

**ANALISIS KINERJA *VISION TRANSFORMER* BERBASIS *TRANSFER*
LEARNING UNTUK KLASIFIKASI MULTIKELAS PADA
GAMBAR FASE PERTUMBUHAN PADI**



**ROSA SALSABILA AZARINE
NIM 22337013**

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

2026

**ANALISIS KINERJA *VISION TRANSFORMER* BERBASIS *TRANSFER*
LEARNING UNTUK KLASIFIKASI MULTIKELAS PADA
GAMBAR FASE PERTUMBUHAN PADI**

SKRIPSI

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



**Oleh
ROSA SALSABILA AZARINE
NIM 22337013**

**PROGRAM STUDI SARJANA STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2026**

PERSETUJUAN SKRIPSI

ANALISIS KINERJA *VISION TRANSFORMER* BERBASIS *TRANSFER LEARNING* UNTUK KLASIFIKASI MULTIKELAS PADA GAMBAR FASE PERTUMBUHAN PADI

Nama : Rosa Salsabila Azarine
NIM : 22337013
Program Studi : S1 Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Mengetahui:
Kepala Departemen Statistika



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

Padang, 2 Februari 2026
Disetujui Oleh:
Pembimbing



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

PENGESAHAN LULUS UJIAN SKRIPSI

Nama : Rosa Salsabila Azarine
NIM : 22337013
Program Studi : S1 Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

ANALISIS KINERJA *VISION TRANSFORMER* BERBASIS *TRANSFER LEARNING* UNTUK KLASIFIKASI MULTIKELAS PADA GAMBAR FASE PERTUMBUHAN PADI

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

Padang, 2 Februari 2026

Tim Penguji	Nama
Ketua	: Dr. Yenni Kurniawati, S.Si., M.Si.
Anggota	: Dr. Syafriandi, M.Si.
Anggota	: Dwi Sulistiowati, S.Si., M.Stat.

Tanda Tangan



The image shows three handwritten signatures in blue ink, each placed on a horizontal line. The first signature is at the top, the second in the middle, and the third at the bottom.

SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini :

Nama : Rosa Salsabila Azarine
NIM : 22337013
Program Studi : S1 Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan, bahwa skripsi saya dengan judul "**Analisis Kinerja Vision Transformer Berbasis Transfer Learning Untuk Klasifikasi Multikelas Pada Gambar Fase Pertumbuhan Padi**" adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan cara-cara yang tidak sesuai dengan 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 di masyarakat dan negara.

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

Padang, 06 Februari 2026

Diketahui Oleh:

Ketua Departemen Statistika



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

Saya yang menyatakan,



Rosa Salsabila Azarine
NIM. 22337013

Analisis Kinerja *Vision Transformer* Berbasis *Transfer learning* Untuk Klasifikasi Multikelas Pada Gambar Fase Pertumbuhan Padi

Rosa Salsabila Azarine

ABSTRAK

Klasifikasi gambar berbasis *deep learning* merupakan salah satu teknik yang banyak digunakan di berbagai bidang, termasuk sektor pertanian. Di Indonesia, identifikasi fase pertumbuhan padi pada kegiatan Kerangka Sampel Area (KSA) masih dilakukan secara manual, sehingga membutuhkan waktu yang cukup lama dan berisiko menimbulkan kesalahan pengamatan. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model otomatis untuk mengklasifikasikan fase pertumbuhan padi yang meliputi vegetatif awal, vegetatif akhir, generatif dan panen menggunakan arsitektur *Vision Transformer* (ViT). Model yang dikembangkan ViT-B/16 berbasis *transfer learning* guna meningkatkan akurasi dan efisiensi proses klasifikasi.

Dataset yang digunakan terdiri atas 1.417 gambar tanaman padi yang diperoleh dari BPS Kabupaten Padang Pariaman selama periode November 2023 hingga Februari 2024. Tahapan praproses mencakup penyesuaian dimensi gambar menjadi 224×224 piksel dan normalisasi ke rentang $[0,1]$. Dataset dibagi menjadi data *training* (80%) dan data *testing* (20%). Augmentasi data dilakukan pada data *training* untuk meningkatkan variasi gambar serta mengurangi risiko *overfitting*. Pelatihan data dilakukan melalui dua pendekatan yaitu *transfer learning* dan *from scratch* sebagai *baseline* menggunakan beberapa pengaturan *hyperparameter* seperti *epoch* 10, *batch size* berukuran 32, *learning rate* dengan nilai 0,00005, *optimizer* AdamW dengan *weight decay* 0,01 dan fungsi *cross-entropy loss*.

Hasil evaluasi menunjukkan bahwa model yang dilatih dengan pendekatan *transfer learning* berhasil mencapai akurasi 90,46% dan mampu meningkatkan akurasi sebesar 14,84% dibandingkan dengan model yang dilatih dari awal (*from scratch*) yang hanya memiliki akurasi sebesar 75,62%. Hal ini menunjukkan bahwa implementasi *transfer learning* efektif dalam meningkatkan kinerja ViT dalam klasifikasi multikelas pada gambar fase pertumbuhan padi serta termasuk kategori *excellent classification*.

Kata Kunci: Fase Pertumbuhan Padi, Kerangka Sampel Area, Klasifikasi Gambar, *Transfer Learning*, *Vision Transformer*

Performance Analysis of *Transfer learning*-Based *Vision Transformer* for Multiclass Classification of Rice Growth Phase Images

Rosa Salsabila Azarine

ABSTRACT

Deep learning-based image classification is a technique widely used in various fields, including the agricultural sector. In Indonesia, rice growth phase identification in the Area Sample Framework (ASF) is still done manually, which is time-consuming and carries the risk of observation errors. Therefore, this study aims to develop an automated model to classify rice growth phases, including early vegetative, late vegetative, generative, and harvest, using the *Vision Transformer* (ViT) architecture. The developed ViT-B/16 model is based on *transfer learning* to improve the accuracy and efficiency of the classification process.

