

Perbandingan Arsitektur *MobileNetV2* dan *EfficientNet-B0* untuk Klasifikasi Tingkat Kematangan Buah Pepaya berdasarkan Citra RGB

DOI: <http://dx.doi.org/10.35889/jutisi.v15i3.3693>

Creative Commons License 4.0 (CC BY – NC)

Wils Hosea Jaston Kawer¹, Ida Wahyuni^{2*}

Teknik Informatika, Institut Teknologi dan Bisnis Asia Malang, Malang, Indonesia

*e-mail Corresponding Author: idawahyuni@asia.ac.id

Abstract

Manual assessment of papaya ripeness was prone to inconsistency because visual differences between ripeness stages could be affected by color similarity, lighting, and transitional fruit characteristics. This study compared *MobileNetV2* and *EfficientNet-B0* for classifying papaya ripeness from RGB images by integrating dataset deduplication and model validation. The initial dataset contained 500 images in three classes: mature, partially mature, and immature. After duplicate images were removed using SHA-256, 246 unique images were obtained. The models were trained using transfer learning with ImageNet weights, stratified data splitting, and five-fold stratified cross-validation. Testing on the deduplicated dataset showed that *MobileNetV2* achieved 95.92% accuracy, while *EfficientNet-B0* achieved 93.88%. *MobileNetV2* was more efficient in terms of model size and inference time, whereas *EfficientNet-B0* was more stable in cross-validation. These findings showed that dataset deduplication, efficiency evaluation, and visual interpretation produced a more representative model assessment.

Keywords: Papaya ripeness classification; Dataset deduplication; Transfer learning; *MobileNetV2*; *EfficientNet-B0*

Abstrak

Penilaian manual terhadap kematangan buah pepaya dapat menghasilkan ketidakkonsistenan karena perbedaan visual antartingkat kematangan dipengaruhi oleh kemiripan warna, pencahayaan, dan karakter buah pada fase transisi. Penelitian ini membandingkan *MobileNetV2* dan *EfficientNet-B0* untuk klasifikasi tingkat kematangan buah pepaya berbasis citra RGB dengan mengintegrasikan deduplikasi dataset dan validasi model. Dataset awal terdiri dari 500 citra dalam tiga kelas, yaitu *mature*, *partially mature*, dan *immature*. Setelah citra duplikat dihapus menggunakan SHA-256, diperoleh 246 citra unik. Model dilatih menggunakan transfer learning dengan bobot ImageNet, *stratified split*, dan *5-fold stratified cross-validation*. Hasil pengujian pada dataset hasil deduplikasi menunjukkan *MobileNetV2* memperoleh akurasi 95,92%, sedangkan *EfficientNet-B0* memperoleh 93,88%. *MobileNetV2* lebih efisien dari sisi ukuran model dan waktu inferensi, sedangkan *EfficientNet-B0* lebih stabil pada validasi silang. Temuan ini menunjukkan bahwa deduplikasi dataset, evaluasi efisiensi, dan interpretasi visual menghasilkan penilaian model yang lebih representatif.

Kata kunci: Klasifikasi kematangan pepaya; Deduplikasi dataset; Transfer learning; *MobileNetV2*; *EfficientNet-B0*

1. Pendahuluan

Tingkat kematangan buah pepaya menjadi aspek penting dalam kegiatan pascapanen karena berhubungan dengan mutu buah, proses sortasi, distribusi, dan pengelolaan persediaan pangan. Kematangan pepaya umumnya dapat diamati melalui perubahan warna kulit, tekstur permukaan, serta ciri visual lain yang muncul selama proses pematangan. Sutrisna dkk. menunjukkan bahwa deteksi kematangan pepaya berbasis citra memiliki peran penting dalam bidang pertanian dan pengelolaan persediaan makanan, serta dapat diklasifikasikan

menggunakan *Convolutional Neural Network* (CNN) dengan akurasi tertinggi 96,63% [1]. Rizzo dkk. juga menjelaskan bahwa tingkat kematangan buah merupakan faktor penting dalam penentuan kualitas, tetapi penilaian manual masih berpotensi dipengaruhi subjektivitas dan membutuhkan dukungan otomasi berbasis *machine learning* maupun *deep learning* [2]. Berdasarkan hal tersebut, klasifikasi kematangan pepaya berbasis citra digital menjadi relevan untuk membantu proses identifikasi yang lebih objektif, terutama ketika pengamatan manual dipengaruhi oleh pencahayaan, sudut pengambilan gambar, pengalaman pengamat, dan kemiripan warna pada fase transisi.

Objek dalam penelitian ini adalah citra RGB buah pepaya yang dikelompokkan ke dalam tiga kelas, yaitu *mature*, *partially mature*, dan *immature*. Citra RGB digunakan sebagai masukan karena warna kulit merupakan indikator visual utama pada proses pematangan, sedangkan perbedaan antarkelas dipelajari secara otomatis oleh model *deep learning*. Masalah utama pada objek kajian ini muncul karena kelas *partially mature* berada pada fase peralihan, sehingga sebagian cirinya dapat menyerupai kelas *mature* maupun *immature*. Selain itu, penelitian ini juga menemukan persoalan pada kualitas dataset. Dataset awal berjumlah 500 citra, tetapi setelah dilakukan audit duplikasi hanya tersisa 246 citra unik. Artinya, terdapat 254 citra atau sekitar 50,80% data yang terindikasi berulang. Kondisi ini penting diperhatikan karena *near-duplicate image* dapat terjadi akibat kemiripan objek, perubahan pencahayaan, sudut pandang, rotasi, skala, atau transformasi visual lain [3]. Li dkk. menjelaskan bahwa deduplikasi citra dapat mengurangi data berulang dalam pelatihan CNN [4], sedangkan Tampu dkk. menunjukkan bahwa *data leakage* pada pembagian dataset dapat menyebabkan akurasi pengujian terlihat lebih tinggi daripada kemampuan generalisasi model yang sebenarnya [5].

Beberapa penelitian terdahulu telah berupaya menyelesaikan masalah klasifikasi atau prediksi kematangan buah dengan pendekatan yang berbeda. Simbolon dan Baniah menggunakan nilai warna RGB dengan metode City Block dan Canberra Distance untuk mengidentifikasi kematangan pepaya madu, serta memperoleh akurasi 83,33% dan 86,66% [6]. Wiharja, Al Mudzakir, Novita, dan Indra menggunakan fitur tekstur *Gray Level Co-occurrence Matrix* (GLCM) dengan Logistic Regression dan K-Nearest Neighbor untuk klasifikasi kematangan pepaya, dengan evaluasi menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*, serta memperoleh akurasi 73% dan 87% [7]. Pada objek buah apel, Ivani dkk. membandingkan VGG16 dan MobileNetV2 untuk klasifikasi tingkat kematangan buah melalui tahapan pengumpulan data, *preprocessing*, pelatihan model, dan evaluasi menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil penelitian tersebut menunjukkan bahwa VGG16 memperoleh akurasi 98% dan menjadi model yang lebih unggul pada kasus klasifikasi kematangan apel [8]. Santosa dkk. membandingkan MobileNetV2 dan ResNet50V2 untuk klasifikasi tingkat kematangan tomat ke dalam tiga kelas, yaitu mentah, setengah matang, dan matang, dengan evaluasi menggunakan *accuracy*, *precision*, *recall*, *F1-score*, serta waktu pengujian [9]. Fitriani dkk. membandingkan MobileNetV2 dan EfficientNetB0 dalam prediksi kematangan buah Jaboticaba berbasis citra RGB dengan evaluasi menggunakan *Mean Absolute Error* (MAE) [10]. Berdasarkan beberapa penelitian tersebut, klasifikasi dan prediksi kematangan buah telah banyak dilakukan menggunakan fitur manual, *machine learning*, maupun *deep learning*. Namun, penelitian sebelumnya masih cenderung menekankan pencapaian performa model, sementara pengelolaan dataset melalui deduplikasi, validasi model, efisiensi komputasi, dan interpretasi visual belum banyak dipadukan sebagai satu kerangka evaluasi pada klasifikasi kematangan buah pepaya.

Berdasarkan gap tersebut, penelitian ini mengusulkan perbandingan MobileNetV2 dan EfficientNet-B0 untuk klasifikasi tingkat kematangan buah pepaya berbasis citra RGB dengan mengintegrasikan deduplikasi dataset dan validasi model. MobileNetV2 dipilih karena arsitekturnya menggunakan *inverted residual* dan *linear bottleneck* sehingga sesuai untuk kebutuhan model yang ringan [11], sedangkan EfficientNet-B0 dipilih karena menerapkan *compound scaling* yang menyeimbangkan kedalaman, lebar, dan resolusi jaringan [12]. Deduplikasi ditempatkan sebagai tahap penting untuk mengurangi bias evaluasi akibat citra berulang dalam dataset, sehingga performa model lebih mencerminkan kemampuan generalisasi terhadap citra yang lebih representatif. Validasi model digunakan untuk memperoleh pembacaan performa yang lebih stabil, sedangkan evaluasi dilakukan melalui *accuracy*, *precision*, *recall*, *F1-score*, *confusion matrix*, ukuran model, dan waktu inferensi. Selain itu, *Grad-CAM* digunakan untuk melihat area pada citra RGB yang berkontribusi terhadap keputusan model, sehingga hasil klasifikasi tidak hanya dijelaskan melalui metrik numerik, tetapi juga melalui visualisasi yang lebih

mudah dipahami [13]. Tujuan utama penelitian ini adalah menyediakan dasar evaluasi yang lebih representatif dalam pemilihan arsitektur model untuk klasifikasi kematangan buah pepaya berbasis citra RGB. Kontribusi penelitian terletak pada penggabungan perbandingan MobileNetV2 dan EfficientNet-B0 dengan deduplikasi dataset, validasi model, evaluasi efisiensi, serta interpretasi visual, sehingga hasilnya dapat menjadi acuan dalam memilih model yang efisien dan stabil pada klasifikasi citra buah dengan batas visual antarkelas yang tidak tegas.

