

Klasifikasi Citra Topeng Bali Berdasarkan Karakteristik Visual Menggunakan Arsitektur EfficientNetV2

I Gusti Agung Putu Mahendra¹, Afis Julianto², Agus Tedyyyana³

¹Program Studi Teknik Informatika, D4-Rekayasa Perangkat Lunak, Politeknik Negeri Bengkalis

²Program Studi Teknik Informatika, D3-Teknik Informatika, Politeknik Negeri Bengkalis

³Program Studi Teknik Informatika, D2-Administrasi Jaringan Komputer, Politeknik Negeri Bengkalis

Jl. Bathin Alam-Sungai Alam, Bengkalis, Riau

e-mail: agungmahendra@polbeng.ac.id.¹, afisjulianto@polbeng.ac.id², agustedyyana@polbeng.ac.id³

Received : Juli, 2025

Accepted : Desember 2025

Published : Desember, 2025

Abstract

Bali's rich cultural heritage, particularly in the form of traditional mask art, holds profound historical and philosophical value, necessitating digital technology-based preservation efforts. One challenge in this context is the identification and classification of Balinese masks, which still relies on manual observation and limited expertise. This study examines the problem of classifying Balinese pajegan mask images, which exhibit similar visual characteristics between classes, making manual identification difficult. To address this issue, a deep learning approach based on the EfficientNetV2 architecture, incorporating transfer learning and data augmentation strategies, was applied. The dataset comprised 350 images randomly and equally divided into training (70%), validation (15%), and test (15%). Experiments were conducted using the Google Colaboratory platform with the TensorFlow-Keras framework. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. The test results showed that the model achieved maximum accuracy on the test data with consistent evaluation metric values across all classes. The main contribution of this study is to evaluate the effectiveness of the EfficientNetV2 architecture in cultural image classification with a limited dataset size, while providing a critical analysis of its high performance achievements to improve scientific validity and methodological transparency.

Keywords: image classification, Balinese mask, deep learning, EfficientNetV2, data augmentation

Abstrak

Kekayaan budaya Bali, khususnya dalam bentuk seni topeng tradisional, menyimpan nilai historis dan filosofis yang tinggi sehingga perlu didukung oleh upaya pelestarian berbasis teknologi digital. Salah satu tantangan dalam konteks ini adalah proses identifikasi dan klasifikasi topeng Bali yang masih bergantung pada pengamatan manual dan keahlian terbatas. Penelitian ini mengkaji permasalahan klasifikasi citra topeng pajegan Bali yang memiliki kemiripan karakteristik visual antarkelas sehingga menyulitkan proses identifikasi secara manual. Untuk mengatasi permasalahan tersebut, diterapkan pendekatan deep learning berbasis arsitektur EfficientNetV2 dengan strategi transfer learning dan data augmentation. Dataset yang digunakan terdiri dari 350 citra yang terbagi secara acak dan seimbang ke dalam data latih (70%), validasi (15%), dan data uji (15%). Eksperimen dilakukan menggunakan platform Google Colaboratory dengan framework TensorFlow-Keras. Evaluasi performa model dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil pengujian menunjukkan bahwa model mencapai akurasi maksimum pada data uji dengan nilai metrik evaluasi yang konsisten pada seluruh kelas. Kontribusi utama penelitian ini adalah mengevaluasi efektivitas arsitektur EfficientNetV2 dalam klasifikasi

citra budaya dengan dataset berukuran terbatas, sekaligus memberikan analisis kritis terhadap capaian performa tinggi untuk meningkatkan validitas ilmiah dan transparansi metodologi.

Kata Kunci: klasifikasi citra, topeng Bali, deep learning, EfficientNetV2, augmentasi data

1. PENDAHULUAN

Indonesia dikenal sebagai salah satu negara dengan kekayaan budaya paling beragam di dunia. Di antara berbagai warisan budayanya, topeng Bali menempati posisi istimewa karena memadukan nilai-nilai filosofis, spiritual, dan estetika yang tercermin melalui ekspresi wajah, warna, bentuk, serta ornamen hiasannya[1].

Pelestarian topeng Bali kepada generasi muda masih menghadapi kendala, terutama minimnya dokumentasi digital dan ketiadaan sistem klasifikasi yang terstruktur. Akibatnya, proses identifikasi jenis topeng—misalnya topeng keras, bondres, atau raksasa—kerap bergantung pada pengetahuan lisan para seniman dan pemerhati budaya yang jumlahnya semakin terbatas[2].

Citra digital pada dasarnya terdiri dari kumpulan data biner yang membedakan satu objek dengan objek lainnya[3]. Salah satu pendekatan yang umum digunakan dalam analisis ini adalah pendekatan komputasi, (Artificial Intelligence/AI), khususnya Deep Learning (DL), menawarkan solusi inovatif bagi pelestarian budaya visual[4]. Di bidang pengenalan gambar, arsitektur Convolutional Neural Network (CNN) modern telah menunjukkan performa luar biasa.