The dataset used consists of 1,417 rice plant images obtained from the BPS (Statistics Indonesia) of Padang Pariaman Regency between November 2023 and February 2024. Preprocessing steps included adjusting the image dimensions to 224×224 pixels and normalizing to the $[0,1]$ range. The dataset was divided into training data (80%) and testing data (20%). Data augmentation was performed on the training data to increase image variation and reduce the risk of overfitting. Data training was conducted using two approaches: *transfer learning* and from scratch as a baseline, using several hyperparameter settings, including 10 epochs, a batch size of 32, a learning rate of 0.00005, an AdamW optimizer with a weight decay of 0.01, and a cross-entropy loss function.

The evaluation results showed that the model trained with the *transfer learning* approach achieved 90,46% accuracy and improved accuracy by 14,84% compared to the model trained from scratch, which only achieved 75,62% accuracy. This demonstrates that the *transfer learning* implementation is effective in improving ViT's performance in multi-class classification of rice growth phase images and is considered an excellent classification.

Keywords: Rice Growth Phase, Area Sampling Frame, Image Classification, *Transfer learning*, *Vision Transformer*

KATA PENGANTAR

Alhamdulillah, puji dan syukur penulis panjatkan kepada Allah SWT yang telah memberikan rahmat, karunia dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Kinerja *Vision Transformer* Berbasis *Transfer Learning* Untuk Klasifikasi Multikelas Pada Gambar Fase Pertumbuhan Padi”. Shalawat dan salam senantiasa tercurahkan 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. Selama penulisan skripsi ini, penulis banyak mendapat dukungan, bimbingan dan bantuan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang mendalam kepada:

1. Ibu Dr. Yenni Kurniawati., S.Si., M.Si., sebagai Dosen Pembimbing sekaligus Kepala Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Padang. Bimbingan, arahan, nasehat, motivasi dan ilmu yang Ibu berikan selama pembimbingan akademik maupun penyusunan skripsi sangat membantu dalam menyelesaikan studi ini dengan baik.
2. Bapak Dr. Syafriandi, M.Si., dan Ibu Dwi Sulistiowati, S.Si., M.Stat., sebagai Dosen Pembahas Skripsi yang telah memberikan arahan dan masukan selama penyusunan dan penyelesaian skripsi.
3. Seluruh Dosen dan Staf Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang, atas ilmu dan dukungan moral yang selalu diberikan selama masa studi.

4. Teristimewa kepada Ayahanda Taufik Hidayat dan Ibunda Ria Angela. Terima kasih atas doa yang tiada henti, dukungan, semangat, pengorbanan yang telah diberikan, serta kasih sayang yang tak terhingga.
5. Semua pihak yang telah memberikan dukungan, bantuan, dan kontribusi, selama penyusunan skripsi ini sehingga dapat terselesaikan dengan baik.

Semoga segala bantuan, bimbingan, dan dukungan yang telah diberikan mendapatkan balasan kebaikan dari Allah SWT. Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. 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 pembaca dan pihak-pihak yang membutuhkan.

Padang, 25 Januari 2026

Penulis

Rosa Salsabila Azarine

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR.....	iii
DAFTAR ISI	v
DAFTAR TABEL	vi
DAFTAR GAMBAR	vii
DAFTAR LAMPIRAN	viii
BAB I PENDAHULUAN.....	1
A. Latar Belakang Masalah	1
B. Batasan Masalah.....	9
C. Rumusan Masalah	9
D. Tujuan Penelitian.....	10
E. Manfaat Penelitian	10
BAB II KERANGKA TEORITIS.....	11
A. Kajian Teori	11
B. Penelitian Relevan.....	46
BAB III METODOLOGI PENELITIAN	49
A. Jenis Penelitian	49
B. Jenis dan Sumber Data	49
C. Variabel Penelitian.....	49
D. Teknik Analisis Data.....	52
BAB IV HASIL DAN PEMBAHASAN.....	55
A. Deskripsi Data	55
B. Hasil Analisis.....	57
C. Pembahasan	84
BAB V PENUTUP	87
A. Kesimpulan.....	87
B. Saran	87
DAFTAR PUSTAKA.....	89
LAMPIRAN.....	94

DAFTAR TABEL

Tabel	Halaman
1. Fase Amatan KSA	16
2. Jenis Model ViT	37
3. Hyperparameter Arsitektur Model	38
4. Hyperparameter Pelatihan	39
5. Confusion Matrix 4×4	44
6. Tingkatan Akurasi	46
7. Struktur Data	51
8. Pembagian Data Training dan Data Testing	58
9. Jumlah Data Training Setelah Augmentasi Data	60
10. Representasi Numerik Sampel Piksel	64
11. Model ViT-B/16	75
12. Hasil Tuning Hyperparameter	77
13. Kinerja Pelatihan Model ViT-B/16	77
14. <i>Confusion Matrix</i> ViT-B/16 (<i>From Scratch</i>)	80
15. <i>Confusion Matrix</i> ViT-B/16 (<i>Transfer learning</i>)	80
16. Metrik Evaluasi Model	82

