

Klasifikasi Jenis Beras Menggunakan Deep Learning Berbasis Computer Vision dengan Platform Roboflow

Imron Hefni Fauzi ¹, Muhamad Soleh Sulaeman ², Tri Herdiawan Apandi ³, Willy Muhammad Fauzi ⁴

^{1,2,3,4} Teknologi Rekayasa Perangkat Lunak /Jurusan Teknologi Informasi dan Komputer

Politeknik Negeri Subang

imron.h.fauzi@gmail.com; muhamadsoleh871@gmail.com; Tri@polsub.ac.id; Willy@polsub.ac.id;

Blok Kaleng Banteng, Desa Cibogo, Kabupaten Subang, Jawa Barat, 41285

Keywords:

*Rice Variety
Classification,
Computer Vision,
Convolutional
Neural Network
(CNN), Image
Classification,
Machine Learning.*

Abstract

Rice is a fundamental food commodity globally, where accurate variety classification is crucial for pricing, quality control, and food security. Manual classification methods are labor-intensive, time-consuming, and prone to human subjectivity. This research proposes an automated classification system for five rice varieties (Arborio, Basmati, Ipsala, Jasmine, and Karacadag) using a Computer Vision approach with Vision Transformer (ViT) architecture. Unlike Convolutional Neural Networks (CNN) which focus on local features, ViT utilizes self-attention mechanisms to capture global contextual relationships within images. The model was developed using the public "Rice Image Dataset" containing 75,000 images. The methodology includes image preprocessing (resizing and normalization) and training of the ViT Classification model. Model performance was evaluated using standard metrics on a separate test set. The results show that the proposed ViT model achieved an outstanding accuracy of 99.9%. These findings demonstrate that the Transformer-based approach is highly effective and efficient for automating rice variety identification, offering a more robust solution compared to conventional methods.

Kata Kunci:

*Klasifikasi Varietas
Beras, Visi Komputer,
Convolutional Neural
Network (CNN),
Klasifikasi Citra,
Pembelajaran Mesin.*

Abstrak

Beras merupakan komoditas pangan fundamental secara global, di mana klasifikasi varietas yang akurat sangat krusial untuk penentuan harga, kontrol kualitas, dan ketahanan pangan. Metode klasifikasi manual bersifat padat karya, memakan waktu, dan rentan terhadap subjektivitas manusia. Penelitian ini mengusulkan sistem klasifikasi otomatis untuk lima varietas beras (Arborio, Basmati, Ipsala, Jasmine, dan Karacadag) menggunakan pendekatan Computer Vision dengan arsitektur Vision Transformer (ViT). Berbeda dengan Convolutional Neural Network (CNN) yang berfokus pada fitur lokal, ViT memanfaatkan mekanisme self-attention untuk menangkap hubungan kontekstual global pada citra. Model dikembangkan menggunakan "Rice Image Dataset" publik yang berisi 75.000 citra. Metodologi mencakup pra-pemrosesan citra (resizing dan normalisasi) serta pelatihan model ViT Classification. Kinerja model dievaluasi menggunakan metrik standar pada dataset uji terpisah. Hasil pengujian menunjukkan bahwa model ViT yang diusulkan mampu mencapai akurasi luar biasa sebesar 99.9%. Temuan ini membuktikan bahwa pendekatan berbasis Transformer sangat efektif dan efisien untuk otomatisasi identifikasi varietas beras, menawarkan solusi yang lebih tangguh dibandingkan metode konvensional.

1. Pendahuluan

Beras (*Oryza sativa*) merupakan salah satu komoditas pangan terpenting di dunia, menjadi makanan pokok bagi lebih dari separuh populasi global. Keamanan pangan, stabilitas ekonomi, dan kontrol kualitas dalam rantai pasok beras sangat bergantung pada kemampuan untuk mengidentifikasi dan mengklasifikasikan varietasnya secara akurat. Secara tradisional, klasifikasi dan pengukuran dimensi beras masih mengandalkan model pengukuran manual menggunakan alat seperti mikrometer atau kaliper yang memakan waktu dan bersifat subjektif [1]. Untuk mengatasi keterbatasan ini, teknologi pengolahan citra digital (digital image processing) hadir sebagai solusi untuk memperbaiki kualitas interpretasi visual dan memberikan hasil pengukuran yang lebih objektif dan cepat [2]. Pemanfaatan teknologi ini telah meluas ke berbagai bidang biologi dan lingkungan, mulai dari aplikasi perikanan [3] hingga analisis distribusi gelembung udara dalam fluida [4].