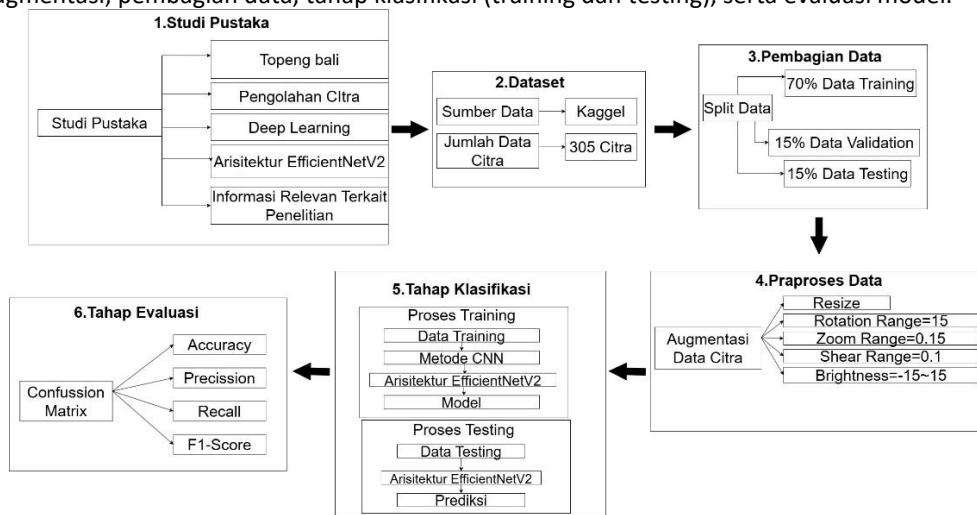
Berdasarkan Penelitian terdahulu, dimulai dari penelitian dengan judul “Klasifikasi Topeng Cirebon Menggunakan Metode Convolutional Neural Network” [5]Topeng Cirebon merupakan salah satu warisan budaya takbenda di Indonesia. Ini adalah salah satu aset budaya unggulan yang dimana pada penelitian ini menggunakan 5 kelas topeng dengan total 1010 citra setiap topeng Cirebon memiliki karakteristik penokohan yang unik dan terkadang kompleks seperti sifat manusia. Hasil eksperimen juga dibandingkan dengan hasil menggunakan first order statistical feature menggunakan KNN menghasilkan 74.00% dan GLCM menggunakan SVM menghasilkan 57.03% dan CNN optimizer yang digunakan adalah adam dengan epoch 100

menghasilkan 99.48%. Metode CNN dengan transfer learning memberikan pengenalan yang lebih baik dibandingkan dengan GLCM, ataupun First Order Feature. Selanjutnya adalah penelitian dengan judul “Klasifikasi Topeng Pandawa dengan SVM”[6] jumlah data citra pada penilitian ini berjumlah 500 citra dengan rasio 250 data untuk *trainingset* dan 250 data untuk *testingset*. Hasil pengujian menggunakan metode SVM Radial Basis Function (RBF) memperoleh akurasi 82.00%. Selanjutnya terdapat penelitian dengan judul “Analisis Klasifikasi Citra Karakteristik Topeng Bali Menggunakan Model InceptionV3 Dan MobileNetV2” [7]pada penelitian ini menggunakan dataset yang sama yaitu topeng bali yang menghasilkan akurasi 88.57% untuk model InceptionV3 dan 74.28% untuk model MobileNetV2 secara keseluruhan metode InceptionV3 dan MobileNetV2 berhasil melakukan klasifikasi namun masih bisa ditingkatkan dengan menerapkan EfficientNetV2, sebagai salah satu arsitektur CNN mutakhir yang dikembangkan Google Research, dirancang dengan teknik compound scaling sehingga mampu mencapai akurasi tinggi dengan kebutuhan komputasi lebih efisien dibandingkan model-model pendahulunya[8]. Walaupun transfer learning memungkinkan pemanfaatan bobot pre-trained EfficientNetV2 pada dataset kecil, tantangan jumlah data tetap signifikan[9]. Koleksi digital topeng Bali yang tersedia umumnya terbatas, tidak seimbang antarkelas, dan bervariasi dalam kualitas pencahayaan maupun sudut pengambilan gambar.

Oleh karena itu, penelitian ini menekankan strategi seperti augmentasi data konvensional (rotasi, flip, penyesuaian kecerahan) dan fine-tuning selektif untuk memaksimalkan kinerja model meski dengan data yang relatif terbatas[10]. Dengan demikian, hasil penelitian diharapkan tidak hanya memperkaya khazanah ilmu pengetahuan di bidang pengolahan citra budaya, tetapi juga mendorong pelestarian warisan budaya Nusantara melalui pemanfaatan teknologi AI yang tepat guna

2. METODE PENELITIAN

Alur penelitian ditunjukkan pada Gambar 1 yang mencakup tahapan: pengumpulan data, pra-pemrosesan dan augmentasi, pembagian data, tahap klasifikasi (training dan testing), serta evaluasi model.



Gambar 1. Alur Penelitian

2.1 Pengumpulan Data

Dataset yang digunakan adalah dataset topeng bali yang dapat diakses pada platform kaggle pada url: <https://www.kaggle.com/datasets/suryapradana/balinese-mask>.

No	Nama Topeng	Gambar
1	Bujuh	
2	Dalem	
3	Keras	
4	Penasar	
5	Sidakarya	
6	Tua	
7	Wijil	

Dataset yang digunakan dalam penelitian ini terdiri dari kumpulan citra topeng pajegan Bali yang dikategorikan ke dalam tujuh kelas, yaitu: topeng bujuh, topeng dalem, topeng keras, topeng penasar, topeng sidikarya, topeng tua, dan topeng wijil. Setiap kelas memuat sebanyak 50 gambar, sehingga total keseluruhan dataset berjumlah 350 citra. Beberapa contoh visual dari topeng-topeng tersebut ditampilkan pada Tabel 1. Secara kasat mata, sejumlah topeng memiliki kemiripan bentuk, seperti antara topeng dalem dengan topeng keras, serta topeng penasar dengan topeng wijil. Meskipun tampak serupa, masing-masing topeng merepresentasikan karakter dan penokohan yang berbeda dalam pertunjukan. Oleh karena itu, dibutuhkan pendekatan komputasional yang mampu mengidentifikasi dan membedakan secara objektif serta akurat antara jenis-jenis topeng tersebut.

