BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394

Implementasi Convolutional Neural Network untuk Klasifikasi Kualitas Sarang Burung Walet Menggunakan MobileNetV2 dan EfficientNet-B3

Implementation of Convolutional Neural Network for Swiftlet Nest Quality Classification Using MobileNetV2 and EfficientNet-B3

I Komang Aditya*, Muhammad Fachrie

Fakultas Sains dan Teknologi, Program Studi Informatika, Universitas Teknologi Yogyakarta, Sleman, Indonesia
Email: [1,*]ikomangadtya04@gmail.com, [2]muhammad.fachrie@uty.ac.id
Email Penulis Korespondensi: ikomangadtya04@gmail.com
Received: **03/10/2026**, Revised: **15/10/2026**, Accepted: **29/10/2026**, Published: **31/10/2026**

Abstrak−Kualitas sarang burung walet menjadi salah satu faktor utama yang memengaruhi harga jualnya, mengingat komoditas ini memiliki nilai ekonomi tinggi. Penilaian kualitas yang masih dilakukan secara visual dan manual rentan terhadap subjektivitas dan tidak konsisten. Karena itu, diperlukan sistem yang dapat mengklasifikasikan kualitas sarang burung walet secara otomatis dan objektif. Penelitian ini mengembangkan sistem klasifikasi berbasis *Convolutional Neural Network* (CNN) dengan membandingkan dua arsitektur, yaitu MobileNetV2 dan EfficientNet-B3 melalui *transfer learning* dengan strategi pelatihan dua fase. *Dataset* terdiri dari 1.200 citra sarang burung walet pada tiga kategori kualitas, yaitu tinggi, sedang, dan rendah (400 citra per kategori) berupa data primer, kecuali kategori tinggi yang terdiri dari 300 data primer dan 100 data sekunder. Data dibagi menjadi 70% data latih, 15% data validasi, dan 15% data uji, kemudian augmentasi diterapkan hanya pada data latih pada kedua arsitektur sehingga data validasi dan data uji hanya berisi citra asli tanpa citra induk yang sama antar subset. Sistem diimplementasikan dalam *website* bernama Qualet yang memungkinkan pengguna mengunggah gambar sarang dan memperoleh hasil klasifikasi. Hasil pengujian menunjukkan MobileNetV2 memperoleh akurasi uji 78,89% dengan *loss* 0,6767, sedangkan EfficientNet-B3 memperoleh akurasi uji 80,56% dengan *loss* 0,6954. Kesalahan klasifikasi pada kedua model terutama terjadi pada kategori sedang, yang diduga karena kategori ini merupakan kelas peralihan dengan batas visual yang kurang tegas serta jumlah citra asli yang terbatas. Kedua arsitektur menunjukkan performa setara, dengan selisih akurasi 1,67 poin persentase atau setara 3 dari 180 citra uji. Penelitian ini berkontribusi pada perbandingan dua arsitektur CNN dan penyediaan *website* Qualet.

Kata Kunci: Convolutional Neural Network; EfficientNet-B3; MobileNetV2; Sarang Burung Walet; Transfer Learning

Abstract−The quality of swiftlet nests is a main factor affecting their selling price, given its high economic value. Quality assessment still performed visually and manually is prone to subjectivity and inconsistency. Therefore, a system that can classify swiftlet nest quality automatically and objectively is needed. This study develops a Convolutional Neural Network (CNN)-based classification system by comparing MobileNetV2 and EfficientNet-B3 through transfer learning with a two-phase training strategy. The dataset consists of 1,200 swiftlet nest images in three quality categories, namely high, medium, and low (400 images per category), all primary data except the high category, which comprises 300 primary and 100 secondary images. The data were split into 70% training, 15% validation, and 15% test sets, after which augmentation was applied only to the training data for both architectures, so the validation and test sets contain only original images with no shared parent images across subsets. The system is implemented as a website, Qualet, that lets users upload nest images and obtain classification results. MobileNetV2 achieved a test accuracy of 78.89% with a loss of 0.6767, while EfficientNet-B3 achieved 80.56% with a loss of 0.6954. Misclassifications in both models occurred mainly in the medium category, presumably because it is a transitional class with less distinct visual boundaries and a limited number of original images. Both architectures showed comparable performance, with an accuracy difference of 1.67 percentage points, equivalent to 3 of the 180 test images. This study contributes a comparison of two CNN architectures and the Qualet website.

Keywords: Convolutional Neural Network; EfficientNet-B3; MobileNetV2; Swiftlet Nests; Transfer Learning

1. PENDAHULUAN

Burung walet merupakan salah satu jenis burung yang banyak dijumpai di wilayah pesisir dan kawasan permukiman dengan kondisi lingkungan yang lembap, seperti gua maupun bangunan. Burung ini menghasilkan sarang yang terbentuk dari air liur yang mengeras dan memiliki nilai ekonomi tinggi karena banyak dimanfaatkan sebagai bahan dalam produk makanan, obat-obatan, kosmetik, serta perawatan kulit [1]. Indonesia termasuk salah satu negara penghasil sarang burung walet utama di Asia Tenggara [2]. Sarang dengan kualitas tinggi umumnya memiliki warna putih, tingkat kebersihan yang tinggi dengan sedikit atau tanpa bulu, bentuk mangkok, ukuran yang utuh, serta kadar air yang sesuai [3]. Karakteristik sarang burung walet juga dapat diamati melalui aspek fisik dan kimia, seperti warna, bentuk, morfologi, dan kadar air [4]. Oleh karena itu, diperlukan proses penilaian yang mampu mengidentifikasi kualitas sarang berdasarkan karakteristik tersebut secara tepat dan konsisten.

Pada praktiknya, penilaian kualitas sarang burung walet umumnya masih dilakukan secara manual melalui pengamatan terhadap karakteristik fisik sarang [5]. Proses penilaian tersebut dapat menghasilkan perbedaan hasil karena dipengaruhi oleh pengalaman, pengetahuan, dan ketelitian masing-masing penilai [6]. Penilaian kualitas juga menjadi cukup kompleks karena melibatkan beberapa karakteristik, seperti warna, kebersihan, bentuk, dan ukuran sarang yang digunakan sebagai dasar dalam proses *grading*[3]. Variasi dalam menilai karakteristik tersebut dapat menyebabkan hasil penilaian yang kurang seragam dan menyulitkan penerapan standar kualitas secara konsisten [3], [6]. Kondisi tersebut menunjukkan bahwa proses penilaian secara manual masih memiliki keterbatasan dalam menghasilkan penilaian yang

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

objektif dan konsisten. Oleh karena itu, diperlukan suatu pendekatan yang dapat membantu mengurangi subjektivitas dalam proses penilaian kualitas sarang burung walet.

Permasalahan dalam proses penilaian kualitas sarang burung walet mendorong pemanfaatan berbagai metode komputasi untuk membantu proses klasifikasi dan penentuan kualitas. Parhusip dkk. (2024) menerapkan *deep learning* menggunakan arsitektur SSD MobileNetV2 untuk mengklasifikasikan sarang burung walet ke dalam tujuh kategori berdasarkan intensitas bulu dan memperoleh akurasi sebesar 85% [7]. Indrajaya dkk. (2024) mengembangkan arsitektur MobileNetV2-D dengan menggunakan kombinasi empat kamera untuk mengklasifikasikan sarang ke dalam tujuh kelas berdasarkan intensitas bulu dan memperoleh akurasi sebesar 94,07% pada data pengujian [8]. Selain itu, Ahmad dan Supatman (2025) menggunakan *Convolutional Neural Network* (CNN) dengan arsitektur DenseNet-121 untuk mengklasifikasikan kebersihan sarang berdasarkan 900 citra ke dalam tiga kategori dan memperoleh akurasi sebesar 84% [6]. Agus dkk. (2025) menggunakan Pendekatan *Fuzzy Tsukamoto* untuk menentukan kualitas sarang berdasarkan beberapa variabel karakteristik, seperti warna, bulu, jenis, kondisi, dan kadar air, dengan tingkat akurasi sebesar 97% [1]. Sementara itu, Medina dan Rahayu (2026) menerapkan metode *Support Vector Machine* (SVM) untuk memprediksi hasil produksi sarang burung walet berdasarkan data historis dengan variabel jumlah sarang dan jumlah patahan [9]. Secara khusus, pendekatan *transfer learning* juga telah diterapkan untuk klasifikasi kualitas sarang burung walet. Ilmiati dkk. (2026) mengklasifikasikan kualitas sarang berdasarkan SNI 8998:2021 dengan membandingkan MobileNetV2, ResNet50, dan YOLOv8n-cls pada tiga kelas kualitas, yaitu baik, sedang, dan buruk. Hasil penelitian menunjukkan bahwa YOLOv8n-cls memperoleh akurasi 99,5%, sedangkan MobileNetV2 mencapai 98,37% dengan efisiensi komputasi yang baik [10]. Berbagai penelitian tersebut menunjukkan bahwa metode kecerdasan buatan telah digunakan dalam pengelolaan sarang burung walet, baik untuk klasifikasi karakteristik, penentuan kualitas, maupun prediksi produksi.

