

Model CNN Untuk Klasifikasi Spesies Burung Dengan Ciri Serupa Akibat Evolusi Konvergen

Nino Andera¹, Ghofar Taufik^{2*}, Giatika Chrisnawati³

^{1,2,3}Fakultas Teknik dan Informatika, Informatika, Universitas Bina Sarana Informatika, Jakarta, Indonesia
Jl. Kramat Raya No. 98 Senen, Jakarta Pusat-10420, Indonesia
E-mail: ¹nanderaa54@gmail.com, ^{2*}ghofar.gft@bsi.ac.id, ³giatika.gcw@bsi.ac.id
(*: corresponding author)

Abstrak— Kemiripan morfologi pada spesies burung akibat evolusi konvergen menjadi tantangan dalam klasifikasi otomatis berbasis citra. Penelitian ini bertujuan untuk mengembangkan dan menganalisis kinerja model Convolutional Neural Network (CNN) dalam membedakan spesies burung yang memiliki ciri visual serupa. CNN digunakan karena kemampuannya dalam mengekstraksi fitur citra secara efisien, bahkan pada spesies dengan perbedaan morfologi yang tipis. Evaluasi dilakukan terhadap empat spesies burung melalui pengujian utama, serta dua kelompok pasangan burung dengan kemiripan morfologi tinggi, yaitu *Icterus graduacauda* & *Oriolus larvatus*, serta *Sturnella magna* & *Macronyx croceus*. Pengujian utama menunjukkan akurasi sebesar 78%, sedangkan pada dua pasangan spesies serupa, akurasi meningkat masing-masing menjadi 80,42% dan 85%, dengan F1-score yang stabil di kisaran 80–85%. Hasil ini menunjukkan bahwa model CNN mampu membedakan spesies burung yang serupa secara visual meskipun terdapat variasi latar belakang dan posisi objek dalam dataset. Kesimpulan dari penelitian ini mengindikasikan bahwa pendekatan CNN efektif digunakan dalam klasifikasi spesies burung dengan morfologi serupa akibat evolusi konvergen, serta memiliki potensi untuk diterapkan dalam bidang konservasi, pemantauan keanekaragaman hayati, dan sistem identifikasi spesies otomatis berbasis citra.

Kata Kunci— Burung, CNN, Evolusi Konvergen, Klasifikasi

Abstract— *Morphological similarity in bird species due to convergent evolution is a challenge in image-based automatic classification. This study aims to develop and analyze the performance of a Convolutional Neural Network (CNN) model in distinguishing bird species that have similar visual features. CNNs are used because of their ability to efficiently extract image features, even in species with subtle morphological differences. Four bird species were evaluated through the main test, as well as two groups of bird pairs with high morphological similarity, namely *Icterus graduacauda* & *Oriolus larvatus*, and *Sturnella magna* & *Macronyx croceus*. The primary test showed an accuracy of 78%, while on the two similar species pairs, the accuracy increased to 80.42% and 85%, respectively, with a stable F1-score in the 80-85% range. These results show that the CNN model is able to visually distinguish similar bird species despite variations in background and object position in the dataset. The conclusion of this study indicates that the CNN approach is effective in the classification of bird species with similar morphology due to convergent evolution, and has the potential to be applied in the fields of conservation, biodiversity monitoring, and image-based automatic species identification systems.*

Keyword— *Birds, Convergent Evolution, Classification, CNN*

I. PENDAHULUAN

Evolusi konvergen merupakan fenomena biologis di mana spesies berbeda mengembangkan karakteristik morfologi serupa sebagai respons terhadap tekanan lingkungan yang sama [1]. Pada burung, kemiripan ini dapat terjadi antar spesies yang masih berada dalam kelas yang sama, sehingga perbedaan antar spesies menjadi lebih sulit dikenali secara visual. Tantangan ini menjadikan proses klasifikasi burung secara otomatis berbasis citra menjadi semakin kompleks.

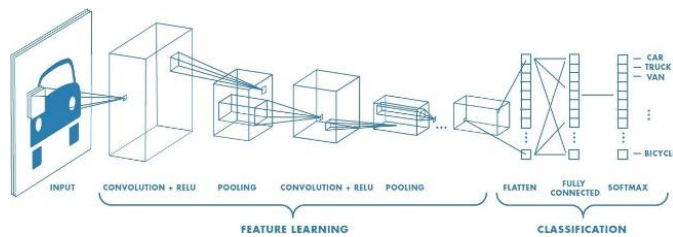
Di era digital, pendekatan pengenalan visual dengan memanfaatkan kecerdasan buatan, khususnya Deep Learning, telah berkembang pesat. Salah satu bagian dari machine learning adalah Deep Learning, yang menggunakan algoritma yang dibuat berdasarkan hukum matematik dan berfungsi seperti otak manusia. Convolutional Neural Network (CNN), algoritma Deep Learning yang dikembangkan dari Multilayer Perceptron (MLP), digunakan dalam pengolahan gambar untuk memproses suara atau gambar [2]. Salah satu metode yang banyak digunakan dalam klasifikasi citra adalah Convolutional Neural Network (CNN), yang mampu mengenali pola dan fitur visual secara otomatis tanpa proses pemrograman eksplisit [3]. CNN telah digunakan dalam berbagai bidang, seperti identifikasi hewan laut hingga analisis kualitas ikan, dan terbukti efektif dalam mengekstraksi ciri visual dari objek yang diamati [4]. Melihat keberhasilan pendekatan ini, penelitian ini dilakukan untuk mengklasifikasikan spesies burung yang memiliki kemiripan morfologi akibat evolusi konvergen menggunakan CNN dengan arsitektur yang efisien dan ringan, yaitu MobileNetV2 [5].