2. Metodologi

2.1. Pengumpulan Dataset

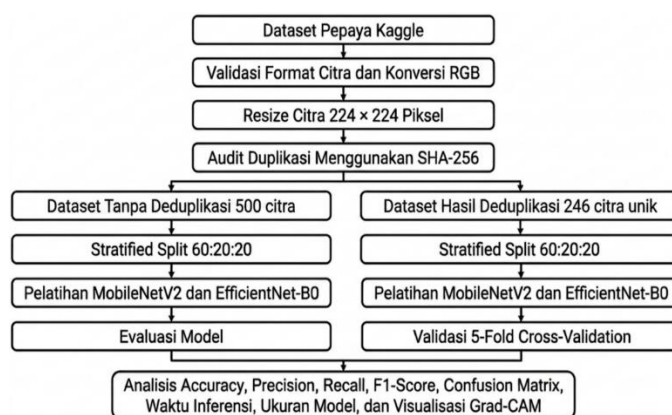
Pada tahap awal, penelitian ini menggunakan dataset citra buah pepaya yang diperoleh dari platform Kaggle dengan nama *Dataset Pepaya* melalui alamat <https://www.kaggle.com/datasets/nurhafiizah/dataset-pepaya>. Dataset tersebut berisi 500 citra buah pepaya yang terbagi ke dalam tiga kelas kematangan, yaitu *mature*, *partially mature*, dan *immature*. Dataset ini dipilih karena struktur kelasnya sesuai dengan kebutuhan penelitian, yaitu klasifikasi tingkat kematangan buah pepaya berbasis citra RGB. Komposisi dataset awal ditampilkan pada Tabel 1.

Tabel 1. Dataset Citra Buah Pepaya

Kategori	Jumlah	Contoh Gambar
<i>Mature</i>	168	
<i>Partially Mature</i>	142	
<i>Immature</i>	190	

2.2. Preprocessing dan Deduplikasi Dataset

Seluruh citra dikonversi ke format RGB berukuran 224×224 piksel. Pada data latih, augmentasi terbatas diterapkan melalui *flip horizontal*, rotasi, dan *zoom* untuk menambah variasi visual tanpa mengubah makna kelas.



Gambar 1. Alur Penelitian Klasifikasi Kematangan Buah Pepaya

Dataset awal digunakan sebagai skenario tanpa deduplikasi, sedangkan dataset hasil deduplikasi diperoleh melalui penghapusan citra duplikat identik berdasarkan nilai *hash SHA-256*. Dengan pendekatan ini, deduplikasi dibatasi pada citra yang benar-benar identik, sehingga

citra yang hanya mirip secara visual tetapi memiliki nilai hash berbeda tidak dihapus pada tahap ini. Melalui dua skenario tersebut, kinerja MobileNetV2 dan EfficientNet-B0 dapat dievaluasi secara lebih objektif, sebagaimana ditunjukkan pada **Gambar 1**.

2.3. Pembagian Dataset dan Validasi

Dataset dibagi menjadi data latih, data validasi, dan data uji dengan rasio 60:20:20 menggunakan stratified split agar proporsi kelas *mature*, *partially mature*, dan *immature* tetap terjaga pada setiap subset. Pembagian ini diterapkan pada dua skenario, yaitu dataset tanpa deduplikasi sebanyak 500 citra dan dataset hasil deduplikasi sebanyak 246 citra unik. Setelah proses deduplikasi, evaluasi diperkuat dengan *5-fold stratified cross-validation* agar performa model tidak hanya bergantung pada satu pembagian data. Audit antarsplit juga dilakukan untuk memastikan tidak ada citra identik yang muncul pada data latih, validasi, dan uji. Rincian pembagian dataset disajikan pada **Tabel 2**.

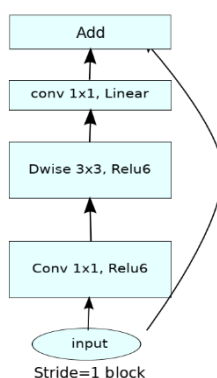
Tabel 2. Komposisi Pembagian Dataset Tanpa Deduplikasi dan Hasil Deduplikasi

Kondisi Dataset	Split	Mature	Partially Mature	Immature	Total
Tanpa Deduplikasi	Training	101	85	114	300
Tanpa Deduplikasi	Validation	34	28	38	100
Tanpa Deduplikasi	Testing	33	29	38	100
Hasil Deduplikasi	Training	52	34	62	148
Hasil Deduplikasi	Validation	17	11	21	49
Hasil Deduplikasi	Testing	18	11	20	49

2.4. Pembuatan Model

Model yang digunakan dalam penelitian ini adalah MobileNetV2 dan EfficientNet-B0. Kedua model dibuat menggunakan *transfer learning* dengan bobot awal ImageNet. Citra masukan berukuran 224×224 piksel dengan tiga kanal RGB, kemudian diklasifikasikan ke dalam tiga kelas, yaitu *mature*, *partially mature*, dan *immature*.

MobileNetV2 digunakan karena memiliki arsitektur ringan melalui konsep *inverted residual* dan *linear bottleneck*. Konsep ini mendukung ekstraksi fitur dengan beban komputasi yang lebih rendah, sehingga sesuai untuk penelitian yang mempertimbangkan ukuran model dan waktu inferensi [11]. Pada penelitian ini, MobileNetV2 digunakan sebagai *backbone* pengekstrak fitur, kemudian lapisan akhir disesuaikan untuk klasifikasi tiga kelas kematangan pepaya. Ilustrasi blok MobileNetV2 ditampilkan pada **Gambar 2**.



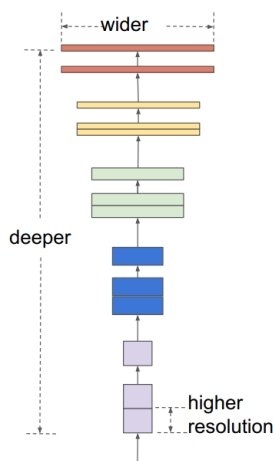
Gambar 2. Blok Inverted Residual MobileNetV2 dengan Stride 1

EfficientNet-B0 digunakan sebagai pembanding karena menerapkan *compound scaling*, yaitu penskalaan jaringan yang menyeimbangkan kedalaman, lebar, dan resolusi input. Konsep ini membantu EfficientNet menjaga keseimbangan antara performa klasifikasi dan efisiensi model [12]. Pada penelitian ini, EfficientNet-B0 juga digunakan sebagai *backbone* pengekstrak fitur, kemudian lapisan akhir disesuaikan dengan tiga kelas keluaran. Ilustrasi *compound scaling* EfficientNet ditampilkan pada **Gambar 3**.

Keluaran dari *backbone* diteruskan ke lapisan klasifikasi dengan aktivasi *softmax*. Aktivasi ini digunakan karena penelitian ini merupakan klasifikasi multikelas, sehingga keluaran model dinyatakan sebagai probabilitas pada setiap kelas. Fungsi loss yang digunakan adalah

sparse categorical crossentropy karena label kelas berbentuk indeks integer dan sesuai untuk menghitung kesalahan prediksi pada klasifikasi multikelas. Untuk mengurangi risiko *overfitting*, digunakan *dropout* sebesar 0,4 dan regularisasi L2 sebesar 0,0001.

Pelatihan dilakukan menggunakan *optimizer Adam* dengan *batch size 16*. Tahap awal pelatihan dilakukan selama *8 epoch* dengan *learning rate 0,001*, kemudian dilanjutkan *fine-tuning* selama *12 epoch* dengan *learning rate 0,00001*. *Fine-tuning* dilakukan agar bobot model pra-latih menyesuaikan pola visual citra pepaya. Dengan rancangan ini, MobileNetV2 dan EfficientNet-B0 dibandingkan berdasarkan performa klasifikasi, stabilitas, ukuran model, dan waktu inferensi.



Gambar 3. Konsep Compound Scaling pada EfficientNet

2.5. Pengujian dan Evaluasi Model

Pengujian dilakukan menggunakan data uji yang telah dipisahkan pada tahap pembagian dataset. Model yang diuji adalah MobileNetV2 dan EfficientNet-B0 pada dua skenario, yaitu tanpa deduplikasi dan hasil deduplikasi. Evaluasi ini bertujuan untuk menilai kemampuan model dalam mengklasifikasikan tingkat kematangan buah pepaya serta melihat pengaruh deduplikasi terhadap performa model.