2.2 Pembagian Data

Langkah terakhir sebelum memasuki tahap klasifikasi adalah membagi data menjadi data *training*, *validasi*, dan *testing*. Dimulai dengan membagi 70% dari dataset menjadi data untuk proses *training*, 15% untuk data *testing* dan 15 % data validasi.

2.3 Prapemrosesan Data

Dalam penelitian ini, beberapa teknik augmentasi yang digunakan dapat dijelaskan sebagai berikut.

1. *Preprocessing_function=efficientnet_preprocess*, menyesuaikan rentang piksel sesuai skema pra-proses bawaan EfficientNetV2 .
2. *rotation_range=15*, citra diputar hingga $\pm 15^\circ$.
3. *zoom_range=0.15*, melakukan pembesaran atau pengecilan citra sampai 15 %.
4. *shear_range=0.1*, menggeser sudut citra hingga 10 % secara diagonal.
5. *Brightness*, data citra akan teraumentasi secara acak dengan limit brightness yang digunakan antara -15 hingga 15.
6. *horizontal_flip=True*, membalik citra secara horizontal secara acak.
7. *fill_mode='reflect'*, piksel kosong akibat transformasi diisi dengan pantulan piksel terdekat sehingga tepi gambar tampak alami [11], [12].

Contoh Hasil Augmentasi untuk Kelas 'dalem'



Gambar 2. Hasil Augmentasi Citra

2.4 Tahap Klasifikasi

Tahap klasifikasi merupakan tahapan utama dalam penelitian ini dan ditampilkan secara eksplisit pada Gambar 2.1 sebagai Tahap Klasifikasi. Tahapan ini terdiri atas dua proses utama, yaitu proses training (pelatihan model) dan proses testing (pengujian model), sebagaimana ditunjukkan pada diagram alur.

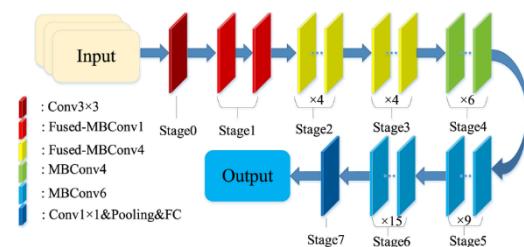
a. Proses Training

Pada proses training, data citra yang telah melalui tahap pra-pemrosesan dan augmentasi digunakan sebagai data training. Data tersebut diproses menggunakan metode Convolutional

Neural Network (CNN) dengan arsitektur EfficientNetV2 untuk membangun model klasifikasi.

EfficientNetV2 merupakan salah satu arsitektur CNN yang dirancang untuk menangani tugas pengenalan dan klasifikasi citra dengan efisiensi komputasi dan akurasi yang tinggi. Arsitektur ini merupakan penyempurnaan dari EfficientNet yang dikembangkan oleh Google Research. Dibandingkan pendahulunya, EfficientNetV2 menawarkan peningkatan signifikan dalam kecepatan pelatihan serta efisiensi sumber daya komputasi tanpa mengorbankan performa akurasi[13].

Salah satu keunggulan utama EfficientNetV2 adalah penerapan compound scaling, yaitu teknik penskalaan yang menyeimbangkan kedalaman jaringan (depth), lebar jaringan (width), dan resolusi input citra (resolution) secara optimal [14]. Selain itu, EfficientNetV2 memperkenalkan penggunaan blok Fused-MBConv, yang merupakan pengembangan dari blok MBConv pada versi sebelumnya. Blok ini mengintegrasikan beberapa operasi konvolusi ke dalam satu struktur yang lebih efisien, sehingga mempercepat proses pelatihan model. Struktur arsitektur EfficientNetV2 ditunjukkan pada Gambar 3.



Gambar 3. Arsitektur EfficientNetV2

Dalam penelitian ini, EfficientNetV2 diterapkan menggunakan pendekatan transfer learning, di mana model dasar yang telah dilatih (pre-trained) pada dataset ImageNet digunakan sebagai feature extractor. Bobot hasil pelatihan awal tersebut dimanfaatkan untuk mengekstraksi fitur visual umum, kemudian disesuaikan (fine-tuning) agar mampu mengenali karakteristik visual spesifik citra topeng Bali [15].

1. EfficientNetV2 (tanpa lapisan atas):

Bagian inti model yang bertugas mengekstraksi fitur-fitur penting dari citra masukan.

2. Global Average Pooling 2D:

Setelah melalui ekstraksi fitur, hasil dari *feature map* dua dimensi diringkas menjadi satu dimensi melalui operasi global *average pooling*. Teknik ini mengurangi jumlah parameter dan membantu menghindari *overfitting*.

3. Dropout Layer:

Untuk meningkatkan kemampuan generalisasi dan mengurangi *overfitting*, ditambahkan *dropout layer* dengan rasio 0,2 (20%). Lapisan ini secara acak mematikan sejumlah neuron selama pelatihan.

