

# KLASIFIKASI TOKOH WAYANG KULIT DALAM CERITA MAHABHARATA VERSI PEWAYANGAN JAWA MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) BERBASIS CONVNEXT-SMALL

Faozan Ganes Setiawan, Eddy Muntina Dharma, Nengah Widya Utami

Sistem Informasi, Teknologi Informasi & Design Universitas Primakara

Jl. Tukad Badung No. 135, Renon, Denpasar Selatan, Kota Denpasar, Bali 80226, Indonesia

fauzanganes04@gmail.com

## ABSTRAK

Wayang kulit merupakan warisan budaya Indonesia yang memiliki nilai filosofis tinggi, namun kompleksitas visual dan kemiripan antar tokoh dalam cerita Mahabharata sering menyulitkan masyarakat, khususnya generasi muda, dalam mengenalinya. Permasalahan utama dalam penelitian ini adalah rendahnya kemampuan manusia membedakan tokoh wayang secara visual serta belum optimalnya pemanfaatan arsitektur CNN modern untuk klasifikasi tokoh wayang Mahabharata. Penelitian ini bertujuan membangun dan mengevaluasi model klasifikasi citra tokoh wayang kulit Mahabharata versi pewayangan Jawa menggunakan Convolutional Neural Network (CNN) dengan arsitektur ConvNeXt-Small, serta membandingkannya dengan ResNet-18 sebagai model baseline. Metode yang digunakan meliputi pendekatan transfer learning, pelatihan dua tahap (head-only dan fine-tuning), augmentasi data offline, serta evaluasi menggunakan Stratified K-Fold Cross Validation. Dataset yang digunakan berasal dari Kaggle “Indonesian Wayang (Traditional Puppet Show)” yang terdiri dari 6.576 citra 22 tokoh dan diperluas menjadi 13.200 citra. Hasil penelitian menunjukkan bahwa model ConvNeXt-Small mampu mencapai akurasi 59,50%, macro precision 64,42%, macro recall 59,50%, dan macro F1-score 57,70%, serta menunjukkan kemampuan generalisasi yang lebih baik dibandingkan ResNet-18, meskipun masih mengalami kesulitan pada tokoh dengan kemiripan visual tinggi.

**Kata kunci :** Wayang Kulit, Mahabharata, Klasifikasi citra, CNN, ConvNeXt-Small, ResNet-18

## 1. PENDAHULUAN

Indonesia merupakan negara yang kaya akan keberagaman budaya yang diwariskan secara turun-temurun dan terus dilestarikan hingga saat ini. Salah satu warisan budaya tersebut adalah kesenian wayang, yang sejak awal berkembang sebagai media pertunjukan berbasis bayangan dan berfungsi sebagai sarana spiritual serta penyampaian nilai-nilai kehidupan [1]. Wayang terbagi menjadi beberapa jenis, di antaranya wayang orang dan wayang boneka, dengan wayang kulit sebagai salah satu bentuk yang paling dikenal. Wayang kulit memiliki ciri visual khas yang dibuat dari kulit dengan berbagai ornamen dan simbol filosofis yang mendalam [2]. Cerita yang dibawakan umumnya berasal dari kisah klasik seperti Ramayana dan Mahabharata yang telah mengalami proses akulturasi budaya Hindu-India dengan budaya Jawa. Tokoh-tokoh dalam cerita Mahabharata, seperti Pandawa dan Kurawa, memiliki karakter visual yang khas sekaligus kompleks, sehingga kerap sulit dibedakan, terutama oleh generasi muda [1].

Permasalahan muncul ketika kompleksitas visual dan kemiripan antar tokoh wayang kulit menyebabkan rendahnya kemampuan masyarakat, khususnya generasi muda, dalam mengenali dan membedakan karakter wayang. Setiap tokoh memang memiliki ciri visual dan makna simbolis tersendiri, namun kesamaan atribut, pola, dan warna sering kali menimbulkan kebingungan. Kondisi ini diperparah oleh menurunnya minat generasi muda terhadap seni tradisional, yang disebabkan oleh keterbatasan bahasa dalam pertunjukan wayang serta dominasi teknologi

modern seperti gawai dalam kehidupan sehari-hari [3]. Selain itu, ketertarikan generasi muda terhadap budaya asing turut menggeser perhatian terhadap warisan budaya lokal, sehingga pengenalan tokoh wayang menjadi semakin terbatas [4]. Oleh karena itu, diperlukan pendekatan berbasis teknologi yang mampu membantu mengenali dan mengklasifikasikan tokoh wayang kulit secara akurat.

Salah satu solusi yang dapat diterapkan adalah pemanfaatan teknik klasifikasi citra berbasis computer vision dan pembelajaran mesin. Melalui pengolahan citra digital, karakteristik visual tokoh wayang seperti struktur wajah, busana, dan ornamen dapat dianalisis secara sistematis [5]. Dalam bidang image classification, Convolutional Neural Network (CNN) merupakan metode yang banyak digunakan karena kemampuannya dalam mengekstraksi fitur visual dan mengenali pola kompleks pada citra secara otomatis [6]. CNN juga mampu memodelkan hubungan nonlinier pada data citra sehingga memberikan hasil klasifikasi yang lebih akurat dibandingkan metode konvensional [7].