Perkembangan kecerdasan buatan (*artificial intelligence*/AI) dalam beberapa tahun terakhir turut meningkatkan perhatian terhadap *machine learning* dan *deep learning* [11], yang kemudian memperluas pemakaian CNN untuk tugas visi komputer (*computer vision*), termasuk klasifikasi citra [12]. CNN memanfaatkan struktur spasial pada citra sehingga banyak digunakan di bidang tersebut [13]. Salah satu pendekatan yang banyak digunakan adalah *transfer learning*, yaitu memanfaatkan pengetahuan model dari tugas sebelumnya untuk tugas baru, sehingga mengurangi kebutuhan data pelatihan dan biaya komputasi [14]. Isnan dkk. (2023) menerapkan *transfer learning* pada klasifikasi citra pertanian dengan berbagai arsitektur CNN dan augmentasi data untuk mengurangi risiko *overfitting* [15]. Halloum dan Zahraouy (2025) menguji empat pengaturan praproses pada klasifikasi citra medis otak, yaitu dengan dan tanpa *Contrast Limited Adaptive Histogram Equalization* (CLAHE) serta dengan dan tanpa augmentasi data, dan menemukan bahwa kombinasi CLAHE dan augmentasi data memberikan kinerja terbaik pada pendekatan *transfer learning* [16].

Perbedaan arsitektur CNN dapat memengaruhi hasil klasifikasi citra. Penelitian pada penilaian kualitas buah menunjukkan bahwa *transfer learning* MobileNetV2 dapat mempercepat konvergensi dan memberikan kinerja yang lebih baik dibandingkan pelatihan dari awal, serta memiliki ukuran model yang ringkas dan efisien [17]. Pengembangan MobileNetV2 dengan modul Res-Inception dan EMA pada klasifikasi produk pertanian juga meningkatkan akurasi sebesar 1,86% dibandingkan model aslinya [18]. Sementara itu penelitian identifikasi cacat butir beras yang membandingkan beberapa arsitektur, termasuk MobileNetV2 dan EfficientNet-B0, menunjukkan adanya perbedaan kinerja antar arsitektur [19]. Penerapan CNN berbasis EfficientNet-B3 untuk mengklasifikasikan fertilitas telur bebek Cihateup menunjukkan bahwa peningkatan kualitas citra dapat meningkatkan akurasi model dari 49% menjadi 77% dan 80% pada dua tahap inkubasi [20]. Perbandingan MobileNetV2 dan EfficientNet-B3 pada klasifikasi penyakit daun jagung menunjukkan adanya perbedaan kinerja, dengan akurasi pengujian masing-masing sebesar 96,50% dan 93,00% [21]. Temuan tersebut diperkuat oleh penelitian lain pada objek yang sama, yang menunjukkan bahwa kedua arsitektur juga berbeda dalam hal waktu prediksi, ukuran model, dan kebutuhan komputasi secara keseluruhan [22].

Berbagai karakteristik visual, seperti bulu, warna, dan kebersihan sarang, telah banyak dimanfaatkan dalam penelitian terdahulu untuk mengklasifikasikan kualitas sarang burung walet [6]. Namun, penelitian yang secara khusus membandingkan MobileNetV2 dan EfficientNet-B3 untuk mengklasifikasikan kualitas sarang burung walet berdasarkan citra pada tiga kategori kualitas masih terbatas. Selain itu, penelitian pada objek lain menunjukkan adanya perbedaan kinerja ketika menggunakan arsitektur CNN yang berbeda, seperti pada identifikasi cacat butir beras [19] maupun deteksi citra katarak yang menunjukkan perbedaan kinerja antara MobileNetV2 dan EfficientNet-B4 [23]. Variasi hasil pada berbagai objek tersebut mengindikasikan bahwa kinerja arsitektur CNN sangat bergantung pada karakteristik data yang diolah, sehingga arsitektur yang paling optimal untuk mengenali karakteristik visual sarang burung walet belum dapat dipastikan. Berdasarkan kesenjangan tersebut, penelitian ini dirancang untuk mengembangkan sistem klasifikasi kualitas sarang burung walet berbasis CNN dengan membandingkan arsitektur MobileNetV2 dan EfficientNet-B3 pada tiga kategori kualitas, yaitu tinggi, sedang, dan rendah. Kedua arsitektur diuji menggunakan *dataset*, pembagian data, dan metrik evaluasi yang sama guna mengungkap perbedaan kemampuannya dalam mengenali karakteristik visual sarang, seperti bulu, warna, dan kebersihan. Penelitian ini berfokus pada augmentasi data untuk meningkatkan variasi data latih, sedangkan penggunaan CLAHE pada klasifikasi kualitas sarang burung walet tidak dieksplorasi dan dapat menjadi arah penelitian selanjutnya.

Penelitian ini bertujuan untuk membandingkan kinerja arsitektur MobileNetV2 dan EfficientNet-B3 dalam mengklasifikasikan kualitas sarang burung walet ke dalam tiga kategori, yaitu tinggi, sedang, dan rendah, serta mengimplementasikannya ke dalam *website*. Perbandingan dilakukan menggunakan *dataset* dan pembagian data yang sama, dengan augmentasi data yang hanya diterapkan pada data latih, serta dievaluasi menggunakan metrik klasifikasi yang identik pada data uji. Hasil penelitian diharapkan menjadi dasar dalam menentukan model yang sesuai untuk

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

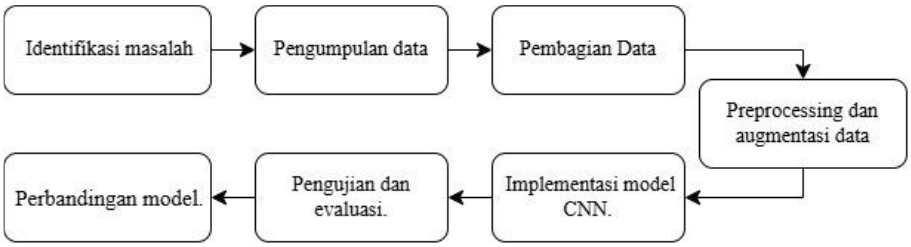
BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394

klasifikasi kualitas sarang burung walet secara lebih objektif dan konsisten. Kontribusi penelitian ini adalah penyajian perbandingan kinerja MobileNetV2 dan EfficientNet-B3 pada klasifikasi tiga kelas kualitas sarang burung walet berdasarkan foto ponsel dalam kondisi pemotretan lapangan, serta penyediaan *website* Qualet sebagai alat bantu penilaian awal.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini terdiri atas tujuh tahap yang ditampilkan pada Gambar 1, dimulai dari identifikasi masalah dan diakhiri dengan perbandingan model. Gambar 1 menunjukkan alur yang berurutan, dengan pembagian *dataset* dilakukan sebelum *preprocessing* dan augmentasi agar tidak terjadi kebocoran data.



Gambar 1. Tahapan penelitian

Penelitian dimulai dari identifikasi masalah, yakni proses penentuan kualitas sarang burung walet yang hingga kini masih mengandalkan penilaian manual sehingga hasilnya sering berbeda antar penilai. Persoalan tersebut kemudian dikaji lebih jauh lewat penelusuran literatur tentang studi terdahulu, pemanfaatan CNN dalam klasifikasi citra, serta ciri khas kedua arsitektur yang akan dibandingkan. Setelah itu, dikumpulkan foto sarang burung walet sebagai data primer yang dilengkapi data sekunder pada kategori tinggi, yang mewakili tiga kelas kualitas, yaitu tinggi, sedang, dan rendah. *Dataset* lalu dibagi secara berstrata menjadi data latih, data validasi, dan data uji dengan perbandingan 70:15:15, dan pembagian ini dilakukan sebelum proses lainnya. Pada tahap *preprocessing* dan augmentasi, seluruh citra melalui *preprocessing*, sedangkan augmentasi hanya diterapkan pada data latih agar data validasi dan uji tetap berupa citra asli dan kebocoran data (*data leakage*) dapat dicegah. MobileNetV2 dan EfficientNet-B3 kemudian dilatih menggunakan data latih dan dipantau dengan data validasi melalui *transfer learning* dua fase, yaitu *feature extraction* dan *fine-tuning*. Kedua model diuji pada data uji menggunakan akurasi, *loss*, *precision*, *recall*, *F1-score*, dan *confusion matrix*, lalu hasilnya dibandingkan untuk menentukan arsitektur yang lebih sesuai. Model hasil pelatihan selanjutnya diimplementasikan dalam *website* Qualet.

