

Klasifikasi Real-Time Tingkat Kematangan Buah Carica Menggunakan Convolutional Neural Network dan Grad-CAM

Informasi Naskah

Wahyu Suhardiyono¹, Nahar Mardiyantoro², Saifu Rohman³

Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Sains Al-Qur'an, Wonosobo, Indonesia

wahyusuhardiyono@mhs.unsiq.ac.id¹,
nahar.mardiyantoro@unsiq.ac.id²,
saifu.rohman@unsiq.ac.id³

Abstrak

Buah Carica (*Carica pubescens*) merupakan komoditas unggulan dataran tinggi Dieng, Wonosobo, yang kualitas produk olahannya sangat ditentukan oleh tingkat kematangan saat panen. Penentuan kematangan secara manual melalui pengamatan warna kulit dan perabaan tekstur bersifat subjektif, tidak konsisten, dan rentan kelelahan. Penelitian ini bertujuan membangun sistem klasifikasi tingkat kematangan buah Carica ke dalam tiga kelas (Mentah, Mengkal, Matang) secara real-time menggunakan Convolutional Neural Network (CNN) berarsitektur MobileNetV2 melalui pendekatan transfer learning dan fine-tuning dua fase, dilengkapi visualisasi Explainable Artificial Intelligence (XAI) berbasis Gradient-weighted Class Activation Mapping (Grad-CAM) untuk mengatasi sifat black box model. Dataset berjumlah 1.800 citra hasil akuisisi mandiri (600 citra per kelas), dibagi menjadi 1.440 citra latih dan 360 citra validasi. Model diintegrasikan ke dalam aplikasi web berbasis Flask untuk klasifikasi real-time melalui kamera. Hasil evaluasi menunjukkan model TensorFlow Lite mencapai akurasi 90,56% pada data validasi, sementara pengujian dinamis real-time terhadap 90 sampel buah menghasilkan akurasi rata-rata 95,5% dengan kecepatan pemrosesan stabil 60 FPS, melampaui ambang kebutuhan real-time sebesar 24 FPS. Visualisasi Grad-CAM membuktikan model secara konsisten memfokuskan perhatian pada gradasi warna kulit buah sebagai fitur penentu kematangan, sehingga sistem tidak hanya akurat tetapi juga transparan dan dapat dipertanggungjawabkan secara visual kepada pengguna, khususnya petani dan pelaku UMKM pengolahan Carica.

Kata Kunci:

Buah Carica, Convolutional Neural Network, MobilenetV2, Grad-CAM, Explainable Artificial Intelligence.

Abstract

Carica fruit (*Carica pubescens*) is a signature commodity of the Dieng highlands, Wonosobo, whose processed-product quality is largely determined by ripeness at harvest. Manual ripeness assessment through visual and tactile inspection is subjective, inconsistent, and prone to fatigue. This study aims to build a real-time ripeness classification system for Carica fruit into three classes (Unripe, Semi-ripe, Ripe) using a Convolutional Neural Network (CNN) based on the MobileNetV2 architecture through a two-phase transfer learning and fine-tuning approach, equipped with Explainable Artificial Intelligence (XAI) visualization based on Gradient-weighted Class Activation Mapping (Grad-CAM) to address the model's black-box nature. The dataset consists of 1,800 self-acquired images (600 per class), divided into 1,440 training images and 360 validation images. The model is integrated into a Flask-based web application for real-time classification via camera. Evaluation results show the TensorFlow Lite model achieved 90.56% accuracy on the validation set, while dynamic real-time testing on 90 fruit samples achieved an average accuracy of 95.5% with a stable processing speed of 60 FPS, exceeding the 24 FPS real-time requirement threshold. Grad-CAM visualization confirms that the model consistently focuses on the fruit skin's color gradation as the dominant ripeness-determining feature, making the system both accurate and visually accountable to users, particularly farmers and Carica processing small businesses.

Key Word:

Carica Fruit, Convolutional Neural Network, MobileNetV2, Grad-CAM, Explainable Artificial Intelligence.

Information Production:

© 2026 by the authors of this article. This work is licensed under an Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)



PENDAHULUAN

Buah Carica (*Carica pubescens*) merupakan komoditas hortikultura bernilai ekonomis tinggi dan menjadi ikon produk oleh-oleh khas dataran tinggi Dieng, Wonosobo. Dalam industri pengolahan makanan seperti manisan, sirup, maupun selai Carica, tingkat kematangan buah merupakan parameter utama yang menentukan kualitas produk akhir. Buah yang dipetik dalam kondisi terlalu mentah menghasilkan tekstur keras

dan rasa kurang manis, sedangkan buah yang terlalu matang (overripe) mempercepat kerusakan atau pembusukan produk, sehingga proses sortasi dan klasifikasi tingkat kematangan menjadi tahap krusial dalam menjaga standar mutu (Minarno, 2015).