Beberapa penelitian sebelumnya telah membuktikan efektivitas CNN dalam klasifikasi citra wayang kulit dengan berbagai arsitektur, seperti ResNet, MobileNet, dan DenseNet, yang mampu mencapai tingkat akurasi tinggi pada objek tertentu [8]–[10]. Namun demikian, sebagian besar penelitian tersebut masih berfokus pada jenis wayang tertentu atau kelompok tokoh terbatas, serta belum banyak mengeksplorasi arsitektur CNN modern seperti ConvNeXt untuk klasifikasi tokoh wayang

Mahabharata secara menyeluruh. Selain itu, perbandingan performa antara ConvNeXt dan arsitektur CNN yang lebih ringan seperti ResNet-18 dalam konteks wayang kulit masih jarang dibahas.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk membangun dan mengevaluasi model klasifikasi tokoh wayang kulit dalam cerita Mahabharata versi pewayangan Jawa menggunakan Convolutional Neural Network dengan arsitektur ConvNeXt-Small, serta membandingkan performanya dengan ResNet-18 sebagai model pembanding. Rencana penyelesaian masalah dilakukan melalui penerapan pendekatan transfer learning, augmentasi data citra, serta evaluasi model menggunakan metrik akurasi, precision, recall, F1-score, dan confusion matrix. Dataset yang digunakan berasal dari platform Kaggle dan mencakup 22 tokoh wayang Mahabharata. Dengan pendekatan ini, diharapkan model ConvNeXt-Small mampu memberikan kemampuan klasifikasi dan generalisasi yang lebih baik, sehingga dapat mendukung pengembangan media edukasi digital dan upaya pelestarian budaya wayang kulit secara modern dan berkelanjutan.

## 2. TINJAUAN PUSTAKA

### 2.1. Wayang dan Wayang Mahabharata

Wayang kulit merupakan kesenian tradisional yang menjadi identitas budaya Indonesia. Meskipun cerita yang dibawakan banyak diadaptasi dari kisah India, pertunjukan wayang di Jawa berkembang dengan gaya tutur dan pesan kontekstual yang relevan dengan kondisi masyarakat pada masanya [8]. Asal-usul wayang masih diperdebatkan, namun catatan tertua mengenai pertunjukan wayang ditemukan pada prasasti abad ke-9 di Kerajaan Mataram Kuno dan terus berkembang hingga masa Majapahit [9]. Karena nilai historis dan budayanya, UNESCO menetapkan wayang sebagai warisan budaya dunia [10].

Mahabharata merupakan kisah suci Hindu yang kemudian diadaptasi ke dalam pewayangan Jawa sebagai bagian dari proses akulturasi budaya. Kisah ini dianggap sebagai Weda kelima dalam kepercayaan Hindu dan mengisahkan konflik antara keluarga Pandawa dan Kurawa. Adaptasi Mahabharata diterima luas oleh masyarakat Jawa dan berkembang menjadi bagian penting dalam pertunjukan wayang kulit [11]. Tokoh-tokoh Pandawa seperti Yudhistira, Bima, Arjuna, Nakula, dan Sadewa menjadi representasi utama dalam penelitian ini.

### 2.2. Klasifikasi

Klasifikasi merupakan proses pengelompokan objek ke dalam kelas tertentu berdasarkan karakteristik yang dimiliki. Proses ini bertujuan membentuk fungsi pemetaan dari vektor fitur ke label kelas yang telah ditentukan [12]. Model klasifikasi bekerja layaknya *black box* yang memproses input berdasarkan pola yang dipelajari dan menghasilkan prediksi kelas sebagai output [13].

### 2.3. Artificial Intelligence dan Machine Learning

Artificial Intelligence (AI) adalah bidang ilmu komputer yang berfokus pada pengembangan sistem yang mampu melakukan tugas-tugas yang membutuhkan kecerdasan manusia, seperti pengambilan keputusan dan pembelajaran [14]. AI juga didefinisikan sebagai sistem yang mampu memahami lingkungan, mempelajari data, dan beradaptasi untuk mencapai tujuan tertentu [15]. Salah satu cabang utama AI adalah Machine Learning, yaitu metode yang memungkinkan sistem belajar dari data tanpa diprogram secara eksplisit [16], [17].

### 2.4. Computer Vision

Computer vision merupakan bidang yang memungkinkan komputer memahami dan menafsirkan informasi visual dari citra atau video secara otomatis, dengan meniru cara kerja penglihatan manusia [18]. Teknologi ini berkaitan erat dengan pengolahan citra dan sering dikombinasikan dengan jaringan saraf tiruan untuk meningkatkan akurasi pengenalan pola visual [18].

### 2.5. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) adalah pengembangan dari Multilayer Perceptron (MLP) yang termasuk dalam jaringan saraf *feed forward* dan banyak digunakan untuk pengolahan citra [19], [20]. CNN memproses data melalui convolution layer, pooling layer, dan fully connected layer untuk mengekstraksi serta mengklasifikasikan fitur visual secara bertahap [21].

### 2.6. ConvNeXt dan ResNet-18

ConvNeXt merupakan arsitektur CNN modern yang dikembangkan dari ResNet dengan pendekatan desain yang terinspirasi dari vision transformer. Model ini mengadopsi pembaruan pada struktur layer, normalisasi, dan efisiensi komputasi sehingga mampu menghasilkan performa tinggi dalam tugas klasifikasi citra [22], [23].