Berbagai pendekatan pengolahan citra konvensional telah diterapkan untuk menangani masalah di sektor agrikultur dan identifikasi objek. Dalam konteks beras, penelitian terdahulu berfokus pada pengukuran dimensi fisik (panjang dan lebar) menggunakan teknik segmentasi, konversi ruang warna, dan binerisasi [1]. Selain itu, teknik perbaikan kualitas citra seperti penggunaan Median Filter untuk mereduksi noise [5] dan metode Image Inpainting untuk merestorasi bagian citra yang hilang [6] sering digunakan sebagai tahapan pra-pemrosesan. Untuk identifikasi fitur yang lebih kompleks, metode seperti operasi morfologi citra telah digunakan untuk mendeteksi penyakit pada daun padi [7], sementara metode pengelompokan warna (clustering) seperti Hierarchical Agglomerative Clustering digunakan untuk ekstraksi fitur objek berdasarkan warna [8]. Meskipun metode-metode ini efektif untuk tugas-tugas spesifik, mereka sangat bergantung pada feature engineering manual yang kaku dan rentan terhadap variasi kondisi citra.

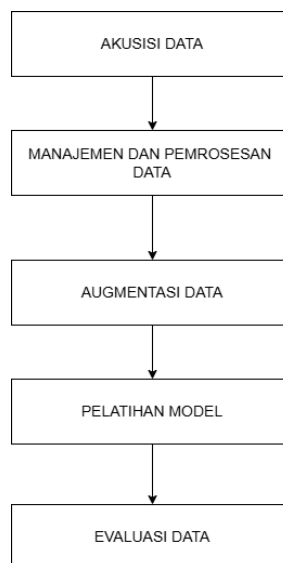
Perkembangan kecerdasan buatan, pendekatan Deep Learning telah menjadi standar baru karena kemampuannya mempelajari fitur secara otomatis. Salah satu arsitektur yang paling dominan adalah Convolutional Neural Network (CNN). Efektivitas CNN telah dibuktikan dalam berbagai studi kasus, seperti identifikasi jenis penyakit kulit yang mencapai akurasi tinggi [9] serta sistem identifikasi wajah pada keamanan yang terbukti mengungguli metode klasik seperti PCA dan SVM [10]. Keunggulan CNN terletak pada kemampuannya mengekstraksi fitur visual secara hierarkis tanpa intervensi manual yang ekstensif.

Namun, dalam beberapa tahun terakhir, arsitektur Vision Transformer (ViT) telah muncul sebagai terobosan baru yang menantang dominasi CNN. Berbeda dengan CNN yang memproses citra melalui kernel lokal (konvolusi), ViT membagi citra menjadi patch dan menggunakan mekanisme self-attention untuk memahami hubungan global antar bagian citra secara holistik. Kemampuan ViT untuk menangkap konteks global ini menjadikannya kandidat yang sangat potensial untuk membedakan varietas beras yang memiliki kemiripan visual tinggi namun memiliki detail tekstur yang halus, yang mungkin terlewatkan oleh metode konvensional maupun CNN standar.

Meskipun CNN telah banyak diterapkan, terdapat kesenjangan penelitian dalam penerapan arsitektur state-of-the-art seperti Vision Transformer (ViT) untuk klasifikasi varietas beras pada dataset berskala besar. Oleh karena itu, tujuan utama penelitian ini adalah untuk merancang, mengimplementasikan, dan mengevaluasi model Vision Transformer (ViT) untuk klasifikasi otomatis lima varietas beras (Arborio, Basmati, Ipsala, Jasmine, dan Karacadag). Kebaruan (novelty) penelitian ini terletak pada penerapan arsitektur ViT yang menggantikan pendekatan pengukuran manual [1], metode morfologi konvensional [7], maupun ketergantungan pada CNN [5], dengan tujuan mencapai tingkat akurasi dan generalisasi yang lebih tinggi dalam sistem kontrol kualitas beras otomatis.