Penelitian ini bertujuan untuk menguji kemampuan model CNN dalam mengklasifikasikan spesies burung yang memiliki kemiripan morfologi akibat evolusi konvergen, termasuk saat citra mengandung gangguan visual seperti latar belakang dan posisi objek yang bervariasi. Diharapkan penelitian ini dapat memberikan kontribusi terhadap pemahaman hubungan antara biologi evolusi dan kecerdasan buatan, serta mendorong penerapan teknologi visual dalam studi biodiversitas. Penelitian difokuskan pada empat spesies burung dengan total 2.000 citra dari Macaulay Library, menggunakan arsitektur MobileNetV2, dengan harapan model mampu mencapai akurasi $\geq 80\%$ dalam klasifikasi.

Penelitian ini menghadirkan pendekatan baru dalam klasifikasi citra berbasis *Deep Learning* dengan menitikberatkan pada kemampuan *CNN* dalam mengenali

spesies burung yang memiliki kemiripan morfologi akibat evolusi konvergen. Berbeda dengan studi sebelumnya yang lebih banyak berfokus pada klasifikasi objek dengan karakter visual yang berbeda secara jelas, seperti identifikasi spesies ubur ubur atau penilaian kesegaran ikan, penelitian ini menantang model *CNN* untuk membedakan spesies burung yang secara visual sangat mirip meskipun berasal dari famili dan wilayah geografis yang berbeda. Hal ini memperkenalkan tantangan baru dalam pengolahan citra, khususnya dalam konteks taksonomi visual dan konservasi biodiversitas.

CNN sendiri merupakan jenis jaringan saraf dalam yang banyak digunakan untuk pengolahan data visual karena kemampuannya dalam mengekstraksi fitur secara otomatis dari gambar tanpa memerlukan intervensi manual dalam proses ekstraksi fitur [6]. *CNN* bekerja melalui serangkaian lapisan, mulai dari convolution layer yang bertugas mendeteksi fitur lokal seperti tepi dan pola dalam gambar dengan menggunakan filter atau kernel, lalu dilanjutkan dengan fungsi aktivasi yang memperkenalkan non-linearitas sehingga model mampu menangkap hubungan kompleks antar fitur [7]. Selanjutnya, pooling layer berperan dalam mengurangi dimensi spasial fitur yang dihasilkan, sehingga memperkecil beban komputasi tanpa kehilangan informasi penting, dan akhirnya fully connected layer yang menggabungkan fitur-fitur tersebut untuk menghasilkan klasifikasi akhir (Gambar 1).



Gambar 1. Convolutional Neural Network

Sementara itu, MobileNetV2 adalah arsitektur jaringan CNN yang dirancang khusus untuk kebutuhan komputasi ringan pada perangkat dengan sumber daya terbatas [8]. Model ini diperkenalkan sebagai pengembangan dari MobileNetV1 dengan dua inovasi utama, yaitu penggunaan inverted residuals yang memungkinkan aliran informasi langsung dari input ke output melalui koneksi shortcut, serta linear bottlenecks yang menjaga informasi penting agar tidak hilang akibat fungsi aktivasi non-linear seperti ReLU. Dengan demikian, MobileNetV2 mampu mempertahankan representasi fitur secara efektif sambil mengurangi beban komputasi, sehingga sangat cocok untuk aplikasi mobile dan embedded [9].

Input	Operator	t	c	n	s
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	-

Gambar 2. Arsitektur MobileNetV2

Pada Gambar 2 terlihat dengan menggunakan arsitektur MobileNetV2 yang ringan namun efisien, serta dataset yang beragam secara latar belakang dan posisi objek, penelitian ini membuktikan bahwa klasifikasi berbasis AI dapat diterapkan pada kasus biologis yang kompleks. Pendekatan ini tidak hanya memperluas aplikasi CNN dalam domain biologi evolusi tetapi juga membuka jalan bagi pengembangan sistem klasifikasi cerdas yang lebih adaptif dan kontekstual di masa depan.

II. METODE PENELITIAN

A. Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining berbasis algoritma Convolutional Neural Network (CNN), yang merupakan metode pembelajaran mendalam berbasis supervisi (supervised learning) [10]. Tujuannya adalah untuk mengklasifikasikan spesies burung yang memiliki kemiripan morfologi akibat evolusi konvergen berdasarkan data citra. Empat spesies burung dari famili berbeda digunakan sebagai objek penelitian dengan total 2.000 citra. Hasil klasifikasi ini diharapkan dapat menjadi kontribusi dalam pengembangan teknologi pengenalan visual berbasis kecerdasan buatan, sekaligus mendukung edukasi biodiversitas melalui pendekatan interdisipliner antara biologi dan ilmu komputer.