ResNet-18 adalah arsitektur residual network dengan 18 lapisan utama yang menggunakan *shortcut connection* untuk menjaga stabilitas pelatihan dan efisiensi komputasi, sehingga cocok digunakan sebagai model pembanding (baseline) [24], [25].

### 2.7. Python

Python merupakan bahasa pemrograman tingkat tinggi yang memiliki sintaks sederhana dan fleksibel. Bahasa ini mendukung berbagai paradigma pemrograman dan banyak digunakan dalam pengembangan sistem berbasis kecerdasan buatan dan pembelajaran mesin [26], [27].

### 2.8. Evaluasi Model

Confusion matrix digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan label aktual dan hasil prediksi, serta menjadi dasar

perhitungan metrik evaluasi seperti precision, recall, dan F1-score [28].

Akurasi mengukur persentase prediksi yang benar terhadap keseluruhan data uji [29]. Precision menunjukkan ketepatan prediksi kelas positif, sedangkan recall mengukur kemampuan model dalam mengenali data positif yang sebenarnya [2]. F1-score merupakan rata-rata harmonik dari precision dan recall yang memberikan evaluasi seimbang terhadap performa model, terutama pada data yang tidak seimbang [30].

## 2.9. Penelitian terdahulu

Penelitian terdahulu menunjukkan bahwa CNN efektif digunakan dalam klasifikasi citra wayang kulit dengan berbagai arsitektur seperti VGG-16, ResNet, EfficientNet, dan MobileNetV2 [2], [4], [5], [6], [30], [31]. Beberapa penelitian juga membuktikan bahwa optimasi dan teknik *deep learning* modern mampu meningkatkan akurasi klasifikasi citra budaya secara signifikan [7], [1].

Namun demikian, sebagian besar penelitian sebelumnya belum secara spesifik mengkaji klasifikasi tokoh wayang dalam cerita Mahabharata serta belum mengeksplorasi arsitektur CNN modern seperti ConvNeXt. Oleh karena itu, penelitian ini mengusulkan penggunaan ConvNeXt sebagai model utama dengan ResNet-18 sebagai pembanding, guna meningkatkan efisiensi dan akurasi klasifikasi citra tokoh wayang kulit Mahabharata versi pewayangan Jawa. Penelitian ini diharapkan dapat berkontribusi pada pengembangan teknologi pelestarian budaya berbasis kecerdasan buatan.

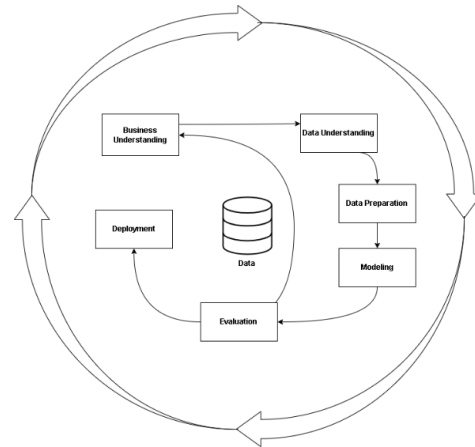
## 3. METODE PENELITIAN

Penelitian ini menggunakan metode eksperimen kuantitatif untuk mengevaluasi kinerja model Convolutional Neural Network (CNN) dengan arsitektur ConvNeXt dalam mengklasifikasikan citra tokoh wayang kulit. Proses eksperimen meliputi tahap preprocessing citra berupa *resize*, normalisasi piksel, pelabelan data, serta augmentasi data untuk menyeimbangkan jumlah citra pada setiap kelas. Dataset kemudian dibagi menjadi data latih dan data uji dengan beberapa rasio pembagian (80:20, 70:30, dan 60:40) guna menguji konsistensi performa model.

Penelitian ini menerapkan metodologi Cross Industry Standard Process for Data Mining (CRISP-DM) yang terdiri dari enam tahap utama [33]. Namun, ruang lingkup penelitian dibatasi hingga tahap Evaluation, tanpa mencakup proses *deployment* [34]. Tahapan CRISP-DM yang digunakan meliputi:

- Business Understanding, untuk merumuskan tujuan dan permasalahan penelitian.
- Data Understanding, untuk memahami karakteristik dan kualitas dataset citra wayang.
- Data Preparation, mencakup pembersihan, transformasi, dan persiapan data.

- Modeling, yaitu pelatihan model CNN ConvNeXt dengan pengaturan parameter tertentu.
- Evaluation, untuk menilai performa model menggunakan metrik evaluasi yang telah ditentukan.



Gambar 1. Metode CRISP-DM [32]

Pendekatan ini memastikan proses penelitian berjalan sistematis serta menghasilkan model klasifikasi citra wayang kulit yang akurat dan konsisten.

## 4. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengevaluasi kinerja arsitektur ConvNeXt-Small dalam melakukan klasifikasi tokoh wayang kulit Mahabharata versi pewayangan Jawa. Evaluasi dilakukan melalui beberapa skenario, yaitu pelatihan tanpa cross validation, penerapan Test-Time Augmentation (TTA), penggunaan Stratified K-Fold Cross Validation, serta perbandingan dengan model baseline ResNet-18.