2. Metodologi Penelitian

Penelitian ini mengadopsi pendekatan kuantitatif berbasis Computer Vision dan Deep Learning untuk mengotomatisasi identifikasi varietas beras. Metode yang digunakan mengintegrasikan dua pendekatan utama, yaitu Image Classification dan Object Detection, yang dikelola melalui platform end-to-end Roboflow. Untuk tugas klasifikasi, penelitian ini memanfaatkan arsitektur Convolutional Neural Network (CNN), yang telah terbukti memiliki kemampuan superior dalam mengekstraksi fitur visual kompleks dan mencapai akurasi tinggi pada dataset varietas beras (seperti Arborio, Basmati, dll), sebagaimana divalidasi dalam studi komprehensif oleh Hazra et al. (2023) [11]. Di sisi lain, untuk kebutuhan deteksi objek yang cepat dan efisien, penelitian ini menerapkan algoritma YOLO (You Only Look Once). Penggunaan YOLO didasarkan pada efektivitasnya dalam aplikasi real-time dan kemampuannya untuk berjalan pada perangkat dengan sumber daya terbatas (mobile), sejalan dengan temuan terbaru dari Zheng et al. (2025) yang mengembangkan model YOLO ringan untuk deteksi fenotipe padi. Kombinasi ini dipilih untuk memastikan sistem tidak hanya akurat dalam membedakan jenis beras tetapi juga efisien dalam proses komputasi [12].



Gambar 1. Metode Penelitian

Proses penelitian ini dibagi menjadi empat tahapan utama yang berurutan, memastikan seluruh proses pengolahan citra dan pembangunan model terlaksana secara sistematis:

2.1. Tahap pengumpulan dan akusisi data

a. Akusisi citra

Tahap ini dimulai dengan mengumpulkan citra bulir beras dari berbagai jenis (kelas) yang menjadi objek penelitian. Citra diambil dalam kondisi pencahayaan dan latar belakang (background) yang konsisten untuk meminimalisir noise

b. Organisasi dataset

Data citra diorganisasi berdasarkan kelas jenis beras yang telah ditentukan. Volume data harus mencukupi (direkomendasikan minimal ratusan hingga ribuan sampel per kelas) untuk memenuhi kebutuhan pelatihan model Deep Learning.

2.2. Tahap pra-pemrosesan dan anotasi data

a. Unggah data

Seluruh citra diunggah ke workspace Roboflow.

b. Anotasi

Setiap citra diberi label (anotasi) sesuai dengan jenis berasnya. Dalam konteks klasifikasi, seluruh gambar diberi label kelas (Image Classification).

c. Pembagian dataset

Dataset dibagi secara otomatis menjadi Data Pelatihan (Training), Data Validasi (Validation), dan Data Pengujian (Testing), dengan rasio pembagian standar (misalnya 70:15:15).

d. Pra-pemrosesan

Sumber data yang digunakan dalam penelitian ini adalah Rice Image Dataset yang bersumber dari Kaggle, terdiri dari lima varietas beras: Arborio, Basmati, Ipsala, Jasmine, dan Karacadag. Total citra yang berhasil diakuisisi dan dikelola ke dalam platform Roboflow adalah sebanyak 74.779 citra [1].

Dataset dibagi secara otomatis menggunakan rasio 70:20:10, dengan rincian:

- Training Set: 52.351 gambar (70%).
- Validation Set: 14.955 gambar (20%).
- Testing Set: 7.473 gambar (10%).

Sebelum dilakukan pelatihan, seluruh citra melewati tahapan pra-pemrosesan (preprocessing) standar untuk memastikan konsistensi data masukan terhadap arsitektur model. Langkah-langkah ini diterapkan pada seluruh subset data (Train, Valid, Test):

1. Auto-Orient: Diterapkan untuk mengoreksi orientasi gambar secara otomatis berdasarkan metadata EXIF, memastikan posisi objek tegak lurus.
2. Resize: Semua gambar diubah ukurannya (resizing) dari dimensi asli menjadi 640×640 piksel dengan metode Stretch. Resolusi ini dipilih untuk menyesuaikan dengan input layer dari model Transformer yang membutuhkan resolusi lebih tinggi untuk menangkap detail tekstur beras.

2.3. Augmentasi data

Augmentasi Data (misalnya, rotasi, flipping, perubahan kecerahan, cropping) melalui fitur Roboflow. Augmentasi (Augmentations). Berdasarkan konfigurasi pada dataset versi 1 (v1 2025-10-29), tidak ada teknik augmentasi tambahan (No augmentations applied) yang diterapkan pada dataset pelatihan. Model dilatih menggunakan data murni hasil pra-pemrosesan untuk menguji kemampuan model dalam mengenali fitur asli citra beras tanpa modifikasi sintetik.