Selama ini, penentuan tingkat kematangan buah Carica pada industri skala kecil hingga menengah masih dilakukan secara manual oleh tenaga manusia, yaitu dengan mengamati perubahan warna kulit buah dan meraba tingkat kekerasannya. Cara ini sangat bergantung pada penglihatan dan kepekaan fisik pekerja yang bersifat subjektif, sehingga standar kematangan antar pekerja dapat bervariasi. Proses ini juga rentan terhadap kelelahan, membutuhkan waktu lebih lama, dan sering menyebabkan ketidakkonsistenan hasil sortasi terutama saat volume produksi meningkat. Oleh karena itu, diperlukan sebuah sistem otomatisasi yang dapat mengklasifikasikan tingkat kematangan buah secara cepat, tepat, dan objektif.

Perkembangan teknologi computer vision dan kecerdasan buatan menawarkan solusi efektif atas permasalahan tersebut. Convolutional Neural Network (CNN) terbukti sangat baik dalam pengolahan citra digital karena mampu mengekstrak fitur spasial dari sebuah citra dan melakukan klasifikasi secara otomatis dengan akurasi tinggi tanpa memerlukan ekstraksi fitur manual (Purwono et al., 2023; Li et al., 2022). Di antara berbagai arsitektur CNN, MobileNetV2 dipilih secara luas untuk aplikasi real-time karena efisiensi parameternya yang jauh lebih kecil dibandingkan arsitektur konvensional seperti VGG16, tanpa mengorbankan akurasi klasifikasi (Sandler et al., 2018).

Meskipun CNN memiliki akurasi tinggi, model deep learning sering dikritik sebagai black box karena pengguna tidak dapat mengetahui secara pasti fitur visual apa yang menjadi dasar pengambilan keputusan model (Adadi & Berrada, 2018). Untuk mengatasi hal ini, diperlukan teknik Explainable Artificial Intelligence (XAI), salah satunya Gradient-weighted Class Activation Mapping (Grad-CAM). Grad-CAM bekerja dengan memetakan area piksel yang paling berpengaruh terhadap keputusan klasifikasi model ke dalam bentuk heatmap, sehingga sistem tidak hanya memberikan hasil klasifikasi tetapi juga dapat memvalidasi apakah fokus model telah sesuai dengan karakteristik kematangan buah yang sebenarnya (Selvaraju et al., 2017).

Penelitian terdahulu mengenai klasifikasi kematangan buah berbasis CNN telah banyak dilakukan. Hidayat dan Sari (2021) mengembangkan aplikasi deteksi kematangan buah berbasis Android yang dapat digunakan langsung di lapangan, namun sering mengalami misclassification akibat pantulan cahaya matahari. Kumar dan Singh (2022) menunjukkan CNN mampu mendeteksi kematangan berbagai buah tropis dengan akurasi di atas 95%, tetapi belum dilengkapi metode visualisasi untuk menjelaskan alasan klasifikasi. Saputra et al. (2022) menerapkan MobileNetV2 untuk identifikasi kematangan pepaya dengan akurasi 92%, namun performanya menurun pada latar belakang yang kompleks. Nugroho (2022) membandingkan VGG16 dan ResNet50 untuk klasifikasi buah dan menemukan model cenderung overfitting akibat keterbatasan dataset. Wibowo dan Arifianto (2023) menerapkan Grad-CAM pada deteksi penyakit tanaman dan menunjukkan efektivitasnya dalam menyoroti area terinfeksi, tetapi komputasi Grad-CAM menambah beban proses sehingga sulit berjalan stabil secara real-time. Zhang et al. (2023) berhasil memproses 30 FPS untuk grading buah pada jalur conveyor, namun membutuhkan spesifikasi GPU tinggi untuk menjaga kestabilan frame rate.

Berbagai penelitian juga menunjukkan bahwa arsitektur CNN ringan seperti keluarga MobileNet mampu menghasilkan performa klasifikasi yang baik dengan kebutuhan komputasi yang jauh lebih hemat dibandingkan arsitektur CNN yang lebih dalam, sehingga sangat sesuai diterapkan pada perangkat dengan sumber daya terbatas maupun aplikasi berbasis web browser yang mengandalkan inferensi sisi server (Kosasih & Fahrurrozi, 2023). Selain efisiensi komputasi, aspek kepercayaan pengguna terhadap sistem kecerdasan buatan juga menjadi pertimbangan penting, terutama pada sektor pertanian dan pengolahan hasil panen yang melibatkan keputusan ekonomis langsung bagi petani dan pelaku UMKM. Sistem yang hanya memberikan label prediksi tanpa penjelasan visual berisiko menimbulkan keraguan pengguna terhadap validitas keputusan model, sehingga integrasi XAI menjadi semakin relevan untuk diterapkan pada domain aplikasi praktis semacam ini.