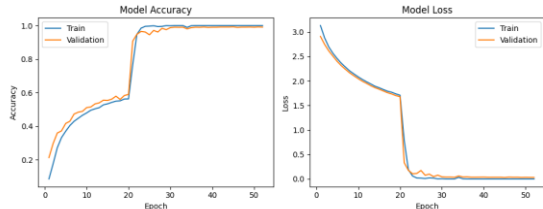
### 4.1. Hasil Pelatihan Model ConvNeXt-Small

Model ConvNeXt-Small dilatih menggunakan pendekatan *transfer learning* dengan bobot awal ImageNet. Proses pelatihan dilakukan secara bertahap, dimulai dengan pelatihan pada *classification head*, kemudian dilanjutkan dengan *fine-tuning* seluruh backbone untuk menyesuaikan model terhadap karakteristik visual tokoh wayang kulit.

Tabel 1. Parameter Pelatihan Model ConvNeXt-Small

| Parameter               | Nilai                            |
|-------------------------|----------------------------------|
| Epoch (Head-only)       | 20                               |
| Epoch (Fine-tuning)     | 200                              |
| Ukuran Citra            | 224 × 224                        |
| Batch Size              | 16                               |
| Optimizer (Head)        | Adam ( $lr = 3 \times 10^{-4}$ ) |
| Optimizer (Fine-tuning) | Adam ( $lr = 3 \times 10^{-5}$ ) |
| Loss Function           | Sparse Categorical Crossentropy  |
| Dropout                 | 0.2                              |

Tabel ini berisi konfigurasi parameter pelatihan yang digunakan pada model ConvNeXt-Small, meliputi jumlah epoch, learning rate, batch size, optimizer, dan fungsi loss. Parameter ini dirancang untuk menjaga keseimbangan antara kecepatan konvergensi dan stabilitas pembelajaran.

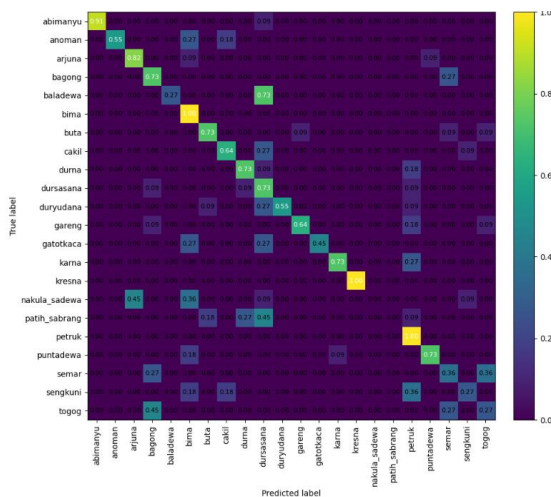


Gambar 2. Grafik Akurasi dan Loss Model ConvNeXt-Small

Gambar ini menunjukkan perubahan nilai akurasi dan loss pada data pelatihan dan validasi selama proses training. Terlihat bahwa akurasi meningkat secara konsisten hingga mencapai kondisi stabil, sementara nilai loss mengalami penurunan, yang mengindikasikan proses pembelajaran berjalan dengan baik dan model tidak mengalami overfitting yang signifikan.

#### 4.2. Evaluasi Model ConvNeXt-Small

Evaluasi performa model dilakukan menggunakan metrik precision, recall, dan F1-score, serta divisualisasikan melalui confusion matrix untuk melihat distribusi kesalahan klasifikasi antar kelas.



Gambar 3. Confusion Matrix Model ConvNeXt-Small

Gambar ini menggambarkan hasil klasifikasi model ConvNeXt-Small terhadap data uji. Dominasi nilai pada diagonal utama menunjukkan bahwa sebagian besar citra tokoh wayang berhasil diklasifikasikan dengan benar. Kesalahan klasifikasi masih ditemukan pada beberapa kelas yang memiliki kemiripan visual tinggi, seperti Nakula-Sadewa dan tokoh pendukung dengan atribut visual serupa.

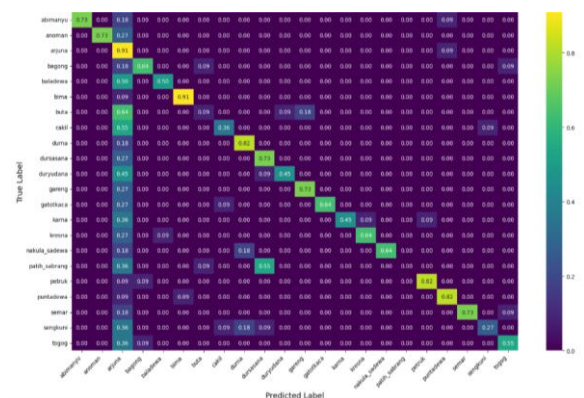
Tabel 2. Hasil Evaluasi Model ConvNeXt-Small