B. Langkah Penelitian

Penelitian ini dilakukan dengan langkah berurut dimulai dari tahap:

1) Pengumpulan Data

Dataset diperoleh dari situs **Macaulay Library** melalui proses pengajuan unduhan berbasis file .csv yang berisi daftar spesies target. Setelah verifikasi, tautan unduhan citra dikirim melalui email. Data yang dikumpulkan terdiri dari 2.000 citra berformat JPG dan JPEG, masing-masing 500 citra untuk empat spesies burung (Gambar 3, Gambar 4, Gambar 5).

a) *Sturnella magna (eastme)*



Gambar 3. *Sturnella magna*

b) *Macronyx croceus* (longclaw)



Gambar 4. *Macronyx croceus*

c) *Icterus Graduacauda* (auriole)



Gambar 5. *Icterus graduacauda*

d. *Oriolus Larvatus* (oriole)



Gambar 6. *Oriolus larvatus*

2. Pembagian Data

Setelah data berhasil dikumpulkan, selanjutnya gambar dibagi secara acak menggunakan Google Colab yang merupakan tools untuk melakukan penelitian ini. Data dibagi menjadi 1400 data latih, 300 data validasi, dan 300 data uji dengan 350 data latih untuk masing-masing kelas serta 75 data validasi dan data uji

3. Pra-pemrosesan Citra

Tahap ini dilakukan untuk menyesuaikan data dengan standar arsitektur MobileNetV2. Semua gambar diubah ukurannya menjadi 224×224 piksel dan disesuaikan saluran warnanya agar seragam. Nilai piksel dinormalisasi ke dalam rentang yang sesuai agar hasil prediksi lebih stabil [11]. Selama proses pelatihan, dilakukan augmentasi untuk memperkaya variasi gambar secara sintetis, seperti rotasi, pergeseran posisi, perubahan ukuran, penyesuaian kecerahan, dan pencerminan horizontal. Area kosong yang muncul akibat transformasi diisi dengan nilai piksel terdekat agar tidak mengganggu proses klasifikasi.

4. Ekstraksi Fitur

Setelah tahap praproses, fitur citra diekstraksi menggunakan model MobileNetV2 yang telah dilatih sebelumnya (pre-trained) [12]. Model digunakan tanpa bagian klasifikasi akhirnya agar hanya menghasilkan representasi fitur visual seperti tepi, warna, dan tekstur. Bobot model dibekukan agar pelatihan hanya dilakukan pada lapisan klasifikasi tambahan. Hasil ekstraksi ini berupa vektor fitur yang digunakan sebagai input untuk proses klasifikasi.

5. Klasifikasi

Setelah fitur diekstraksi, tahap klasifikasi dilakukan dengan menambahkan beberapa lapisan fully connected dan lapisan output softmax untuk memetakan fitur ke dalam empat kelas spesies burung. Proses ini dilakukan dalam tiga tahap:

a) Pelatihan

Model dilatih menggunakan data latih untuk menyesuaikan bobot melalui backpropagation, sehingga dapat mengenali pola visual tiap spesies [13]. Akurasi dan loss dipantau selama proses ini.

b) Validasi

Data validasi digunakan secara paralel untuk mengevaluasi kinerja model dan mendeteksi overfitting. Hasilnya menjadi dasar penyesuaian parameter model [3].

c) Testing

Pengujian dilakukan dengan data yang benar-benar baru (belum dilatih/validasi) untuk mengukur kemampuan generalisasi model secara objektif [14].

C. Analisa Hasil

Pada tahap ini dilakukan analisis performa model klasifikasi yang menggunakan arsitektur MobileNetV2 dengan mengevaluasi metrik seperti akurasi, loss, precision, recall, dan F1-score pada data uji [15]. Dengan analisis berfokus pada dua pasang spesies burung yang memiliki kemiripan morfologi, yaitu :

- S. magna dan M. croceus
- I. graduacauda dan O. larvatus

Yang bertujuan untuk melihat seberapa baik model membedakan spesies yang memiliki ciri visual serupa akibat evolusi konvergen.

D. Penarikan Kesimpulan

Merupakan tahap akhir yang merangkum hasil penelitian secara menyeluruh dan menjawab rumusan masalah. Pada bagian ini, dipaparkan efektivitas model CNN dalam

mengklasifikasikan spesies burung dari data citra, sekaligus mempertimbangkan validitas dan keterbatasan penelitian. Selain menyimpulkan temuan utama, tahap ini juga memberikan saran sebagai arahan untuk penelitian selanjutnya.

III. HASIL DAN PEMBAHASAN

A. Persiapan Model

Sebelum melakukan pelatihan dan validasi pada model, perlu dibuat konfigurasi untuk mempersiapkan target penelitian menggunakan data dari direktori latih dan validasi yang sudah disesuaikan dengan metodologi penelitian. Konfigurasi ini mencakup penyesuaian ukuran gambar, format label, serta penyusunan data ke dalam batch agar sesuai dengan kebutuhan pemrosesan model. Berikut adalah konfigurasi yang digunakan dalam penelitian ini yang terlihat pada Gambar 7.

```

train_generator = train_datagen.flow_from_directory(
    train_dir, # folder berisi 4 subfolder kelas
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical'
)

val_generator = val_datagen.flow_from_directory(
    validation_dir,
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical',
    shuffle=False
)
    
```

Found 1400 images belonging to 4 classes.
Found 300 images belonging to 4 classes.