Berdasarkan tinjauan tersebut, penelitian yang secara khusus mengklasifikasikan tingkat kematangan buah Carica menggunakan pendekatan computer vision masih sangat terbatas, dan penelitian pada komoditas serupa umumnya belum memadukan klasifikasi real-time dengan validasi visual XAI sekaligus tanpa memerlukan perangkat keras GPU khusus. Kebaruan (novelty) penelitian ini terletak pada integrasi tiga komponen ke dalam satu sistem berbasis web browser: (1) model CNN ringan berbasis MobileNetV2 yang dioptimasi untuk komoditas Carica, (2) visualisasi Grad-CAM secara real-time tanpa GPU khusus, dan (3) arsitektur inferensi hybrid server-side yang menjaga kestabilan frame rate di atas ambang batas real-time. Berdasarkan uraian tersebut, penelitian ini bertujuan merancang dan membangun sistem yang dapat mengklasifikasikan tingkat kematangan buah Carica secara real-time menggunakan CNN dan dilengkapi visualisasi Grad-CAM sebagai bentuk transparansi pengambilan keputusan model.

METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif eksperimental untuk mengklasifikasikan tingkat kematangan buah Carica (kelas Mentah, Mengkal, dan Matang) berdasarkan citra digital, dengan tahapan meliputi pengumpulan dataset, praproses dan augmentasi data, pembangunan serta pelatihan model, evaluasi kinerja, integrasi model ke dalam sistem, dan pengujian aplikasi.

Pengumpulan dan Praproses Data

Seluruh dataset diperoleh melalui akuisisi mandiri tanpa mengandalkan dataset dari platform daring, dengan cara membeli buah Carica secara langsung di tempat produksi Excotic Carica Wonosobo dan mengabadikan citra menggunakan kamera ponsel serta webcam dari berbagai sudut. Total dataset berjumlah 1.800 citra yang terbagi merata ke dalam tiga kelas (600 citra per kelas) untuk menghindari ketimpangan data (data imbalance). Kelas Mentah dicirikan warna kulit hijau tua dominan dan tekstur keras; kelas Mengkal ditandai gradasi hijau ke kuning; kelas Matang dicirikan warna kuning cerah merata. Dataset dibagi dengan proporsi 80% data latih (1.440 citra) dan 20% data validasi (360 citra) menggunakan parameter `validation_split` pada `ImageDataGenerator`. Setiap citra diubah ukurannya menjadi 224×224 piksel dan dinormalisasi menggunakan fungsi `preprocess_input` bawaan `MobileNetV2` ke rentang $[-1, 1]$.

Augmentasi data diterapkan untuk memperkaya variasi sampel dan menekan risiko overfitting, meliputi rotasi hingga 40° , pergeseran horizontal/vertikal hingga 30%, shear hingga 30%, zoom hingga 30%, variasi kecerahan pada rentang 70%-130%, pergeseran kanal warna hingga 20 satuan intensitas, serta flip horizontal dan vertikal.

Arsitektur Model dan Pelatihan

Model dibangun menggunakan arsitektur `MobileNetV2` dengan bobot awal `ImageNet` (`include_top=False`) sebagai ekstraktor fitur dasar, ditambah lapisan `Global Average Pooling 2D`, `Dense 256` neuron dengan aktivasi `ReLU`, `Dropout 0,5`, dan lapisan keluaran `Dense 3` neuron dengan aktivasi `softmax`. Pelatihan dilakukan melalui strategi dua fase: Fase 1 (transfer learning) selama 6 epoch dengan bobot dasar `MobileNetV2` dibekukan dan learning rate 1×10^{-4} ; Fase 2 (fine-tuning) selama 5 epoch tambahan dengan 20 lapisan konvolusi teratas dibuka kembali dan learning rate diperkecil menjadi 1×10^{-5} . Tiga mekanisme callback diterapkan selama pelatihan: `EarlyStopping` (patience 4 pada Fase 1, patience 10 pada Fase 2, memantau `val_loss`), `ModelCheckpoint` (menyimpan bobot terbaik berdasarkan `val_accuracy` tertinggi), dan `ReduceLROnPlateau` (menurunkan learning rate sebesar faktor 0,2 apabila `val_loss` stagnan selama 3 epoch, dengan batas minimum 1×10^{-6}).

