSA memiliki beberapa kelemahan yang perlu diperhatikan. Salah satu kelemahan KSA adalah kesalahan dalam melakukan klasifikasi fase pertumbuhan padi yang dilakukan secara manual oleh petugas KSA. Menurut Hutami (2024), hal ini disebabkan oleh kurangnya pengalaman dan pengetahuan petugas dalam memahami fase pertumbuhan padi sehingga dapat mempengaruhi hasil pengukuran dan mengakibatkan kesalahan perhitungan estimasi luas panen padi.

Kelemahan lain dari KSA terletak pada proses pengumpulan dan validasi data. Proses ini memakan waktu cukup lama, sehingga efisiensi operasional KSA menjadi berkurang. Untuk meningkatkan efisiensi KSA, diperlukan langkah-langkah perbaikan. Salah satu langkah penting adalah melatih petugas lapangan secara berkelanjutan. Selain itu, penggunaan teknologi yang lebih canggih dalam pengukuran perlu dipertimbangkan (Ruslan, 2019). Dengan cara ini, data yang dihasilkan dapat menjadi lebih akurat dan dapat diandalkan.

Pemanfaatan teknologi *machine learning* dapat meningkatkan efisiensi dan akurasi dalam KSA. *Machine learning*, sebagai cabang kecerdasan buatan, memungkinkan sistem untuk belajar dari data tanpa diprogram secara eksplisit (Khatibi & Ali, 2024). Teknologi ini dapat mengotomatiskan berbagai tugas yang sebelumnya memerlukan intervensi manusia, sehingga meningkatkan efisiensi operasional, mengurangi biaya, dan meningkatkan akurasi data.

Machine learning menganalisis gambar tanaman padi untuk mengidentifikasi fase pertumbuhan secara otomatis. Algoritma *machine learning* membangun model klasifikasi dari gambar berlabel. Salah satu metode *machine learning* yang digunakan untuk klasifikasi adalah *decision tree*. Menurut Breiman dkk., (1993), *decision tree* merupakan metode yang membagi data menjadi beberapa subset berdasarkan atribut paling informatif, kemudian membentuk struktur seperti pohon. Metode ini banyak digunakan untuk klasifikasi karena memudahkan interpretasi dan visualisasi, tetapi rentan terhadap *overfitting* jika model yang dibangun terlalu kompleks.

Sebagai upaya untuk mengatasi keterbatasan *decision tree* dalam hal *overfitting*, metode *ensemble learning* dapat dimanfaatkan untuk meningkatkan performa model. Menurut Hastie dkk., (2009), *ensemble learning* adalah teknik yang menggabungkan beberapa model untuk meningkatkan akurasi dan stabilitas prediksi. Salah satu teknik yang populer dalam *ensemble learning* adalah *gradient boosting*. Teknik ini membangun model prediktif dengan mengoptimalkan *loss function* dan melatih model baru untuk memperbaiki kesalahan model sebelumnya.

Extreme Gradient Boosting (XGBoost) merupakan pengembangan dari metode *gradient boosting*. XGBoost dikembangkan oleh Chen dan Guestrin pada tahun 2016. XGBoost mengoptimalkan kecepatan dan efisiensi melalui teknik seperti regulasi, pemrosesan paralel, dan penanganan data hilang secara otomatis. Keunggulan-keunggulan ini menjadikan XGBoost sangat tepat dan efektif digunakan untuk melakukan klasifikasi gambar.

XGBoost memanfaatkan ekstraksi fitur sebagai input utama dalam melakukan klasifikasi gambar. Fitur-fitur yang diekstraksi mencakup tekstur, warna, dan bentuk. Proses ekstraksi fitur dilakukan untuk mengubah data gambar mentah menjadi representasi numerik. Ekstraksi fitur dapat dilakukan secara otomatis menggunakan teknik *deep learning* yang merupakan metode pembelajaran mesin. *Deep learning* merupakan bagian dari kecerdasan buatan yang menggunakan jaringan saraf tiruan dengan banyak lapisan. Setiap lapisan dalam jaringan saraf bertugas untuk mengekstrak fitur dari gambar (Sarraf, 2021).