DAFTAR GAMBAR

Gambar	Halaman
1. Tahap Pembangunan Kerangka Sampel Area	11
2. Peta Stratifikasi Sawah	12
3. Ilustrasi Pembagian Wilayah dalam Blok dan Segmen	13
4. Ilustrasi Replikasi Model Random Sampling	14
5. Segmen Terpilih Hasil Seleksi	15
6. Fase Tumbuh Padi yang Diamati dalam KSA	18
7. Lapisan pada Deep Learning	20
8. Arsitektur <i>Vision Transformer</i> (ViT).....	30
9. Perbedaan Traditional Learning dan <i>Transfer learning</i>	39
10. Data Penelitian	50
11. Diagram Alir Penelitian	54
12. Potret Amatan Lahan Sawah (a) Vegetatif Awal, (b) Vegetatif Akhir, (c) Generatif, (d) Panen	55
13. Distribusi Proporsi Kelas Fase Pertumbuhan Padi	56
14. Hasil Preprocessing Data (a) Sebelum Resize, (b) Sesudah Resize	57
15. Hasil Augmentasi Data (a) Asli, (b) Rotasi, (c) Horizontal Flipping, (c) Brightness Adjustments.....	59
16. Arsitektur Model ViT-B/16	62
17. <i>Patch</i> Gambar	63
18. (a) Grafik Accuracy, (b) Grafik Loss	79
19. Hasil Klasifikasi Gambar	83

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Gambar Padi Fase Vegetatif Awal	94
2. Data Gambar Padi Fase Vegetatif Akhir	95
3. Data Gambar Padi Fase Generatif.....	96
4. Data Gambar Padi Fase Panen	97
5. Syntax Python: Ekstraksi Bobot ImageNet-1k	98
6. Syntax Python: <i>Vision Transformer</i>	100
7. Output Syntax	109

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Klasifikasi adalah proses atau metode untuk memprediksi label kelas kategori dari suatu data berdasarkan pola karakteristik atau atributnya (Han dkk., 2012). Tujuan utama klasifikasi adalah mendukung analisis data, membantu dalam pengambilan keputusan serta meningkatkan pemahaman terhadap pola data. Peran klasifikasi menjadi sangat penting karena mampu menyederhanakan data yang kompleks sehingga lebih terstruktur dan mudah dipahami. Dengan berbagai metode klasifikasi, data yang beragam dapat disusun secara sistematis sehingga memungkinkan diperolehnya informasi dan kesimpulan yang tepat dan akurat.

Dalam ilmu statistika dan *machine learning*, metode klasifikasi tradisional banyak diterapkan pada data numerik atau data tabular, dimana variabel-variabel telah ditentukan secara jelas dan kelas muncul sebagai hasil analisis model. Namun, kemajuan teknologi digital di era *big data* berdampak pada peningkatan volume data tidak terstruktur seperti data teks, data audio, data video dan data gambar. Studi literatur menunjukkan bahwa sekitar 80% dari total volume data digital saat ini merupakan data tidak terstruktur yang sulit diolah menggunakan basis data relasional biasa (Balducci & Marinova, 2018). Hal ini mendorong kebutuhan akan teknik klasifikasi yang lebih adaptif, akurat dan mampu menangani kompleksitas pola visual dalam skala besar.

Seiring dengan meningkatnya dominasi data visual dalam berbagai bidang, kebutuhan akan metode klasifikasi tidak lagi terbatas pada data tabular, tetapi juga meluas pada data berbasis citra yang memiliki struktur dan karakteristik jauh lebih

kompleks. Hal ini mendorong perkembangan metode klasifikasi gambar yang merupakan strategi paling sederhana untuk mengekstrak informasi yang terkandung pada sebuah gambar dan memprediksi kelas dari sejumlah kelas terbatas yang telah ditentukan sebelumnya, berdasarkan gambar input (Fleuret, 2024). Menurut Chollet (2021), klasifikasi gambar adalah proses memberikan satu atau lebih label pada sebuah gambar sesuai dengan kategori yang relevan. Oleh karena itu, metode klasifikasi gambar terus dikembangkan dengan tujuan untuk membantu dalam mengenali objek, menganalisis kondisi visual, hingga mengidentifikasi kategori tertentu secara otomatis.

Perkembangan metode klasifikasi gambar dimulai dari metode tradisional yang mengandalkan ekstraksi fitur manual hingga metode modern berbasis *deep learning* yang mampu mempelajari fitur secara otomatis. Metode klasifikasi gambar tradisional merupakan pendekatan awal yang menggunakan ekstraksi fitur manual seperti *Histogram of Oriented Gradients* (HOG), *Scale-Invariant Feature Transform* (SIFT), dan *Gray Level Co-occurrence Matrix* (GLCM) (Upadhyay & Gupta, 2024; Hallur & Gavade, 2025). Setelah fitur diperoleh, algoritma klasifikasi klasik seperti *Support Vector Machine* (SVM), *K-Nearest Neighbors* (KNN), atau *Random Forest* digunakan untuk menentukan kelas gambar (Nhaila dkk., 2019). Namun, metode klasifikasi gambar tradisional memiliki keterbatasan karena sangat bergantung pada kualitas ekstraksi fitur manual dan kurang mampu menangani variasi visual yang kompleks.

Kebutuhan akan informasi secara *real-time*, menuntut adanya perkembangan teknik pengolahan data secara cepat dengan hasil akurat, termasuk dalam hal klasifikasi gambar. Salah satu kemajuan penting dalam pengembangan metode

klasifikasi gambar adalah teknik *machine learning* modern, khususnya *deep learning*, yang memungkinkan proses identifikasi visual dilakukan secara otomatis dan cepat dengan tingkat akurasi yang tinggi. *Deep learning* merupakan cabang *machine learning* yang terdiri dari algoritma pemodelan abstraksi data tingkat lanjut yang menggunakan serangkaian fungsi transformasi nonlinear yang diatur berlapis-lapis dan mendalam (Cholissodin dkk., 2020).