Integrasi Grad-CAM dan Deployment

Grad-CAM diimplementasikan dengan menghitung gradien skor kelas target terhadap peta fitur pada lapisan konvolusi terakhir, menghasilkan bobot kepentingan neuron yang dikombinasikan secara linear dan difilter fungsi `ReLU` untuk menghasilkan heatmap area fokus model. Model final dikonversi ke format `TensorFlow Lite` dengan kuantisasi `float16` untuk mempercepat inferensi, sementara model Keras (.h5) tetap digunakan secara asinkron khusus untuk perhitungan Grad-CAM. Sistem dideploy sebagai aplikasi web berbasis Flask dengan arsitektur hybrid server-side inference, dilengkapi filter heuristik `OpenCV` berbasis analisis warna HSV dan kerapatan tepi untuk menstabilkan deteksi objek. Untuk mengurangi fluktuasi prediksi akibat variasi pencahayaan atau pergerakan buah, diterapkan mekanisme `rolling majority voting` terhadap 15 frame berurutan serta `stable zone filtering` yang membatasi inferensi hanya pada area tengah layar.

Evaluasi

Kinerja model dievaluasi menggunakan metrik `accuracy`, `precision`, `recall`, dan `F1-score` yang diturunkan dari confusion matrix pada 360 citra data validasi. Pengujian tambahan dilakukan secara dinamis melalui aplikasi web menggunakan 90 sampel buah Carica baru (45 melalui webcam laptop dan 45 melalui kamera smartphone) untuk mengukur akurasi sistem serta kecepatan pemrosesan real-time dalam satuan `Frame Per Second (FPS)`, dengan ambang batas kebutuhan real-time ditetapkan pada ≥ 24 FPS.

HASIL DAN PEMBAHASAN

Hasil Eksperimen Konfigurasi Model (Ablation Study)

Untuk menentukan konfigurasi arsitektur optimal, dilakukan eksperimen bertahap dengan melatih ulang model CNN sebanyak tiga kali pada dataset identik (1.440 citra latih dan 360 citra validasi) dengan kombinasi hyperparameter berbeda. Hasil pengujian ketiga skenario disajikan pada Tabel 1.

Tabel 1. Rekapitulasi Hasil Pengujian Skenario Konfigurasi Model

Skenario	Konfigurasi Parameter	Augmentasi	Akurasi Latih	Akurasi Validasi	Keterangan
1	Dense 128, LR 0,0001 statis, Dropout 0,3	Tidak Ada	98,15%	72,40%	Overfitting parah
2	Dense 256, LR 0,0001 statis, Dropout 0,5	Ringan	94,50%	85,30%	Overfitting berkurang
3	Dense 256, LR Dinamis (1e-4→1e-5), Dropout 0,5	Agresif	96,15%	97,92%	Terbaik, dipilih final

Skenario 1 tanpa augmentasi data mengalami overfitting parah dengan selisih akurasi latih-validasi mencapai 25,75%, karena model menghafal detail data latih secara spesifik. Skenario 2 dengan peningkatan Dropout menjadi 0,5 dan augmentasi ringan berhasil menekan overfitting, namun konvergensi loss belum stabil akibat learning rate statis. Skenario 3 dengan kombinasi Dense 256, learning rate dinamis dua fase, dan augmentasi agresif menghasilkan generalisasi model paling kuat, dengan akurasi validasi puncak 97,92% pada epoch ke-6, dan skenario ini dipilih sebagai konfigurasi final.

Kurva akurasi pelatihan pada skenario terbaik menunjukkan akselerasi pemahaman model yang konstan sejak fase awal iterasi hingga mencapai puncaknya pada epoch ke-6, dengan akurasi validasi 97,92% dan akurasi data latih 96,15%. Namun, setelah epoch ke-6 hingga epoch ke-11 (akhir Fase 2), terdapat tren penurunan akurasi validasi yang stabil di kisaran 88%-92%, sementara akurasi data latih terus meningkat hingga 97,06%. Pelebaran celah (gap) antara akurasi latih dan validasi setelah epoch ke-6 ini mengindikasikan dimulainya gejala overfitting ringan pada Fase 2 (fine-tuning), sebagaimana lazim terjadi ketika lapisan konvolusi bagian atas mulai dibuka kembali. Untuk mengatasi hal tersebut, mekanisme ModelCheckpoint dengan parameter `save_best_only=True` secara otomatis mengunci dan hanya menyimpan bobot model pada epoch dengan performa validasi terbaik, sehingga bobot yang digunakan pada tahap deployment bukan bobot epoch terakhir, melainkan bobot epoch ke-6 yang memiliki generalisasi paling optimal. Ketidadaan lonjakan validation loss yang ekstrem pada akhir iterasi turut membuktikan bahwa kombinasi Dropout 0,5, learning rate dinamis, dan EarlyStopping berhasil menekan gejala overfitting dalam batas yang aman.