Metrik evaluasi yang digunakan meliputi *accuracy*, *precision*, *recall*, *F1-score*, dan *AUC-ROC*. Metrik tersebut dipilih agar penilaian tidak hanya bertumpu pada akurasi, tetapi juga memperhatikan kualitas prediksi pada setiap kelas. Pendekatan evaluasi seperti ini sejalan dengan penelitian klasifikasi kematangan pepaya yang memanfaatkan *confusion matrix* dan ukuran performa turunan untuk menilai hasil prediksi model [1], [7].

Selain metrik klasifikasi, penelitian ini mengukur waktu inferensi dan ukuran model untuk menilai efisiensi komputasi. Pengukuran ini diperlukan karena MobileNetV2 dan EfficientNet-B0 memiliki karakter arsitektur yang berbeda dalam menyeimbangkan performa dan efisiensi model [11], [12]. Untuk memperjelas analisis, digunakan *confusion matrix* dan *Grad-CAM*. *Confusion matrix* digunakan untuk melihat pola kesalahan prediksi antarkelas, sedangkan *Grad-CAM* digunakan untuk menunjukkan area pada citra RGB yang paling memengaruhi keputusan model [13]. Hasil evaluasi disajikan dan dibahas pada bagian Hasil dan Pembahasan.

3. Hasil dan Pembahasan

3.1. Realisasi Deduplikasi dan Pembagian Dataset

Hasil awal penelitian menunjukkan bahwa deduplikasi mengubah komposisi dataset secara signifikan. Dataset awal berjumlah 500 citra pepaya yang terbagi ke dalam tiga kelas, yaitu *mature*, *partially mature*, dan *immature*. Setelah audit duplikasi menggunakan SHA-256, jumlah data berkurang menjadi 246 citra yang unik. Kondisi ini menunjukkan adanya pengulangan citra pada dataset awal, sehingga pembersihan data diperlukan agar evaluasi model tidak dipengaruhi oleh pola yang sama. Ringkasan perubahan dataset ditampilkan pada Tabel 3.

Tabel 3. Ringkasan Dataset Sebelum dan Sesudah Deduplikasi

Kondisi Dataset	<i>Mature</i>	<i>Partially Mature</i>	<i>Immature</i>	Total	Keterangan
Sebelum Deduplikasi	168	142	190	500	Dataset awal
Sesudah Deduplikasi	87	56	103	246	Dataset unik
Reduksi Data	81	86	87	254	Dataset duplikat

Berdasarkan Tabel 3, proses deduplikasi mengurangi 254 citra atau sekitar 50,80% dari dataset awal. Secara proporsional, reduksi terbesar terjadi pada kelas *partially mature*, sehingga kelas tersebut memiliki tingkat pengulangan data paling tinggi dibandingkan dengan jumlah awal kelasnya. Temuan ini memperkuat bahwa deduplikasi bukan sekadar pembersihan data, melainkan bagian penting dalam menjaga validitas evaluasi. Data berulang dapat membuat performa tampak tinggi, tetapi belum tentu mencerminkan kemampuan generalisasi model pada citra yang tidak berulang. Hal ini sejalan dengan penelitian terdahulu yang menekankan pentingnya penghapusan *near-duplicate image*, deduplikasi data pelatihan, dan pencegahan *data leakage* agar evaluasi model tidak menghasilkan pembacaan performa yang terlalu optimistis [3], [4], [5].

Setelah deduplikasi, dataset dibagi dengan rasio 60:20:20 menggunakan pendekatan *stratified split*. Pembagian ini dilakukan agar proporsi kelas tetap terjaga pada data latih, data validasi, dan data uji. Pada skenario deduplikasi, audit antarsplit juga dilakukan untuk memastikan tidak ada citra identik yang berpindah antarsubset.

3.2. Skenario Pengujian Model

Pengujian dilakukan terhadap MobileNetV2 dan EfficientNet-B0 pada dua kondisi dataset, yaitu tanpa deduplikasi dan dengan deduplikasi. Skenario tanpa deduplikasi digunakan sebagai pembandingan terhadap dataset awal, sedangkan skenario hasil deduplikasi digunakan untuk menilai performa model setelah data duplikat dihilangkan. Dataset hasil deduplikasi juga diuji menggunakan *5-fold stratified cross-validation* untuk memperoleh pembacaan performa yang lebih stabil.

Antarmuka pengujian berbasis Streamlit memiliki dua mode, yaitu Mode Penelitian dan Mode Uji Gambar. Mode Penelitian menampilkan metrik evaluasi, grafik pelatihan, dan *confusion matrix* pada skenario tanpa deduplikasi serta pada hasil deduplikasi. Mode Uji Gambar menampilkan prediksi citra tunggal, nilai *confidence*, probabilitas kelas, rekomendasi model, serta *Grad-CAM*. Ringkasan skenario pengujian ditampilkan pada Tabel 4.

Tabel 4. Skenario Pengujian Model

Skenario	Kondisi Dataset	Model	Bentuk Evaluasi
1	Tanpa Deduplikasi	MobileNetV2	Evaluasi performa dan visualisasi
2	Tanpa Deduplikasi	EfficientNet-B0	Evaluasi performa dan visualisasi
3	Hasil Deduplikasi	MobileNetV2	Evaluasi performa, visualisasi, dan validasi silang
4	Hasil Deduplikasi	EfficientNet-B0	Evaluasi performa, visualisasi, dan validasi silang
5	Hasil Deduplikasi	MobileNetV2 dan EfficientNet-B0	Uji citra tunggal dan interpretasi visual

Tabel 4 menunjukkan bahwa pengujian tidak hanya berfokus pada akurasi, tetapi juga mempertimbangkan stabilitas model, efisiensi komputasi, pola kesalahan klasifikasi, serta interpretasi visual. Evaluasi metrik mencakup *accuracy*, *precision*, *recall*, *F1-score*, *AUC-ROC*, waktu inferensi, dan ukuran model.

3.3. Hasil Pengujian Model pada Dataset Tanpa Deduplikasi dan Hasil Deduplikasi

Hasil pengujian utama diperoleh dari perbandingan MobileNetV2 dan EfficientNet-B0 pada dataset tanpa deduplikasi dan dataset hasil deduplikasi. Evaluasi dilakukan menggunakan *accuracy*, *precision*, *recall*, *F1-score*, *AUC-ROC*, waktu inferensi, dan ukuran model. Hasil pengujian ditampilkan pada Tabel 5.

Tabel 5. Hasil Pengujian Model pada Dataset Tanpa Deduplikasi dan Hasil Deduplikasi

Kondisi Dataset	Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC	Waktu Inferensi	Ukuran Model
Tanpa Deduplikasi	MobileNetV2	96,00%	95,89%	96,49%	96,01%	0,9971	70,22 ms	20,89 MB
Tanpa Deduplikasi	EfficientNet-B0	99,00%	99,15%	98,85%	98,98%	0,9999	116,32 ms	33,92 MB
Hasil Deduplikasi	MobileNetV2	95,92%	95,22%	95,30%	95,21%	0,9970	71,09 ms	20,89 MB
Hasil Deduplikasi	EfficientNet-B0	93,88%	93,33%	92,27%	92,63%	0,9900	97,05 ms	33,92 MB

Tabel 5 menunjukkan bahwa EfficientNet-B0 memperoleh akurasi tertinggi pada skenario tanpa deduplikasi, yaitu 99,00%, sedangkan MobileNetV2 memperoleh akurasi 96,00%. Hasil tersebut menunjukkan bahwa kedua model mampu mengenali pola visual kematangan pepaya dengan baik. Namun, hasil tanpa deduplikasi perlu dibaca dengan saksama karena dataset awal masih mengandung citra berulang yang dapat membuat evaluasi tampak lebih tinggi.

Pada skenario hasil deduplikasi, MobileNetV2 memperoleh akurasi 95,92%, sedangkan EfficientNet-B0 memperoleh 93,88%. Penurunan performa setelah deduplikasi menunjukkan bahwa pengujian menjadi lebih ketat karena model tidak lagi memperoleh keuntungan dari pengulangan citra yang identik. Dari sisi efisiensi, MobileNetV2 memiliki ukuran model yang lebih kecil, yaitu 20,89 MB, serta waktu inferensi yang lebih rendah dibandingkan dengan EfficientNet-B0. Hasil ini selaras dengan karakter MobileNetV2 sebagai arsitektur ringan berbasis *inverted residual* dan *linear bottleneck* [11]. Sementara itu, EfficientNet-B0 menggunakan *compound scaling* untuk menyeimbangkan skala jaringan dan performa model [12]. Sehingga hasilnya perlu dibaca tidak hanya dari akurasi, tetapi juga dari ukuran model, waktu inferensi, dan stabilitas performa.

3.4. Hasil Validasi 5-Fold Cross-Validation pada Dataset Hasil Deduplikasi

Validasi tambahan dilakukan menggunakan *5-fold stratified cross-validation* pada dataset hasil deduplikasi. Pengujian ini digunakan karena jumlah data setelah deduplikasi berkurang cukup besar, sehingga evaluasi tidak cukup bergantung pada satu kali pembagian data. Hasil validasi silang ditampilkan pada **Tabel 6**.