| Kelas         | Precision | Recall | F1-score | Accuracy |
|---------------|-----------|--------|----------|----------|
| Abimanyu      | 1.00      | 0.91   | 0.95     | 0.91     |
| Anoman        | 1.00      | 0.55   | 0.71     | 0.55     |
| Arjuna        | 0.64      | 0.82   | 0.72     | 0.82     |
| Bagong        | 0.44      | 0.73   | 0.55     | 0.73     |
| Baladewa      | 1.00      | 0.27   | 0.43     | 0.27     |
| Bima          | 0.42      | 1.00   | 0.59     | 1.00     |
| Buta          | 0.73      | 0.73   | 0.73     | 0.73     |
| Cakil         | 0.64      | 0.64   | 0.64     | 0.64     |
| Durna         | 0.67      | 0.73   | 0.70     | 0.73     |
| Dursasana     | 0.24      | 0.27   | 0.36     | 0.27     |
| Duryudana     | 1.00      | 0.55   | 0.71     | 0.55     |
| Gareng        | 0.88      | 0.64   | 0.74     | 0.64     |
| Gatotkaca     | 1.00      | 0.45   | 0.63     | 0.45     |
| Karna         | 0.89      | 0.73   | 0.80     | 0.73     |
| Kresna        | 1.00      | 1.00   | 1.00     | 1.00     |
| Nakula_Sadewa | 0.00      | 0.00   | 0.00     | 0.00     |
| Patih_Tab     | 0.00      | 0.00   | 0.00     | 0.00     |
| Petruk        | 0.44      | 1.00   | 0.61     | 1.00     |
| Puntadewa     | 0.89      | 0.73   | 0.80     | 0.73     |
| Semar         | 0.36      | 0.36   | 0.36     | 0.36     |
| Sengkuni      | 1.00      | 0.27   | 0.38     | 0.27     |
| Togog         | 0.33      | 0.27   | 0.30     | 0.27     |

Tabel ini menyajikan nilai precision, recall, dan F1-score untuk setiap kelas tokoh wayang. Hasil evaluasi menunjukkan bahwa model memiliki performa tinggi pada kelas dengan ciri visual yang kuat dan konsisten, seperti Bima dan Arjuna, sementara performa sedikit menurun pada kelas dengan perbedaan visual yang tipis.

#### 4.3. Pengaruh Test-Time Augmentation (TTA)

Untuk meningkatkan kestabilan prediksi, model ConvNeXt-Small diuji menggunakan pendekatan Test-Time Augmentation. Teknik ini menghasilkan beberapa variasi citra uji yang kemudian digabungkan untuk memperoleh prediksi akhir.



Gambar 4. Confusion Matrix Model ConvNeXt-Small dengan TTA

Gambar ini menunjukkan confusion matrix hasil pengujian model dengan penerapan TTA. Terlihat adanya peningkatan prediksi benar pada beberapa kelas dibandingkan model tanpa TTA, meskipun peningkatan tersebut belum merata pada seluruh kelas.

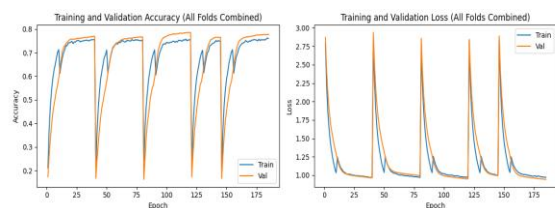
Tabel 3. Hasil Evaluasi Model ConvNeXt-Small dengan TTA

| Kelas         | Precision | Recall | F1-score | Accuracy |
|---------------|-----------|--------|----------|----------|
| Abimanyu      | 1.00      | 0.73   | 0.84     | 0.73     |
| Anoman        | 1.00      | 0.73   | 0.84     | 0.73     |
| Arjuna        | 0.17      | 0.82   | 0.28     | 0.82     |
| Bagong        | 0.57      | 0.73   | 0.64     | 0.73     |
| Baladewa      | 0.33      | 0.09   | 0.14     | 0.09     |
| Bima          | 0.64      | 0.82   | 0.72     | 0.82     |
| Buta          | 0.71      | 0.45   | 0.56     | 0.45     |
| Cakil         | 0.40      | 0.18   | 0.25     | 0.18     |
| Durna         | 0.19      | 0.64   | 0.30     | 0.64     |
| Dursasana     | 0.31      | 0.45   | 0.37     | 0.45     |
| Duryudana     | 0.44      | 0.36   | 0.40     | 0.36     |
| Gareng        | 0.55      | 0.55   | 0.55     | 0.55     |
| Gatotkaca     | 0.75      | 0.27   | 0.40     | 0.27     |
| Karna         | 0.00      | 0.00   | 0.00     | 0.00     |
| Kresna        | 1.00      | 0.36   | 0.53     | 0.36     |
| Nakula_Sadewa | 1.00      | 0.27   | 0.43     | 0.27     |
| Patih_Sabrang | 0.33      | 0.18   | 0.24     | 0.18     |
| Petruk        | 0.83      | 0.45   | 0.59     | 0.45     |
| Puntadewa     | 0.78      | 0.64   | 0.70     | 0.64     |
| Semar         | 0.67      | 0.55   | 0.60     | 0.55     |
| Sengkuni      | 0.29      | 0.18   | 0.22     | 0.18     |
| Togog         | 0.40      | 0.36   | 0.38     | 0.36     |

Tabel ini menampilkan hasil evaluasi model ConvNeXt-Small dengan TTA. Secara umum, penerapan TTA mampu meningkatkan robustnes model terhadap variasi citra, namun belum sepenuhnya mengatasi kesalahan klasifikasi pada kelas dengan tingkat kemiripan visual yang sangat tinggi.

#### 4.4. Evaluasi Menggunakan Stratified K-Fold Cross Validation

Untuk memperoleh evaluasi yang lebih stabil dan mengurangi bias pembagian data, penelitian ini menerapkan Stratified K-Fold Cross Validation dengan jumlah fold sebanyak lima.



Gambar 5. Grafik Akurasi dan Loss pada Setiap Fold

Gambar ini menunjukkan pola perubahan akurasi dan loss pada setiap fold selama proses pelatihan. Pola yang relatif konsisten pada seluruh fold menunjukkan bahwa model ConvNeXt-Small memiliki kemampuan generalisasi yang baik.