2.2 Pengumpulan dan Pembagian Dataset

Dataset terdiri atas 1.200 citra sarang burung walet pada tiga kelas, yaitu Tinggi, Sedang, dan Rendah, masing-masing 400 citra. Data primer dikumpulkan melalui pemotretan langsung di peternakan walet di Desa Bukit Makmur, Kecamatan Kaliorang, Kabupaten Kutai Timur, Kalimantan Timur, menggunakan kamera ponsel pada latar yang beragam dan pencahayaan yang bervariasi. Kelas kualitas ditentukan langsung oleh peternak setempat melalui penilaian visual dan lisan berdasarkan pengalaman, tanpa dokumen tertulis, dengan mengacu pada warna, kebersihan bulu, bentuk, dan ukuran. Data sekunder berasal dari dataset Kaggle "Swiftlet's Nest Shape (3 Classes)" oleh IOT Lab Sensor [24]. Sebanyak 100 citra kelas bentuk mangkok dipakai untuk melengkapi kelas Tinggi (300 primer dan 100 sekunder), sedangkan dua kelas lainnya seluruhnya primer. Contoh citra ditampilkan pada Gambar 2.



Tinggi Sedang Rendah

Gambar 2. Contoh citra sarang burung walet

Gambar 2 menunjukkan variasi warna, kebersihan bulu, bentuk, dan ukuran sarang walet yang menjadi dasar penilaian peternak. Perbedaan latar belakang dan pencahayaan mencerminkan kondisi nyata, tetapi juga meningkatkan kompleksitas klasifikasi. Selain itu, penentuan label berdasarkan penilaian lisan tanpa kriteria tertulis menyebabkan batas antar kelas, khususnya kelas Sedang, cenderung subjektif.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394

Dataset dibagi secara berstrata dengan rasio 70:15:15 sebelum augmentasi, sehingga tiap kelas memiliki 280 citra latih, 60 validasi, dan 60 uji. Augmentasi hanya diterapkan pada data latih sehingga mencapai 800 citra per kelas (termasuk 280 citra asli), sedangkan data validasi dan uji tetap citra asli. Rincian pembagian disajikan pada Tabel 1.

Tabel 1. Pembagian Data

Kelas	Data Latih Awal	Data Latih Augmentasi	Data Validasi	Data Uji	Total Awal
Tinggi	280	800	60	60	400
Sedang	280	800	60	60	400
Rendah	280	800	60	60	400
Total	840	2.400	180	180	1.200

Tabel 1 menunjukkan jumlah data yang seimbang pada setiap kelas di semua subset, sehingga tidak ada kelas yang mendominasi pelatihan dan metrik rata-rata *macro* dapat ditafsirkan secara adil. Augmentasi menaikkan data latih dari 840 menjadi 2.400 citra, sedangkan data validasi dan uji tetap berupa citra asli sehingga mengukur kemampuan generalisasi yang sebenarnya.

2.3 *Preprocessing* dan Augmentasi Data

Sebelum pelatihan, ukuran setiap citra disesuaikan dengan input masing-masing arsitektur, yaitu 224 × 224 piksel untuk MobileNetV2 dan 300 × 300 piksel untuk EfficientNet-B3, lalu diproses menggunakan fungsi *preprocess_input* bawaan masing-masing model. Augmentasi hanya dikenakan pada data latih dan dilakukan dalam dua tahap. Tahap pertama adalah augmentasi *offline* setelah pembagian data, menggunakan *rotation* hingga 40°, *width shift* dan *height shift* 20%, *shear* 20%, *zoom* hingga 30%, *horizontal flip*, dan *brightness* 0,7-1,3 dengan *fill mode nearest*. Tahap ini menambah 520 citra pada tiap kelas sehingga data latih menjadi 800 citra per kelas (termasuk 280 citra asli), atau 2.400 citra seluruhnya, dengan contoh pada Gambar 3.



Gambar 3. Hasil augmentasi offline citra sarang burung walet

Gambar 3 menunjukkan bahwa transformasi mengubah posisi, sudut, skala, dan pencahayaan sarang, sehingga model tidak bergantung pada kondisi pengambilan gambar tertentu.

Tahap kedua adalah augmentasi *real-time* melalui ImageDataGenerator selama proses pelatihan, dengan parameter yang dirinci pada Tabel 2. Augmentasi ini tidak menambah jumlah citra, melainkan menghasilkan variasi acak pada setiap *epoch*. Berbeda dengan augmentasi *offline* yang sama untuk kedua arsitektur, parameter augmentasi *real-time* disesuaikan pada masing-masing arsitektur. Sementara itu, data validasi dan data uji tidak diaugmentasi dan hanya melalui *preprocessing*.

Tabel 2. Parameter augmentasi real-time

Parameter	MobileNetV2	EfficientNet-B3
Rotation	22°	15°
Width/height shift	15%	5%
Shear	10%	-
Zoom	25%	10%

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2615
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394

Parameter	MobileNetV2	EfficientNet-B3
Horizontal flip	Ya	Ya
Vertical flip	Ya	Tidak
Brightness	0,8-1,2	-
Fill mode	reflect	reflect

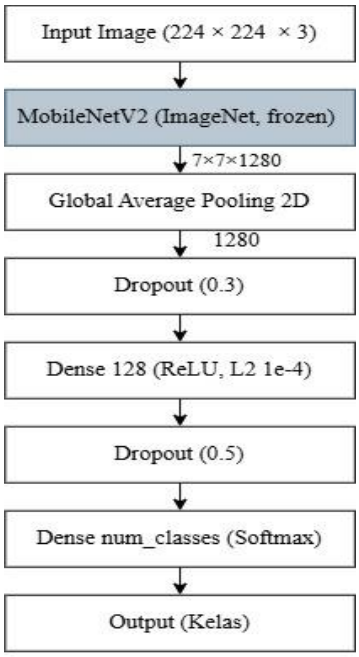
Tabel 2 menunjukkan augmentasi *real-time* yang lebih agresif pada MobileNetV2 dibandingkan EfficientNet-B3, misalnya rotasi 22° dan 15°, serta *shear* dan *brightness* yang hanya dipakai pada MobileNetV2. Akibatnya, perbandingan akhir mencerminkan konfigurasi masing-masing model, bukan konfigurasi yang identik.

2.4 Arsitektur dan Pelatihan Model

Penelitian ini membandingkan MobileNetV2 dan EfficientNet-B3 melalui *transfer learning* dengan bobot awal ImageNet. Susunan *classification head* disesuaikan untuk masing-masing model.

2.4.1 MobileNetV2

Gambar 4 memperlihatkan struktur MobileNetV2 yang dipakai dalam penelitian ini.



Gambar 4. Model MobileNetV2

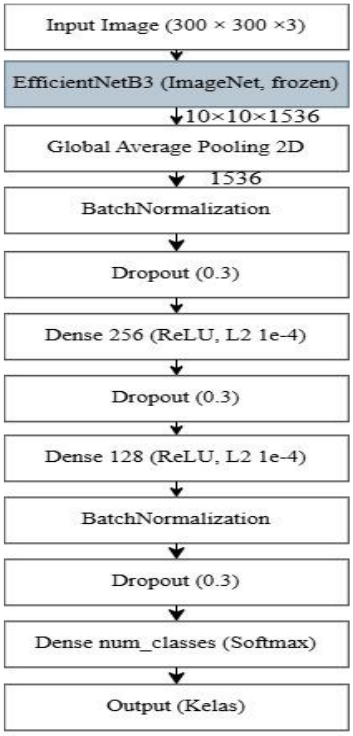
Gambar 4 menunjukkan dua bagian model, yaitu *base model* ImageNet sebagai ekstraktor fitur dan *classification head* yang memetakan fitur tersebut ke tiga kelas.MobileNetV2 merupakan CNN ringan yang dibangun dari blok *bottleneck* berbasis *inverted residual* dan *linear bottleneck*, dengan *depthwise separable convolution* untuk mengurangi beban komputasi [17]. Pada tiap blok, jumlah kanal diperbesar terlebih dahulu melalui konvolusi 1×1, fitur diambil dengan konvolusi *depthwise* 3×3, kemudian kanal dipadatkan kembali lewat konvolusi 1×1. Koneksi residual hanya dipakai ketika dimensi masukan dan keluaran blok sama. Masukan model berupa citra 224 × 224 piksel hasil *preprocessing*. Lapisan klasifikasi bawaan tidak dipakai (*include_top = False*) dan digantikan oleh *classification head* yang terdiri atas *Global Average Pooling*, *Dropout* 0,3, lapisan *Dense* 128 neuron dengan aktivasi ReLU dan regularisasi L2 sebesar 10^{-4}, *Dropout* 0,5, serta lapisan *Dense* 3 neuron dengan aktivasi *softmax*. Dengan susunan ini, keluaran model terdiri atas tiga kelas, yaitu Tinggi, Sedang, dan Rendah.