4. Dense Layer dengan Aktivasi *Softmax*:

Lapisan akhir terdiri dari neuron sebanyak jumlah kelas (kategori topeng Bali) dengan fungsi aktivasi *softmax*. Fungsi ini menghasilkan probabilitas untuk masing-masing kelas dan memungkinkan model melakukan klasifikasi multikelas.

Model ini kemudian dikompilasi menggunakan optimizer Adam dengan *learning rate* sebesar 0.00001 dan menggunakan fungsi *loss categorical crossentropy*[16]. Parameter utama yang digunakan pada tahap training dirangkum dalam Tabel 2.

Tabel 2. Parameter

Parameter	Nilai
Arsitektur	EfficientNetV2
Learning Rate	0.00001
Optimizer	Adam
Dropout	0.2
Epoch	30
Loss Function	Categorical Crossentropy

b. Proses Testing

Setelah proses training menghasilkan sebuah model klasifikasi, tahap selanjutnya adalah proses testing sebagaimana ditampilkan pada bagian bawah Tahap Klasifikasi pada Gambar 2.1. Pada tahap ini, data testing yang tidak pernah digunakan selama pelatihan dimasukkan ke dalam model EfficientNetV2 yang telah dilatih sebelumnya.

Model kemudian menghasilkan prediksi kelas untuk setiap citra uji berdasarkan probabilitas tertinggi yang dihasilkan oleh fungsi aktivasi softmax. Hasil prediksi inilah yang selanjutnya digunakan dalam proses

evaluasi performa model menggunakan confusion matrix dan metrik evaluasi klasifikasi.

2.5 Evaluasi Model

Confusion matrix merupakan matriks yang berisi informasi tentang nilai aktual dan prediksi dari sebuah klasifikasi. Confusion matrix sangat berguna untuk mengukur *Recall*, *Precision*, *Accuracy* dan *F1-Score* [17]. Gambar 4. menunjukkan contoh *confusion matrix* untuk pengklasifikasi dengan multikelas.

		Predicted		
		CLASS A	CLASS B	CLASS C
Actual	CLASS A	TP_A	FP_B	FP_C
	CLASS B	FN_A	TP_B	FP_C
	CLASS C	FN_A	FN_B	TP_C

Gambar 4. *Confusion matrix* multikelas

Adapun perhitungan confusion matrix dalam permasalahan klasifikasi sebagai berikut:

1. *Recall (Sensitivity)*

Recall menyatakan dari semua kelas positif, seberapa banyak model memprediksi dengan benar. *Recall* dapat dihitung dengan persamaan (1) berikut.

$$Recall = \frac{(TP)}{(TP)+(FN)} \quad (1)$$

2. *Precision*

Precision menyatakan bahwa dari semua kelas positif yang telah diprediksi model, seberapa banyak yang sebenarnya positif. *Precision* dapat dihitung dengan persamaan (2) berikut.

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

3. *Accuracy*

Accuracy (error rate) merupakan tingkat prediksi benar yang dihasilkan oleh model atas kumpulan data. *Accuracy* menyatakan seberapa banyak data yang diprediksi benar dan tidak benar dari keseluruhan data atau seberapa dekat nilai prediksi yang dihasilkan model

dengan nilai aktual. *Accuracy* dapat dihitung menggunakan persamaan (3) berikut.

$$\text{Accuracy} = \frac{(TP + TN)}{TP + FP + FN + TN} \quad (3)$$

4. *F1-Score*

Sulit untuk membandingkan model dengan *precision* yang rendah dan *recall* tinggi atau sebaliknya. Sehingga untuk membuatnya sebanding, dapat digunakan *F-Score*. *F1-score* membantu mengukur *recall* dan *precision* pada saat yang bersamaan (Narkhede, 2018). *F-Score* dapat dihitung menggunakan persamaan (4) berikut.

$$F - \text{measure} = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \quad (4)$$

Keterangan:

True Positive (TP) : Data bernilai positif yang diprediksi positif.

False Positive (FP) : Data bernilai negatif yang diprediksi positif.

True Negative (TN) : Data bernilai negatif yang diprediksi negatif.

False Negative (FN) : Data bernilai positif yang diprediksi negatif.

3. HASIL DAN PEMBAHASAN

Eksperimen dalam penelitian ini dilaksanakan pada platform Google Colaboratory dengan memanfaatkan library Keras berbasis Python untuk pengembangan dan pelatihan model deep learning. Evaluasi performa dilakukan menggunakan data uji yang telah dipisahkan sebelumnya, dengan total **53 citra uji** (15% dari 350 citra). Nilai akurasi diperoleh berdasarkan perhitungan confusion matrix pada masing-masing konfigurasi epoch.

3.1 Hasil Akurasi Model

Tabel 2 menyajikan nilai akurasi yang diperoleh menggunakan arsitektur EfficientNetV2 dengan learning rate sebesar 0,00001 pada dua konfigurasi epoch yang berbeda.