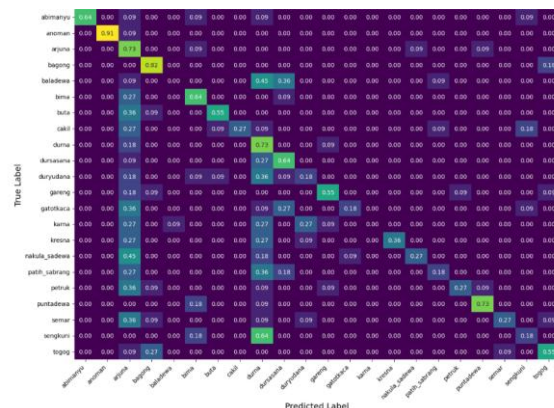
Tabel 4. Hasil Evaluasi Model ConvNeXt-Small dengan K-Fold Cross Validation

| Kelas         | Precision | Recall | F1-score | Accuracy |
|---------------|-----------|--------|----------|----------|
| Abimanyu      | 1.00      | 0.73   | 0.84     | 0.73     |
| Anoman        | 1.00      | 0.73   | 0.84     | 0.73     |
| Arjuna        | 0.13      | 0.91   | 0.23     | 0.91     |
| Bagong        | 0.78      | 0.64   | 0.70     | 0.64     |
| Baladewa      | 0.83      | 0.50   | 0.62     | 0.50     |
| Bima          | 0.91      | 0.91   | 0.91     | 0.91     |
| Buta          | 0.33      | 0.09   | 0.14     | 0.09     |
| Cakil         | 0.67      | 0.36   | 0.47     | 0.36     |
| Durna         | 0.69      | 0.82   | 0.75     | 0.82     |
| Dursasana     | 0.59      | 0.73   | 0.59     | 0.73     |
| Duryudana     | 0.83      | 0.45   | 0.59     | 0.45     |
| Gareng        | 0.80      | 0.73   | 0.76     | 0.73     |
| Gatotkaca     | 1.00      | 0.64   | 0.78     | 0.64     |
| Karna         | 0.00      | 0.00   | 0.00     | 0.00     |
| Kresna        | 0.88      | 0.64   | 0.74     | 0.64     |
| Nakula_Sadewa | 1.00      | 0.64   | 0.78     | 0.64     |
| Patih_Sabrang | 0.00      | 0.00   | 0.00     | 0.00     |
| Petruk        | 0.90      | 0.82   | 0.86     | 0.82     |
| Puntadewa     | 0.82      | 0.82   | 0.82     | 0.82     |
| Semar         | 0.60      | 0.73   | 0.64     | 0.73     |
| Sengkuni      | 0.75      | 0.27   | 0.40     | 0.27     |
| Togog         | 0.75      | 0.55   | 0.63     | 0.55     |

Tabel ini menyajikan hasil evaluasi rata-rata dari seluruh fold. Pendekatan K-Fold Cross Validation menghasilkan performa yang lebih stabil antar kelas dibandingkan pelatihan tanpa cross validation, meskipun nilai akurasi keseluruhan sedikit lebih rendah.

#### 4.5. Perbandingan dengan Model Baseline ResNet-18

Sebagai pembandingan, penelitian ini menggunakan ResNet-18 sebagai model baseline untuk mengukur efektivitas ConvNeXt-Small dalam menangani klasifikasi citra wayang kulit yang kompleks.



Gambar 6. Confusion Matrix Model ResNet-18

Gambar ini menggambarkan confusion matrix hasil pengujian model ResNet-18. Terlihat bahwa jumlah kesalahan klasifikasi lebih tinggi dibandingkan

ConvNeXt-Small, terutama pada kelas dengan jumlah data terbatas dan karakteristik visual yang mirip.

Tabel 5. Hasil Evaluasi Model ResNet-18

| Kelas         | Precision | Recall | F1-score | Accuracy |
|---------------|-----------|--------|----------|----------|
| Abimanyu      | 1.00      | 0.27   | 0.43     | 0.27     |
| Anoman        | 0.45      | 0.45   | 0.45     | 0.45     |
| Arjuna        | 0.45      | 0.45   | 0.45     | 0.45     |
| Bagong        | 0.28      | 0.73   | 0.40     | 0.73     |
| Baladewa      | 0.00      | 0.00   | 0.00     | 0.00     |
| Bima          | 0.30      | 0.55   | 0.39     | 0.55     |
| Buta          | 0.67      | 0.18   | 0.29     | 0.18     |
| Cakil         | 0.21      | 0.27   | 0.24     | 0.27     |
| Durna         | 0.24      | 0.82   | 0.38     | 0.82     |
| Dursasana     | 0.50      | 0.09   | 0.15     | 0.09     |
| Duryudana     | 0.20      | 0.73   | 0.31     | 0.73     |
| Gareng        | 0.75      | 0.27   | 0.40     | 0.27     |
| Gatotkaca     | 0.25      | 0.09   | 0.13     | 0.09     |
| Karna         | 0.00      | 0.00   | 0.00     | 0.00     |
| Kresna        | 1.00      | 0.09   | 0.17     | 0.09     |
| Nakula_Sadewa | 0.00      | 0.00   | 0.00     | 0.00     |
| Patih_Sabrang | 0.00      | 0.00   | 0.00     | 0.00     |
| Petrus        | 0.00      | 0.00   | 0.00     | 0.00     |
| Puntadewa     | 0.00      | 0.00   | 0.00     | 0.00     |
| Semar         | 0.04      | 0.09   | 0.06     | 0.09     |
| Sengkuni      | 0.17      | 0.64   | 0.27     | 0.64     |
| Togog         | 0.50      | 0.09   | 0.15     | 0.09     |