Evaluasi Akhir Model

Terdapat selisih antara akurasi validasi puncak saat pelatihan (97,92%) dengan akurasi evaluasi akhir model TensorFlow Lite (90,56%). Selisih ini bersifat metodologis, bukan disebabkan augmentasi: nilai 97,92% merupakan akurasi puncak pada satu epoch terbaik yang dihitung terhadap data validasi yang masih melalui alur augmentasi yang sama dengan data latih, sedangkan nilai 90,56% diperoleh dari evaluasi tunggal menggunakan model yang telah dikonversi ke format TensorFlow Lite dengan kuantisasi float16 pada citra validasi bersih tanpa augmentasi, sehingga nilai ini lebih representatif terhadap performa model yang sesungguhnya di-deploy ke sistem scanner. Rincian hasil klasifikasi terhadap 360 citra uji validasi (120 sampel per kelas) disajikan pada Tabel 2, sedangkan metrik precision, recall, dan F1-score per kelas disajikan pada Tabel 3.

Tabel 2. Confusion Matrix Evaluasi Model TensorFlow Lite (360 Citra Validasi)

Kelas Aktual	Prediksi Matang	Prediksi Mengkal	Prediksi Mentah	Total
Matang	111	9	0	120
Mengkal	1	117	2	120
Mentah	0	22	98	120

Tabel 3. Precision, Recall, dan F1-Score per Kelas

Kelas	Precision	Recall	F1-Score
Matang	99,11%	92,50%	95,69%
Mengkal	79,05%	97,50%	87,31%
Mentah	98,00%	81,67%	89,09%

Berdasarkan Tabel 2, model berhasil mengklasifikasikan 326 dari 360 citra dengan benar, menghasilkan akurasi global sebesar 90,56%. Kelas Matang menunjukkan precision tertinggi (99,11%) karena karakteristik warna kuning cerahnya sangat khas dan jarang tertukar dengan kelas lain. Kesalahan klasifikasi paling banyak terjadi antara kelas Mentah dan Mengkal, dengan 22 dari 120 citra Mentah salah diprediksi sebagai Mengkal, sehingga recall kelas Mentah menjadi paling rendah (81,67%) dan precision kelas Mengkal ikut tertekan (79,05%). Hal ini disebabkan kemiripan gradasi warna kuning-hijau tipis pada kulit buah yang menjadi transisi alami antara kedua kelas tersebut, sebuah tantangan yang juga dilaporkan pada penelitian klasifikasi buah berbasis warna lainnya (Nugroho, 2022).

Secara matematis, akurasi global diperoleh dari perbandingan jumlah prediksi benar terhadap total sampel uji, yaitu $(111+117+98)/360 = 326/360 = 0,9056$ atau 90,56%. Nilai precision, recall, dan F1-score pada Tabel 3 diturunkan langsung dari confusion matrix pada Tabel 2 menggunakan rumus baku $\text{Precision} = \text{TP}/(\text{TP}+\text{FP})$, $\text{Recall} = \text{TP}/(\text{TP}+\text{FN})$, dan $\text{F1} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$. Kelas Mengkal memperoleh recall tertinggi (97,50%) karena karakter warnanya yang berada di antara dua kelas lain membuat model cenderung mengklasifikasikan sampel ambigu ke kelas ini, namun hal yang sama menurunkan precision-nya karena turut menampung kesalahan prediksi dari kelas Mentah. Secara keseluruhan, nilai F1-score yang berada pada rentang 87%-96% untuk ketiga kelas menunjukkan bahwa model memiliki performa yang seimbang dan tidak bias secara signifikan terhadap kelas tertentu, meskipun tantangan pada batas transisi warna Mentah-Mengkal tetap menjadi area yang perlu diperkuat melalui penambahan variasi data pada penelitian selanjutnya.