Convolutional Neural Network (CNN) merupakan salah satu metode *deep learning* yang dapat digunakan untuk mendeteksi dan mengenali sebuah objek pada sebuah gambar digital (Maulana & Rochmawati, 2020). CNN merupakan jenis jaringan saraf tiruan yang dirancang untuk melakukan pemrosesan data gambar. Algoritma CNN menggunakan teknik konvolusi untuk mengekstrak fitur-fitur dari gambar, sehingga dapat mengenali pola-pola dan objek-objek dalam gambar. CNN terdiri dari beberapa lapisan yaitu, lapisan konvolusi, lapisan *pooling*, dan lapisan *fully connected* (Sarraf & Tofighi, 2016).

Terdapat beberapa model dalam CNN yang digunakan untuk klasifikasi gambar. Beberapa contoh model yang umum digunakan adalah LeNet, AlexNet,

VGGNet, ResNet, dan Inception. Diantara model tersebut, VGGNet merupakan model yang menggunakan teknik konvolusi dan *pooling* yang lebih dalam untuk meningkatkan akurasi. VGG16 merupakan salah satu varian dari VGGNet dengan 16 lapisan yang dapat dilatih, yaitu 13 lapisan konvolusi dan 3 lapisan *fully connected*. Model ini menggunakan filter konvolusi berukuran 3×3 yang diatur dalam tumpukan, yang memungkinkan model untuk belajar fitur-fitur yang lebih kompleks dari gambar.

Penelitian yang dilakukan oleh Bagas & Viriri (2023) mengusulkan metode *hybrid* yang menggabungkan VGG16 sebagai ekstraktor fitur dan XGBoost sebagai pengklasifikasi. Penelitian ini menunjukkan bahwa model mencapai akurasi 0.97 dalam klasifikasi gambar kanker pankreas. Hasil ini menunjukkan bahwa pendekatan *hybrid* antara *deep learning* dan XGBoost dapat meningkatkan kinerja model secara signifikan. Berlandaskan pada efektivitas algoritma XGBoost dalam klasifikasi gambar dan potensi peningkatannya melalui ekstraksi fitur, penelitian ini mengambil judul **“Klasifikasi Gambar Fase Pertumbuhan Padi Menggunakan Algoritma *Extreme Gradient Boosting*”**.

B. Batasan Masalah

Berdasarkan latar belakang yang telah dijelaskan, batasan masalah dalam penelitian ini adalah data yang digunakan merupakan data gambar empat fase pertumbuhan padi yaitu vegetatif awal, vegetatif akhir, generatif, dan panen yang diperoleh dari *website* KSA BPS Kabupaten Padang Pariaman selama empat periode waktu, yaitu November 2023 hingga Februari 2024. Ekstraktor fitur yang digunakan untuk mengekstrak fitur gambar pada penelitian ini adalah VGG16.

C. Rumusan Masalah

Berdasarkan batasan masalah yang telah diuraikan, maka penelitian ini difokuskan pada beberapa permasalahan utama yang dirumuskan sebagai berikut.

1. Bagaimana hasil klasifikasi gambar fase pertumbuhan padi menggunakan algoritma *Extreme Gradient Boosting*?
2. Bagaimana tingkat akurasi yang diperoleh dalam melakukan klasifikasi gambar fase pertumbuhan padi menggunakan algoritma *Extreme Gradient Boosting*?

D. Tujuan Penelitian

Berdasarkan permasalahan yang telah diidentifikasi, penelitian ini dirancang dengan tujuan sebagai berikut.

1. Untuk mengetahui bagaimana hasil klasifikasi gambar fase pertumbuhan padi menggunakan algoritma *Extreme Gradient Boosting*.
2. Untuk mengetahui akurasi yang diperoleh dalam melakukan klasifikasi gambar fase pertumbuhan padi menggunakan algoritma *Extreme Gradient Boosting*.

E. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat praktis dan teoritis, antara lain:

1. Bagi Penulis, penelitian ini dapat memberikan pengetahuan yang lebih mendalam tentang dinamika pertumbuhan padi dan implementasi algoritma *Extreme Gradient Boosting* dalam melakukan klasifikasi gambar fase pertumbuhan tanaman padi.

2. Bagi BPS Kabupaten Padang Pariaman, penelitian ini memberikan kontribusi dalam proses operasional kegiatan Survei Kerangka Sampel Area (KSA) dalam pemantauan luas panen, sehingga meminimalisir kesalahan dan efisiensi waktu dalam melakukan klasifikasi gambar fase pertumbuhan padi.
3. Bagi Pembaca, dapat memberikan kontribusi pada pengembangan keterampilan dan pengetahuan pembaca dalam bidang pengolahan data gambar pada sektor pertanian.

BAB V

PENUTUP

A. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, didapatkan kesimpulan sebagai berikut.

1. Hasil klasifikasi gambar fase pertumbuhan padi yaitu vegetatif awal, vegetatif akhir, generatif, dan panen dengan menggunakan algoritma *Extreme Gradient Boosting* (XGBoost) disertai ekstraktor fitur VGG16 dan menggunakan parameter $n_estimator=500$, $max_depth=5$, $learning_rate=0,01$, $subsample=0,5$, dan $colsample_bytree=0,9$ terbukti efektif dalam mengklasifikasikan fase pertumbuhan padi,.
2. Model yang dikembangkan mencapai akurasi klasifikasi sebesar 80,63%, menunjukkan kemampuan yang baik dalam mengenali dan membedakan setiap fase pertumbuhan padi. Hasil ini membuktikan bahwa kombinasi ekstraksi fitur berbasis CNN dengan metode klasifikasi XGBoost dapat menjadi pendekatan yang efisien dan andal dalam mendukung analisis pertumbuhan tanaman berbasis citra digital

B. Saran

Penelitian selanjutnya disarankan mengeksplorasi penggunaan metode klasifikasi alternatif yang lebih dalam untuk meningkatkan akurasi dan kemampuan generalisasi model dalam mengklasifikasikan fase pertumbuhan padi. Selain itu, pengumpulan data dari lokasi dan kondisi yang lebih beragam, termasuk variasi dalam jenis padi dan kondisi lingkungan, akan sangat bermanfaat untuk memperkuat model yang dikembangkan. Teknik augmentasi data yang

lebih kompleks, seperti penggunaan *Generative Adversarial Networks* (GANs), dapat dipertimbangkan untuk meningkatkan keragaman data pelatihan, sehingga model dapat belajar dari variasi yang lebih luas.

Evaluasi berkelanjutan terhadap model yang dikembangkan dengan menggunakan data baru secara berkala sangat penting untuk memastikan relevansi dan akurasi model seiring dengan perubahan kondisi pertanian. Peningkatan infrastruktur teknologi di sektor pertanian, termasuk akses ke perangkat keras dan perangkat lunak yang memadai, juga perlu diperhatikan untuk mendukung penerapan teknologi ini.