Gambar 7. Penyesuaian Target Dataset

Kode “train_generator = train_datagen.flow_from_directory (...)” dan “val_generator = val_datagen.flow_from_directory (...)” digunakan untuk memuat dan mempersiapkan data gambar dari direktori yang telah diklasifikasikan ke dalam subdirektori berdasarkan kelas masing-masing. Fungsi “flow_from_directory” secara otomatis membaca gambar dari direktori utama yang di dalamnya terdapat beberapa subdirektori, di mana nama nya tersebut akan dianggap sebagai label kelas. Data disusun menjadi batch sejumlah 32 gambar melalui parameter “batch_size=32”, sehingga model tidak perlu memproses seluruh data sekaligus, melainkan dalam kelompok kecil yang efisien.

Parameter (class_mode='categorical') menunjukkan bahwa label akan diubah menjadi format one-hot encoding, yang sesuai untuk kasus klasifikasi multi-kelas. Untuk data validasi, parameter “shuffle=False” digunakan agar data tidak diacak, sehingga urutan data tetap konsisten, terutama saat proses evaluasi model (Gambar 8).

```

[ ] model.compile(loss = 'categorical_crossentropy',
                 optimizer = 'Adam',
                 metrics = ['accuracy'])

[ ] from tensorflow.keras.callbacks import EarlyStopping

early_stop = EarlyStopping(monitor='val_loss', patience=5, restore_best_weights=True)

history = model.fit(
    train_generator,
    epochs=30,
    validation_data=val_generator,
    callbacks=[early_stop]
)
    
```

Gambar 8. Persiapan Model

Model dikompilasi menggunakan fungsi kerugian categorical_crossentropy, optimizer Adam, dan metrik evaluasi accuracy karena tugas yang dijalankan adalah klasifikasi multi-kelas. Proses pelatihan dilakukan selama maksimal 30 epoch menggunakan data dari “train_generator” dan divalidasi dengan “val_generator”. Untuk mencegah overfitting, digunakan callback “EarlyStopping” yang akan menghentikan pelatihan jika val_loss tidak membaik selama 5 epoch, serta mengembalikan bobot model ke kondisi terbaik. Riwayat pelatihan disimpan dalam variabel “history” untuk analisis lebih lanjut.

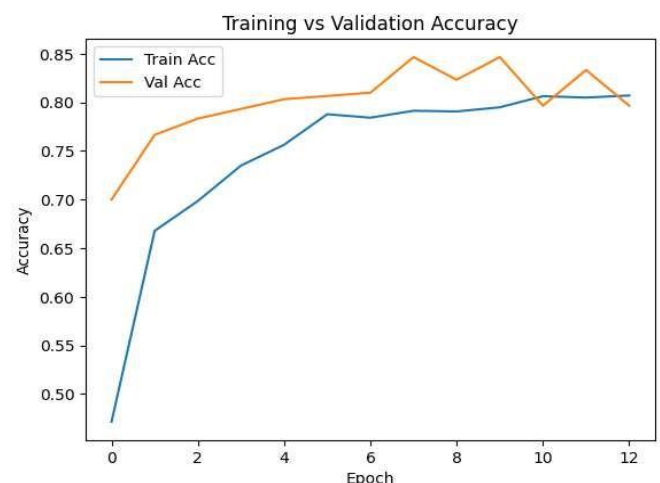
B. Pelatihan (Training) dan Validasi

Dengan penyesuaian dan konfigurasi yang sudah ditentukan, pelatihan berhenti pada epoch ke-13 dengan total waktu pelatihan selama 29 menit dan 25 detik. Penghentian pelatihan ini terjadi karena fungsi “EarlyStopping” mendeteksi bahwa tidak ada perbaikan nilai val_loss dalam lima epoch terakhir. Hal ini menunjukkan bahwa model telah mencapai performa optimal sebelum mencapai batas maksimum epoch yang ditentukan. Hasil proses pelatihan ditampilkan pada Gambar 9 berikut:

Epoch	Time	Step	Accuracy	Loss	Val Accuracy	Val Loss
44/44	132s	3s/step	0.8881	0.4870	0.8233	0.5130
Epoch 10/30	44/44	133s	0.7905	0.5220	0.8467	0.4941
Epoch 11/30	44/44	130s	0.7971	0.5059	0.7967	0.5054
Epoch 12/30	44/44	131s	0.7868	0.5277	0.8333	0.5011
Epoch 13/30	44/44	141s	0.7967	0.4945	0.7967	0.5722

Gambar 9. Hasil Pelatihan (Training)

Kemudian, oleh karena tahap pelatihan dan validasi dilakukan secara bersamaan, maka pada Gambar 10 disajikan gambar mempresentasikan grafik akurasi dari pelatihan dan validasi, disertai dengan Tabel 1 yang menunjukkan hasil nilai akurasi dan loss dari proses Training dan Validation. Data ini digunakan untuk mengevaluasi kinerja model selama proses pembelajaran agar dapat dilanjutkan ke tahap pengujian.