Tabel 2. Hasil Akurasi

Epoch	Akurasi
25	97.00%
30	100%

3.2 Akurasi pada Epoch ke-25

Akurasi dihitung menggunakan persamaan berikut:

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Seluruh Data Uji}} \times 100$$

Berdasarkan confusion matrix pada epoch ke-25 (Gambar 5), diperoleh hasil sebagai berikut:

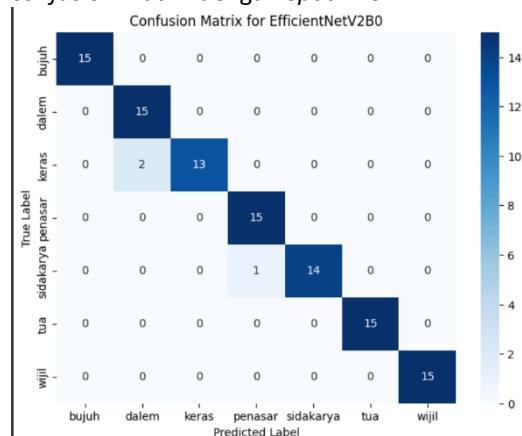
- Jumlah data uji = **53 citra**
- Jumlah prediksi benar = **51 citra**
- Jumlah prediksi salah = **2 citra**

Dengan demikian, perhitungan akurasi secara manual adalah:

$$\text{Akurasi}_{\text{epoch 25}} = \frac{51}{53} \times 100\% = 96,23$$

Nilai tersebut dibulatkan oleh sistem evaluasi Keras menjadi **97.00%**, sebagaimana ditampilkan pada hasil eksperimen dan Tabel 2. Kesalahan klasifikasi terjadi pada kelas *keras* dan *sidakarya*, yang memiliki kemiripan karakteristik visual dengan kelas lain, sehingga mempengaruhi hasil prediksi model pada tahap ini.

Berdasarkan dari table diatas percobaan dengan nilai *epoch* 25 dilihat dari gambar *confusion matrix* masih ada kesalahan pengklasifikasian berikut dibawah ini gambar 5. *confusion matrix* dengan *epoch* 25.



Gambar 5. *Confusion Matrix* Epoch 25

Berdasarkan confusion matrix yang ditampilkan, sebagian besar kelas berhasil diklasifikasikan dengan sempurna. Lima kelas—yakni bujuh, dalem, penasar, tua, dan wijil—mencapai akurasi 100%, yang berarti seluruh citra dalam masing-masing kelas berhasil dikenali dengan tepat oleh model. Sementara itu, dua kelas lainnya, yaitu keras dan sidakarya, menunjukkan adanya kesalahan klasifikasi. Untuk kelas keras, 13 citra berhasil diprediksi dengan benar, sedangkan 2 citra lainnya salah diklasifikasikan

sebagai topeng dalem. Pada kelas sidakarya, terdapat satu kesalahan klasifikasi di mana satu citra diprediksi sebagai topeng penasar. Kemungkinan besar, kesalahan ini disebabkan oleh kemiripan karakteristik visual antar jenis topeng, baik dari ekspresi wajah, warna dominan, hingga detail ornamen, yang menjadi tantangan dalam klasifikasi citra budaya.

3.3 Akurasi pada Epoch ke-30

Pada konfigurasi epoch ke-30, confusion matrix (Gambar 6) menunjukkan bahwa seluruh data uji berhasil diklasifikasikan dengan benar tanpa adanya kesalahan prediksi.

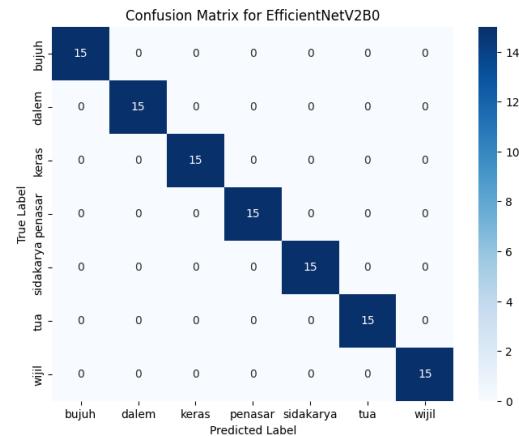
- Jumlah data uji = 53 citra
- Jumlah prediksi benar = 53 citra
- Jumlah prediksi salah = 0 citra

Sehingga perhitungan akurasi manual pada epoch ke-30 adalah sebagai berikut:

$$\text{Akurasi}_{\text{epoch 30}} = \frac{53}{53} \times 100\% = 100$$

Hasil ini menunjukkan bahwa model telah mencapai performa maksimum pada data uji setelah proses pelatihan hingga epoch ke-30.

Gambar 6 memperlihatkan confusion matrix yang dihasilkan dari model terbaik pada epoch ke-30.



Gambar 6. Confusion Matrix Epoch 30

Model berhasil mencapai akurasi sempurna sebesar 100%, sebagaimana ditunjukkan pada confusion matrix yang memperlihatkan tidak adanya kesalahan klasifikasi. Seluruh gambar pada data uji, masing-masing terdiri dari 15 sampel per kelas, berhasil diklasifikasikan dengan tepat. Capaian ini secara deskriptif menunjukkan efektivitas kombinasi antara arsitektur model dan penerapan strategi *data augmentation*. Teknik *data augmentation* terbukti mampu meningkatkan kinerja dan kemampuan generalisasi model *deep learning*

secara signifikan, terutama dalam konteks tugas klasifikasi citra [18]. Data augmentation memanfaatkan teknik seperti transformasi spasial dan spektral untuk mensintesis variasi data yang realistik, sehingga memperkuat ketahanan model terhadap overfitting dan memperluas cakupan fitur yang dapat dipelajari [19]. Dampak positif dari data augmentation sangat menonjol ketika diterapkan pada dataset dengan jumlah terbatas. Teknik ini secara signifikan dapat meningkatkan akurasi prediksi dan memperkuat ketahanan model terhadap variasi data yang belum pernah dilihat sebelumnya. Selain itu, data augmentation juga berfungsi sebagai bentuk regularisasi yang efektif, mencegah model dari sekadar menghafal data pelatihan, dan mendorong kemampuan model untuk mengenali pola secara lebih umum, terutama dalam konteks klasifikasi citra dengan jumlah data yang relatif terbatas.[20].