Tabel 6. Hasil 5-Fold Cross-Validation pada Dataset Hasil Deduplikasi

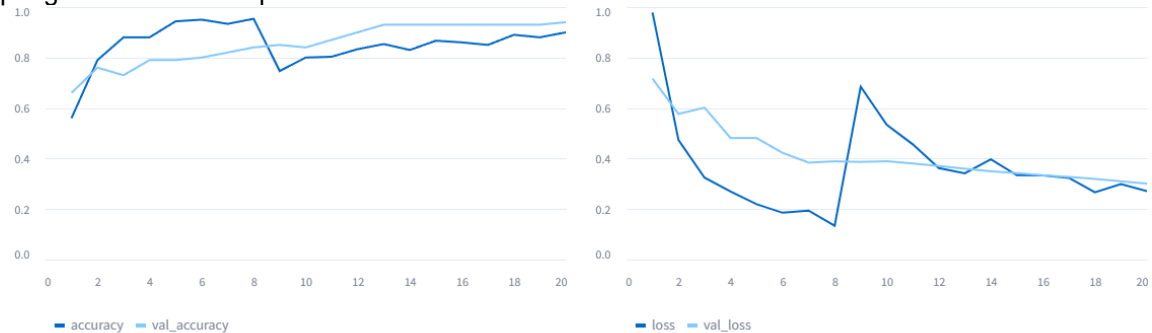
Model	Fold	Accuracy	Precision	Recall	F1-Score
MobileNetV2	1	97,50%	96,67%	98,04%	97,24%
MobileNetV2	2	95,00%	94,07%	94,34%	94,14%
MobileNetV2	3	89,74%	89,74%	91,37%	89,26%
MobileNetV2	4	97,44%	97,78%	96,30%	96,89%
MobileNetV2	5	92,31%	91,67%	94,12%	92,01%
MobileNetV2	Rata-rata	94,40%	93,99%	94,83%	93,91%
EfficientNet-B0	1	100,00%	100,00%	100,00%	100,00%
EfficientNet-B0	2	100,00%	100,00%	100,00%	100,00%
EfficientNet-B0	3	97,44%	96,67%	97,92%	97,17%
EfficientNet-B0	4	100,00%	100,00%	100,00%	100,00%
EfficientNet-B0	5	97,44%	96,67%	98,04%	97,24%
EfficientNet-B0	Rata-rata	98,97%	98,67%	99,19%	98,88%

Berdasarkan **Tabel 6**, EfficientNet-B0 memperoleh rata-rata akurasi 98,97% dan *F1-score* 98,88%, lebih tinggi dibandingkan MobileNetV2 yang memperoleh akurasi rata-rata 94,40% dan *F1-score* 93,91%. Hasil ini menunjukkan bahwa EfficientNet-B0 lebih stabil ketika diuji pada beberapa variasi pembagian data pada dataset hasil deduplikasi. Perbedaan antara hasil data uji pada **Tabel 5** dan hasil validasi silang pada **Tabel 6** perlu dibaca secara metodologis. Data uji hasil deduplikasi hanya berjumlah 49 citra, sehingga satu kesalahan prediksi dapat memengaruhi persentase akhir secara signifikan. Sebaliknya, *5-fold cross-validation* memberikan gambaran performa yang lebih luas karena seluruh data hasil deduplikasi memperoleh kesempatan menjadi data uji pada *fold* yang berbeda.

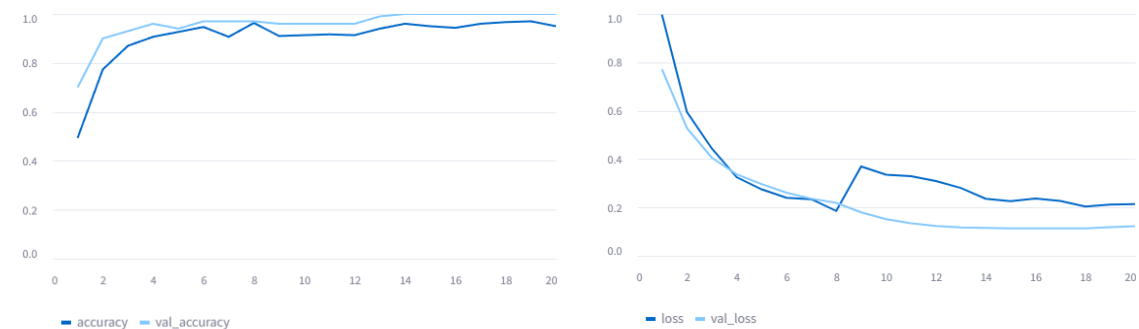
3.5. Hasil Visualisasi Grafik Pelatihan pada Dataset Tanpa Deduplikasi

Grafik pelatihan dataset tanpa deduplikasi menampilkan akurasi dan *loss* MobileNetV2 serta EfficientNet-B0, sehingga memperlihatkan performa akhir dan stabilitas proses pelatihan. Berdasarkan **Gambar 4**, MobileNetV2 menunjukkan peningkatan akurasi dan penurunan *loss* selama pelatihan. Pola ini mendukung hasil pada **Tabel 5**, di mana MobileNetV2 memperoleh akurasi 96,00% pada skenario tanpa deduplikasi. Namun, hasil tersebut tetap perlu dibaca dengan saksama karena dataset masih mengandung citra berulang.

Gambar 5 menunjukkan pola pelatihan EfficientNet-B0 yang kuat pada dataset tanpa deduplikasi. Hal ini sejalan dengan **Tabel 5**, di mana EfficientNet-B0 memperoleh akurasi sebesar 99,00%. Meskipun demikian, performa tinggi pada skenario ini belum sepenuhnya merepresentasikan kemampuan generalisasi karena masih terdapat kemungkinan adanya pengaruh dari citra duplikat.



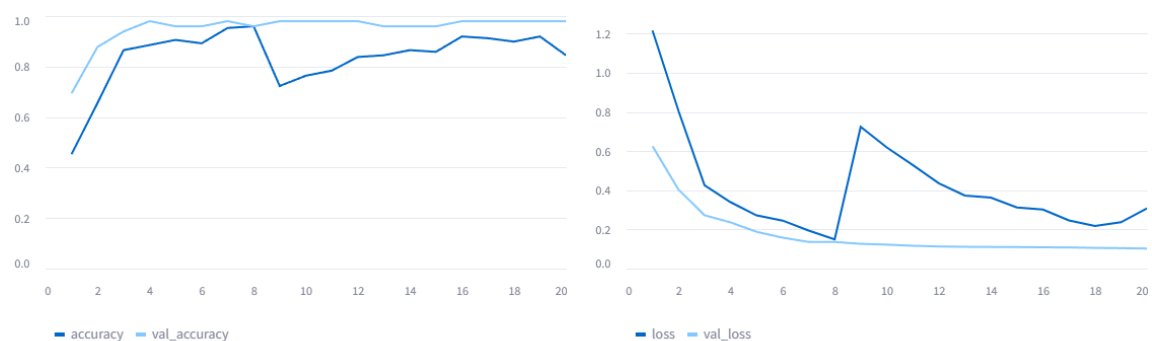
Gambar 4. Grafik Akurasi dan Loss MobileNetV2 pada Dataset Tanpa Deduplikasi



Gambar 5. Grafik Akurasi dan Loss EfficientNet-B0 pada Dataset Tanpa Deduplikasi

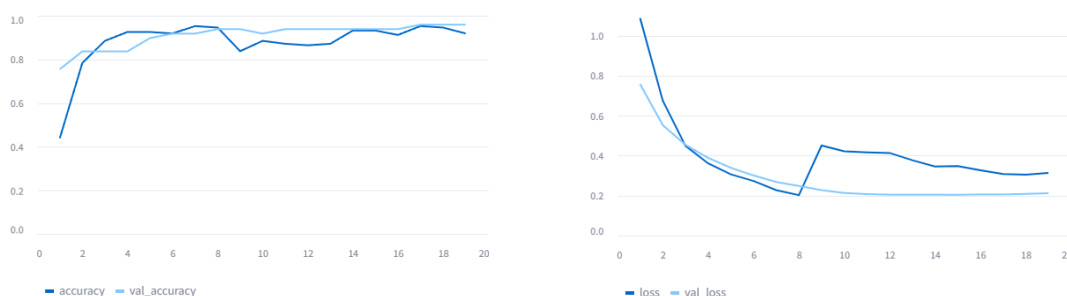
3.6. Hasil Visualisasi Grafik Pelatihan pada Dataset Hasil Deduplikasi

Grafik pelatihan dataset hasil deduplikasi menampilkan akurasi dan *loss* kedua model, menegaskan efektivitas deduplikasi dalam mengevaluasi performa dan kestabilan MobileNetV2 dan EfficientNet-B0 secara lebih realistis.