Gambar 10. Grafik Akurasi Pelatihan (Training)

Setelah melalui rangkaian langkah penelitian mulai dari pra-pemrosesan citra, pada langkah klasifikasi dilakukan tahap validasi (validation) bersamaan dengan tahap pelatihan (training) guna memastikan apakah model siap dilanjutkan ke

tahap selanjutnya atau tahap pengujian (testing). Pada Tabel I berikut disajikan nilai akurasi dan nilai loss dari tahap validasi pada model.

TABEL I
HASIL PROSES PENELITIAN

Tahap	Accuracy	Loss
Training	83.29%	46.92%
Validation	84.67%	47.12%

Untuk memastikan bahwa nilai akurasi adalah valid, berikutnya disajikan confusion matrix pada tahap validation guna menentukan apakah penelitian pada model bisa dilanjutkan pada tahap selanjutnya atau tahap pengujian model (Tabel II).

TABEL II
CONFUSION MATRIX VALIDASI

		Prediksi			
		Auriole	Eastme	Longclaw	Oriole
Data Asli	Auriole	61	3	3	8
	Eastme	4	64	6	1
	Longclaw	1	6	65	3
	Oriole	8	1	2	64

Berdasarkan confusion matrix, didapatkan nilai akurasi sebesar 84.67% yang membuktikan validitas pelatihan model sehingga penelitian bisa dilanjutkan ke tahap pengujian berdasarkan perhitungan berikut.

$$\frac{\text{jumlah data benar}}{\text{jumlah data}} \times 100\%$$

$$\frac{61 + 64 + 65 + 64}{300} \times 100\% = 84,67\%$$

C. Pengujian (Testing)

Dikarenakan data uji yang cukup banyak, pengujian secara manual satu persatu akan memakan waktu. Oleh karena itu, pengujian dilakukan otomatis menggunakan kodepe muatan data “test_generator=test_datagen.flow_from_directory(...)” dengan normalisasi dan konfigurasi yang sama dengan penyesuaian untuk data validasi seperti terlihat pada Gambar 11 berikut:

```
[ ] test_datagen = ImageDataGenerator(rescale=1./255)

[ ] test_generator = test_datagen.flow_from_directory(
    testing_dir,
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical',
    shuffle=False
)
```

Found 300 images belonging to 4 classes.

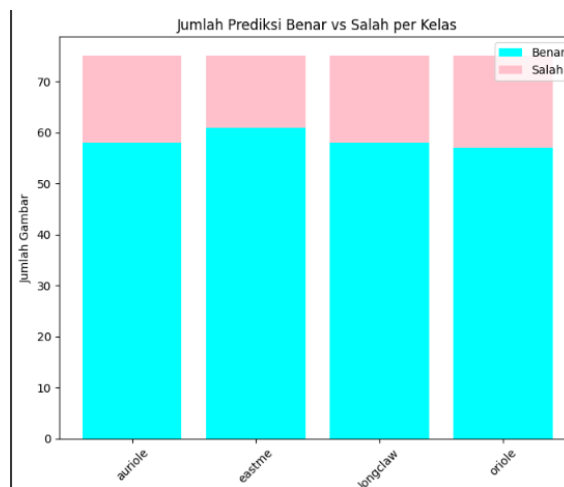
Gambar 11. Penyesuaian Data Tes

Setelah kode dijalankan, model hasil pelatihan akan digunakan untuk melakukan prediksi terhadap dataset yang disimpan dalam subdirektori "uji". Proses ini bertujuan untuk mengevaluasi performa model dalam mengenali citra dari kelas-kelas yang belum pernah dilihat sebelumnya. Nilai akurasi dan loss pada tahap pengujian disajikan pada Tabel III:

TABEL III
NILAI HASIL PENGUJIAN

Accuracy	Loss
78%	52.34%

Selain itu, hasil pengujian (testing) model dalam melakukan klasifikasi pada keempat kelas direpresentasikan dengan diagram batang pada Gambar 12 berikut:



Gambar 12. Diagram hasil prediksi pengujian

Berdasarkan diagram bertumpuk hasil prediksi, banyaknya prediksi benar dan salah pada tiap kelas hampir sama rata. Dengan eastme (Sturnella Magna) memiliki nilai precision, recall, dan F1-score tertinggi setinggi 84%, 81%, dan 82% berurutan dan oriole (Oriolus Larvatus) dengan nilai terendah setinggi 75%, 76%, dan 75%. Untuk memastikan berapa banyak gambar benar dan salah serta validitas akurasi, selanjutnya disajikan confusion matrix pada Tabel IV berikut.

TABEL IV
CONFUSION MATRIX TESTING

		Prediksi			
		Auriole	Eastme	Longclaw	Oriole
Data Asli	Auriole	58	2	2	13
	Eastme	2	61	11	1
	Longclaw	2	10	58	5
	Oriole	15	0	3	57

Setelah didapatkan hasil penelitian, sekarang dapat ditentukan nilai akurasi dan kesalahan prediksi pengujian keseluruhan untuk memastikan validitas model.

$$\frac{58 + 61 + 58 + 57}{300} \times 100\% = 78\%$$

Berikutnya disajikan hasil perhitungan confusion matrix secara manual, untuk membandingkan dengan perhitungan

python sebagai pertimbangan apakah model melakukan pengujian atau testing sesuai dengan jumlah dataset yang ada berdasarkan nilai *TP* (True Positive), *FP* (False Positive), *FN* (False Positive), *TN* (True Negative).