Tabel ini menunjukkan nilai precision, recall, dan F1-score dari model ResNet-18. Hasil evaluasi mengindikasikan bahwa meskipun ResNet-18 memiliki arsitektur yang lebih sederhana dan efisien, performanya masih berada di bawah ConvNeXt-Small dalam klasifikasi tokoh wayang kulit.

#### 4.6. Ringkasan Pembahasan

Berdasarkan seluruh skenario pengujian, ConvNeXt-Small menunjukkan performa terbaik dalam klasifikasi tokoh wayang kulit Mahabharata, baik dari segi akurasi maupun kestabilan prediksi. Penerapan Test-Time Augmentation meningkatkan ketahanan model terhadap variasi citra, sementara Stratified K-Fold Cross Validation memberikan evaluasi yang lebih konsisten antar kelas. Dibandingkan dengan ResNet-18, ConvNeXt-Small terbukti lebih efektif dalam mengekstraksi fitur visual kompleks pada citra wayang kulit, sehingga lebih sesuai digunakan untuk sistem pengenalan citra budaya berbasis deep learning.

#### 5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa arsitektur ConvNeXt-Small mampu digunakan secara efektif untuk klasifikasi tokoh wayang kulit Mahabharata versi pewayangan Jawa, dengan performa yang relatif baik pada kelas-kelas yang memiliki ciri visual jelas, namun masih menghadapi kendala pada tokoh dengan kemiripan visual tinggi. Dibandingkan ResNet-18, ConvNeXt-Small menghasilkan performa yang lebih stabil dan merata antar kelas, khususnya ketika dikombinasikan dengan Stratified K-Fold Cross

Validation dan Test Time Augmentation, sehingga lebih layak dijadikan arsitektur utama dalam konteks penelitian ini. Meski demikian, penelitian ini memiliki keterbatasan terkait penggunaan dataset sekunder, cakupan objek yang terbatas, belum dilibatkannya pakar wayang dan pengguna akhir, serta keterbatasan teknis pada keragaman data, strategi augmentasi, eksplorasi arsitektur, dan evaluasi generalisasi. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset primer, memperluas cakupan pewayangan, melibatkan ahli domain, mengeksplorasi arsitektur dan hyperparameter secara sistematis, serta menguji model pada data eksternal agar kinerja dan generalisasi model dapat ditingkatkan..

#### DAFTAR PUSTAKA

- [1] M. Muhathir, M. H. Santoso, and D. A. Larasati, "Wayang Image Classification Using SVM Method and GLCM Feature Extraction," *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 4, no. 2, pp. 373–382, Jan. 2021, doi: 10.31289/jite.v4i2.4524.
- [2] M. R. A. Yudianto, Kusriani, and H. Al Fatta, "ANALISIS PENGARUH TINGKAT AKURASI KLASIFIKASI CITRA WAYANG DENGAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK," 2020.
- [3] R. Generasi Muda Suku Jawa Terhadap Pemaknaan Pesan Seni Pertunjukan Wayang Kulit Di Desa Kali Cinta, I. Yogi Pratama, B. Huri, and I. Komunikasi Universitas Tulang Bawang, "RESPON GENERASI MUDA SUKU JAWA TERHADAP PEMAKNAAN PESAN SENI PERTUNJUKAN WAYANG KULIT DI DESA KALI CINTA," 2024.
- [4] H. K. Maulana, H. E. Wahanani, and M. M. Al Haromani, "PENERAPAN ARSITEKTUR CNN-EFFICIENTNETB2 DENGAN TRANSFER LEARNING PADA KLASIFIKASI GAMBAR TOKOH WAYANG KULIT," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5626.
- [5] W. Nurhasanah, W. Witanti, and H. Ashaury, "Klasifikasi Citra Pada Wayang Kulit Menggunakan Convolutional Neural Network," vol. 11, no. 1, pp. 112–119, 2025, doi: 10.33050/cerita.v11i1.3700.
- [6] A. Setya, S. Pratama, A. Prasetya Wibawa, and A. N. Handayani, "CONVOLUTIONAL NEURAL NETWORK (CNN) UNTUK MENENTUKAN GAGRAK WAYANG KULIT," 2022.
- [7] R. K. Vasanthakumari, R. V. Nair, and V. G. Krishnappa, "Improved learning by using a modified activation function of a Convolutional Neural Network in multi-spectral image classification," *Machine Learning with*