Gambar 6. Grafik Akurasi dan Loss MobileNetV2 pada Dataset Hasil Deduplikasi

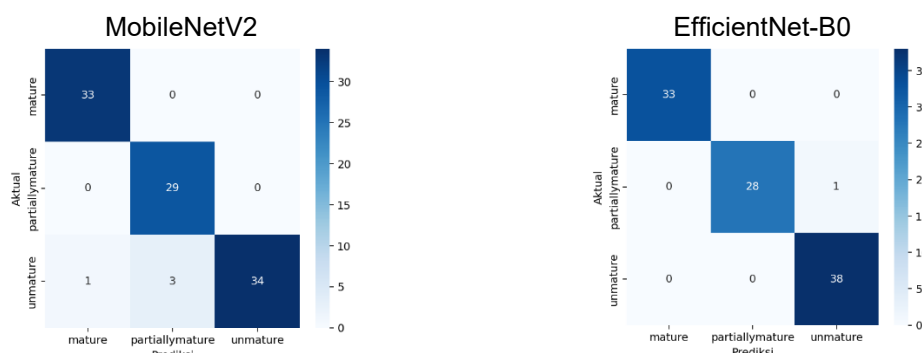
Gambar 6 menunjukkan bahwa MobileNetV2 tetap mampu mempelajari pola visual pada dataset hasil deduplikasi. Fluktuasi masih dapat terjadi karena jumlah data setelah deduplikasi lebih kecil daripada dataset awal. Namun, pola akurasi dan *loss* tetap menunjukkan bahwa model dapat mengenali karakter visual dari tiga kelas kematangan pepaya. Selain itu, berdasarkan **Gambar 7**, EfficientNet-B0 menunjukkan pola pelatihan yang konsisten pada dataset hasil deduplikasi. Grafik ini mendukung hasil *5-fold cross-validation* pada **Tabel 6**, di mana EfficientNet-B0 menunjukkan performa yang lebih stabil dibandingkan MobileNetV2. Dengan demikian, evaluasi model pada skenario hasil deduplikasi tidak hanya bergantung pada satu pembagian data, tetapi juga diperkuat melalui validasi silang.



Gambar 7. Grafik Akurasi dan Loss EfficientNet-B0 pada Dataset Hasil Deduplikasi

3.7. Hasil Confusion Matrix pada Dataset Tanpa Deduplikasi

Confusion matrix pada dataset tanpa deduplikasi digunakan untuk melihat distribusi prediksi benar dan salah sebelum data duplikat dihilangkan. Visualisasi ini melengkapi **Tabel 5** karena akurasi dan *F1-score* hanya merangkum performa, sedangkan *confusion matrix* menunjukkan pola kesalahan pada setiap kelas. Hasilnya ditampilkan pada **Gambar 8**.

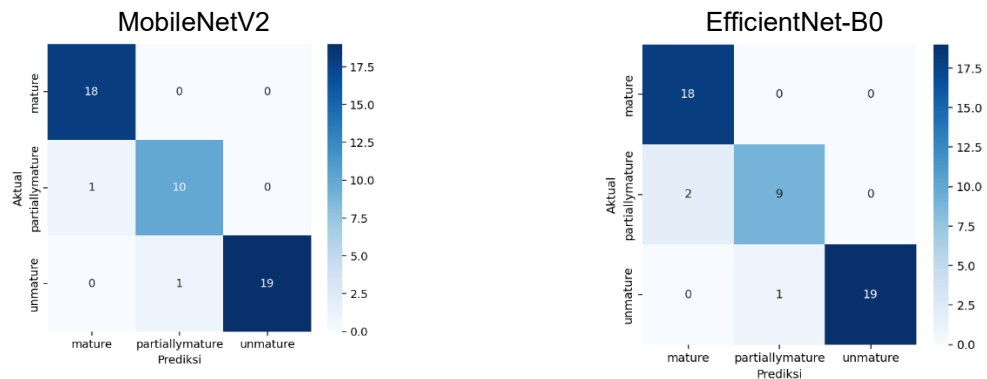


Gambar 8. Confusion Matrix MobileNetV2 dan EfficientNet-B0 pada Dataset Tanpa Deduplikasi

Berdasarkan **Gambar 8**, kedua model mampu mengklasifikasikan tiga kelas kematangan pepaya pada dataset tanpa deduplikasi. Namun, hasil ini lebih tepat dibaca sebagai pembandingan, bukan sebagai dasar utama generalisasi model, karena dataset tanpa deduplikasi masih mengandung citra berulang yang dapat memengaruhi jumlah prediksi yang benar.

3.8. Hasil Confusion Matrix pada Dataset Hasil Deduplikasi

Confusion matrix pada dataset hasil deduplikasi digunakan untuk membaca pola kesalahan setelah citra duplikat dihapus. Pada skenario ini, hasil prediksi lebih representatif karena data telah melalui pembersihan dan audit antarsplit. Dengan demikian, kesalahan yang muncul lebih tepat dibaca sebagai tantangan klasifikasi antarkelas, bukan akibat pengulangan citra. Berdasarkan **Gambar 9**, kesalahan model terutama muncul pada perpindahan kelas yang berdekatan secara visual, yaitu *partially mature* ke *mature* dan *immature* ke *partially mature*. Kondisi ini wajar karena perubahan warna pepaya pada fase transisi tidak selalu terlihat jelas. Oleh sebab itu, tantangan utama penelitian bukan hanya membedakan pepaya matang dan mentah, tetapi juga mengenali batas visual pada kelas peralihan.



Gambar 9. Confusion Matrix MobileNetV2 dan EfficientNet-B0 pada Dataset Hasil Deduplikasi

Berdasarkan *confusion matrix* pada dataset hasil deduplikasi, akurasi kedua model belum mencapai 100%. MobileNetV2 menghasilkan 2 kesalahan prediksi, sedangkan EfficientNet-B0 menghasilkan 3 kesalahan prediksi. Oleh karena itu, analisis dilanjutkan dengan melihat sampel citra yang berhasil dikenali serta citra yang mengalami kesalahan klasifikasi. Analisis ini diperlukan agar hasil evaluasi tidak hanya dinilai dari nilai akurasi, tetapi juga dari karakter visual citra yang memengaruhi keputusan model.

3.9. Analisis Citra yang Berhasil Diklasifikasikan dan Terlibat dalam Kesalahan Klasifikasi

Analisis pada subbab ini dilakukan untuk menjelaskan hasil pengujian pada dataset hasil deduplikasi yang belum mencapai akurasi 100%. *Confusion matrix* pada **Gambar 9** telah menunjukkan pola prediksi benar dan salah secara keseluruhan, sedangkan subbab ini menampilkan sampel citra untuk memperjelas karakter visual yang memengaruhi keputusan model. Citra yang berhasil dikenali ditampilkan secara representatif, yaitu satu citra dari setiap kelas. Sementara itu, citra yang terlibat dalam kesalahan klasifikasi ditampilkan secara menyeluruh berdasarkan hasil pengujian pada dataset uji hasil deduplikasi.

Tabel 7. Sampel Citra yang Berhasil Diklasifikasikan dengan Tepat

Kelas Aktual	Nama Citra	Prediksi MobileNetV2	Prediksi EfficientNet-B0	Karakter Citra
Matang	mature_016.jpg	Matang (79,91%)	Matang (94,29%)	Ciri kematangan terlihat cukup jelas sehingga kedua model mampu mengenali kelas dengan tepat.
Setengah Matang	partiallymature_130.jpeg	Setengah Matang (99,51%)	Setengah Matang (98,27%)	Citra menunjukkan fase peralihan yang masih dapat dibedakan dari kelas matang dan mentah.
Mentah	unmature_186.jpg	Mentah (96,29%)	Mentah (94,43%)	Ciri mentah masih dominan sehingga kedua model memberikan prediksi sesuai kelas aktual.

Berdasarkan **Tabel 7**, citra yang berhasil diklasifikasikan dengan tepat umumnya memiliki karakter visual yang cukup tegas. Citra matang menunjukkan ciri kematangan yang jelas, citra setengah matang masih memiliki pola transisi yang dapat dibedakan, sedangkan citra mentah memperlihatkan ciri visual mentah yang dominan. Kondisi tersebut menunjukkan bahwa MobileNetV2 dan EfficientNet-B0 mampu mengenali tingkat kematangan pepaya dengan baik ketika karakter visual pada citra konsisten dengan kelas aktualnya.

Tabel 8. Citra yang Terlibat dalam Kesalahan Klasifikasi pada Dataset Uji

Kelas Aktual	Nama Citra	Prediksi MobileNetV2	Prediksi EfficientNet-B0	Karakter Citra
Setengah Matang	partiallymature_021.jpg	Matang (54,21%)	Matang (56,31%)	Citra setengah matang memiliki karakter visual yang mendekati kelas matang, sehingga kedua model salah mengenali kelasnya.
Setengah Matang	partiallymature_097.jpg	Setengah Matang (72,09%)	Matang (69,56%)	MobileNetV2 masih mengenali kelas dengan tepat, sedangkan EfficientNet-B0 membaca citra ini sebagai matang.
Mentah	unmature_001.jpg	Setengah Matang (75,29%)	Setengah Matang (72,37%)	Citra mentah memiliki ciri yang menyerupai fase awal pematangan, sehingga kedua model memprediksinya sebagai setengah matang.

Tabel 8 memuat seluruh citra unik yang terlibat dalam kesalahan klasifikasi pada dataset uji hasil deduplikasi. MobileNetV2 menghasilkan dua kesalahan prediksi, yaitu pada *partiallymature_021.jpg* dan *unmature_001.jpg*. EfficientNet-B0 menghasilkan tiga kesalahan prediksi, yaitu pada *partiallymature_021.jpg*, *partiallymature_097.jpg*, dan *unmature_001.jpg*. Jumlah citra yang ditampilkan sebanyak tiga karena *partiallymature_097.jpg* hanya salah pada EfficientNet-B0, sedangkan MobileNetV2 masih mengklasifikasikannya dengan tepat.