3.4 Perbandingan dengan penelitian terdahulu

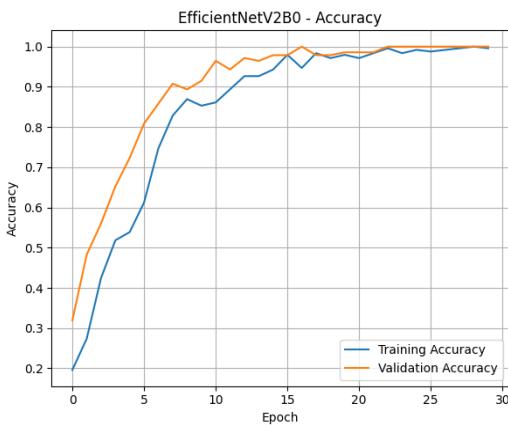
Berdasarkan perbandingan yang disajikan pada Tabel 3, dapat dilihat bahwa usulan penelitian dengan menggunakan arsitektur EfficientNetV2 menunjukkan kinerja yang lebih unggul dibandingkan beberapa penelitian terdahulu yang mengkaji klasifikasi topeng tradisional berbasis citra.

Tabel 3. Perbandingan Dengan Penlitian Terdahulu

Judul	Metode	Hasil
Klasifikasi Topeng Pandawa dengan	SVM	82.00%
Klasifikasi Topeng Cirebon Menggunakan Metode	CNN	74.00%
Analisis Klasifikasi Topeng Bali	InceptionV3 Dan MobileNetV2	88,57% Dan 74,28%
Usulan Penelitian	EfficientNetV2	100%

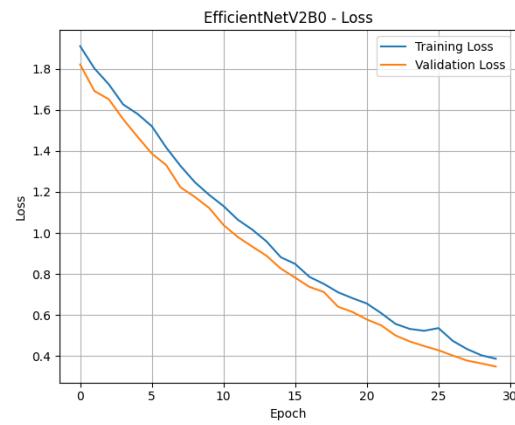
3.5 Pembahasan

Grafik akurasi menggambarkan bagaimana kemampuan model dalam menebak kelas dengan benar meningkat seiring waktu. Garis biru (*training accuracy*) dan garis oranye (*validation accuracy*) keduanya menunjukkan tren positif yang signifikan.



Gambar 7. Grafik Accuracy

Grafik akurasi menggambarkan bagaimana kemampuan model dalam menebak kelas dengan benar meningkat seiring waktu. Garis biru (*training accuracy*) dan garis oranye (*validation accuracy*) keduanya menunjukkan tren positif yang signifikan. Pada awal pelatihan, kedua akurasi dimulai dari nilai yang sangat rendah, sekitar 20%, yang menandakan model belum memiliki pengetahuan apa-apa. Seiring bertambahnya jumlah epoch, kurva akurasi menunjukkan peningkatan yang signifikan. Menarik untuk dicermati bahwa akurasi pada data validasi (ditandai dengan warna oranye) secara konsisten sedikit lebih tinggi dibandingkan dengan akurasi pada data pelatihan (warna biru) di sebagian besar tahap pelatihan. Fenomena ini seringkali mengindikasikan penerapan teknik augmentasi data yang efektif pada data pelatihan, yang membuat data tersebut lebih kompleks dibandingkan data validasi yang cenderung lebih sederhana. Pada tahap akhir pelatihan, kedua kurva menunjukkan pola konvergen dengan nilai mendekati 1.0 atau 100%, menandakan bahwa model tidak hanya mampu mempelajari pola pada data pelatihan, tetapi juga memiliki kemampuan generalisasi yang sangat baik terhadap data validasi.



Gambar 8. Grafik Loss

Grafik loss merepresentasikan tingkat "kesalahan" model; semakin rendah nilainya, semakin baik performa model. Garis biru (Training Loss) dan garis oranye (Validation Loss) keduanya dimulai dari nilai yang tinggi, sekitar 1.9, dan secara konsisten menurun tajam di awal, kemudian melandai seiring berjalannya waktu.

Penurunan yang stabil pada kedua kurva loss ini mengonfirmasi bahwa model terus-menerus memperbaiki dirinya dan menjadi lebih baik dalam membuat prediksi. Kurva *training loss* dan *validation loss* saling mengikuti dengan sangat dekat, tanpa ada tanda-tanda salah satu kurva mulai naik kembali. Ini adalah ciri utama dari model yang tidak mengalami *overfitting*, di mana model mampu belajar tanpa hanya "menghafal" data latih. Model dilatih dengan 25 dan 30 epoch. Pada epoch ke-25, model mencapai akurasi sebesar 97%. Pada epoch ke-30, model mencapai akurasi maksimum pada data uji. Confusion matrix menunjukkan bahwa seluruh citra pada data uji berhasil diklasifikasikan dengan benar pada konfigurasi terbaik. Nilai precision, recall, dan F1-score untuk seluruh kelas mencapai nilai maksimum. Berikut tabel 4. menunjukkan rangkuman hasil evaluasi citra topeng bali.