1) *AURIOLE*

TP : Prediksi benar AURIOLE = 58
 FP : Prediksi AURIOLE tapi salah = 2 (dari EASTME) + 2 (dari LONGCLAW) + 15 (dari ORIOLE) = 19
 FN : Seharusnya AURIOLE tapi salah prediksi = 2 (ke EASTME) + 2 (ke LONGCLAW) + 13 (ke ORIOLE) = 17
 TN : Bukan AURIOLE dan prediksi benar bukan AURIOLE = 61+11+1 + 10+58+5 + 0+3+57 = 206

Total data : TP + FP + FN + TN = 58 + 19 + 17 + 206 = 300

2) *EASTME*

TP : Prediksi benar EASTME = 61
 FP : Prediksi EASTME tapi salah = 2 (dari AURIOLE) + 10 (dari LONGCLAW) + 0 (dari ORIOLE) = 12
 FN : Seharusnya EASTME tapi salah prediksi = 2 (ke AURIOLE)+ 11 (ke LONGCLAW) + 1 (ke ORIOLE) = 14
 TN : Bukan EASTME dan prediksi benar bukan EASTME = 58+2+13 + 2+58+5 + 15+3+57 = 213

Total data : 61 + 12 + 14 + 213 = 300

3) *LONGCLAW*

TP : Prediksi benar LONGCLAW = 58
 FP : Prediksi LONGCLAW tapi salah = 2 (dari AURIOLE) + 11 (dari EASTME) + 3 (dari ORIOLE) = 16
 FN : Seharusnya LONGCLAW tapi salah prediksi = 2 (ke AURIOLE) + 10 (ke EASTME) + 5 (ke ORIOLE) = 17
 TN : Bukan LONGCLAW dan prediksi benar bukan LONGCLAW = 58+2+13 + 2+61+1 + 15+0+57 = 209

Total data : 58 + 16 + 17 + 209 = 300

4) *ORIOLE*

TP : Prediksi benar ORIOLE = 57
 FP : Prediksi ORIOLE tapi salah = 13 (dari AURIOLE) + 1 (dari EASTME) + 5 (dari LONGCLAW) = 19
 FN : Seharusnya ORIOLE tapi salah prediksi = 15 (ke AURIOLE) + 0 (ke EASTME) + 3 (ke LONGCLAW) = 18
 TN : Bukan ORIOLE dan prediksi benar bukan ORIOLE = 58+2+2+ 2+61+11 + 2+10+58 = 206

Total data : 57 + 19 + 18 + 206 = 300

Hasil akurasi menunjukkan 78%, yang berarti pengujian pada model adalah valid. Perbandingan antara perhitungan python dengan perhitungan manual pada confusion matrix juga menunjukkan hasil yang konsisten, sehingga semua dataset testing berhasil diklasifikasikan oleh model. Jika dianalisis lebih lanjut, terdapat beberapa nilai kesalahan prediksi yang relatif besar dibanding kelas lain. Hal tersebut disebabkan oleh ciri morfologi burung yang sangat mirip, oleh karena penelitian ini berfokus pada klasifikasi berdasarkan kemiripan morfologi, maka akan dilakukan subset analisis lanjutan dengan mengelompokkan dua kelas sebagaimana yang telah dijabarkan pada metodologi penelitian.

D. Pengujian (*I. graduacauda* – *O.larvatus*)

Confusion Matrix yang didapatkan pada klasifikasi *Icterus Graduacauda* (*Auriole*) dan *Oriolus Larvatus* (*Oriole*) setelah menyaring dua kelas lain dapat dilihat pada Tabel V berikut:

TABEL V
CONFUSION MATRIX AURIOLE-ORIOLE

		Prediksi	
		Auriole	Oriole
Data Asli	Auriole	58	13
	Oriole	15	57

Berdasarkan tabel, terdapat 143 gambar yang terklasifikasi dalam kelas auriole dan oriole dengan 115 prediksi benar dan 28 prediksi salah. Sehingga untuk mendapatkan nilai akurasi prediksi dari kedua kelas ini adalah dengan perhitungan berikut:

$$\frac{58 + 57}{143} \times 100\% = 80,42\%$$

Didapatkan nilai akurasi prediksi sebesar 80,42% dengan *F1-score* atau keseimbangan antara *precision* dan *recall* pada kelas auriole sebesar 81% dan kelas oriole sebesar 80% di mana nilai skor kedua kelas mendekati satu sama lain, artinya performa model seimbang di kedua kelas (Gambar 13).

	precision	recall	f1-score	support
auriole	0.79	0.82	0.81	71
oriole	0.81	0.79	0.80	72

Gambar 13. Matrik Auriole–Oriole

Nilai *precision* oriole yang lebih tinggi menunjukkan bahwa model relatif yakin saat mengklasifikasi gambar sebagai oriole dibanding auriole. Nilai *recall* yang lebih tinggi pada auriole berarti model lebih “sensitif” terhadap auriole daripada oriole, dalam artian lebih banyak auriole yang terdeteksi dengan benar. Sedangkan *F1-score* merupakan rata-rata harmonik antara *precision* dan *recall*, dalam hal ini nilai skor kedua kelas mendekati satu sama lain, artinya performa model seimbang di kedua kelas.