Kesalahan pada *partiallymature_021.jpg* dan *partiallymature_097.jpg* menunjukkan bahwa kelas setengah matang dapat memiliki karakter visual yang mendekati kelas matang. Hal ini terjadi karena sebagian area buah telah menunjukkan warna yang lebih menyerupai fase matang, sehingga model dapat membaca citra tersebut sebagai kelas matang. Sementara itu, kesalahan pada *unmature_001.jpg* menunjukkan bahwa sebagian citra mentah dapat memiliki pola warna atau tekstur yang mendekati fase awal pematangan. Akibatnya, kedua model memprediksi citra tersebut sebagai setengah matang.

Perbandingan antara **Tabel 7** dan **Tabel 8** menunjukkan bahwa kesalahan klasifikasi terutama terjadi pada batas visual antarkelas yang berdekatan. Pola kesalahan tidak muncul pada kelas yang berjauhan, seperti mentah langsung menjadi matang, tetapi terjadi pada perpindahan kelas mentah ke setengah matang dan setengah matang ke matang. Kondisi ini sesuai dengan karakter objek penelitian karena proses pematangan pepaya berlangsung bertahap, sehingga batas visual antara kelas *immature*, *partially mature*, dan *mature* tidak selalu tegas. Dengan demikian, kelas *partially mature* menjadi kelas yang paling menantang karena memiliki ciri visual campuran antara mentah dan matang.

**Gambar 10.** Sampel Citra yang Berhasil Diklasifikasikan dan Terlibat dalam Kesalahan Klasifikasi

Untuk memperjelas perbedaan visual antara citra yang berhasil diklasifikasikan dan citra yang terlibat dalam kesalahan klasifikasi, sampel citra disajikan pada **Gambar 10**. Baris pertama

menampilkan citra yang berhasil diklasifikasikan dengan tepat, sedangkan baris kedua menampilkan citra yang mengalami kesalahan klasifikasi.

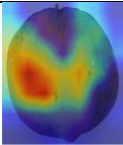
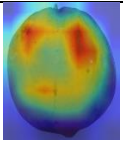
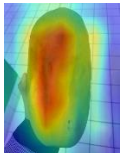
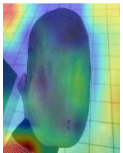
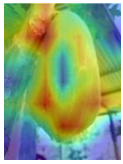
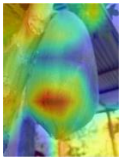
Baris pertama pada Gambar 10 menampilkan *mature_016.jpg*, *partiallymature_130.jpeg*, dan *unmature_186.jpg* sebagai sampel citra yang berhasil diklasifikasikan dengan tepat. Baris kedua menampilkan *partiallymature_021.jpg*, *partiallymature_097.jpg*, dan *unmature_001.jpg* sebagai citra yang terlibat dalam kesalahan klasifikasi.

3.10. Visualisasi Grad-CAM pada Citra Uji

Grad-CAM digunakan untuk menampilkan area citra yang menjadi perhatian model saat menghasilkan prediksi. Visualisasi ini ditampilkan pada citra uji dari skenario hasil deduplikasi sebagai pendukung interpretasi, bukan sebagai metrik utama. Hasil prediksi dan nilai *confidence* telah disajikan pada Tabel 7 dan Tabel 8, sedangkan Gambar 10 memperlihatkan contoh citra yang berhasil diklasifikasikan dan citra yang terlibat dalam kesalahan klasifikasi. Oleh karena itu, bagian ini berfokus pada peta perhatian visual untuk melihat apakah model mengambil keputusan berdasarkan area buah yang relevan dengan ciri kematangan. Visualisasi *Grad-CAM* ditampilkan pada Tabel 9.

Pada Tabel 9 ditampilkan citra asli dan peta *Grad-CAM* dari MobileNetV2 serta EfficientNet-B0. Apabila perhatian model berada pada permukaan buah, keputusan klasifikasi dapat dinilai relevan secara visual karena kematangan pepaya ditentukan oleh perubahan warna dan tekstur kulitnya. Sebaliknya, apabila perhatian model mengarah pada latar belakang, bayangan, atau area *non*-buah, kondisi tersebut menjadi catatan keterbatasan model dalam membaca ciri visual citra tertentu. Dengan demikian, *Grad-CAM* digunakan untuk memperkuat pembahasan hasil klasifikasi, terutama dalam menjelaskan perbedaan antara citra yang berhasil dikenali dan citra yang terlibat dalam kesalahan klasifikasi.

Tabel 9. Visualisasi Grad-CAM pada Citra Uji Berdasarkan Model Hasil Deduplikasi

Kelas Citra Uji	Citra Asli	Grad-CAM MobileNetV2	Grad-CAM EfficientNet-B0
<i>Mature</i>			
<i>Partially Mature</i>			
<i>Immature</i>			

3.11. Pembahasan

Hasil penelitian menunjukkan bahwa deduplikasi dataset berpengaruh langsung terhadap performa pembacaan model klasifikasi kematangan buah pepaya. Pada skenario tanpa deduplikasi, EfficientNet-B0 memperoleh akurasi 99,00%, sedangkan MobileNetV2 memperoleh akurasi 96,00%. Nilai tersebut menunjukkan bahwa kedua model mampu mengenali pola visual kematangan pepaya dengan baik. Namun, hasil pada dataset awal belum dapat langsung dipandang sebagai gambaran kemampuan generalisasi karena dataset tersebut masih mengandung citra yang berulang. Setelah deduplikasi dilakukan, jumlah citra berkurang dari 500 menjadi 246 citra unik, sehingga evaluasi menjadi lebih ketat dan lebih representatif terhadap kemampuan model dalam mengenali citra yang tidak berulang.

Temuan tersebut memperkuat pentingnya audit kualitas dataset dalam penelitian klasifikasi citra. Shah dan Shrivastava menjelaskan bahwa deduplikasi citra dapat membantu meningkatkan kualitas data dan mengurangi pengaruh kemiripan data dalam sistem pengenalan berbasis citra [14]. Pada penelitian ini, deduplikasi tidak hanya berperan sebagai tahap pembersihan data, tetapi juga menjadi bagian dari rancangan evaluasi. Perbandingan antara skenario tanpa deduplikasi dan hasil deduplikasi menunjukkan bahwa performa model perlu dibaca bersama dengan kondisi dataset yang digunakan. Dengan demikian, akurasi tinggi pada dataset awal harus dipahami dengan saksama karena dapat dipengaruhi oleh pengulangan pola visual.

Jika dikaitkan dengan penelitian terdahulu pada objek pepaya, hasil penelitian ini mendukung temuan bahwa citra buah pepaya dapat digunakan sebagai dasar untuk mengklasifikasikan tingkat kematangan. Nurmalasari dkk. menggunakan CNN untuk klasifikasi tingkat kematangan buah pepaya dan menunjukkan bahwa model dapat mempelajari ciri visual untuk membedakan kelas kematangan [15]. Penelitian ini memperluas temuan tersebut dengan membandingkan MobileNetV2 dan EfficientNet-B0, serta menambahkan deduplikasi dataset, *5-fold stratified cross-validation*, ukuran model, waktu inferensi, *confusion matrix*, analisis sampel citra, dan *Grad-CAM*. Dengan demikian, penelitian ini tidak hanya menguatkan penggunaan *deep learning* untuk klasifikasi kematangan pepaya, tetapi juga memperluas cara evaluasi agar tidak hanya bertumpu pada akurasi.

Hasil penelitian ini juga relevan dengan beberapa penelitian klasifikasi citra buah yang menggunakan *transfer learning*. Amanda dkk. menerapkan MobileNetV2 untuk klasifikasi kondisi buah alpukat mentega [16], Nastitie dan Handayani membandingkan MobileNetV2 dan EfficientNet-B0 pada klasifikasi jenis pisang [17], sedangkan Hartono dan Muslikh menunjukkan bahwa MobileNetV2 dapat digunakan pada klasifikasi citra jenis buah-buahan [18]. Sejalan dengan penelitian tersebut, MobileNetV2 pada penelitian ini mampu memperoleh akurasi 95,92% pada dataset hasil deduplikasi. Namun, penelitian ini memberikan pembacaan yang lebih luas karena performa model juga dianalisis melalui ukuran model, waktu inferensi, stabilitas validasi silang, serta interpretasi visual.

Perbedaan performa MobileNetV2 dan EfficientNet-B0 menunjukkan bahwa pemilihan model tidak cukup hanya berdasarkan nilai akurasi. Pada data uji hasil deduplikasi, MobileNetV2 memperoleh akurasi 95,92%, ukuran model 20,89 MB, dan waktu inferensi 71,09 ms. EfficientNet-B0 memperoleh akurasi 93,88%, ukuran model 33,92 MB, dan waktu inferensi 97,05 ms. Hasil ini menunjukkan bahwa MobileNetV2 lebih unggul dari sisi efisiensi pada skenario data uji. Sebaliknya, pada *5-fold stratified cross-validation*, EfficientNet-B0 memperoleh rata-rata akurasi 98,97%, sedangkan MobileNetV2 memperoleh rata-rata akurasi 94,40%. Dengan demikian, MobileNetV2 lebih kuat dalam hal efisiensi dan hasil uji deduplikasi, sedangkan EfficientNet-B0 lebih kuat dalam hal kestabilan performa lintas pembagian data.