Visualisasi Grad-CAM

Visualisasi Grad-CAM menunjukkan bahwa model secara konsisten memfokuskan aktivasi tertinggi pada area permukaan kulit buah yang mengandung gradasi warna hijau hingga kuning-oranye, sesuai dengan karakteristik biologis yang sebenarnya menjadi indikator kematangan Carica. Pada skenario deteksi objek asing seperti tangan atau logam, Grad-CAM membuktikan model mengabaikan objek tersebut karena tidak terdapat aktivasi positif pada area yang tidak sesuai dengan karakteristik warna Carica. Temuan ini menunjukkan bahwa model tidak mengambil keputusan berdasarkan artefak latar belakang atau fitur yang tidak relevan secara biologis, melainkan berdasarkan fitur kulit buah yang sesungguhnya, sehingga mengatasi kekhawatiran black box yang umum ditemukan pada model deep learning (Adadi & Berrada, 2018; Selvaraju et al., 2017).

Secara lebih spesifik, pada sampel kelas Matang, area aktivasi Grad-CAM terpusat pada permukaan kulit buah yang mengandung pigmen kuning hingga oranye pekat, sedangkan pada sampel kelas Mentah, aktivasi terfokus pada area berpigmen hijau tua. Pada sampel kelas Mengkal yang secara visual berada pada zona transisi warna, pola aktivasi Grad-CAM cenderung tersebar pada beberapa titik gradasi warna sekaligus, yang secara visual menjelaskan mengapa kelas ini paling rentan mengalami kesalahan klasifikasi sebagaimana ditunjukkan pada Tabel 2 dan Tabel 3. Konsistensi pola aktivasi ini memberikan bukti kuat bahwa fitur yang dipelajari model bukan artefak acak, melainkan representasi warna kulit buah yang secara agronomis memang menjadi indikator utama kematangan Carica, sehingga hasil klasifikasi model dapat dipertanggungjawabkan secara ilmiah kepada pengguna akhir.

Pengujian Real-Time pada Aplikasi Web

Pengujian dinamis dilakukan terhadap 90 sampel buah Carica baru menggunakan kamera laptop (45 sampel) dan kamera smartphone (45 sampel) pada kondisi pencahayaan mini studio yang stabil. Hasil pengujian menunjukkan akurasi 95,2% melalui webcam laptop dan 95,8% melalui kamera smartphone, dengan rata-rata keseluruhan 95,5%, serta kecepatan pemrosesan stabil pada 60 FPS untuk kedua jenis perangkat—jauh melampaui ambang kebutuhan real-time sebesar 24 FPS. Menariknya, akurasi pengujian dinamis ini lebih tinggi dibandingkan evaluasi statis pada model TFLite (90,56%). Peningkatan ini disebabkan oleh tiga faktor: pertama, mekanisme rolling majority voting yang mengambil suara terbanyak dari 15 frame berurutan sehingga meredam fluktuasi tebakan sesaat akibat blur atau pantulan cahaya; kedua, stable zone filtering yang mengabaikan citra buah pada tepi frame yang rentan distorsi lensa; ketiga, lingkungan pengujian mini studio yang lebih terkontrol dibandingkan variasi citra pada dataset statis.

Dari sisi pengalaman pengguna, kedua jenis perangkat pengujian menunjukkan performa yang setara, mengindikasikan bahwa sistem cukup fleksibel untuk diakses baik melalui perangkat laptop di lokasi pengolahan tetap maupun perangkat smartphone yang lebih portabel untuk digunakan langsung di lapangan atau titik pengumpulan hasil panen. Konsistensi kecepatan 60 FPS pada kedua platform membuktikan bahwa strategi optimasi berupa kuantisasi model ke format TensorFlow Lite float16 dan pemrosesan Grad-CAM secara asinkron pada thread terpisah berhasil memangkas beban komputasi CPU secara signifikan, sehingga aplikasi tetap responsif tanpa mengalami lag visual meskipun dijalankan tanpa akselerasi GPU khusus.

Perbandingan dengan Penelitian Terdahulu

Dibandingkan dengan penelitian sejenis, sistem yang dibangun menawarkan tiga keunggulan utama. Pertama, dari segi efisiensi model, MobileNetV2 yang digunakan hanya berukuran sekitar 5 MB dalam format TFLite, jauh lebih ringan dibandingkan arsitektur VGG16 yang memiliki sekitar 138 juta parameter (Sandler et al., 2018), sehingga memungkinkan inferensi di bawah 30 milidetik per frame tanpa GPU khusus. Kedua, dari segi transparansi, visualisasi Grad-CAM diterapkan secara real-time langsung pada antarmuka web

browser, mengatasi kelemahan penelitian Wibowo dan Arifianto (2023) yang melaporkan beban komputasi Grad-CAM menyulitkan operasi real-time yang stabil. Ketiga, dari segi kelayakan implementasi, sistem mencapai kecepatan 60 FPS yang melampaui standar industri 24 FPS tanpa memerlukan spesifikasi perangkat keras tinggi, berbeda dengan sistem grading buah pada penelitian Zhang et al. (2023) yang membutuhkan GPU kelas atas untuk menjaga kestabilan frame rate. Sistem ini juga menjadi kontribusi baru karena penelitian klasifikasi kematangan pada komoditas Carica secara spesifik menggunakan pendekatan CNN dan XAI belum banyak dilakukan sebelumnya.