Menurut Mohamed dkk. (2022), *deep learning* adalah metode untuk melakukan klasifikasi, pengelompokan dan prediksi melalui pemanfaatan jaringan saraf tiruan yang dilatih menggunakan data dalam jumlah besar. Berbeda dengan metode klasifikasi gambar tradisional, *deep learning* menggabungkan proses ekstraksi fitur gambar dan klasifikasi dalam satu jaringan sehingga model dapat belajar secara *end-to-end* tanpa memerlukan perancangan fitur secara manual (Aurisano & Whitehead, 2022). Dengan demikian, *deep learning* mampu menghasilkan representasi visual yang lebih kaya dan akurat sehingga meningkatkan performa dalam tugas klasifikasi gambar, khususnya pada gambar dengan pola yang kompleks dan data berskala besar.

Klasifikasi gambar banyak diaplikasikan dalam berbagai bidang, mulai dari bidang kesehatan, otomotif, hingga bidang keamanan. Dalam bidang kesehatan, klasifikasi gambar digunakan untuk membantu proses identifikasi penyakit melalui citra medis, seperti *Magnetic Resonance Imaging* (MRI), *Computed Tomography* (CT) *scan*, dan proyeksi pencitraan *X-rays* (Mohamed dkk., 2022). Dengan memanfaatkan algoritma pembelajaran mesin, berbagai kelainan pada organ tubuh dapat dikenali secara otomatis dengan akurasi tinggi, sehingga mendukung tenaga medis melakukan evaluasi dan diagnosis (Litjens dkk., 2017). Dalam industri

otomotif, teknik klasifikasi gambar memiliki peran penting dalam mendukung pengembangan kendaraan otonom, khususnya untuk mengenali berbagai objek di lingkungan sekitar seperti pejalan kaki, kendaraan lain, maupun rambu-rambu lalu lintas (Janai dkk., 2020). Sementara itu, dalam bidang keamanan, klasifikasi gambar dimanfaatkan pada sistem keamanan berbasis pengenalan wajah (Wang & Deng, 2020), mendeteksi perilaku mencurigakan guna membantu meningkatkan keamanan publik.

Selain itu, klasifikasi gambar juga memberikan kontribusi dalam bidang pertanian. Di sektor pertanian, klasifikasi gambar membantu dalam memantau kondisi dan kesehatan tanaman serta mendeteksi hama atau penyakit melalui analisis visual. Hal ini bertujuan untuk meningkatkan produktivitas dan kualitas hasil pertanian serta diharapkan mampu mengurangi biaya operasional (Fracarolli dkk., 2020). Klasifikasi gambar juga berperan dalam pengelolaan sumber daya pertanian, salah satu contohnya adalah kegiatan Survei Kerangka Sampel Area (KSA). Survei ini dilakukan untuk mengumpulkan data pertanian yang akurat dengan cara mengamati langsung kondisi lahan pertanian. Penerapan KSA dalam bidang pertanian merupakan salah satu bentuk upaya pemerintah dalam mendukung tujuan pembangunan nasional dan tujuan pembangunan berkelanjutan (SDGs), khususnya tujuan kedua yaitu Tanpa Kelaparan (*Zero Hunger*) (BPS, 2024a).

Survei KSA merupakan metode pendataan berbasis area yang dilakukan oleh Badan Pusat Statistik (BPS) secara rutin setiap bulan. KSA dilaksanakan dengan cara melakukan pengamatan langsung terhadap titik-titik sampel tertentu dan bertujuan untuk mengestimasi luasan panen dengan ekstrapolasi dari sampel ke populasi dalam periode waktu yang relatif singkat (BPS, 2024c). Metode ini

diterapkan sebagai solusi atas masalah akurasi data luas panen yang lemah karena pengumpulan data yang dilakukan melalui pengukuran secara subjektif. Oleh karena itu, KSA hadir untuk memperbaiki metode pengumpulan data konvensional (*eye estimate*) menjadi lebih objektif, ilmiah dan modern, serta memanfaatkan peran teknologi agar data pertanian yang dikumpulkan lebih akurat dan tepat waktu. Output dari survei KSA diharapkan dapat dijadikan acuan untuk memprediksi waktu panen sekaligus untuk mengestimasi luas panen tanaman pangan, seperti padi dan jagung (Ruslan, 2019).

Dalam pelaksanaannya, KSA memberikan informasi terkait perkembangan fase tumbuh tanaman padi setiap bulannya. Tanaman padi menjadi fokus utama dalam survei KSA karena perannya yang penting sebagai komoditas pangan strategis di Indonesia (Mubekti & Sumargana, 2016). Padi merupakan komoditas utama pertanian di Indonesia karena memegang peranan krusial dalam perekonomian Indonesia. Selain itu, mayoritas penduduk Indonesia menjadikan padi sebagai sumber karbohidrat utama. Oleh karena itu, untuk menjaga ketahanan pangan nasional pemerintah berupaya untuk memastikan produksi padi di Indonesia mencukupi kebutuhan masyarakat Indonesia.

Produksi padi dalam skala besar memperoleh dukungan kuat di Indonesia, termasuk di Provinsi Sumatera Barat. Pada tahun 2024, produksi padi di Sumatera Barat diproyeksikan mencapai total produksi sekitar 1.352.049 ton padi, sehingga menempatkannya sebagai salah satu daerah dengan kontribusi produksi padi terbesar secara nasional (BPS, 2024b). Di tingkat kabupaten, Padang Pariaman menjadi salah satu lumbung pangan penting bagi Sumatera Barat. Sepanjang tahun 2024, wilayah ini memberikan kontribusi yang cukup besar terhadap total produksi

provinsi dengan total panen sekitar 132 ribu ton padi (BPS, 2024b). Selain memenuhi kebutuhan masyarakat setempat, produksi padi dari Padang Pariaman juga berfungsi menjaga pasokan beras tetap stabil, yang merupakan komponen kunci dalam menunjang ketahanan pangan Sumatera Barat secara berkelanjutan. Untuk memastikan akurasi perhitungan data hasil produksi padi, survei KSA dilakukan sebagai pendekatan yang lebih objektif dibandingkan survei konvensional.