2.4.2 EfficientNet-B3

EfficientNet-B3 dibangun dari blok MBConv (*mobile inverted bottleneck*) dan mengatur keseimbangan kedalaman, lebar, serta resolusi jaringan melalui *compound scaling* [21]. Masukan model berupa citra 300 × 300 piksel hasil *preprocessing*, yang kemudian diproses oleh rangkaian blok MBConv untuk mengekstraksi fitur. Lapisan klasifikasi bawaan tidak dipakai (*include_top = False*) dan digantikan oleh *classification head* yang terdiri atas *Global Average Pooling*, *Batch Normalization*, *Dropout* 0,3, lapisan *Dense* 256 neuron dengan aktivasi ReLU dan regularisasi L2 sebesar 10^{-4}, *Dropout* 0,3, lapisan *Dense* 128 neuron dengan aktivasi ReLU dan regularisasi L2 sebesar 10^{-4}, *Batch Normalization*, *Dropout* 0,3, serta lapisan *Dense* 3 neuron dengan aktivasi *softmax*. Dengan susunan ini, keluaran model terdiri atas tiga kelas, yaitu Tinggi, Sedang, dan Rendah. Struktur EfficientNet-B3 yang dipakai dalam penelitian ini ditampilkan pada Gambar 5.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2616
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394



Gambar 5. Model EfficientNet-B3

Dibandingkan Gambar 4, Gambar 5 menunjukkan susunan classification head yang lebih kompleks serta ukuran masukan citra yang lebih besar. Perbedaan tersebut memungkinkan model memproses fitur yang lebih beragam, tetapi juga berpotensi meningkatkan kebutuhan komputasi selama pelatihan dan prediksi.

2.4.3 Strategi Pelatihan Dua Fase

Kedua model dilatih dalam dua fase. Pada fase pertama (*feature extraction*), seluruh lapisan model dasar dibekukan dan hanya *classification head* yang dilatih dengan *learning rate* 0,001. Pada fase kedua (*fine-tuning*), sebagian lapisan atas model dasar dibuka, yaitu lapisan mulai indeks ke-100 pada MobileNetV2 serta blok 5, 6, 7, dan *top_conv* pada EfficientNet-B3, sedangkan *Batch Normalization* bawaan model dasar tetap dibekukan. Pelatihan menggunakan *ReduceLROnPlateau* (faktor 0,5, *patience* 4), *EarlyStopping* (*patience* 8), dan *ModelCheckpoint* yang memantau *val_loss*, sehingga bobot terbaik dipilih berdasarkan data validasi. Pengaturan pelatihan dirangkum pada Tabel 3.

Tabel 3. Hiperparameter pelatihan

Parameter	MobileNetV2	EfficientNet-B3
Optimizer	Adam	Adam
Loss	*Categorical crossentropy* (*label smoothing* 0,1)	*Categorical crossentropy* (*label smoothing* 0,05
Batch size	32	16
Learning rate fase 1	0,001	0,001
Learning rate fase 2	0,00003	0,00001
Epoch fase 1 / fase 2	20/40	20/40

Tabel 3 menunjukkan bahwa kedua model menerapkan *optimizer*, *learning rate* pada fase pertama, dan jumlah *epoch* yang sama, sedangkan nilai *label smoothing*, *batch size*, dan *learning rate* pada fase kedua berbeda. Penggunaan *learning rate* yang lebih rendah saat *fine-tuning* bertujuan menyesuaikan bobot pralatih ImageNet secara perlahan agar fitur yang telah dipelajari tidak mengalami perubahan besar. Sementara itu, mekanisme *EarlyStopping* dapat menghentikan pelatihan sebelum mencapai batas maksimum *epoch* apabila kinerja pada data validasi tidak lagi meningkat.

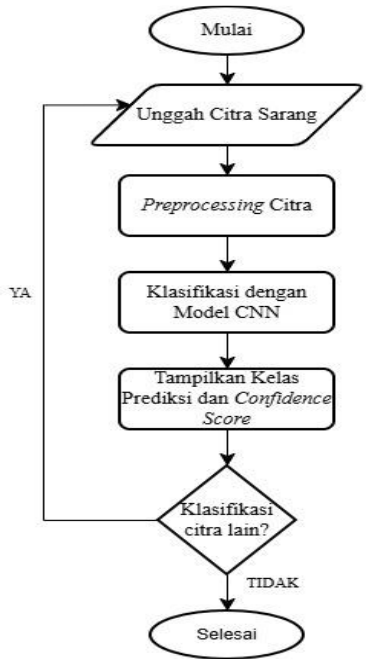
2.5 Perancangan Sistem dan Evaluasi Model

2.5.1 Perancangan Sistem

Model hasil pelatihan diimplementasikan dalam *website* Qualet menggunakan model MobileNetV2 dan EfficientNet-B3 untuk mengklasifikasikan kualitas sarang burung walet. Alur proses klasifikasi ditunjukkan pada Gambar 6.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2617
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394



Gambar 6. Alur klasifikasi pada *website* Qualet

Gambar 6 memperlihatkan tahapan klasifikasi pada website Qualet, yang diawali dengan pengunggahan citra sarang burung walet oleh pengguna. Citra tersebut kemudian melalui proses preprocessing yang disesuaikan dengan tahap pelatihan sebelum diproses oleh model CNN untuk menentukan kategori kualitas, yaitu Tinggi, Sedang, atau Rendah, beserta nilai confidence score. Hasil prediksi ditampilkan pada halaman website, sehingga pengguna dapat melakukan klasifikasi terhadap citra lainnya.

2.5.2 Evaluasi Model

Pengujian dilakukan pada 180 citra data uji asli (60 citra per kelas) tanpa augmentasi, dengan data uji yang sama untuk kedua model. Kinerja model diukur menggunakan *accuracy*, *precision*, *recall*, dan *F1-score* yang dihitung dari *confusion matrix*, serta nilai *loss* pada data uji. *Precision*, *recall*, dan *F1-score* dihitung pada masing-masing kelas dengan membandingkan satu kelas terhadap kelas lainnya, lalu dirata-ratakan secara *macro*.
Accuracy menunjukkan persentase prediksi yang benar terhadap seluruh data dan dihitung menggunakan Persamaan (1).

$$Accuracy = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \quad (1)$$

Precision mengukur ketepatan prediksi suatu kelas dan dihitung menggunakan Persamaan (2).

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \quad (2)$$

Recall mengukur kemampuan model dalam mengenali data pada suatu kelas dan dihitung menggunakan Persamaan (3).

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad (3)$$

F1-score mengukur keseimbangan antara *precision* dan *recall* dan dihitung menggunakan Persamaan (4).

$$F1\text{-}score = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (4)$$

TP, TN, FP, dan FN masing-masing adalah *true positive*, *true negative*, *false positive*, dan *false negative* yang dihitung untuk setiap kelas, sedangkan *accuracy* keseluruhan diperoleh dari jumlah prediksi benar dibagi jumlah citra uji. Nilai *loss* dihitung menggunakan *categorical cross-entropy* dengan *label smoothing* sesuai pengaturan masing-masing model, yaitu 0,1 pada MobileNetV2 dan 0,05 pada EfficientNet-B3.

Kedua model dibandingkan berdasarkan metrik yang sama pada data uji yang sama, serta pola kesalahan pada *confusion matrix*. Karena data uji berjumlah 180 citra, selisih satu citra setara dengan 0,56 poin persentase, sehingga selisih akurasi yang sangat kecil dinyatakan sebagai kinerja yang setara.