Implikasi Praktis dan Arsitektur Deployment

Dari sisi implementasi, sistem dibangun dengan arsitektur hybrid server-side inference berbasis Flask, di mana model dimuat dalam dua format secara paralel: TensorFlow Lite untuk inferensi klasifikasi berkecepatan tinggi, dan model Keras (.h5) yang dijalankan secara asinkron khusus untuk perhitungan Grad-CAM, mengingat format TFLite tidak menyimpan grafik gradien yang diperlukan untuk visualisasi tersebut. Agar aplikasi web lokal dapat diakses secara aman melalui kamera smartphone dari luar jaringan lokal, sistem diintegrasikan dengan layanan secure tunneling yang menyediakan URL berbasis protokol HTTPS, sebuah syarat mutlak yang diwajibkan oleh browser seluler modern agar application programming interface (API) akses kamera perangkat dapat diaktifkan secara aman. Pendekatan ini memungkinkan pengguna mengakses sistem klasifikasi hanya melalui tautan browser tanpa perlu memasang aplikasi tambahan pada perangkat masing-masing.

Dari perspektif kebermanfaatan, arsitektur berbasis web yang ringan ini relevan diterapkan pada konteks UMKM pengolahan Carica di kawasan Dieng yang umumnya memiliki keterbatasan akses terhadap infrastruktur komputasi khusus. Dengan hanya mengandalkan satu unit komputer atau laptop sebagai server lokal, beberapa perangkat pengguna dapat mengakses layanan klasifikasi secara bersamaan melalui jaringan yang sama, sehingga biaya investasi teknologi dapat ditekan tanpa mengorbankan kecepatan maupun akurasi sistem. Kombinasi antara efisiensi model, transparansi keputusan melalui Grad-CAM, dan kemudahan akses berbasis web ini menjadikan sistem yang dibangun berpotensi lebih mudah diadopsi secara luas dibandingkan pendekatan yang memerlukan aplikasi khusus maupun perangkat keras tambahan.

Meskipun hasil yang diperoleh cukup menjanjikan, beberapa keterbatasan tetap perlu dicermati sebagai bahan evaluasi. Pertama, pengujian real-time hanya dilakukan pada kondisi pencahayaan mini studio yang relatif terkontrol, sehingga ketahanan model terhadap kondisi lapangan yang lebih ekstrem seperti sinar matahari langsung atau bayangan tajam belum sepenuhnya teruji. Kedua, posisi kamera yang digunakan pada tahap klasifikasi bersifat tegak lurus (top-down) terhadap objek, sehingga sistem hanya menganalisis kondisi permukaan atas buah dan belum mempertimbangkan kemungkinan perbedaan tingkat kematangan pada sisi buah yang tidak terjangkau sudut pandang kamera. Ketiga, meskipun akurasi keseluruhan tergolong tinggi, tingkat kesalahan pada batas transisi kelas Mentah-Mengkal menunjukkan bahwa model masih dapat ditingkatkan lebih lanjut melalui penambahan variasi data pada kedua kelas tersebut. Ketiga catatan ini menjadi dasar penting bagi arah pengembangan pada penelitian selanjutnya, khususnya untuk memperluas cakupan pengujian pada kondisi lingkungan yang lebih beragam dan representatif terhadap skenario penggunaan nyata di lapangan. Secara keseluruhan, capaian akurasi, kecepatan pemrosesan, dan transparansi visual yang diperoleh dalam penelitian ini menunjukkan bahwa pendekatan yang diusulkan layak dipertimbangkan sebagai alternatif solusi sortasi kematangan buah Carica yang lebih objektif dibandingkan metode manual konvensional yang selama ini masih banyak digunakan oleh petani dan pelaku usaha pengolahan Carica di kawasan Dieng, Wonosobo.