Apabila ditinjau dari penelitian citra pertanian dan pangan, temuan ini menunjukkan arah yang sejalan dengan penggunaan *deep learning* untuk membaca ciri visual objek melalui warna, tekstur, bentuk, maupun area objek yang relevan. Beberapa studi terbaru menunjukkan bahwa arsitektur berbasis *convolutional neural network* dapat diterapkan pada berbagai objek visual, seperti penilaian kualitas pangan, deteksi kondisi tanaman, identifikasi penyakit tanaman, dan segmentasi area pertanian berbasis citra [19], [20], [21], [22]. Keterkaitan studi-studi tersebut dengan penelitian ini tidak terletak pada kesamaan objek pepaya secara langsung, melainkan pada pendekatan teknis yang digunakan untuk mengekstraksi pola visual dari citra. Dalam konteks tersebut, penelitian ini memperluas penerapan *computer vision* untuk klasifikasi kematangan buah pepaya dengan menambahkan audit dataset melalui deduplikasi, pengujian validasi silang, evaluasi efisiensi model, analisis kesalahan klasifikasi, serta interpretasi visual menggunakan *Grad-CAM*.

Analisis *confusion matrix* menunjukkan bahwa kesalahan klasifikasi terutama terjadi pada kelas yang berdekatan secara visual. Pada dataset hasil deduplikasi, MobileNetV2 menghasilkan 2 kesalahan prediksi, sedangkan EfficientNet-B0 menghasilkan 3 kesalahan prediksi. Kesalahan tersebut muncul pada citra *partiallymature_021.jpg*, *partiallymature_097.jpg*, dan *unmature_001.jpg*. Pola ini menunjukkan bahwa kesalahan model tidak terjadi secara acak, tetapi berada pada batas visual antarkelas. Citra setengah matang dapat terbaca sebagai matang ketika warna kuning mulai dominan, sedangkan citra mentah dapat terbaca sebagai setengah matang ketika sebagian cirinya mulai menyerupai fase awal pematangan.

Kondisi tersebut menunjukkan bahwa klasifikasi kematangan pepaya berbeda dari klasifikasi objek yang memiliki batas visual yang tegas. Proses pematangan pepaya berlangsung bertahap, sehingga kelas *partially mature* memiliki karakter campuran antara *immature* dan *mature*. Oleh karena itu, kelas *partially mature* menjadi kelas yang paling menantang dalam penelitian ini. Citra yang berhasil diklasifikasikan dengan tepat umumnya memiliki ciri visual yang lebih kuat dan konsisten dengan kelasnya, sedangkan citra yang terlibat dalam kesalahan klasifikasi berada pada fase transisi. Dengan demikian, tantangan utama klasifikasi kematangan pepaya bukan hanya membedakan pepaya matang dan mentah, tetapi juga menentukan batas visual pada kelas peralihan.

Penggunaan *confusion matrix*, analisis sampel citra, dan *Grad-CAM* memberikan pembacaan hasil yang lebih lengkap dibandingkan dengan evaluasi berbasis metrik numerik saja. *Accuracy*, *precision*, *recall*, *F1-score*, dan *AUC-ROC* dapat menunjukkan ringkasan performa model, tetapi belum menjelaskan pola pertukaran kelas dan area citra yang memengaruhi keputusan prediksi. *Confusion matrix* digunakan untuk melihat kecenderungan kesalahan antarkelas, sedangkan analisis sampel citra membantu menjelaskan karakter visual yang menyebabkan kesalahan tersebut. *Grad-CAM* melengkapi analisis ini dengan menampilkan area pada citra yang menjadi perhatian model saat melakukan klasifikasi. Apabila area perhatian model berada pada permukaan buah, keputusan klasifikasi dapat dinilai lebih relevan dengan indikator kematangan. Sebaliknya, apabila perhatian model lebih banyak mengarah pada latar belakang atau area *non-buah*, kondisi tersebut dapat menunjukkan adanya keterbatasan data atau potensi bias visual. Pendekatan ini sejalan dengan konsep *Grad-CAM* yang digunakan untuk menghasilkan penjelasan visual terhadap keputusan model berbasis CNN [13]. Selain itu, penerapan *transfer learning* pada citra pertanian juga terlihat dalam penelitian Sujaka dkk. yang menggunakan MobileNetV3 untuk membedakan tanaman cabai berdasarkan citra [23]. Pada objek buah lainnya, Fadhilatuzzahro dkk. menerapkan EfficientNet-B0 dengan *fine-tuning* untuk klasifikasi penyakit buah jeruk dan menunjukkan bahwa konfigurasi pelatihan serta pembagian data dapat memengaruhi performa model [24]. Dengan demikian, hasil penelitian ini memperkuat bahwa evaluasi model klasifikasi citra perlu memperhatikan performa numerik, pola kesalahan antarkelas, karakteristik visual data, serta interpretasi area perhatian pada model.

Kontribusi utama penelitian ini terhadap pengembangan konsep terletak pada penyusunan kerangka evaluasi yang lebih menyeluruh untuk klasifikasi kematangan buah pepaya. Penelitian ini tidak hanya membandingkan MobileNetV2 dan EfficientNet-B0 berdasarkan akurasi, tetapi juga menggabungkan deduplikasi dataset, *stratified split*, *5-fold cross-validation*, ukuran model, waktu inferensi, *confusion matrix*, analisis sampel citra, dan *Grad-CAM*. Kerangka ini penting karena penelitian klasifikasi citra sering kali menekankan performa akhir model, sementara kualitas dataset, kestabilan hasil, efisiensi komputasi, dan interpretabilitas belum selalu dibahas secara terpadu. Dengan pendekatan ini, performa model dapat dipahami secara lebih utuh dari sisi data, model, hasil, dan interpretasi visual.

Kontribusi berikutnya adalah penegasan bahwa dataset publik perlu diaudit sebelum digunakan dalam eksperimen *deep learning*. Temuan bahwa dataset awal 500 citra berkurang menjadi 246 citra unik menunjukkan bahwa dataset yang tampak besar belum tentu sepenuhnya merepresentasikan variasi data yang sebenarnya. Jika citra duplikat dibiarkan, hasil evaluasi dapat terlihat lebih optimis. Oleh sebab itu, penelitian ini memberikan kontribusi metodologis bahwa deduplikasi perlu ditempatkan sebagai bagian dari validasi performa model, terutama dalam penelitian klasifikasi citra dengan jumlah data yang terbatas.

Secara praktis, hasil penelitian ini dapat menjadi dasar dalam memilih model yang sesuai dengan kebutuhan. MobileNetV2 lebih sesuai untuk kebutuhan yang menekankan ukuran model kecil dan waktu inferensi rendah, sedangkan EfficientNet-B0 lebih sesuai ketika prioritas utamanya adalah kestabilan performa berdasarkan validasi silang. Namun, penelitian ini masih memiliki beberapa keterbatasan. Pertama, dataset hasil deduplikasi hanya berjumlah 246 citra unik dan berasal dari satu sumber dataset, sehingga variasi visual yang diuji masih terbatas. Kedua, ruang lingkup penelitian ini merupakan eksperimen klasifikasi citra berbasis dataset, bukan pengembangan aplikasi *real-time* yang melakukan klasifikasi secara langsung melalui kamera. Prototipe pengujian yang digunakan dalam penelitian ini berfungsi untuk menampilkan hasil prediksi dan interpretasi visual pada skenario data uji, sehingga belum divalidasi pada citra eksternal seperti foto langsung dari kamera, gambar dari internet, atau citra yang diambil dalam kondisi lapangan yang berbeda. Oleh karena itu, pengembangan berikutnya dapat diarahkan pada perluasan dataset dengan variasi pencahayaan, sudut pengambilan, jarak kamera, latar

belakang, serta kondisi permukaan buah yang lebih beragam. Setelah validasi terhadap citra eksternal dilakukan, penelitian selanjutnya juga dapat dikembangkan menjadi sistem klasifikasi yang lebih aplikatif, termasuk pengujian berbasis kamera secara langsung. Selain itu, penelitian berikutnya dapat mempertimbangkan segmentasi area buah sebelum klasifikasi agar model lebih fokus pada permukaan pepaya dan tidak terpengaruh oleh latar belakang citra.

4. Simpulan

Berdasarkan hasil penelitian, deduplikasi dataset terbukti berpengaruh terhadap evaluasi performa model dalam klasifikasi tingkat kematangan buah pepaya berbasis citra RGB. Dataset awal berjumlah 500 citra dan berkurang menjadi 246 citra unik setelah proses deduplikasi menggunakan SHA-256. Kondisi ini menunjukkan bahwa evaluasi pada dataset hasil deduplikasi menjadi lebih ketat karena model tidak lagi memperoleh keuntungan dari pengulangan citra yang identik. Pada skenario tanpa deduplikasi, EfficientNet-B0 memperoleh akurasi tertinggi sebesar 99,00%, sedangkan pada skenario deduplikasi, MobileNetV2 memperoleh akurasi lebih tinggi, yaitu 95,92%, dibandingkan dengan EfficientNet-B0 sebesar 93,88%. Dari sisi efisiensi, MobileNetV2 lebih unggul karena memiliki ukuran model yang lebih kecil, yaitu 20,89 MB, serta waktu inferensi yang lebih rendah dibandingkan EfficientNet-B0. Namun, hasil *5-fold cross-validation* pada dataset hasil deduplikasi menunjukkan bahwa EfficientNet-B0 lebih stabil dengan rata-rata akurasi 98,97%, sedangkan MobileNetV2 memperoleh rata-rata akurasi 94,40%. Dengan demikian, MobileNetV2 lebih sesuai untuk kebutuhan yang menekankan efisiensi komputasi, sedangkan EfficientNet-B0 lebih sesuai ketika prioritas utamanya adalah kestabilan performa berdasarkan validasi silang. Hasil *confusion matrix*, uji citra tunggal, dan *Grad-CAM* juga menunjukkan bahwa prototipe pengujian mampu menampilkan hasil prediksi pada kelas *mature*, *partially mature*, dan *immature*, serta menyediakan interpretasi visual terhadap area citra yang memengaruhi keputusan model. Penelitian ini menyimpulkan bahwa kombinasi deduplikasi dataset, pembagian data terstratifikasi, validasi silang, evaluasi efisiensi, dan interpretasi visual dapat menghasilkan evaluasi model yang lebih representatif. Pengembangan berikutnya dapat diarahkan pada perluasan dataset dengan variasi pencahayaan, sudut pengambilan, jarak kamera, latar belakang, dan kondisi permukaan buah yang lebih beragam, serta pengujian pada citra eksternal agar kemampuan generalisasi model dapat dinilai dengan lebih kuat.