3. HASIL DAN PEMBAHASAN

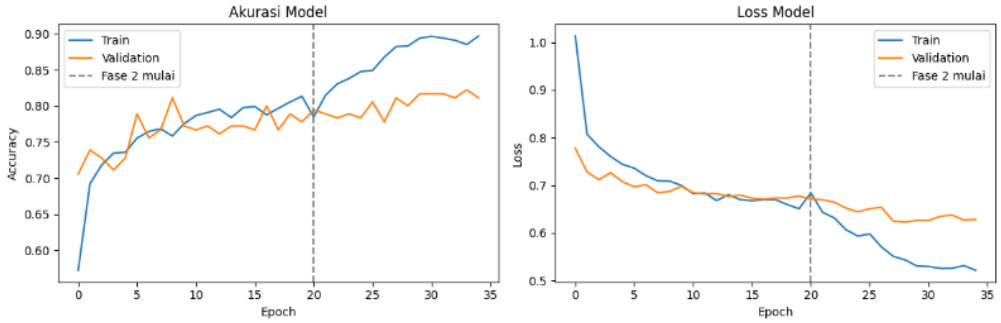
Bagian ini menyajikan hasil pelatihan, pengujian, dan analisis kesalahan klasifikasi MobileNetV2 dan EfficientNet-B3, implementasi sistem Qualet, serta pembahasan hasil dibandingkan dengan penelitian terdahulu.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394

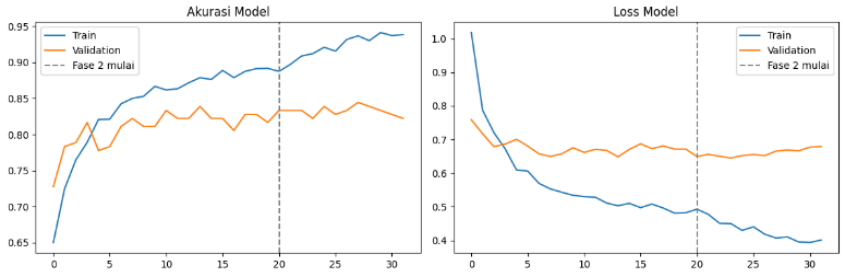
3.1 Hasil Pelatihan Model

Pelatihan kedua model berlangsung dalam dua fase, dengan data validasi dipakai untuk memantau kinerja sekaligus memilih bobot terbaik. Gambar 7 memperlihatkan perubahan akurasi dan *loss* pada data latih dan validasi untuk MobileNetV2, sedangkan Gambar 8 menampilkan hal yang sama untuk EfficientNet-B3.



Gambar 7. Kurva loss dan akurasi latih serta validasi MobileNetV2

Gambar 7 menunjukkan bahwa akurasi validasi MobileNetV2 mencapai 81,67% pada fase pertama dan meningkat menjadi 82,22% setelah penerapan fine-tuning. Di sisi lain, akurasi data latih mendekati 89% pada akhir pelatihan. Jarak antara kurva latih dan validasi menunjukkan bahwa model menyesuaikan diri dengan data latih lebih baik daripada dengan data validasi.



Gambar 8. Kurva loss dan akurasi latih serta validasi EfficientNet-B3

Gambar 8 menunjukkan EfficientNet-B3 mencapai akurasi validasi tertinggi 83,89% pada fase pertama dan 84,44% pada fase kedua, dengan akurasi latih akhir 93,83%. Jarak kurva latih dan validasi tampak lebih lebar daripada Gambar 7, sehingga kecenderungan *overfitting* lebih besar pada EfficientNet-B3.

Pada kedua model, akurasi validasi hanya naik 0,56 poin persentase pada fase kedua, atau setara satu dari 180 citra validasi, sehingga manfaat *fine-tuning* tergolong tipis. Hasil ini mengindikasikan bahwa bobot ImageNet diduga sudah memberi representasi yang cukup baik untuk citra sarang, sehingga fase *feature extraction* menghasilkan sebagian besar kinerja akhir. Di sisi lain, jarak antara akurasi latih dan validasi menandakan *overfitting* ringan pada MobileNetV2 (sekitar 7 poin) dan ringan hingga sedang pada EfficientNet-B3 (sekitar 9 poin). Hal ini diduga berkaitan dengan data latih yang sebagian besar berasal dari augmentasi 280 citra asli per kelas, sehingga variasi independen yang dapat dipelajari model terbatas. Untuk menekan *overfitting*, pelatihan memakai augmentasi, *Dropout*, regularisasi L2, *label smoothing*, dan *EarlyStopping* yang mengembalikan bobot dengan *loss* validasi terendah.

3.2 Hasil Pengujian

Pengujian memakai 180 citra data uji asli yang tidak terlibat dalam pelatihan maupun pemilihan bobot. Pengujian ini bertujuan menilai seberapa baik model mengenali citra baru ke dalam tiga kelas kualitas, yaitu Tinggi, Sedang, dan Rendah. Hasil kedua model dirangkum pada Tabel 4.

Tabel 4. Hasil pengujian pada data uji

Model	Akurasi uji (%)	*Loss* uji
MobileNetV2	78,89	0,6767
EfficientNet-B3	80,56	0,6954

Menurut Tabel 4, MobileNetV2 mencapai akurasi 78,89% dengan *loss* 0,6767, yakni 142 citra terklasifikasi benar dan 38 citra salah. EfficientNet-B3 mencapai akurasi 80,56% dengan *loss* 0,6954, yakni 145 citra benar dan 35 citra salah. Akurasi EfficientNet-B3 lebih tinggi 3 citra (1,67 poin persentase), sedangkan *loss* MobileNetV2 sedikit lebih rendah. Pada 180 citra uji, rentang kepercayaan 95% untuk akurasi berada di sekitar 72-84% pada MobileNetV2 dan 74-86% pada EfficientNet-B3, sehingga selisih kedua model belum cukup kuat untuk menyatakan salah satunya lebih unggul. Akurasi uji kedua model juga sekitar 3-4 poin lebih rendah daripada akurasi validasi terbaiknya, sesuai dengan *overfitting*

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2619
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394

ringan yang telah dibahas pada subbab sebelumnya. *Loss* kedua model tidak dapat dibandingkan langsung karena *label smoothing* berbeda (0,1 dan 0,05), sehingga perbandingan bertumpu pada akurasi dan metrik per kelas pada Tabel 5.

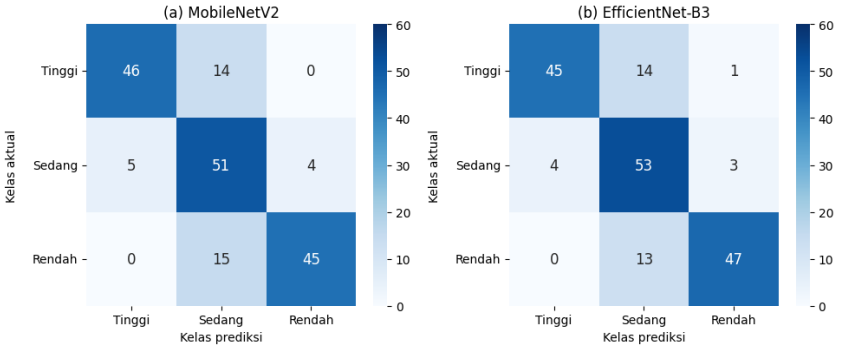
Tabel 5. Precision, recall, dan F1-score per kelas (%)

Model	Kelas	Precision	Recall	*F1-score*
MobileNetV2	Tinggi	90,20	76,67	82,88
	Sedang	63,75	85,00	72,86
	Rendah	91,84	75,00	82,57
	Rata-rata *macro*	81,93	78,89	79,44
EfficientNet-B3	Tinggi	91,84	75,00	82,57
	Sedang	66,25	88,33	75,71
	Rendah	92,16	78,33	84,68
	Rata-rata *macro*	83,42	80,56	80,99

Kinerja pada tiap kelas dirinci pada Tabel 5. Rata-rata *macro* untuk *precision*, *recall*, dan *F1-score* berturut-turut adalah 81,93%, 78,89%, dan 79,44% pada MobileNetV2, serta 83,42%, 80,56%, dan 80,99% pada EfficientNet-B3. Kedua model menunjukkan pola yang mirip antar kelas. Kelas Tinggi dan Rendah memiliki *precision* di atas 90% dengan *recall* sekitar 75-78%. Sebaliknya, kelas Sedang memperoleh *recall* paling tinggi (85,00% pada MobileNetV2 dan 88,33% pada EfficientNet-B3) tetapi *precision* paling rendah (63,75% dan 66,25%). Hal ini menandakan bahwa kedua model cenderung memberi label Sedang, sehingga sebagian citra dari kelas lain ikut masuk ke kelas tersebut. Karena itu, *F1-score* terendah pada kedua model ada di kelas Sedang (72,86% dan 75,71%), sementara kelas lainnya berkisar 82-85%. Pola ini berkaitan langsung dengan harga jual sarang. Citra kelas Tinggi yang diprediksi sebagai kelas Sedang berpotensi membuat sarang bernilai tinggi dihargai lebih rendah, sedangkan citra kelas Rendah yang diprediksi sebagai kelas Sedang berpotensi membuat sarang berkualitas rendah dihargai lebih tinggi. Karena kelas Tinggi dan Rendah hampir tidak pernah tertukar secara langsung, kesalahan yang terjadi umumnya hanya berselisih satu tingkat kualitas, bukan dua tingkat.