Dalam prosedurnya, KSA membagi wilayah ke dalam blok, segmen dan subsegmen untuk kemudian dilakukan sampling acak. Segmen yang digunakan berukuran $300 \text{ m} \times 300 \text{ m}$ dengan lokasi yang telah ditetapkan (BPS, 2023). Pengamatan dilakukan pada setiap titik pusat subsegmen dalam segmen terpilih untuk mengidentifikasi fase pertumbuhan padi. Setiap bulannya petugas KSA memotret hasil amatan dengan menggunakan *handphone* berbasis Android, dimana hasil potret tersebut akan diklasifikasikan secara manual oleh petugas survei dan dikirimkan ke *server* pusat untuk diolah (Ruslan, 2019).

Proses klasifikasi gambar padi secara manual oleh petugas memiliki keterbatasan dalam hal efisiensi waktu, kecepatan dan ketepatan. Selain proses validasi data di tingkat kabupaten/kota yang memakan waktu, resiko terjadinya kesalahan klasifikasi juga cukup tinggi. 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.

Untuk meningkatkan efisiensi KSA, diperlukan langkah-langkah perbaikan. Salah satu langkah penting adalah melatih petugas lapangan secara berkelanjutan.

Selain itu, penerapan teknologi yang lebih canggih dalam identifikasi fase tumbuh tanaman padi menjadi solusi efektif yang perlu dipertimbangkan. Sehingga, data dapat diperoleh secara cepat dengan hasil yang lebih akurat dan dapat diandalkan. Penelitian sebelumnya telah berupaya mengoptimalkan pemantauan fenologi padi ini menggunakan teknik klasifikasi *ensemble* untuk menangani data yang tidak seimbang (Kurniawati dkk., 2024). Salah satu cara untuk meningkatkan efisiensi dan akurasi dalam KSA yaitu dengan memanfaatkan teknologi *deep learning*. Hal ini sejalan dengan penelitian Kurniawati dkk. (2025) yang menerapkan konsep *smart farming* berbasis *deep learning* untuk pemantauan klasifikasi lahan pertanian tanaman padi, di mana teknologi ini terbukti mampu membantu proses identifikasi fase pertumbuhan padi secara lebih otomatis dan presisi.

Deep learning sebagai cabang dari *machine learning* dan kecerdasan buatan mampu menganalisis gambar tanaman padi dan mengidentifikasi fase pertumbuhan padi secara otomatis. Salah satu arsitektur *deep learning* yang paling berpengaruh dalam perkembangan klasifikasi gambar adalah *Convolutional Neural Network* (CNN). CNN menggunakan konsep konvolusi, yaitu teknik pemrosesan data dengan menggunakan filter atau kernel pada seluruh bagian citra (Sukandar dkk., 2024). Kelebihan dari metode CNN adalah dapat belajar dan mengekstrak fitur kompleks dalam keadaan jaringan yang tinggi pada data gambar (Rahmadani dkk., 2024). Meskipun efektif, CNN memiliki keterbatasan dalam menangkap hubungan global pada gambar karena *receptive field* yang bersifat lokal dan ketergantungannya pada desain kernel. Hal ini membuat CNN kurang optimal untuk memahami struktur visual secara menyeluruh. Untuk mengatasi keterbatasan

tersebut, *Vision Transformer* (ViT) muncul sebagai pilihan yang menjanjikan dibandingkan CNN (Omer, 2024).

ViT merupakan arsitektur pengolahan gambar yang menggunakan dasar dari arsitektur *Transformer*. ViT menggunakan mekanisme *self-attention* dari *Transformer* untuk menggantikan pengguna konvolusi umum dalam CNN. Pendekatan ini memungkinkan ViT mencapai kinerja maksimal dengan waktu yang singkat (Uthama dkk., 2024). Model ViT terdiri dari beberapa komponen, yaitu *patch embedding*, *transformer encoder*, dan *classification head* (Dosovitskiy dkk., 2021). Sehingga ViT mampu menangkap hubungan antar bagian gambar secara langsung dan menghasilkan performa yang lebih baik dibanding CNN. Meskipun ViT menunjukkan kinerja unggul, model ini membutuhkan jumlah data yang sangat besar agar dapat dilatih secara optimal. Namun, dalam bidang pertanian seperti gambar fase pertumbuhan padi, ketersediaan data sangat terbatas. Oleh karena itu, *transfer learning* menjadi pendekatan penting guna meningkatkan kinerja ViT.

Transfer learning adalah suatu teknik dalam pembelajaran mesin di mana pengetahuan yang telah diperoleh dari suatu domain atau tugas digunakan kembali sebagai titik awal untuk meningkatkan performa model pada domain atau tugas lain yang terkait (Pan & Yang, 2009). *Transfer learning* bekerja dengan mengambil parameter atau representasi fitur yang telah dipelajari pada tugas sumber dan kemudian digunakan kembali pada tugas target melalui strategi seperti *feature extraction* dan *fine-tuning* (Zhuang dkk., 2020). Dengan penerapan *transfer learning*, model dapat dilatih lebih cepat dan efisien dengan tujuan untuk meningkatkan performa atau mengurangi kebutuhan data pada tugas baru.