E. Pengujian (*S. Magna* – *M. Croceus*)

Confusion Matrix yang didapatkan pada klasifikasi *Sturnella Magna* (*Eastme*) dan *Macronyx Croceus* (*Longclaw*) setelah menyaring dua kelas lain dapat dilihat pada Tabel VI berikut:

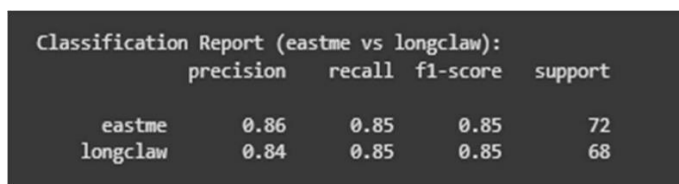
TABEL VI
CONFUSION MATRIX EASTME-LONGCLAW

		Prediksi	
		Eastme	Longclaw
Data Asli	Eastme	61	11
	Longclaw	10	58

Terdapat 140 gambar yang terklasifikasi dalam kelas eastme dan longclaw dengan 119 prediksi benar dan 21 prediksi salah. Sehingga untuk mendapatkan nilai akurasi prediksi dari kedua kelas ini adalah dengan perhitungan berikut:

$$\frac{61 + 58}{140} \times 100\% = 85\%$$

Didapatkan nilai akurasi prediksi sebesar 85% dengan F1-score pada kedua kelas sebesar 85% yang berarti performa model juga seimbang di kedua kelas.



	precision	recall	f1-score	support
eastme	0.86	0.85	0.85	72
longclaw	0.84	0.85	0.85	68

Gambar 13. Matrik Eastme-Longclaw

Nilai precision eastme yang lebih tinggi menunjukkan bahwa model relatif yakin saat mengklasifikasi gambar sebagai eastme dibanding longclaw. Nilai recall dan F1-score yang sama tinggi menandakan bahwa model tidak hanya mampu mengenali sebagian besar data dari kelas tersebut, tetapi juga melakukannya dengan tingkat ketepatan yang seimbang, sehingga performa klasifikasi pada kedua kelas ini dapat terbilang stabil.

F. Hasil Pengujian

Berdasarkan hasil pengujian yang telah diperoleh dari penelitian ini, model CNN dengan arsitektur MobileNetV2 mampu melakukan klasifikasi empat spesies burung dengan tingkat akurasi prediksi sebesar 78%, mendekati hipotesis awal penelitian. Namun, performa model menunjukkan peningkatan ketika klasifikasi difokuskan pada pasangan spesies dengan morfologi serupa.

Dengan keempat kelas dibagi menjadi dua kelompok subset pengujian untuk spesies burung dengan warna morfologi serupa akibat evolusi konvergen, tingkat akurasi prediksi mengalami kenaikan, dengan catatan beberapa support atau data yang tidak terklasifikasi pada masing-masing kelompok dihilangkan. Pada pasangan I, *graduacauda*–*O. larvatus*, tingkat akurasi naik hingga 80,42% dengan nilai F1-score untuk auriole 81% dan 80% untuk oriole di mana kedua skor saling mendekati, sehingga performa model dalam melakukan klasifikasi dapat dikatakan seimbang.

Pada kelompok kedua *S. Magna*–*M. Croceus*, tingkat akurasi naik hingga 85% dengan kedua nilai F1-score yang sama sebesar 85%, menandakan performa model dapat melakukan klasifikasi dengan stabil dan seimbang. Dalam penelitian ini yang berfokus pada klasifikasi burung dengan warna morfologi serupa, maka hasil penelitian yang dijadikan acuan adalah analisa dari pembagian kedua kelompok.

Dengan capaian tingkat akurasi prediksi pada kedua kelompok sebesar 80,42% dan 85%, maka hasil ini positif dengan hipotesis awal yang sudah dirumuskan. Model mampu mencapai tingkat akurasi hingga $\geq 80\%$ serta dapat mengklasifikasikan spesies dengan warna morfologi serupa dengan hanya sedikit kesalahan, meskipun dataset memiliki variasi posisi objek dan latar belakang secara acak.

IV. PENUTUP

Penelitian ini membuktikan bahwa model Convolutional Neural Network (CNN) dengan arsitektur MobileNetV2 mampu memenuhi tujuan yang telah dirumuskan dalam pendahuluan, yaitu melakukan klasifikasi spesies burung

berdasarkan kemiripan morfologi. Hasil pengujian menunjukkan bahwa model tidak hanya mampu mengenali empat spesies burung secara umum, tetapi juga menunjukkan performa yang semakin baik ketika diuji pada kelompok spesies dengan ciri morfologi yang sangat mirip akibat evolusi konvergen. Hal ini menunjukkan bahwa pendekatan yang digunakan dalam penelitian ini relevan dan efektif untuk mengatasi tantangan visual dalam klasifikasi gambar burung.