3.3 Analisis Kesalahan Klasifikasi

Gambar 9 memperlihatkan *confusion matrix* kedua model pada 180 citra uji, dengan 38 kesalahan klasifikasi pada MobileNetV2 dan 35 kesalahan pada EfficientNet-B3. Pada MobileNetV2, sebagian besar kesalahan, yakni 29 dari 38 (76%), terjadi ketika citra Rendah (15 citra) dan Tinggi (14 citra) diprediksi sebagai Sedang. Adapun citra Sedang yang keliru diklasifikasikan hanya 9. EfficientNet-B3 menunjukkan kecenderungan yang sama: 27 dari 35 kesalahan (77%) merupakan citra Rendah (13 citra) dan Tinggi (14 citra) yang diprediksi sebagai Sedang, sementara citra Sedang yang salah berjumlah 7. Kelas Tinggi dan Rendah nyaris tidak pernah saling tertukar. Hal ini tidak terjadi sama sekali pada MobileNetV2, dan pada EfficientNet-B3 hanya satu citra Tinggi yang diprediksi sebagai Rendah.



Gambar 9. Confusion matrix pada data uji

Jika dilihat per kelas, tingkat kesalahan pada MobileNetV2 adalah 23,33% untuk kelas Tinggi (14 dari 60 citra), 15,00% untuk kelas Sedang (9 dari 60), dan 25,00% untuk kelas Rendah (15 dari 60). Pada EfficientNet-B3, tingkat kesalahannya 25,00% untuk kelas Tinggi (15 dari 60), 11,67% untuk kelas Sedang (7 dari 60), dan 21,67% untuk kelas Rendah (13 dari 60). Artinya, kelas Sedang paling jarang salah dikenali, tetapi paling sering menjadi tujuan salah prediksi dari dua kelas lainnya.

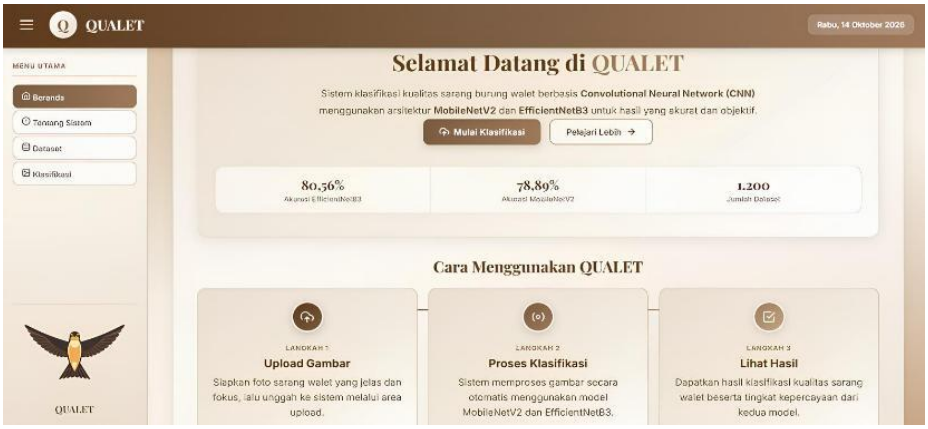
Temuan tersebut selaras dengan *precision* kelas Sedang yang paling rendah pada Tabel 5, yang menunjukkan bahwa kedua model lebih sering memberi label Sedang pada citra dari kelas lain. Kecenderungan ini diduga terkait dengan posisi kelas Sedang yang berada di antara kelas Tinggi dan Rendah sehingga batas visualnya kurang jelas, serta dengan terbatasnya citra asli (280 citra latih per kelas) yang membatasi variasi yang dapat dipelajari model. Karena kelas Tinggi dan Rendah hampir tidak pernah tertukar, kedua model tampak mampu membedakan kelas yang ekstrem, sedangkan kesulitan utamanya ada pada penentuan batas kelas tengah. Temuan ini menunjukkan bahwa peningkatan kinerja berikutnya kemungkinan lebih bergantung pada kejelasan definisi kelas Sedang dan variasi citra asli kelas tersebut daripada pada pergantian arsitektur.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2620
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394

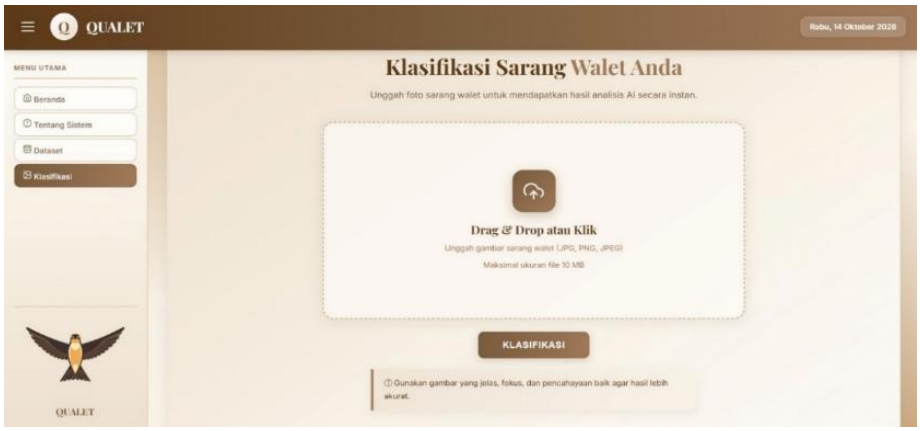
3.4 Implementasi Sistem Qualet

Qualet adalah *website* berbasis *framework* Flask yang menjalankan MobileNetV2 dan EfficientNet-B3 untuk mengklasifikasikan kualitas sarang burung walet melalui halaman beranda (Gambar 10), unggah (Gambar 11), dan hasil (Gambar 12).



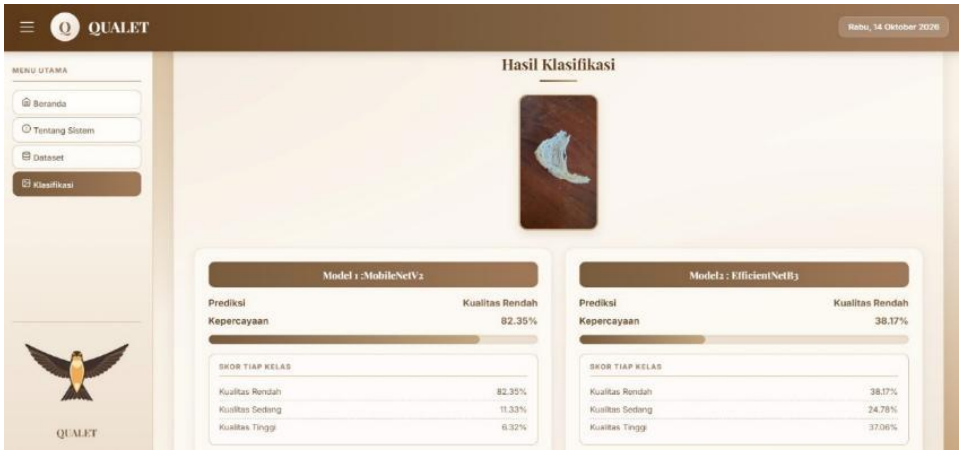
Gambar 10. Tampilan website Qualet

Gambar 10 menunjukkan tampilan halaman beranda website Qualet yang menjadi halaman awal ketika pengguna mengakses sistem. Halaman ini memuat informasi singkat tentang sistem dan panduan penggunaannya. Panduan tersebut membantu pengguna memahami langkah-langkah sebelum mengunggah citra sarang burung walet. Selanjutnya, pengguna dapat menuju halaman unggah citra untuk memulai proses klasifikasi kualitas sarang.



Gambar 11. Halaman unggah

Gambar 11 memperlihatkan halaman unggah yang digunakan untuk memilih citra sarang burung walet yang akan diklasifikasikan. Antarmuka dirancang sederhana agar mudah digunakan tanpa memerlukan pengetahuan teknis. Setelah citra dipilih, sistem memprosesnya dan menampilkan hasil klasifikasi pada halaman hasil yang ditunjukkan pada Gambar 12.



Gambar 12. Halaman hasil

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2621
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2612-2624
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1394

Gambar 12 menampilkan kelas prediksi, *confidence score*, dan skor tiap kelas dari kedua model secara berdampingan, beserta informasi apakah prediksi keduanya sama. Jika hasil berbeda atau *confidence score* rendah, citra disarankan diperiksa ulang oleh penilai manusia. Dengan demikian, Qualet menjadi alat bantu penilaian, bukan pengganti penilai.