2.4. Tahap pelatihan model

a. Pemilihan arsitektur model

Pemilihan Arsitektur Model: Dipilih arsitektur Deep Learning yang relevan (misalnya, model YOLO untuk deteksi objek atau CNN berbasis MobileNet/ResNet untuk klasifikasi) yang tersedia atau dapat diimplementasikan melalui Roboflow.

b. Inisiasi dan pelatihan

Inisiasi dan Pelatihan: Model dilatih menggunakan Data Pelatihan. Parameter pelatihan (*Hyperparameter*) seperti *epochs*, *batch size*, dan *learning rate* ditentukan untuk mengoptimalkan proses konvergensi model.

c. Validasi

Kinerja model dipantau selama pelatihan menggunakan Data Validasi untuk memastikan model belajar dengan efektif dan mendeteksi potensi *overfitting*.

2.5. Tahap pengujian dan evaluasi kinerja

a. Pengujian akhir

Model yang telah selesai dilatih diuji menggunakan Data Pengujian yang belum pernah dilihat oleh model.

b. Perhitungan metrik

Kinerja model diukur menggunakan metrik evaluasi standar Deep Learning, meliputi Akurasi (Accuracy), Presisi (Precision), Recall, dan F1-Score. Hasil ini direpresentasikan dalam Matriks Kebingungan (Confusion Matrix) untuk analisis detail.

c. Analisa hasil

Hasil kinerja model dikaji dan dibandingkan dengan studi literatur sebelumnya untuk menentukan keberhasilan model dalam mengklasifikasikan jenis beras menggunakan Roboflow.

3. Hasil dan Pembahasan

Analisis mendalam ini menyajikan dan menginterpretasikan hasil pengujian model Deep

Learning yang dikembangkan untuk tujuan klasifikasi jenis beras berbasis citra, menggunakan platform Roboflow. Pembahasan dimulai dengan penyajian data kuantitatif dari metrik kinerja model – termasuk Akurasi, Presisi, Recall, dan F1-Score – yang diperoleh dari pengujian pada Test Set independen. Kami menganalisis kinerja model secara spesifik, khususnya melalui Matriks Kebingungan, untuk mengidentifikasi keberhasilan dan tantangan dalam membedakan antar jenis beras yang memiliki kemiripan visual. Hasil temuan ini selanjutnya akan dibandingkan dengan pencapaian yang telah dilaporkan dalam studi literatur terdahulu, sehingga dapat dijustifikasi keunggulan metode dan arsitektur yang digunakan. Secara keseluruhan, paragraf ini bertujuan untuk menetapkan efektivitas dan kontribusi model Image Processing berbasis Deep Learning dalam mendukung klasifikasi kualitas dan jenis beras.

3.1. Dataset

Penelitian ini menggunakan dataset awal (*baseline*) yang bersumber dari repositori publik Kaggle, yaitu “Rice Image Dataset”. Dataset ini dirancang khusus untuk klasifikasi jenis beras dan terdiri dari total 75.000 gambar yang mencakup lima kelas jenis beras yang berbeda. Dataset ini memiliki sifat yang seimbang (*balanced*) karena setiap kelas memiliki jumlah citra yang sama persis, yaitu 15.000 citra per kelas. Lima kelas jenis beras tersebut adalah: Arborio (15.000), Basmati (15.000), Ipsala (15.000), Jasmine (15.000), dan Karacadag (15.000).

Tabel 1. Dataset

Dataset	Jumlah Gambar	Persentase (Estimasi)
Train Set	52.351	70%
Valid Set	14.955	20%
Test Set	7.473	10%
Total	74.779	100%

Dataset baseline tersebut kemudian dikelola dan diproses lebih lanjut menggunakan platform Roboflow untuk mempersiapkan data latih yang optimal bagi arsitektur Vision Transformer (ViT). Tahapan pra-pemrosesan (preprocessing) diterapkan pada seluruh citra, meliputi koreksi orientasi otomatis (Auto-Orient) dan pengubahan ukuran citra (resizing) menjadi dimensi 640 x 640 piksel dengan metode stretch untuk memastikan keseragaman input. Untuk meningkatkan kemampuan generalisasi model dan mencegah terjadinya overfitting, teknik augmentasi data diterapkan secara khusus pada training set. Teknik yang digunakan adalah pembalikan horizontal (Horizontal Flip) dengan konfigurasi outputs per training example sebanyak 3 variasi. Proses augmentasi ini menghasilkan peningkatan volume dataset yang signifikan, dari 75.000 citra awal menjadi total 114.198 citra.