Kompabilitas antara tujuan awal dan hasil yang dicapai menunjukkan bahwa metode yang diterapkan dapat dijadikan dasar untuk pengembangan sistem identifikasi otomatis di bidang konservasi, pemantauan keanekaragaman hayati, atau edukasi lingkungan. Prospek penerapan penelitian juga cukup luas, terutama untuk aplikasi berbasis perangkat mobile yang membutuhkan model ringan namun akurat.

Untuk pengembangan lebih lanjut, disarankan dilakukan studi perbandingan dengan berbagai arsitektur CNN lainnya untuk mengevaluasi konsistensi performa serta mengidentifikasi arsitektur yang paling sesuai dan tahan terhadap keragaman visual dalam dataset, termasuk variasi latar belakang dan posisi objek. Selain itu, pengumpulan dataset yang lebih besar dan representatif juga direkomendasikan untuk meningkatkan generalisasi model dalam skenario dunia nyata.

REFERENSI

- [1] A. E. Morales *et al.*, "Distinct Genes with Similar Functions Underlie Convergent Evolution in Myotis Bat Ecomorphs," *Molecular Biology and Evolution*, vol. 41, no. 9, pp. 1-34, 2024, doi: 10.1093/molbev/msae165.
- [2] U. S. Rahmadhani and N. L. Marpaung, "Klasifikasi Jamur Berdasarkan Genus Dengan Menggunakan Metode CNN," *Jurnal Informatika: Jurnal pengembangan IT (JPIT)*, vol. 8, no. 2, pp. 169-173, 2023, doi: 10.30591/jpit.v8i2.5229.
- [3] U. Khultsum and G. Taufik, "Komparasi Kinerja DenseNet 121 dan MobileNet untuk Klasifikasi Citra Penyakit Daun Kentang," *JURIKOM (Jurnal Riset Komputer)*, vol. 10, no. 2, pp. 558-565, 2023, doi: 10.30865/jurikom.v10i2.6047.
- [4] C. Cakra, *et al.*, "Analisis Kesegaran Ikan Mujair dan Ikan Nila Dengan Metode Convolutional Neural Network," *Jurnal Sistem Informasi dan Teknik Komputer*, vol. 7, no. 2, pp. 74-79, 2022, doi: 10.51876/simtek.v7i2.138.
- [5] R. Tadjudin and D. Rosmala, "Implementasi Mobilenetv2 Dan Frame Difference untuk Penentuan Kecepatan Kendaraan," *JITTER*, vol. 7, no. 3, pp. 193-204, 2021.
- [6] R. Gunawan, *et al.*, "Klasifikasi Penyakit Daun Kentang dengan Transfer Learning Menggunakan CNN Optimalisasi Arsitektur MobileNetV2 Potato Leaf Disease Classification with Transfer Learning Using CNN Optimization of MobileNetV2 Architecture," *Jurnal Computer Science and Information Technology (CoSciTech)*, vol. 6, no. 2, pp. 254-258, 2025, doi: 10.37859/coscitech.v6i2.8599.
- [7] M. M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," *Computation*, vol. 11, no. 3, pp. -23, 2023, doi: 10.3390/computation11030052.
- [8] Y. A. Sitorus, Y. A. Sari, and S. Adinugroho, "Evaluasi Komparatif Arsitektur Lightweight CNN, MobileNetV2, dan EfficientNetB0 dalam Deteksi Penyakit Daun Jagung," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 8, pp. 2548-964, 2025, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [9] M. Sandler, *et al.*, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
- [10] H. Farman, S. Ahmed, M. Imran, Z. Noureen, and M. Ahmed, "Deep Learning Based Bird Species Identification and Classification Using

- Images.” *Journal of Computing & Biomedical Informatics*, vol. 06, no. 01, 2023, doi: 10.56979/601/2023.
- [11] D. D. Parsaulian, *et al.*, “Klasifikasi Empat Tanaman Obat Menggunakan Arsitektur MobileNetV2,” *Jurnal TIMES*, vol. 13, no. 2, pp. 135-141, 2024, doi: 10.51351/jtm.13.2.2024780
- [12] P. F. Sarah and N. Matondang, “Deteksi Citra Digital Penyakit Cacar Monyet menggunakan Algoritma Convolutional Neural Network dengan Arsitektur MobileNetV2,” *Jurnal Ilmu Komputer Agri-Informatika*, vol. 10, pp. 200–211, 2023, doi: 10.29244/jika.10.2.200-211
- [13] W. N. Kencana, R. Umar, and Murinto, “Implementasi Transfer Learning Untuk Klasifikasi Jenis Ras Ayam Menggunakan Arsitektur MobileNetV2,” *JIP (Jurnal Informatika Polinema)*, vol. 11, no. 2, 2025.
- [14] S. A. Maulana, *et al.*, “Penerapan Metode CNN (Convolutional Neural Network) Dalam Mengklasifikasi Jenis Ubur-Ubur,” *Jurnal Penelitian Rumpun Ilmu Teknik*, vol. 2, no. 4, pp. 122–130, 2023, doi: 10.55606/juprit.v2i4.3084.
- [15] Y. H. Puspita and A. Sabri, “Transfer Learning Model Pralatih MobileNetV2 dan DenseNet121 untuk Klasifikasi Tanaman Rempah,” *Jurnal Ilmiah Komputasi*, vol. 23, no. 1, 2024, doi: 10.32409/jikstik.23.1.3502.