KESIMPULAN

Penelitian ini berhasil membangun sistem klasifikasi tingkat kematangan buah Carica ke dalam tiga kelas (Mentah, Mengkal, Matang) secara real-time menggunakan arsitektur MobileNetV2 melalui pendekatan transfer learning dan fine-tuning dua fase, yang dilengkapi visualisasi Grad-CAM sebagai bentuk transparansi pengambilan keputusan model. Model final mencapai akurasi 90,56% pada evaluasi validasi menggunakan format TensorFlow Lite, sementara pengujian dinamis real-time melalui aplikasi web menghasilkan akurasi rata-rata 95,5% dengan kecepatan pemrosesan stabil 60 FPS, melampaui ambang kebutuhan real-time. Visualisasi Grad-CAM membuktikan model mengambil keputusan berdasarkan fitur biologis buah yang tepat, yaitu gradasi warna kulit, sehingga sistem tidak lagi bersifat black box. Kesalahan klasifikasi paling banyak terjadi antara kelas Mentah dan Mengkal akibat kemiripan warna transisi pada kulit buah, sehingga penambahan variasi dataset pada kedua kelas tersebut menjadi arah perbaikan utama bagi penelitian selanjutnya. Keterbatasan penelitian ini terletak pada posisi kamera yang tegak lurus (top-down) pada tahap

klasifikasi, sehingga sistem hanya menganalisis permukaan atas buah. Penelitian selanjutnya disarankan mengeksplorasi mekanisme pengambilan citra multi-sudut, perbandingan dengan arsitektur CNN ringan lainnya, serta konversi sistem ke perangkat edge computing agar dapat beroperasi secara mandiri tanpa bergantung pada inferensi sisi server. Secara praktis, sistem ini berpotensi diadopsi oleh kelompok tani dan pelaku UMKM pengolahan Carica di kawasan Dieng untuk mendukung proses sortasi yang lebih objektif, konsisten, dan efisien, tanpa memerlukan investasi perangkat keras khusus maupun keahlian teknis tambahan dari penggunanya.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Sains Al-Qur'an Wonosobo atas dukungan fasilitas penelitian, serta kepada pengelola Excotic Carica Wonosobo atas ketersediaan bahan baku buah Carica yang digunakan dalam penyusunan dataset penelitian ini.

REFERENSI

- Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Hidayat, & Sari. (2021). Deteksi kematangan buah berbasis Android secara real-time. *Jurnal Sistem Informasi*, 5(2), 88–95.
- Ikhsan, Kurniawan, D., & Prasetyo, B. (2020). Pengolahan citra digital untuk perbaikan kualitas citra. *Jurnal Teknik Informatika*, 8(1), 45–52.
- Kosasih, & Fahrurrozi. (2023). Klasifikasi citra dalam konteks visi komputer: Ekstraksi fitur visual dan probabilitas kelas. *Jurnal Ilmu Komputer dan Informasi*, 16(1), 33–41.
- Kumar, & Singh. (2022). Fruit maturity classification using deep learning CNN. *IEEE Conference on Computer Vision*, 112–118.
- Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2022). A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12), 6999–7019. <https://doi.org/10.1109/TNNLS.2021.3084827>
- Minarno, A. E. (2015). Morfologi buah *Carica pubescens*. *Jurnal Biologi Terapan*, 3(1), 12–19.
- Nugroho. (2022). Perbandingan arsitektur VGG16 dan ResNet50 untuk klasifikasi buah. *Jurnal Teknik Elektro*, 14(3), 201–210.
- Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
- Purwono, P., Ma'arif, A., Rahmiani, W., Fathurrahman, H. I. K., Frisky, A. Z. K., & ul Haq, Q. M. (2023). Understanding of convolutional neural network (CNN): A review. *International Journal of Robotics and Control Systems*, 2(4), 739–748. <https://doi.org/10.31763/ijrcs.v2i4.888>
- Rahmanda, Fadillah, R., & Setiawan, T. (2023). Klasifikasi jenis buah menggunakan convolutional neural network (CNN). *Jurnal ELKOMIKA*, 11(2), 310–320.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211–252. <https://doi.org/10.1007/s11263-015-0816-y>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4510–4520.
- Saputra, Wibisono, A., & Kurniawan, F. (2022). Identifikasi kematangan pepaya menggunakan MobileNetV2. *Jurnal Informatika*, 9(2), 155–163.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618–626.
- Stankovic, J. A. (1988). Misconceptions about real-time computing: A serious problem for next-generation systems. *Computer*, 21(10), 10–19. <https://doi.org/10.1109/2.7053>
- Wibowo, & Arifianto. (2023). Implementasi explainable AI (XAI) pada deteksi penyakit tanaman. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 7(3), 455–463.

- Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems*, 27, 3320–3328.
- Zhang, Liu, Y., & Chen, H. (2023). Real-time fruit detection and grading using CNN. *Sensors*, 23(4), 1892–1901. <https://doi.org/10.3390/s23041892>