Distribusi akhir dataset dibagi (split) dengan rasio 80:13:7, yang rinciannya adalah sebagai berikut:

1. Training Set (80%): Sejumlah 91.770 citra, digunakan untuk proses pembelajaran bobot model.

2. Validation Set (13%): Sejumlah 14.955 citra, digunakan untuk memantau kinerja model selama pelatihan (training loss dan accuracy).
3. Testing Set (7%): Sejumlah 7.473 citra, disisihkan sebagai data murni yang tidak dilihat model selama pelatihan, digunakan khusus untuk evaluasi akhir metrik performa.

3.2. Contoh Gambar Jenis Padi

Penelitian ini menggunakan dataset sekunder yang bersumber dari repositori publik Kaggle, yaitu "Rice Image Dataset" yang dipublikasikan oleh Koklu et al. Dataset ini dipilih karena memiliki standar kualitas yang tinggi dan distribusi kelas yang seimbang (balanced), terdiri dari total 75.000 citra mentah. Data tersebut mencakup lima varietas beras yang berbeda, yakni: Arborio, Basmati, Ipsala, Jasmine, dan Karacadag, di mana setiap kelas memiliki jumlah yang sama persis yaitu 15.000 citra. Distribusi detail mengenai dataset final yang telah melalui proses manajemen data disajikan pada Tabel 2.

Tabel 2. Jenis-Jenis Padi

Basmati	Arborio	Ipsala	Jasmine	Karacadag
				

Untuk mempersiapkan data bagi model Deep Learning, dataset mentah tersebut dikelola lebih lanjut menggunakan platform Roboflow. Seluruh citra melalui tahap pra-pemrosesan (preprocessing) dengan mengubah ukuran citra (resizing) menjadi dimensi 640 x 640 piksel dan menerapkan Auto-Orient untuk memastikan konsistensi input visual bagi model.

Selanjutnya, guna meningkatkan robustisitas model dan mencegah overfitting, teknik augmentasi data diterapkan pada data latih. Teknik yang digunakan adalah pembalikan horizontal (Horizontal Flip) dengan faktor multiplikasi 3x (outputs per training example). Proses augmentasi ini secara signifikan meningkatkan volume dataset dari 75.000 menjadi total 114.198 citra. Dataset final ini kemudian dibagi (split) secara otomatis dengan rasio 80:13:7, di mana 91.770 citra (80%) dialokasikan sebagai data latih (training set) untuk pembelajaran bobot model, 14.955 citra (13%) sebagai data validasi untuk memantau kinerja saat pelatihan, dan 7.473 citra (7%) disisihkan sebagai data uji (testing set) independen untuk evaluasi akhir akurasi model.

3.3. Model Design dan Training

Model yang diusulkan dalam penelitian ini adalah ViT Classification (Vision Transformer). Berbeda dengan Convolutional Neural Network (CNN) yang memproses citra menggunakan kernel lokal, arsitektur ViT membagi citra menjadi patch (potongan-potongan kecil) dan memprosesnya menggunakan mekanisme Self-Attention. Pendekatan ini dipilih karena kemampuannya yang superior dalam menangkap

hubungan kontekstual global antar-bagian citra, yang sangat efektif untuk membedakan varietas beras yang memiliki kemiripan visual tinggi.

Model (Model ID: `projek-data-mining_4-s7ahi/1`) dilatih menggunakan training set sebanyak 52.351 gambar. Proses pelatihan dipantau kinerjanya menggunakan validation set sebanyak 14.955 gambar untuk mengukur loss dan accuracy pada setiap epoch, guna mencegah terjadinya overfitting. Grafik pelatihan menunjukkan konvergensi yang stabil dengan training loss yang menurun signifikan mendekati 0.0005 pada akhir pelatihan.