Penelitian yang dilakukan oleh Ayana dkk. (2023) menerapkan *Vision Transformer* (ViT) berbasis *transfer learning* untuk melakukan klasifikasi citra mammogram. Pada penelitian ini model ViT yang telah di latih pada dataset ImageNet digunakan kembali sebagai *feature extractor* utama, kemudian disesuaikan pada dataset *Digital Database for Screening Mammography* (DDSM) untuk klasifikasi. Hasil penelitian menunjukkan bahwa penerapan *transfer learning* pada ViT mencapai performa sangat tinggi dengan nilai akurasi sempurna (100%) pada seluruh model ViT yang diuji. Berlandaskan pada efektivitas ViT dalam klasifikasi gambar dan potensi peningkatannya melalui *transfer learning*, penelitian ini mengambil judul “**Analisis Kinerja Vision Transformer Berbasis Transfer learning Untuk Klasifikasi Multikelas Pada Gambar Fase Pertumbuhan Padi**”.

B. Batasan Masalah

Berdasarkan latar belakang yang telah dijelaskan, batasan masalah dalam penelitian ini adalah data yang digunakan merupakan data gambar fase pertumbuhan padi yang terdiri atas fase vegetatif awal, vegetatif akhir, generatif, dan panen yang diperoleh dari *website* petugas KSA BPS Kabupaten Padang Pariaman selama empat periode waktu, yaitu November 2023 hingga Februari 2024. Model ViT yang digunakan pada penelitian ini adalah ViT-B/16.

C. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut.

1. Bagaimana konfigurasi *hyperparameter* yang memberikan kinerja optimal pada model ViT berbasis *transfer learning* dalam mengklasifikasikan fase pertumbuhan padi?
2. Bagaimana kinerja ViT berbasis *transfer learning* dalam melakukan klasifikasi gambar fase pertumbuhan padi?

D. Tujuan Penelitian

Berdasarkan masalah yang telah dirumuskan, penelitian ini dirancang dengan tujuan sebagai berikut.

1. Untuk menentukan konfigurasi *hyperparameter* yang menghasilkan kinerja optimal pada model ViT berbasis *transfer learning* dalam mengklasifikasikan fase pertumbuhan padi.
2. Untuk menganalisis kinerja ViT berbasis *transfer learning* dalam melakukan klasifikasi gambar fase pertumbuhan padi.

E. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat praktis dan teoritis, antara lain sebagai berikut.

1. Bagi penulis, penelitian ini memberikan pemahaman yang lebih mendalam mengenai penggunaan ViT dan *transfer learning* dalam klasifikasi gambar pertanian, khususnya fase pertumbuhan padi.
2. Bagi BPS, penelitian ini memberikan kontribusi dalam proses operasional kegiatan Survei KSA sehingga meminimalisir kesalahan dan efisiensi waktu dalam melakukan klasifikasi gambar fase pertumbuhan padi.
3. Bagi pembaca, penelitian ini dapat menjadi referensi dalam pengembangan metode klasifikasi gambar berbasis *deep learning* pada sektor pertanian.

BAB V

PENUTUP

A. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan, didapatkan kesimpulan sebagai berikut.

1. Berdasarkan klasifikasi gambar fase pertumbuhan padi (vegetatif awal, vegetatif akhir, generatif, dan panen) menggunakan metode *Vision Transformer* (ViT-B/16) berbasis *transfer learning*, konfigurasi *hyperparameter* yang terbukti memberikan kinerja optimal adalah kombinasi *epoch* sebanyak 10, *batch size* berukuran 32, dan *learning rate* dengan nilai 0,00005, dengan dukungan *optimizer* AdamW dengan *weight decay* standar yaitu 0,01 dan fungsi *cross-entropy loss*.
2. Model ViT yang dikembangkan dengan memanfaatkan pendekatan *transfer learning* memberikan akurasi hingga 90,46%, diikuti dengan *precision* 90,30%, *recall* 90,50%, dan F1-score 90,26%. Sedangkan *baseline* model (*from scratch*) memberikan akurasi hanya sebesar 75,62%, diikuti dengan *precision* 75,11%, *recall* 74,91%, dan F1-score 74,85%. Hal ini menunjukkan bahwa penerapan *transfer learning* efektif dalam meningkatkan akurasi model hingga 14,84%. Penerapan *transfer learning* pada ViT dapat menjadi pendekatan yang efisien dan andal dalam mendukung analisis klasifikasi gambar fase pertumbuhan padi, meskipun dalam kondisi dataset yang terbatas.

B. Saran

Penelitian selanjutnya disarankan untuk menambah jumlah dan variasi dari data yang digunakan. Selain itu, pengumpulan data dari lokasi dan kondisi yang

beragam termasuk variasi dalam jenis padi dan kondisi lingkungan, akan sangat bermanfaat untuk memperkuat akurasi model yang dikembangkan. Teknik augmentasi data yang lebih kompleks, seperti penggunaan *Generative Adversial Networks* (GANs), dapat dipertimbangkan untuk meningkatkan variasi data *training*, sehingga model dapat belajar dari variasi yang lebih luas.

Evaluasi berkelanjutan terhadap model yang dikembangkan dengan menggunakan data baru secara berkala 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 guna mendukung penerapan teknologi ini.