DAFTAR PUSTAKA

- Aubert, G., & Kornprobst, P. (2002). *Mathematical Problems in Gambar Preprocessing: Partial Differential Equations and The Calculus of Variations* (Vol. 147). Springer.
- Bakasa, W., Viriri, S. (2023). VGG16 Feature Extractor with Extreme Gradient Boost Classifier for Pancreas Cancer Prediction. *Journal of Imaging*, 9(138).
<https://doi.org/10.3390/jimaging9070138>
- Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L. D., Monfort, M., Muller, U., Zhang, J., Zhang, X., Zhao, J., & Zieba, K. (2016). End to End Learning for Self-Driving Cars. <https://doi.org/10.48550/arXiv.1604.07316>
- BPS. (2018). *Upaya Perbaikan Data Padi Dengan Metode Kerangka Sampel Area (KSA) 2018*. Jakarta: ISBN: 978-602-438-207-0.
- BPS. (2024, 2 Juli). *Yuk Mengenal Survei Kerangka Sampel Area (KSA)*. <https://blitarkab.bps.go.id/id/news/2024/07/02/342/yuk-mengenal--survei--kerangka-sampel-area--ksa-.html>
- Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1993). *Classification and Regression Trees*. New York: Chapman and Hall.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
<https://doi.org/10.48550/arXiv.1603.02754>
- Claesen, M., & De Moor, B. (2015). Hyperparameter Search in Machine Learning: A review. *Journal of Machine Learning Research*, 16 (1), 1-30.
<https://doi.org/10.48550/arXiv.1502.02127>
- Dzakwan, M. N., & Buana, W. P. (2023). Estimasi Luas Fase Tanaman Padi dengan KSA-Hybrid (Studi Kasus di Kabupaten Lamongan, Jawa Timur). *Seminar Nasional Official Statistics 2023*.
- Gonzalez, R. C., & Woods, R. E. (2017). *Digital Image Processing*. Pearson Education India.
- Haba, D. (2023). *Data Augmentation with Python: Enhance Deep Learning Accuracy with Data Augmentation Methods for Gambar, Text, Audio, and Tabular Data*. Packt Publishing Ltd.
- Hanif, I. (2019). *Implementing Extreme Gradient Boosting (XGBoost) Classifier to Improve Customer Churn Prediction*. *International Conference on Statistics and Analytics*.
<http://dx.doi.org/10.4108/eai.2-8-2019.2290338>

- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Hutami, R. T. (2024). *Analisis Perbandingan Model Klasifikasi Foto Fase Pertumbuhan Padi dengan Algoritma CNN, ResNet, dan VGG (Studi Kasus: Survei KSA Badan Pusat Statistik)*. Tesis, Institut Teknologi Bandung.
- Khasanah, I. N., Triscowati, D. W., Handoyo, A., Buana, W. P., & Kadir. (2020). Kombinasi Survei Kerangka Sampel Area dan Remote Sensing Untuk Estimasi Luas Tanaman Padi di Masa Pandemi (KSA-Hybrid): Studi Kasus di Provinsi Lampung. *Jurnal ASKSD*, 12(2), 53-62.
<http://dx.doi.org/10.34123/jurnalasks.v12i2.284>
- Khatibi, S. M. H., & Ali, J. (2024). Harnessing the Power of Machine Learning for Crop Improvement and Sustainable Production. *Frontiers in Plant Science*, 15, Article 1417912.
<https://doi.org/10.3389/fpls.2024.1417912>
- Litjens, G., Kooi, T., Bejnordi, B. E., dkk. (2017). A Survey on Deep Learning in Medical Gambar Analysis. *Medical Gambar Analysis*, 42, 60-88.
<https://doi.org/10.1016/j.media.2017.07.005>.
- Malik, S., Harode, R., & Kunwar, A. S. (2020). *XGBoost: A Deep Dive into Boosting*.
<http://dx.doi.org/10.13140/RG.2.2.15243.64803>
- Maulana, F. F., & Rochmawati, N. (2020). Klasifikasi Citra Buah Menggunakan Convolutional Neural Network. *Journal of Informatics and Computer Science (JINACS)*, I(02), 104-108.
<http://dx.doi.org/10.26740/jinacs.v1n02.p104-108>
- Mehmood, M., Shahzad, A., Zafar, B., Shabbir, A., & Ali, N. (2022). Remote Sensing Image Classification: A Comprehensive Review and Applications. *Mathematical Problems in Engineering*.
<http://dx.doi.org/10.1155/2022/5880959>
- Mo, H., Sun, H., Liu, J., & Wei, S. (2019). Developing Window Behavior Models for Residential Buildings Using XGBoost Algorithm. *Energy and Buildings*.
<https://doi.org/10.1016/j.enbuild.2019.109564>
- Mohamed, A. H., Refaat, M., & Hemeida, A. M. (2022). Gambar Classification Based Deep Learning: A Review. *Aswan University Journal of Science and Technology*, 2(1), 11-35.
<https://doi.org/10.21608/aujst.2022.259887>.