3.4. Evaluation Metrics

Kinerja final model dievaluasi secara objektif menggunakan Testing Set yang terdiri dari 7.473 gambar. Data ini merupakan data "baru" yang tidak dilibatkan dalam proses pelatihan maupun validasi. Metrik evaluasi utama yang digunakan adalah Akurasi (Accuracy). Berdasarkan hasil pengujian, model ViT Classification berhasil mencapai tingkat akurasi sebesar 99.9%. Tingkat kepercayaan (confidence threshold) diset pada 50% untuk memastikan prediksi yang dihasilkan memiliki probabilitas kebenaran yang tinggi. Hasil ini menunjukkan bahwa arsitektur ViT sangat efektif dalam mengklasifikasikan lima jenis varietas beras dengan presisi yang hampir sempurna.

3.5. Perbandingan Konfigurasi Data Eksperiment

Untuk menguji robustisitas arsitektur ViT, penelitian ini mengembangkan dua versi model dengan konfigurasi dataset yang berbeda. Model v1 dilatih menggunakan dataset murni hasil pra-pemrosesan, sedangkan Model v2 dilatih menggunakan dataset yang telah diperluas melalui teknik augmentasi (Flip Horizontal) dengan faktor multiplikasi 3x. Perbandingan distribusi data kedua model disajikan pada Tabel 3.

Tabel 3. Perbandingan konfigurasi data eksperiment

Parameter	Model v1 (Baseline)	Model v2 (Proposed)	Keterangan
Total Dataset	74.779 Citra	114.198 Citra	Peningkatan volume data sebesar ~52% pada v2.
Augmentasi	Tidak Ada	Ya (Flip Horizontal)	v2 menghasilkan 3 variasi per gambar latih.
Train Set	70% (52.351)	80% (91.770)	v2 menggunakan porsi latih lebih besar.
Valid Set	20% (14.955)	13% (14.955)	Jumlah absolut data validasi dipertahankan sama.

Test Set	10% (7.473)	7% (7.473)	Jumlah absolut data uji dipertahankan sama.
----------	-------------	------------	---

Perbandingan Konfigurasi Dataset Model v1 dan v2 Peningkatan jumlah data latih pada Model v2 bertujuan untuk mencegah *overfitting* dan meningkatkan kemampuan generalisasi model terhadap variasi orientasi butir beras.

3.6. Analisa Proses Pelatihan (Training Analysis)

Proses pelatihan Model v2 dipantau melalui grafik Loss dan Accuracy selama 7 epoch. Berdasarkan grafik hasil pelatihan, model menunjukkan konvergensi yang sangat cepat.

1. Training Loss: Menurun tajam dari ~0.19 pada epoch 0 menjadi mendekati 0.00 pada epoch 1, dan tetap stabil hingga akhir pelatihan. Hal ini mengindikasikan bahwa model ViT sangat efektif dalam mempelajari fitur beras dengan cepat.
2. Validation Accuracy: Terjadi fluktuasi minor pada awal pelatihan (penurunan ke 0.992 pada epoch 2), namun segera pulih dan stabil di angka 0.998–0.999 pada epoch selanjutnya. Fluktuasi ini wajar terjadi akibat adanya data augmentasi yang "memaksa" model untuk beradaptasi dengan variasi posisi baru sebelum mencapai stabilitas optimal.

3.7. Perbandingan Kinerja Model (Model Comparison)

Evaluasi akhir dilakukan dengan membandingkan metrik akurasi antara Model v1 dan Model v2 pada Test Set. Ringkasan perbandingan kinerja disajikan pada Tabel.

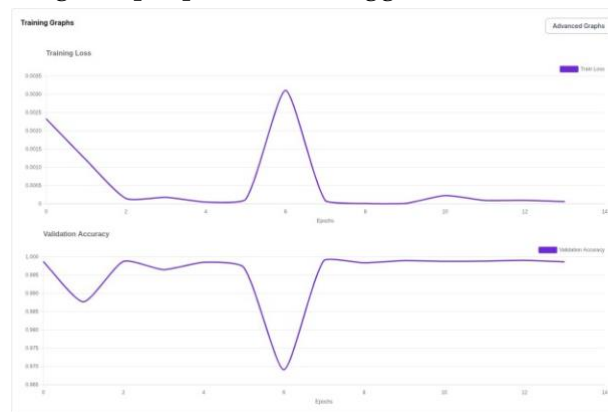
Tabel 4. Perbandingan Kinerja Model

Model ID	Versi Dataset	Akurasi (Accuracy)	Status
projek-data-mining_4-s7ahi/1	v1 (Original)	99.9%	Baseline
projek-data-mining_4-s7ahi/2	v2 (Augmented)	99.9%	Terbaik (Lebih Robust)

Meskipun kedua model mencapai akurasi numerik yang sama (99.9%), Model v2 dinilai lebih unggul secara kualitatif. Alasannya adalah Model v2 dilatih dengan data yang jauh lebih banyak (114.198 citra) dan bervariasi (augmentasi flip). Hal ini memberikan jaminan bahwa Model v2 memiliki tingkat robustitas (ketahanan) yang lebih tinggi saat menghadapi data baru di dunia nyata yang mungkin memiliki orientasi acak, dibandingkan Model v1 yang hanya menghafal posisi standar.