Tabel 4. Hasil Confusion Matrix

Kelas	Precision	Recall	F1-score
bujuh	100%	100%	100%
dalem	100%	100%	100%
keras	100%	100%	100%
penasar	100%	100%	100%
sidakarya	100%	100%	100%
tua	100%	100%	100%
wijil	100%	100%	100%
Akurasi		100%	

Berdasarkan Tabel 3, model EfficientNetV2 menghasilkan nilai precision, recall, dan F1-score sebesar 100% pada seluruh kelas topeng Bali, dengan akurasi keseluruhan data uji mencapai 100%. Data uji yang digunakan dalam evaluasi ini berjumlah 53 citra, yang diperoleh dari 15% total dataset, dengan distribusi per kelas yang relatif seimbang (sekitar 7–8 citra per kelas). Capaian ini menunjukkan bahwa model mampu mengklasifikasikan seluruh citra uji secara tepat pada konfigurasi pelatihan terbaik (epoch ke-30).

Meskipun demikian, hasil dengan nilai evaluasi maksimum perlu ditafsirkan secara kritis. Ukuran data uji yang relatif terbatas berpotensi meningkatkan peluang tercapainya akurasi sempurna, terutama ketika karakteristik visual citra berada dalam kondisi yang relatif terkontrol, seperti latar belakang, pencahayaan, dan sudut pengambilan gambar. Dalam skenario ini, model memiliki kemungkinan lebih besar untuk mempelajari pola visual yang konsisten pada dataset dibandingkan menghadapi variasi ekstrem yang umum dijumpai pada data dunia nyata.

Selain itu, untuk meminimalkan potensi data leakage, proses pembagian data dilakukan sebelum penerapan augmentasi, sehingga citra hasil augmentasi hanya diterapkan pada data latih dan tidak muncul pada data validasi maupun data uji. Dengan demikian, seluruh citra pada data uji bersifat independen dan tidak pernah digunakan dalam proses pelatihan model. Strategi ini bertujuan menjaga validitas evaluasi dan memastikan bahwa performa yang diperoleh benar-benar mencerminkan kemampuan model dalam melakukan generalisasi pada data yang belum pernah dilihat sebelumnya.

Tingginya performa model juga dipengaruhi oleh kompleksitas arsitektur EfficientNetV2 yang mampu mengekstraksi fitur visual secara mendalam. Pada konteks citra topeng Bali, karakteristik visual seperti bentuk wajah, ekspresi mata, warna dominan, serta ornamen khas berperan sebagai fitur pembeda yang cukup kuat antar kelas. Ketika dikombinasikan dengan pendekatan transfer learning dan teknik augmentasi data, model memiliki kapasitas yang memadai untuk mempelajari fitur-fitur diskriminatif tersebut secara optimal.

Namun demikian, capaian ini belum sepenuhnya menjamin kemampuan generalisasi model pada kondisi yang lebih kompleks, seperti variasi pencahayaan ekstrem, latar belakang yang berbeda, atau citra hasil pengambilan di lingkungan luar. Oleh karena itu, penelitian lanjutan disarankan untuk melakukan evaluasi tambahan menggunakan skema cross-validation, pengujian pada dataset eksternal, serta analisis interpretabilitas model (misalnya Grad-CAM) guna memastikan bahwa keputusan klasifikasi benar-benar didasarkan pada karakteristik topeng, bukan pada pola non-ensial di sekitar citra.

4. KESIMPULAN

Penelitian ini berhasil menjawab tantangan dalam pengenalan dan klasifikasi citra topeng pajegan Bali yang memiliki kemiripan visual antarkelas. Dengan memanfaatkan arsitektur EfficientNetV2 dan pendekatan transfer learning, serta didukung oleh teknik data augmentation, model mampu mencapai akurasi 100% pada tahap pengujian. Hasil evaluasi dengan menggunakan metrik presisi, recall, dan f1-score menunjukkan bahwa model mampu mengenali dan membedakan setiap kelas dengan sangat akurat. Temuan ini membuktikan bahwa metode deep learning modern dapat diterapkan secara efektif pada dataset berukuran kecil untuk tugas klasifikasi multikelas citra budaya. Keterbatasan penelitian terletak pada ukuran dan variasi dataset yang relatif terbatas. Penelitian lanjutan disarankan untuk menguji model pada dataset yang lebih besar, tidak seimbang, dan kondisi pencahayaan yang beragam guna meningkatkan generalisasi model.