- Mubekti, & S. (2016). Pendekatan Kerangka Sampel Area untuk Estimasi dan Peramalan Produksi Padi. *Jurnal Pangan*, 25(2), 71-82.
<https://doi.org/10.33964/jp.v25i2.323>
- Muchlis. (2018). *Area Sampling Frame: A New Approach to Reform Agricultural Data Collection*. Asia-Pacific Economic Statistics Week.
- Muller, A. C., & Guido, S. (2016). *Intoduction Machine Learning With Python*. O'Reilly Media.
- Natural Resource Canada. (2025, 8 Januari). *Image Classification and Analysis*. Government of Canada. <https://natural-resources.canada.ca/maps-tools-publications/satellite-elevation-air-photos/gambar-classification-analysis>.
- Rahayu, S., Prasetyo, A., & Widiastuti, T. (2020). Analisis Fase Pertumbuhan Padi dan Pengaruhnya Terhadap Hasil Panen. *Jurnal Agronomi Indonesia*, 48 (2), 123-130.
- Ramadhan, M. M., Sitanggang, I. S., Nasution, F. R., & Ghifari, A. (2017). Parameter Tuning in Random Forest Based on Grid Search Method for Gender Classification Based on Voice Frequency. *DEStech Transactions on Computer Science and Engineering*.
<http://dx.doi.org/10.12783/dtcse/cece2017/14611>
- Ruslan, K. (2019). *Memperbaiki Data Pangan Indonesia Lewat Metode Kerangka Sampel Area*. Center for Indonesian Policy Studies.
- S, I. M., Kusuma, W. A., & Wahjuni, S. (2019). Comparison Analysis of Ensemble Technique with Boosting (Xgboost) and Bagging (Randomforest) for Classify Splice Junction DNA Sequence Category. *Jurnal Penelitian Pos Dan Informatika*, 9(1), 27–36.
<http://repository.ipb.ac.id/handle/123456789/97566>
- Samih, A., Alzubaidi, L., & Alzahrani, A. (2023). A Comprehensive Review of XGBoost: Applications, Challenges, and Future Directions. *Journal of Machine Learning Research* 24 (1), 1-30.
- Sanghvi, K., Aralkar, A., Sanghvi, S., & Saha, I. (2020). A Survey on Image Classification Techniques. *International Journal of Computer Applications*.
<http://dx.doi.org/10.2139/ssrn.3754116>
- Sarraf, A., Azhdari, M., & Sarraf, S. (2021). A Comprehensive Review of Deep Learning Architectures for Computer Vision Applications. *American Scientific Research Journal for Engineering, Technology, and Science*, 77(1), 1-29.
https://asrjetsjournal.org/American_Scientific_Journal/article/view/6712

- Sarraf, S., & Tofighi, G. (2016). Classification of Alzheimer's Disease Using FMRI Data and Deep Learning Convolutional Neural Networks. <https://doi.org/10.48550/arXiv.1603.08631>
- Suhardjo, S. (2018). *Budidaya Padi: Teori dan Praktik*. Yogyakarta: Andi.
- Tamina, S. (2019). Transfer Learning Using VGG-16 With Deep Convolutional Neural Network for Classifying Images. *International Journal of Scientific and Research Publications*, 9(10), 143-150. <http://dx.doi.org/10.29322/IJSRP.9.10.2019.p9420>
- Wigton, W.H. and Bormann, P. (1978). *A Guide to Area Sampling Frame Construction Utilizing Satellite Gamberry. 2nd Int'l Training Course in Remote Sensing Applications for Agriculture: Crop Statistics and Agricultural Census*. UN Outer Space Affairs Division.
- Zhao, W., Zhang, Z., & Chen, Y. (2019). Face Recognition: A Literature Survey. *ACM Computing Surveys*, 52(2), 1-35. <https://doi.org/10.1145/954339.954342>
- Zhao, Y., & Zhang, Y. (2019). Gambar Classification Based on Deep Learning: A Review. *Journal of Computer Science and Technology*, 34(1), 1-20. <http://dx.doi.org/10.21608/aujst.2022.259887>