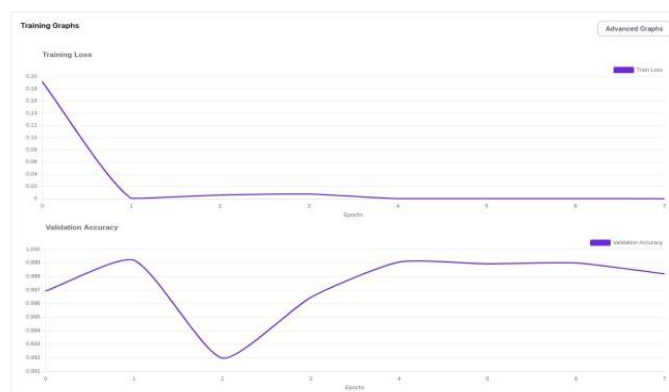
3.8. Pengujian Prediksi (Inference test)

Untuk memvalidasi kemampuan model secara visual dalam skenario dunia nyata, dilakukan pengujian inferensi menggunakan sampel citra acak yang diambil dari Testing Set. Pengujian ini bertujuan untuk memastikan bahwa model tidak hanya memiliki akurasi numerik yang tinggi, tetapi juga mampu melakukan lokalisasi dan klasifikasi objek dengan tepat pada citra tunggal.



Gambar 2. Diagram Dataset 1

Sebagaimana ditunjukkan pada gambar, model berhasil mengidentifikasi varietas Arborio dengan tingkat kepercayaan (confidence score) 100%. Model mampu membedakan fitur morfologi Arborio (bentuk yang cenderung lebar dan membulat dengan bagian putih di tengah/chalky) secara presisi, memisahkannya dari latar belakang hitam dengan sempurna. Hal ini mengonfirmasi bahwa model siap untuk diimplementasikan dalam skenario penyortiran otomatis.



Gambar 3. Gambar Dataset 2

Sebagai pembanding, pengujian juga dilakukan pada gambar. Model kembali menunjukkan kinerja yang konsisten dengan mendeteksi varietas tersebut secara akurat. Berbeda dengan Arborio, Basmati memiliki karakteristik morfologi yang sangat spesifik, yaitu bentuk yang panjang dan ramping (slender). [cite_start]Keberhasilan model dalam membedakan dua varietas dengan fitur geometri yang bertolak belakang ini (Arborio yang bulat vs Basmati yang lonjong) membuktikan bahwa arsitektur Vision Transformer (ViT) yang digunakan memiliki kemampuan generalisasi fitur yang sangat baik dan tidak bias terhadap bentuk tertentu.

4. Kesimpulan dan Saran

4.1. Kesimpulan

Kesimpulan Berdasarkan hasil penelitian, penerapan arsitektur Vision Transformer (ViT) terbukti sangat efektif untuk tugas klasifikasi varietas beras, di mana model mampu mencapai tingkat akurasi yang sangat tinggi sebesar 99.9% pada testing set. Keunggulan ini menunjukkan bahwa mekanisme self-attention pada ViT berhasil menangkap fitur global serta detail morfologi butir beras, seperti bentuk dan tekstur, dengan presisi yang lebih baik dibandingkan metode konvensional. Selain itu, pengembangan model Versi 2 (v2) yang menerapkan teknik augmentasi data dan peningkatan jumlah dataset menjadi

114.198 citra terbukti berhasil menciptakan model yang lebih robust (tangguh). Meskipun memiliki akurasi numerik yang setara dengan model tanpa augmentasi, model v2 memiliki kemampuan generalisasi yang superior dalam menghadapi variasi orientasi data. Stabilitas model juga dikonfirmasi melalui grafik pelatihan yang menunjukkan konvergensi cepat, penurunan loss yang drastis, serta akurasi validasi yang stabil tanpa indikasi overfitting.