3.5 Pembahasan

MobileNetV2 dan EfficientNet-B3 mencapai akurasi uji 78,89% dan 80,56% pada klasifikasi tiga kelas kualitas sarang burung walet (Tabel 4). Angka ini berada di bawah penelitian-penelitian sebelumnya, yaitu Ahmad dan Supatman (2025) dengan CNN untuk tiga kategori kebersihan pada 900 citra (84%) [6], Parhusip dkk. (2024) dengan SSD MobileNetV2 untuk tujuh kelas intensitas bulu (85%) [7], Indrajaya dkk. (2024) dengan MobileNetV2-D dan empat kamera untuk tujuh kelas (94,07%) [8], dan Ilmiati dkk. (2026) pada klasifikasi tiga kelas berdasarkan SNI 8998:2021 (98,37% pada MobileNetV2 dan 99,5% pada YOLOv8n-cls) [10]. Namun, perbandingan secara langsung perlu dilakukan dengan hati-hati karena penelitian-penelitian tersebut mengambil citra dalam kondisi yang lebih terkontrol, misalnya prototipe mesin sortir dengan empat kamera dan pengaturan pencahayaan pada Indrajaya dkk. (2024) [8], serta kamera digital dengan pencahayaan terkontrol dan label yang mengacu pada SNI pada Ilmiati dkk. (2026) [10]. Sebaliknya, penelitian ini menggunakan foto dari ponsel dengan latar dan pencahayaan yang bervariasi, serta label yang berasal dari penilaian lisan peternak, sehingga kondisinya lebih mendekati penggunaan di lapangan. Perbedaan kondisi tersebut kemungkinan turut memengaruhi selisih akurasi, meskipun pengaruhnya tidak diuji dalam penelitian ini.

Hasil menunjukkan bahwa batas antara kelas Sedang dan kelas lainnya kurang tegas, ditandai oleh *precision* kelas Sedang yang paling rendah pada kedua model (Tabel 5) dan sekitar 76-77% kesalahan yang jatuh ke kelas tersebut (Gambar 9), sedangkan kelas Tinggi dan Rendah hampir tidak pernah saling tertukar. Hal ini berkaitan dengan label yang tidak dilengkapi dokumen tertulis maupun uji kesepakatan antar penilai, sehingga batas antar kelas dapat mengandung subjektivitas. Selain itu, data latih hanya memuat 280 citra asli per kelas dan sebagian besar berupa hasil augmentasi, sehingga variasi independen yang dipelajari model diduga terbatas, seperti terlihat dari selisih akurasi latih dan validasi pada Gambar 7 dan Gambar 8. Selisih antara akurasi latih dan uji juga dilaporkan Indrajaya dkk. (2024) [8] (99,99% dan 94,07%), meskipun selisih pada penelitian ini lebih besar, yaitu sekitar 10 poin pada MobileNetV2 dan sekitar 13 poin pada EfficientNet-B3.

Kinerja kedua arsitektur tergolong setara: EfficientNet-B3 unggul 3 citra (1,67 poin persentase), MobileNetV2 memiliki *loss* yang sedikit lebih rendah, dan rentang kepercayaan 95% keduanya saling tumpang tindih. Temuan ini berbeda dengan Tamayasa dan Dewi (2026) pada klasifikasi penyakit daun jagung pada 2.000 citra yang mendapati MobileNetV2 (96,50%) lebih unggul daripada EfficientNet-B3 (93,00%) [21]. Dengan demikian, hasil perbandingan arsitektur dapat bervariasi antar penelitian, kemungkinan akibat perbedaan data dan konfigurasi pelatihan. Selisih hasil pada penelitian ini pun tidak sepenuhnya dapat dikaitkan dengan arsitektur, sebab ukuran input, susunan *classification head*, dan augmentasi *real-time* pada kedua model tidak identik. Selain itu, nilai *loss* sekitar 0,68-0,70 dihitung dengan *label smoothing* 0,1 untuk MobileNetV2 dan 0,05 untuk EfficientNet-B3, sehingga tidak dapat dibandingkan langsung dengan penelitian yang memakai *cross-entropy* biasa.

Karena kinerja kedua model setara, keluaran keduanya ditampilkan berdampingan pada Qualet (subbab 3.4, Gambar 12). Perlu dicatat bahwa *confidence score* merupakan probabilitas keluaran *softmax* yang bersifat indikatif, sehingga skor yang tinggi tidak menjamin prediksi pasti tepat, terutama pada kelas Sedang yang batasnya kurang tegas.

Secara praktis, kedua arsitektur sudah cukup baik dalam memisahkan kelas ekstrem, tetapi belum cukup andal untuk menggantikan penilaian manusia pada kelas tengah. Pada penilaian sarang burung walet, model lebih sesuai digunakan sebagai alat bantu penyaringan awal yang konsisten, sedangkan citra yang diprediksi sebagai kelas Sedang dengan *confidence score* rendah sebaiknya diperiksa ulang secara manual.

Keterbatasan penelitian ini meliputi label yang bersumber dari penilaian lisan tanpa uji kesepakatan antar penilai, data uji yang hanya berjumlah 180 citra (ketidakpastian sekitar ±6 poin), latar pemotretan yang belum tentu merata antar kelas, serta data sekunder Kaggle yang hanya terdapat pada kelas Tinggi. Pada penelitian selanjutnya, disarankan untuk menambah citra asli, terutama pada kelas Sedang, melibatkan beberapa penilai dan menguji kesepakatan antar penilai guna memvalidasi label, menyeragamkan latar dan pencahayaan pemotretan pada setiap kelas, menguji model pada citra dari peternakan lain untuk menilai kemampuan generalisasi, mengukur ukuran model dan waktu inferensi, serta mengeksplorasi CLAHE dan pengujian signifikansi statistik dalam membandingkan kedua arsitektur.

4. KESIMPULAN

Penelitian ini bertujuan membandingkan kinerja MobileNetV2 dan EfficientNet-B3 dalam mengklasifikasikan kualitas sarang burung walet ke dalam tiga kelas, yaitu Tinggi, Sedang, dan Rendah. Kedua arsitektur dilatih dengan strategi *transfer learning* dua fase pada 1.200 citra yang dibagi secara berstrata dengan rasio 70:15:15 sebelum proses augmentasi. Hasil pengujian pada 180 citra memperlihatkan bahwa kinerja keduanya sebanding, dengan akurasi MobileNetV2 sebesar 78,89% dan *loss* 0,6767, sedangkan EfficientNet-B3 memperoleh akurasi 80,56% dan *loss* 0,6954. Selisih tiga citra (1,67 poin persentase) belum cukup kuat untuk menyatakan bahwa salah satu arsitektur lebih unggul karena rentang kepercayaan 95% keduanya saling tumpang tindih. Dengan demikian, kedua arsitektur dinilai setara pada data ini, dan pemilihan di antara keduanya tidak dapat didasarkan pada selisih akurasi semata. Temuan penting lainnya adalah sekitar

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2622
This work is licensed under a Creative Commons Attribution 4.0 International License.

76% kesalahan MobileNetV2 dan 77% kesalahan EfficientNet-B3 berasal dari citra Tinggi atau Rendah yang diprediksi sebagai Sedang, sementara kedua kelas tersebut nyaris tidak pernah saling tertukar. Pola ini mengindikasikan bahwa kelas Sedang berperan sebagai kelas peralihan dengan batas visual yang kurang jelas. Meskipun akurasi penelitian ini lebih rendah daripada beberapa penelitian terdahulu, perbandingan langsung perlu dilakukan dengan hati-hati karena jumlah kelas, acuan penilaian, dan cara pengambilan citra berbeda. Kedua model telah diimplementasikan dalam *website* Qualet yang menampilkan kelas prediksi, *confidence score*, dan skor tiap kelas secara berdampingan. *website* Qualet ini berpotensi membantu penilaian awal, tetapi lebih tepat dimanfaatkan sebagai pendukung keputusan dan bukan sebagai pengganti penilai. Penelitian ini memiliki beberapa keterbatasan, yaitu label yang berasal dari penilaian lisan peternak tanpa uji kesepakatan antar penilai, data uji yang hanya berjumlah 180 citra dengan ketidakpastian sekitar ±6 poin, konfigurasi kedua model yang tidak sepenuhnya sama, latar citra yang beragam dengan persebaran antar kelas yang belum dikendalikan, serta data sekunder dari luar yang hanya terdapat pada kelas Tinggi. Oleh karena itu, penelitian selanjutnya disarankan untuk menambah citra asli kelas Sedang, melibatkan beberapa penilai disertai uji kesepakatan, menyamakan konfigurasi kedua model, mengukur ukuran model dan waktu inferensi, serta menjajaki penggunaan CLAHE agar kinerja klasifikasi dan keandalan sistem semakin baik dan dapat dipertanggungjawabkan.