Daftar Referensi

- [1] N. P. Sutrisna *et al.*, "Deteksi Tingkat Kematangan Buah Pepaya menggunakan Model Convolutional Neural Network," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 3, pp. 569–578, Jul. 2024, doi: 10.25126/jtiik.938119.
- [2] M. Rizzo, M. Marcuzzo, A. Zangari, A. Gasparetto, and A. Albarelli, "Fruit ripeness classification: A survey," *Artif. Intell. Agric.*, vol. 7, pp. 44–57, Mar. 2023, doi: 10.1016/j.aiia.2023.02.004.
- [3] K. K. Thyagarajan and G. Kalaiarasi, "A Review on Near-Duplicate Detection of Images Using Computer Vision Techniques," *Arch. Comput. Methods Eng.*, vol. 28, no. 3, pp. 897–916, May 2021, doi: 10.1007/s11831-020-09400-w.
- [4] X. Li, L. Chang, and X. Liu, "CE-Dedup: Cost-Effective Convolutional Neural Nets Training based on Image Deduplication," in *2021 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCLOUD/SocialCom/SustainCom)*, New York City: IEEE, Sep. 2021, pp. 11–18. doi: 10.1109/ISPA-BDCLOUD-SocialCom-SustainCom52081.2021.00017.
- [5] I. E. Tampu, A. Eklund, and N. Haj-Hosseini, "Inflation of test accuracy due to data leakage in deep learning-based classification of OCT images," *Sci. Data*, vol. 9, no. 1, p. 580, Sep. 2022, doi: 10.1038/s41597-022-01618-6.
- [6] Z. K. Simbolon and B. Baniah, "Perbandingan Metode City Block Dan Canberra Distance Dalam Identifikasi Kematangan Buah Pepaya Madu (Callina Papaya) Berdasarkan Warna RGB," *J. Artif. Intell. Softw. Eng.*, vol. 3, no. 1, pp. 14–17, May 2023, doi: 10.30811/jaise.v3i1.4135.
- [7] W. A. Wiharja, T. Al Mudzakir, H. Y. Novita, and J. Indra, "Perbandingan Algoritma Logistic Regression dan K-Nearest Neighbor Dalam Klasifikasi Kematangan Buah Pepaya," *Bull. Comput. Sci. Res.*, vol. 5, no. 4, pp. 353–360, Jun. 2025, doi: 10.47065/bulletincsr.v5i4.550.
- [8] A. R. R. Ivani, A. Z. Kurniadi, A. B. A. Andira, and I. Wahyuni, "Perbandingan VGG16 dan

- MobileNetV2 untuk Klasifikasi Tingkat Kematangan Buah Apel,” *J. Sist. Komput. Asia*, vol. 3, no. 1, pp. 30–39, 2025, doi: 10.32815/jiskomsia.v3i1.136.
- [9] R. R. Santosa, A. Sasikirana, Z. L. Ramadhan, Y. P. B. Satria, and A. Nugroho, “Perbandingan Kinerja MobileNetV2 dan ResNet50V2 dalam Klasifikasi Tingkat Kematangan Tomat,” *Indones. J. Sci. Technol. Humanit.*, vol. 3, no. 1, pp. 49–56, Jun. 2025, doi: 10.60076/ijstech.v3i1.1335.
- [10] M. Wildan Fitriani, W. Widayani, H. Sismoro, N. Aini, and D. Nurani, “Prediksi Perbandingan Kinerja MobileNetV2 dan EfficientNetB0 dalam Prediksi Kematangan Buah Jaboticaba,” *J. Autom. Comput. Inf. Syst.*, vol. 6, no. 1, pp. 233–246, Apr. 2026, doi: 10.47134/jacis.v6i1.177.
- [11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City: IEEE, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.
- [12] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” in *Proceedings of the 36th International Conference on Machine Learning*, Long Beach: Proceedings of Machine Learning Research, May 2019, pp. 6105–6114. doi: 10.48550/arXiv.1905.11946.
- [13] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, Feb. 2020, doi: 10.1007/s11263-019-01228-7.
- [14] R. Shah and A. K. Shrivastava, “Semantic-Aware Image Deduplication: Leveraging Object Recognition for Enhanced Accuracy,” *Int. J. Intell. Eng. Syst.*, vol. 18, no. 7, pp. 116–130, Aug. 2025, doi: 10.22266/ijies2025.0831.09.
- [15] N. Nurmallasari, Y. A. Setiawan, W. Astuti, M. R. R. Saelan, S. Masturoh, and T. Haryanti, “Classification for Papaya Fruit Maturity Level With Convolutional Neural Network,” *J. Ris. Inform.*, vol. 5, no. 3, pp. 331–338, Jun. 2023, doi: 10.34288/jri.v5i3.225.
- [16] N. Amanda, R. Wulanningrum, and J. Sahertian, “Penerapan MobileNetV2 untuk Klasifikasi Kondisi Buah Alpukat Mentega,” in *Prosiding SEMNAS INOTEK*, Kediri, 2025, pp. 575–582. doi: 10.29407/bp5fen81.
- [17] C. Nastitie and I. Handayani, “Perbandingan Metode MobileNetV2 dan EfficientNetB0 untuk Klasifikasi Jenis Pisang,” *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 8, no. 6, pp. 1990–1999, Nov. 2025, doi: 10.31539/0240cx09.
- [18] A. C. Hartono and A. R. Muslikh, “Penerapan Transfer Learning MobileNetV2 pada Klasifikasi Citra Jenis Buah-Buahan,” *J. Inf. Syst. Appl. Dev.*, vol. 3, no. 2, pp. 102–110, 2025, doi: 10.26905/jisad.v3i2.16187.
- [19] N. R. Fauza, A. S. K. Aprilia, and I. Wahyuni, “MobileNetV2 untuk Deteksi Kualitas Tempe di Sentra Industri Tempe Sanan,” *J. Sist. Komput. Asia*, vol. 3, no. 1, pp. 15–29, 2025, doi: 10.32815/jiskomsia.v3i1.134.
- [20] D. N. Ananda, S. D. Dalimajun, Y. R. Bere, G. R. Talo, and I. Wahyuni, “Perbandingan Kinerja DenseNet dan MobileNetV2 dalam Pengembangan Model Deep Learning untuk Deteksi Tanaman Kacang Tanah Sehat dan Terserang Penyakit,” *J. Sist. Komput. Asia*, vol. 3, no. 1, pp. 63–71, 2025, doi: 10.32815/jiskomsia.v3i1.148.
- [21] W. Febrianto, D. Kurniawan, and I. Wahyuni, “Perbandingan Kinerja Metode VGG-16 dan ResNet50 untuk Deteksi Penyakit pada Tanaman Kubis di Nongkojajar Kabupaten Pasuruan,” *J. Sist. Komput. Asia*, vol. 3, no. 2, pp. 1–9, 2025, doi: 10.32815/jiskomsia.v3i2.144.
- [22] I. Wahyuni, C.-C. Chang, W.-J. Wang, and D. Liang, “Input Variability Analysis of Aerial Imagery in Rice Semantic Segmentation Using UNet-VGG16,” in *2025 9th International Conference On Electrical, Electronics And Information Engineering (ICEEIE)*, IEEE, Sep. 2025, pp. 1–6. doi: 10.1109/ICEEIE66203.2025.11252106.
- [23] T. T. Sujaka, I. N. Switrayana, and I. M. H. Fillah, “Pengaplikasian Convolutional Neural Network (MobileNetV3) Memanfaatkan Transfer Learning Untuk Membedakan Tanaman Cabai Berasal Dari Genus Capsicum Annuum,” *Build. Informatics, Technol. Sci.*, vol. 7, no. 3, pp. 1761–1774, 2025, doi: 10.47065/bits.v7i3.8740.
- [24] M. Fadhilatuzzahro, A. Fazrin, Z. I. Arishandy, T. A. Sukoco, and H. Maulana, “Optimalisasi Model CNN EfficientNet-B0 dengan Fine Tuning untuk Klasifikasi Penyakit Buah Jeruk,” *J. Ilm. Teknol. Inf. dan Robot.*, vol. 7, no. 1, pp. 23–30, 2025, doi: 10.33005/jifti.v7i1.162.