- Applications*, vol. 14, p. 100502, Dec. 2023, doi: 10.1016/j.mlwa.2023.100502.
- [8] E. Saputra, "KONTRIBUSI TOKOH PUNAKAWAN PADA PAGELARAN WAYANG KULIT TERHADAP PENDIDIKAN ISLAM KEPADA MASYARAKAT," *SAP (Susunan Artikel Pendidikan)*, vol. 6, no. 2, 2021.
  - [9] A. Sunaryo, *Rupa Wayang*. Jawa Tengah: CV Kekata Group, 2020.
  - [10] E. Setiawan, "Nilai Filosofi Wayang Kulit Sebagai Media Dakwah," 2020. [Online]. Available: <http://alhikmah.iain-jember.ac.id/>
  - [11] H. Abdul Malik, "(e) 185-200 Islamic dakwah," *Jurnal Ilmu Dakwah*, vol. 44, no. 1, pp. 185–200, 2024, doi: 10.2158/jid.44.1.22066.
  - [12] H. F. Putro, R. T. Vulandari, and W. L. Y. Saptomo, "Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan," *Jurnal Teknologi Informasi dan Komunikasi (TIKomsin)*, vol. 8, no. 2, Oct. 2020, doi: 10.30646/tikomsin.v8i2.500.
  - [13] S. Febriani, "Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5 Siska Febriani Sistem Informasi \* ) Rohmansyah@gmail.com," 2022.
  - [14] E. S. Eriana, S. Kom, M. Kom, and D. A. Zein, "ARTIFICIAL INTELLIGENCE (AI) PENERBIT CV. EUREKA MEDIA AKSARA," 2021.
  - [15] Araf Aliwijaya and Hanny Chairany Suyono, "Peluang Pemanfaatan Big Data di Perpustakaan," *Info Bibliotheca: Jurnal Perpustakaan dan Ilmu Informasi*, vol. 4, no. 2, pp. 1–17, Jun. 2023, doi: 10.24036/ib.v4i2.397.
  - [16] P. Bintoro, Ratnasar, E. Wihardjo, I. P. Putri, and A. Asari, "Pengantar Machine Learning," 2024.
  - [17] T. Thoyyibah, F. Kurniawan, and T. Taryo, "567741-dasar-dasar-machine-learning-pada-google-f62d236f," 2024.
  - [18] T. A. Dompeipen, S. R. U. A. Sompie, and M. E. I. Najoan, "Computer Vision Implementation for Detection and Counting the Number of Humans," 2021.
  - [19] P. Adi Nugroho, I. Fenriana, and R. Arijanto, "IMPLEMENTASI DEEP LEARNING MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK ( CNN ) PADA EKSPRESI MANUSIA," *JURNAL ALGOR*, vol. 2, no. 1, 2020, [Online]. Available: <https://jurnal.buddhidharma.ac.id/index.php/algor/index>
  - [20] A. Arkadia *et al.*, *Klasifikasi Buah Mangga Badami Untuk Menentukan Tingkat Kematangan dengan Metode CNN*. 2021.
  - [21] R. Yohannes and M. E. Al Rivan, "Klasifikasi Jenis Kanker Kulit Menggunakan CNN-SVM," 2022.
  - [22] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," Jan. 2022, [Online]. Available: <http://arxiv.org/abs/2201.03545>
  - [23] S. B. Yengec-Tasdemir, E. Akay, S. Dogan, and B. Yilmaz, "Classification of Colorectal Polyps from Histopathological Images using Ensemble of ConvNeXt Variants," Jul. 12, 2022. doi: 10.21203/rs.3.rs-1791422/v1.
  - [24] W. Setiawan, "Klasifikasi Citra Histopatologi Kanker Payudara menggunakan Data Resampling Random dan Residual Network," *JURNAL SISTEM INFORMASI BISNIS*, vol. 11, no. 1, pp. 70–77, Aug. 2021, doi: 10.21456/vol11iss1pp70-79.
  - [25] F. Ramzan *et al.*, "A Deep Learning Approach for Automated Diagnosis and Multi-Class Classification of Alzheimer's Disease Stages Using Resting-State fMRI and Residual Neural Networks," *J. Med. Syst.*, vol. 44, no. 2, Feb. 2020, doi: 10.1007/s10916-019-1475-2.
  - [26] M. Z. Andrekha and Y. Huda, "Deteksi Warna Manggis Menggunakan Pengolahan Citra dengan Opencv Python," 2021, [Online]. Available: <http://ejournal.unp.ac.id/index.php/voteknika/index>
  - [27] M. Reza *et al.*, "Artificial Intelligence : Image Processing & Application with Python," 2022. [Online]. Available: <http://jurnal.umj.ac.id/index.php/semnaskat>
  - [28] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," 2021.
  - [29] Y. S. Austin *et al.*, "Klasifikasi Penyakit Alzheimer Dari Scan Mri Otak Menggunakan Convnext," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 6, pp. 1223–1232, Dec. 2024, doi: 10.25126/jtiik.2024118117.
  - [30] N. E. Imandayanti, H. E. Wahanani, D. Agung, and M. Rizki, "Klasifikasi Jenis Wayang menggunakan Convolutional Neural Network (CNN) dan Optimasi Adaptive Moment Estimation (ADAM)," vol. 5, no. 2, 2024.
  - [31] W. Supriyanti and A. Anggoro, "Classification of Pandavas Figure in Shadow Puppet Images using Convolutional Neural Networks," 2021. [Online]. Available: <http://tokohwayangpurwa>.
  - [32] M. Fitriani, G. F. Nama, and M. Mardiana, "Implementasi Association Rule Dengan Algoritma Apriori Pada Data Peminjaman Buku UPT Perpustakaan Universitas Lampung Menggunakan Metodologi CRISP-DM," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 10, no. 1, Jan. 2022, doi: 10.23960/jitet.v10i1.2263