DAFTAR PUSTAKA

- Abdullah, F., Imram, S., & Rauf, A. (2022). Analisis Ketersediaan Beras Di Kabupaten Gorontalo Selang Tahun 2021-2030. *AGRINESIA: Jurnal Ilmiah Agribisnis*, 6(3), 187–197. <https://doi.org/10.37046/agr.v6i3.16138>
- Aurisano, A., & Whitehead, L. H. (2022). End-to-End Analysis Using Image Classification. In P. Calafiura, D. Rousseau, & K. Terao (Eds.), *Artificial Intelligence for High Energy Physics* (pp. 313–353). Word Scientific.
- Ayana, G., Dese, K., Dereje, Y., Kebede, Y., Barki, H., Amdissa, D., Husen, N., Mulugeta, F., Habtamu, B., & Choe, S.-W. (2023). Vision-Transformer-Based *Transfer learning* for Mammogram Classification. *Diagnostics*, 13(2), 178. <https://doi.org/10.3390/diagnostics13020178>
- Balducci, B., & Marinova, D. (2018). Unstructured data in marketing. *Journal of the Academy of Marketing Science*, 46, 557–590. <https://doi.org/10.1007/s11747-018-0581-x>
- Bishop, C. M. (2006). *Pettern Recognition and Machine Learning* (M. Jordan, J. Kleinberg, & B. Scholkopf (eds.)). Springer.
- BPS. (2018a). *Luas Panen dan Produksi Beras Di Indonesia 2018*. Badan Pusat Statistik.
- BPS. (2018b). *Upaya Perbaikan Data Padi Dengan Metode Kerangka Sampel Area (KSA) 2018*. Badan Pusat Statistik.
- BPS. (2023). *Survei Kerangka Sampel Area (KSA) Padi 2023*. Badan Pusat Statistik.
- BPS. (2024a). *Kerangka Acuan Kerja Survei Kerangka Sampel Area (KSA) Padi*. Badan Pusat Statistik Purworejo.
- BPS. (2024b). *Luas Panen dan Produksi Padi di Provinsi Sumatera Barat*. Badan Pusat Statistik Sumatera Barat.
- BPS. (2024c). *Yuk Mengenal Survei Kerangka Sampel Area (KSA)*. Badan Pusat Statistik Kabupaten Blitar. <https://blitarkab.bps.go.id/id/news/2024/07/02/342/yuk-mengenal--survei--kerangka-sampel-area--ksa-.html>
- Cholissodin, I., Soebroto, A. A., Hasanah, U., & Febiola, Y. I. (2020). *AI, Machine Learning & Deep Learning (Teori & Implementasi) “from Basic Science to High Scientific Solution for Any Problem.”* Fakultas Ilmu Komputer, Universitas Brawijaya. <http://bit.ly/3piOnnU>
- Chollet, F. (2021). *Deep Learning with Python* (2nd ed.). Manning Publication.
- Cochran, W. G., & Wiley, J. (1977). *Sampling Techniques* (3rd ed.).
- Deng, J., Dong, W., Socher, R., Li, L., Li, K., & Fei-fei, L. (2009). *ImageNet : A Large-Scale Hierarchical Image Database*. 2–9.

- Dhakal, A. S., Amada, T., & Aniya, M. (2000). *Landslide Hazard Mapping and its Evaluation Using GIS : An Investigation of Sampling Schemes for a Grid-Cell Based Quantitative Method*. 66(8), 981–989.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Deghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An Image Is Worth 16x16 Words: Transformer For Image Recognition At Scale. *International Conference on Learning Representations (ICLR)*. <https://doi.org/10.48550/arXiv.2010.11929>
- Fathurrahman, I., Djamaluddin, M., Amri, Z., & Wathani, N. M. (2025). Klasifikasi Motif Batik Nusantara Menggunakan *Vision Transformer* (ViT) Berbasis Deep Learning. *Jurnal Informatika Dan Teknologi*, 8(2), 511–522. <https://doi.org/10.29408/jit.v8i2.31108>
- Fleuret, F. (2024). *The Little Book of Deep Learning*. University of Geneva. <https://fleuret.org/public/lbdl.pdf>
- Fracarolli, J. A., Fernandes, F., Pavarin, A., Castro, W., & Blasco, J. (2020). Computer Vision Applied To Food and Agricultural Products. *Revista Ciência Agronômica*, 51(Special Agriculture 4.0), 1–20. <https://doi.org/10.5935/1806-6690.20200087>
- Geron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (2nd ed.). O'Reilly Media.
- Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Gorunescu, F. (2011). *Data Mining Concept, Models and Techniques*. Springer.
- Haba, D. (2023). *Data Augmentation with Python: enhance accuracy in deep learning with practical data augmentation for image, text, audio and tabular data* (1st ed.). Packt Publishing.
- Hallur, S., & Gavade, A. (2025). Image Feature Extraction Techniques : A Comprehensive Review. *Franklin Open*, 12(January), 100366. <https://doi.org/10.1016/j.fraope.2025.100366>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (3rd ed.). Elsevier Inc.
- Hendrycks, D., & Gimpel, K. (2016). *Gaussian Error Linear Units (GELUs)*. 1–10. <https://doi.org/10.48550/arXiv.1606.08415>
- Hutami, T. R. (2024). *Analisis Perbandingan Model Klasifikasi Foto Fase Pertumbuhan Padi Dengan Algoritma CNN (Studi Kasus: Survei KSA Badan Pusat Statistik)* (Vol. 23220011). Institut Teknologi Bandung.
- Janai, J., Güney, F., Behl, A., & Geiger, A. (2020). *Computer Vision for Autonomous Vehicles Problems, Datasets and State of the*. Foundations and Trends® in