5. PERNYATAAN PENULIS

5.1 Kontribusi Penulis

Konseptualisasi: I. K. A.;
Metodologi: I. K. A.;
Perangkat Lunak: I. K. A.;
Validasi: I. K. A. dan M. F.;
Analisis Formal: I. K. A.;
Investigasi: I. K. A.;
Sumber Daya: I. K. A.;
Kurasi Data: I. K. A.;
Penulisan Draf Awal: I. K. A.;
Penulisan, Peninjauan, dan Penyuntingan: I. K. A. dan M. F.;
Visualisasi: I. K. A.;
Seluruh penulis telah membaca dan menyetujui versi manuskrip yang telah diterbitkan.

5.2 Ketersediaan Data

Data primer yang disajikan dalam penelitian ini tersedia berdasarkan permintaan kepada penulis korespondensi, sedangkan data sekunder tersedia secara publik pada *dataset* Kaggle “Swiftlet’s Nest Shape (3 Classes)” [24]

5.3 Pendanaan

Penelitian ini tidak menerima pendanaan eksternal.

5.4 Pernyataan Dewan Peninjau Institusional

Tidak Berlaku (*Not applicable*).

5.5 Pernyataan Persetujuan Berdasarkan Informasi

Tidak Berlaku (*Not applicable*).

5.6 Pernyataan Konflik Kepentingan

Para penulis menyatakan bahwa mereka tidak memiliki kepentingan keuangan yang bersaing atau hubungan pribadi yang diketahui yang dapat dianggap memengaruhi pekerjaan yang dilaporkan dalam artikel ini.

5.7 Penggunaan AI Generatif dan Teknologi Berbantuan AI dalam Proses Penulisan

Para penulis memanfaatkan AI generatif untuk meningkatkan kejelasan penulisan, memeriksa konsistensi naskah, dan menerjemahkan abstrak. Seluruh konten hasil bantuan AI telah ditinjau dan disunting oleh para penulis, yang bertanggung jawab penuh atas versi akhir artikel.

REFERENCES

[1] F. Agus, E. Sulfika, and G. M. Putra, “Analisis Kualitas Sarang Burung Walet Menggunakan Metode Fuzzy Tsukamoto,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 2, pp. 391–398, 2025, doi: https://doi.org/10.25126/jtiik.2025129441.
[2] P. Pojanalai, P. Akkajit, and A. Sukkuea, “Explainable deep learning for grading of Edible Bird ’s Nest(EBN),” *J. Agric. Food Res.*, vol. 27, p. 102815, 2026, doi: https://doi.org/10.1016/j.jafr.2026.102815.
[3] N. N. Zulkefle *et al.*, “A Review on the Edible Bird’s Nest Quality and Manufacturing Standards of the Three Largest Exporting Countries in the World,” *J. Food Qual.*, vol. 2024, p. 5608357, 2024, doi: https://doi.org/10.1155/2024/5608357.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

[4] B.-H. Yeo, Y. Rukayadi, S.-F. Wong, C.-P. Tan, and O.-M. Lai, "Physicochemical Properties of Raw Cleaned Edible Bird 's Nest after Different Primary Processing Including a New Cleaning Method," *Sains Malaysiana*, vol. 52, no. 12, pp. 3421–3435, 2023, doi: https://doi.org/10.17576/jsm-2023-5212-06.
[5] R. A. Saemani, D. Manongga, and Hendry, "Pendekatan Multi-Hirarki untuk Grading Sarang Burung Walet Berdasarkan Deteksi Bentuk dan Klasifikasi Warna," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 6, no. 2, pp. 462–474, 2026, doi: https://doi.org/10.57152/malcom.v6i2.2594.
[6] S. Ahmad and Supatman, "Analisis Cnn Untuk Klasifikasi Kebersihan Sarang Burung Walet," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 3, p. 8, 2025, doi: https://doi.org/10.23960/jitet.v13i3.6685.
[7] H. Arini Parhusip, S. Trihandaru, K. D. Hartomo, K. Bianca Lewerissa, L. Ariany Mahastanti, and D. Hartanto, "Management of Traditional Business into Modern: from Microsoft Excel to Deep Learning for prototyping classification Swiftlet's nests," *Int. J. Community Serv.*, vol. 4, no. 2, pp. 123–132, 2024, doi: https://doi.org/10.51601/ijcs.v4i2.268.
[8] D. Indrajaya, H. A. Parhusip, S. Trihandaru, and D. Hartanto, "MobileNetV2-D and multiple cameras for swiftlet nest classification based on feather intensity," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 34, no. 2, pp. 1144–1158, 2024, doi: 10.11591/ijeecs.v34.i2.pp1144-1158.
[9] P. Medina and E. Rahayu, "Implementasi Algoritma SVM (Support Vector Machine) Dalam Memprediksi Hasil Produksi Sarang Burung Walet," *Explore*, vol. 16, no. 1, pp. 40–49, 2026, doi: https://doi.org/10.35200/sd.v16i1.188.
[10] Ilmiati, S. Mutmainah, and Khairunnas, "Classification of Swiftlet Nest Quality Based on SNI 8998:2021 Using Deep Learning," *Build. Informatics, Technol. Sci.*, vol. 8, no. 1, pp. 308–317, 2026, doi: https://doi.org/10.47065/bits.v8i1.9901.
[11] F. Chollet, *Deep Learning with Python*, 2nd ed. Shelter Island, NY, USA: Manning Publications, 2021.
[12] X. Zhao, L. Wang, Y. Zhang, X. Han, M. Deveci, and M. Parmar, "A review of convolutional neural networks in computer vision," *Artif. Intell. Rev.*, vol. 57, no. 4, pp. 1–43, 2024, doi: 10.1007/s10462-024-10721-6.
[13] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive into Deep Learning*. Cambridge, U.K.: Cambridge University Press, 2023.
[14] M. Iman, H. R. Arabnia, and K. Rasheed, "A Review of Deep Transfer Learning and Recent Advancements," *Technologies*, vol. 11, no. 2, pp. 1–14, 2023, doi: https://doi.org/10.3390/technologies11020040.
[15] M. Isnan, A. A. Hidayat, and B. Pardamean, "Indonesian Agricultural-crops Classification Using Transfer Learning Model," *Procedia Comput. Sci.*, vol. 227, pp. 128–136, 2023, doi: https://doi.org/10.1016/j.procs.2023.10.510.
[16] K. Halloum and H. E. Zahraouy, "Enhancing Medical Image Classification through Transfer Learning and CLAHE Optimization," *Curr. Med. Imaging*, vol. 21, no. 1, p. e15734056342623, 2025, doi: https://doi.org/10.2174/0115734056342623241119061744.
[17] V. Zárate and D. C. Hernández, "Simplified Deep Learning for Accessible Fruit Quality Assessment in Small Agricultural Operations," *Appl. Sci.*, vol. 14, no. 18, 2024, doi: https://doi.org/10.3390/app14188243.
[18] H. Chen, G. Zhou, W. He, X. Duan, and H. Jiang, "Classification and identification of agricultural products based on improved MobileNetV2," *Sci. Rep.*, vol. 14, no. 1, p. 14, 2024, doi: https://doi.org/10.1038/s41598-024-53349-w.
[19] Y. P. Vera *et al.*, "Deep Learning Models for Defect Identification in Oryza sativa Rice Grains: A Comparative Study," *AgriEngineering*, vol. 8, no. 6, p. 28, 2026, doi: 10.3390/agriengineering8060252.
[20] E. Dewi, S. Mulyani, T. Mufizar, D. Rohpandi, A. Djuliani, and E. Rahmatulloh, "Classification of Cihateup duck egg fertility using convolutional neural network EfficientNet-B3," *IAES Int. J. Artif. Intell.*, vol. 15, no. 1, pp. 798–809, 2026, doi: https://doi.org/10.11591/ijai.v15.i1.pp798-809.
[21] K. A. Tamayasa and L. J. E. Dewi, "Analisis Perbandingan Model Arsitektur MobileNetV2 dan EfficientNetB3 dalam Klasifikasi Penyakit Daun Jagung," *J. Inform. dan Tek. Elektro Terap.*, vol. 14, no. 1, p. 12, 2026, doi: https://doi.org/10.23960/jitet.v14i1.8624.
[22] F. Nan *et al.*, "A novel method for maize leaf disease classification using the image data," *Front. Plant Sci.*, vol. 14, p. 14, 2023, doi: https://doi.org/10.3389/fpls.2023.1268015.
[23] Y. Puspitasari, J. J. Siang, and A. G. Prabawati, "Comparison of Convolutional Neural Network Architectures EfficientNet-B4 and MobileNetV2 in Cataract Disease Detection," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 11, no. 4, pp. 1020–1027, 2026, doi: https://doi.org/10.33480/jitk.v11i4.7502.
[24] IOT Lab Sensor, "Swiftlet's Nest Shape (3 Classes)," 2022, *Kaggle*. [Online]. Available: https://www.kaggle.com/datasets/iotlabsensor/bentuk-sarang-bulu-burung-wallet-3-kelas

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.