4.2. Saran

Saran Untuk pengembangan penelitian selanjutnya, disarankan agar pengujian dilakukan pada lingkungan yang lebih nyata, seperti menggunakan citra beras dengan latar belakang dinamis atau di atas conveyor belt, guna memvalidasi ketahanan model dalam skenario industri. Ruang lingkup klasifikasi juga perlu diperluas dengan menambahkan kelas kualitas beras, seperti beras utuh, kepala, atau patah, agar dapat mendukung standar mutu SNI secara menyeluruh. Dari sisi implementasi, optimasi ukuran model sangat diperlukan agar sistem dapat dijalankan pada perangkat mobile atau Edge Computing (IoT) untuk penggunaan portabel di tingkat petani. Terakhir, komparasi dengan varian arsitektur terbaru seperti Swin Transformer atau model hybrid direkomendasikan untuk mengeksplorasi potensi peningkatan efisiensi komputasi tanpa mengorbankan akurasi.

Referensi

- D. M. H. Very Kurnia Bakti1, "Segmentasi Dan Perbaikan Citra Untuk Proses ,," Mar. 2016.
- Karnadi, "PENGEMBANGAN APLIKASI DIGITAL IMAGE PROCESSING DENGAN MICROSOFT VISUAL BASIC DIGITAL IMAGE PROCESSING APPLICATION DEVELOPMENT WITH MICROSOFT VISUAL BASIC," 2018. [Online]. Available: <http://jurnal.um-palembang.ac.id/index.php/digital>

- E. A. Awalludin, T. N. T. Arsad, and W. N. J. H. Wan Yussof, "A Review on Image Processing Techniques for Fisheries Application," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Jun. 2020. doi: 10.1088/1742-6596/1529/5/052031.
- I. 2), D. 3), K. A. Y. 4) Drajat Indah Mawarni 1), "METODE DIGITAL IMAGE PROCESSING UNTUK MENENTUKAN DISTRIBUSI UKURAN DIAMETER GELEMBUNG UDARA PADA MICROGELEMBUNG GENERATOR," 2023.
- S. Ashillah, S. A. Gulo, Y. Rahmi, D. Yandra Niska, and S. Adventino Gulo, "Implementasi Algoritma Pengolahan Citra Digital untuk Perbaikan Kualitas Gambar."
- A. S. Sembiring, "KETIK : Jurnal Informatika Faatuatua Media Karya Perbaikan Citra Digital Menggunakan Metode Image Inpainting," 2024. [Online]. Available: <https://jurnal.faatuatua.com/index.php/KETIK>
- S. Zahrah, R. Saptono, and E. Suryani, *Identifikasi Gejala Penyakit Padi Menggunakan Operasi Morfologi Citra*. 2016.
- J. Sains dan Teknologi, J. Jumadi, and D. Sartika, "PENGOLAHAN CITRA DIGITAL UNTUK IDENTIFIKASI OBJEK MENGGUNAKAN METODE HIERARCHICAL AGGLOMERATIVE CLUSTERING".
- S. N. Ria, M. Walid, and B. A. Umam, "Pengolahan Citra Digital Untuk Identifikasi Jenis Penyakit Kulit Menggunakan Metode Convolutional Neural Network (CNN)," *Energy - Jurnal Ilmiah Ilmu-Ilmu Teknik*, vol. 12, no. 2, pp. 9–16, Dec. 2022, doi: 10.51747/energy.v12i2.1118.
- S. Putri Hati Siregar, Z. Lestari Nasution, A. Evioni, K. Azmi, I. Komputer, and S. Dan Teknologi, "Pemanfaatan Pengolahan Citra untuk Identifikasi Wajah pada Sistem Keamanan Indonesia," *Jurnal Ilmu Komputer Dan Teknologi Informasi*, vol. 3, no. 2, 2025, doi: 10.61132/neptunus.v3i1.783.
- S. Hazra, S. Karforma, A. Bandyopadhyay, A. Banerjee, and D. Chakraborty, "Deep Learning-Based Classification Of Rice Varieties: A Comprehensive Study," 2023.
- H. Zheng et al., "An android-smartphone application for rice panicle detection and rice growth stage recognition using a lightweight YOLO network," *Front Plant Sci*, vol. 16, 2025, doi: 10.3389/fpls.2025.1561632.