DAFTAR PUSTAKA

- [1] I. G. M. D. S. Putra, I. W. Gulendra, dan N. M. P. Utami, "BALINESE MASKS AS A SOURCE OF INSPIRATION FOR IDEAS IN THE CREATION OF PAINTING WORKS," *CITA KARA : JURNAL PENCPTAAN DAN PENGKAJIAN SENI MURNI*, vol. 5, no. 1, hlm. 73–79, Apr 2025, doi: 10.59997/ctkr.v5i1.4493.
- [2] I. W. Suardana, "STRUKTUR RUPA TOPENG BALI KLASIK," *Imaji*, vol. 4, no. 1, Nov 2015, doi: 10.21831/imaji.v4i1.6703.
- [3] H. Hartatik dan M. K. Anam, "Comparison of Convolutional Neural

- Network Architecture on Detection of Helmet Use by Humans," *Elinvo (Electronics, Informatics, and Vocational Education)*, vol. 8, no. 1, hlm. 44–54, Jun 2023, doi: 10.21831/elinvo.v8i1.52104.
- [4] H. A. Pratiwi, M. Cahyanti, dan M. Lamsani, "IMPLEMENTASI DEEP LEARNING FLOWER SCANNER MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK," *Sebatik*, vol. 25, no. 1, hlm. 124–130, Jun 2021, doi: 10.46984/sebatik.v25i1.1297.
- [5] F. I. Kurniadi, V. Kemala Putri, dan Y. E. Wibawa, "Klasifikasi Topeng Cirebon Menggunakan Metode Convolutional Neural Network," 2021. [Daring]. Tersedia pada: <http://jurnal.mdp.ac.id>
- [6] A. Sanjaya, E. Setyati, dan H. Budianto, "Klasifikasi Topeng Pandawa dengan SVM."
- [7] J. Elektronik Ilmu Komputer Udayana dkk., "Analisis Klasifikasi Citra Karakteristik Topeng Bali Menggunakan Model InceptionV3 Dan MobileNetV2", [Daring]. Tersedia pada: <https://www.kaggle.com/datasets/suryapradana/balinese-mask>.
- [8] M. Hayat, N. Ahmad, A. Nasir, dan Z. Ahmad Tariq, "Hybrid Deep Learning EfficientNetV2 and Vision Transformer (EffNetV2-ViT) Model for Breast Cancer Histopathological Image Classification," *IEEE Access*, vol. 12, hlm. 184119–184131, 2024, doi: 10.1109/ACCESS.2024.3503413.
- [9] D. P. Sidik, F. Utaminingrum, dan L. Muflikhah, "Penggunaan Variasi Model pada Arsitektur EfficientNetV2 untuk Prediksi Sel Kanker Serviks," 2023. [Daring]. Tersedia pada: <http://j-ptiik.ub.ac.id>
- [10] D. Putri dan A. Dkk, "AUGMENTASI DATA PADA IMPLEMENTASI CONVOLUTIONAL NEURAL NETWORK ARSITEKTUR EFFICIENTNET-B3 UNTUK KLASIFIKASI PENYAKIT DAUN PADI."
- [11] T. Kumar, R. Brennan, A. Mileo, dan M. Bendechache, "Image Data Augmentation Approaches: A Comprehensive Survey and Future Directions," *IEEE Access*, vol. 12, hlm. 187536–187571, 2024, doi: 10.1109/ACCESS.2024.3470122.
- [12] C. Shorten dan T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J Big Data*, vol. 6, no. 1, Des 2019, doi: 10.1186/s40537-019-0197-0.
- [13] M. R. M. Kamal, S. Shahbudin, dan F. Y. A. Rahman, "Photovoltaic (PV) Module Defect Image Classification Analysis Using EfficientNetV2 Architectures," dalam *2023 IEEE 14th Control and System Graduate Research Colloquium (ICSGRC)*, IEEE, Agu 2023, hlm. 236–241. doi: 10.1109/ICSGRC57744.2023.10215491.
- [14] L. Jiang, S. Zhu, dan N. Sun, "An Improved Lightweight Variant of EfficientNetV2 Coupled With Sensor Fusion and Transfer Learning Techniques for Motor Fault Diagnosis," *IEEE Access*, vol. 12, hlm. 84470–84487, 2024, doi: 10.1109/ACCESS.2024.3412050.
- [15] H.-J. Jeon dan J. W. Jeon, "Efficient Training of EfficientNetV2-S Using AdaBelief Optimizer," dalam *2022 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, IEEE, Okt 2022, hlm. 1–3. doi: 10.1109/ICCE-Asia57006.2022.9954726.
- [16] A. Tariq, M. M. Iqbal, M. Javed Iqbal, dan Iftikhar Ahmad, "Transforming Brain Tumor Detection Empowering Multi-Class Classification With Vision Transformers and EfficientNetV2," *IEEE Access*, vol. 13, hlm. 63857–63876, 2025, doi: 10.1109/ACCESS.2025.3555638.
- [17] M. Heydarian, T. E. Doyle, dan R. Samavi, "MLCM: Multi-Label Confusion Matrix," *IEEE Access*, vol. 10, hlm. 19083–19095, 2022, doi: 10.1109/ACCESS.2022.3151048.
- [18] C. Shorten dan T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J Big Data*, vol. 6, no. 1, hlm. 60, Des 2019, doi: 10.1186/s40537-019-0197-0.
- [19] F. J. Moreno-Barea, J. M. Jerez, dan L. Franco, "Improving classification accuracy using data augmentation on small data sets," *Expert Syst Appl*, vol. 161, hlm. 113696, Des 2020, doi: 10.1016/j.eswa.2020.113696.
- [20] J. Lemley, S. Bazrafkan, dan P. Corcoran, "Smart Augmentation Learning an

Optimal Data Augmentation Strategy,"
IEEE Access, vol. 5, hlm. 5858–5869,
2017, doi:
10.1109/ACCESS.2017.2696121.