Computer Graphics and Vision. <https://doi.org/10.1561/06000000079>

- Kurniawati, Y., Fitria, D., Permana, D., Martha, Z., & Syafriandi. (2025). *Pemantauan Klasifikasi Lahan Pertanian Tanaman Padi Dengan Konsep Smart Farming Menggunakan Deep Learning*.
- Kurniawati, Y., Wijayanto, H., Kurnia, A., Dirgahayu, D., & Susetyo, B. (2024). Rice phenology monitoring via ensemble classification for an extremely imbalanced multiclass dataset of hybrid remote sensing. *Remote Sensing Applications: Society and Environment*. <https://doi.org/https://doi.org/10.1016/j.rsase.2024.101246>
- Litjens, G., Kooi, T., Bejnordi, B. E., Arindra, A., Setio, A., Ciompi, F., Ghafoorian, M., Laak, J. A. W. M. Van Der, Ginneken, B. Van, & Sánchez, C. I. (2017). *A Survey On Deep Learning In Medical Image Analysis*. 42(December 2012), 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Mohamed, A. H., Refaat, M., & Hemeida, A. M. (2022). Image Classification Based Deep Learning : A Review. *Aswan University Journal of Science and Technology*, 2(1), 11–35. <https://doi.org/10.1109/CVPR.2016.90>
- Mubekti, & Sumargana, L. (2016). Pendekatan Kerangka Sampel Area untuk Estimasi dan Peramalan Produksi Padi (Area Frame Sampling Approach for Estimating and Predicting Rice Production). *Jurnal PANGAN*, 25(2), 71–82.
- Muller, A. C., & Guido, S. (2016). *Introduction to Machine Learning with Python* (D. Schanafelt (ed.)). O'Reilly Media.
- Mumuni, A., & Mumuni, F. (2022). Data augmentation : A Comprehensive Survey of Modern Approaches. *Array*, 16(August), 100258. <https://doi.org/10.1016/j.array.2022.100258>
- Nhaila, H., Elmaizi, A., Sarhrouni, E., & Hammouch, A. (2019). Science Direct Supervised Classification Methods Applied To Airborne Hyperspectral Images : Comparative Study Using Mutual Information. *Intelligent Computing in Data Sciences (ICDS)*, 00, 1–10. <https://doi.org/10.1016/j.procs.2019.01.013>
- Omer, A. A. M. (2024). Image Classification Based on *Vision Transformer*. *Journal of Computer and Communication*, 12, 49–59. <https://doi.org/10.4236/jcc.2024.124005>
- Ouamane, A., Chouchane, A., Himeur, Y., Miniaoui, S., Atalla, S., Mansoor, W., & Al-Ahmad, H. (2025). Optimized *Vision Transformers* for Superior Plant Disease Detection. *IEEE Access*, 13(February), 48552–48570. <https://doi.org/10.1109/ACCESS.2025.3547416>
- Pan, S. J., & Yang, Q. (2009). A Survey on *Transfer learning*. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
- Pangestu, A., Purnama, B., & Risnandar. (2024). *Vision Transformer Untuk Klasifikasi Kematangan Pisang*. *Jurnal Teknologi Informasi Dan Ilmu*

- Komputer (JTIK)*, 11(1), 75–84.
<https://doi.org/10.25126/jtiik.20241117389>
- Patria, D. G., Sukamto, & Sumarji. (2021). *Rice Science and Technology (Ilmu dan Teknologi Beras)*. Literasi Nusantara.
- Rahmadani, R., Kurniawati, Y., Permana, D., & Fitria, D. (2024). Classification of Harvest-Non Harvest in Rice Plant Image Using Convolutional Neural Network Algorithm. *UNP Journal of Statistics and Data Science (UJSDS)*, 2(3), 249–256. <https://doi.org/10.24036/ujsds/vol2-iss3/181>
- Ruslan, K. (2019). *Memperbaiki Data Pangan Indonesia Lewat Metode Kerangka Sampel Area*. <https://doi.org/10.35497/287781>
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(60), 1–48. <https://doi.org/10.1186/s40537-019-0197-0>
- Sukandar, C. I., Anggraeny, T. F., & Swari, P. H. M. (2024). Pengaruh Optimasi Dalam Klasifikasi Tumor Otak Dengan Hibrida CNN-ViT. *ANTIVIRUS: Jurnal Ilmiah Teknik Informatika*, 18(1), 112–124. <https://doi.org/10.35457/antivirus.v18i1.3557> 112
- Szeliski, R. (2022). *Computer Vision : Algorithms and Applications 2nd Edition*. Springer.
- Upadhyay, U., & Gupta, S. (2024). A Survey on Image Feature Extraction Techniques. *International Journal of Scientific Research & Engineering Trends (IJSRET)*, 10(3), 243–255.
- Uthama, R., Hendrik, B., Informatika, M. T., Komputer, F. I., Indonesia, U. P., & Riau, U. M. (2024). Vision Transformer untuk Identifikasi 15 Variasi Citra Ikan Koi. *Jurnal Computer Science and Information Technology*, 5(1), 159–168. <https://doi.org/10.37859/coscitech.v5i1.6711>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, N. A., & Kaiser, L. (2017). Attention Is All You Need. *Neural Information Processing Systems (NIPS)*. <https://doi.org/10.48550/arXiv.1706.03762>
- Wang, M., & Deng, W. (2020). Deep Face Recognition : A Survey. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2020.10.081>
- Xiong, R., Yang, Y., He, D., Zheng, K., Zheng, S., Xing, C., & Zhang, H. (2020). On Layer Normalization in the Transformer Architecture.
- Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks ? *Neural Information Processing Systems (NIPS)*, 27, 3320–3328. <https://doi.org/10.48550/arXiv.1411.1792>
- Zhang, A., Lipton, Z. C., Li, M. U., & Alexander, J. (2023). *Dive into Deep Learning*. Cambridge University Press. <https://doi.org/10.48550/arXiv.2106.11342>
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., & Member, S. (2020). A

Comprehensive Survey on *Transfer learning*. *Proceedings of the IEEE*, 109(1), 43–76. <https://doi.org/10.48550/arXiv.1911.02685>