

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil implementasi, pelatihan, dan evaluasi model *Cross Vision Transformer* (*CrossViT*) dalam klasifikasi karakter wayang, dapat disimpulkan beberapa poin utama sebagai berikut:

1. Keunggulan Arsitektur *CrossViT*

Model ini menerapkan pendekatan multi-skala dengan membagi citra menjadi dua jalur patch berukuran berbeda, yaitu 16×16 piksel dan 32×32 piksel. Pendekatan ini memungkinkan model untuk menangkap pola visual dari berbagai skala, sehingga meningkatkan representasi fitur yang lebih kaya dan mendalam.

2. Performa Model Selama Pelatihan

Model dilatih selama 10 epoch menggunakan optimizer Adam dan fungsi loss categorical *crossentropy*. Hasil pelatihan menunjukkan peningkatan performa seiring waktu, dengan akurasi akhir sebesar 97,56% pada data pelatihan dan 89,12% pada data validasi. Meskipun terjadi fluktuasi pada awal pelatihan, model secara bertahap mampu meningkatkan kemampuannya dalam mengklasifikasikan gambar dengan lebih akurat.

3. Evaluasi Kinerja Model

Beberapa karakter wayang, seperti Gareng, Petruk, Kekayi, Windardi, Resi Gautama, dan beberapa karakter Kapi (Janulen, Arimenda, Arigibas, dan Suweda), menunjukkan hasil sempurna dengan precision, recall, dan f1-score mencapai 1.00, yang berarti model mampu mengenali karakter ini tanpa kesalahan. Karakter seperti Batara Narada, Batara Guru, Janaka, Kapi Anggini, Semar, Rama Bergawa, Sridenta, dan Trijaka memiliki f1-score di atas 0.95, menunjukkan bahwa model bekerja dengan baik dalam mengenali pola visual karakter-karakter ini. Beberapa karakter mengalami kesulitan dalam klasifikasi, seperti Pratalamariam (f1-score 0.59), Resi Subali (0.59), Lesmana (0.63), Ranendra (0.67), dan Saupakenaka (0.67). Precision yang rendah pada beberapa karakter ini menunjukkan bahwa model sering salah mengklasifikasikan karakter lain sebagai karakter tersebut, atau sebaliknya.

4. Analisis Kesalahan Model

Beberapa karakter menunjukkan precision yang sangat tinggi tetapi recall yang rendah, seperti Anjani (precision 1.00, recall 0.62), Batara Brahma (precision 1.00, recall 0.72), dan Batara Indra (precision 0.95, recall 0.66). Hal ini mengindikasikan bahwa model sangat selektif dalam mengklasifikasikan karakter-karakter ini, tetapi sering gagal mengenali mereka dalam beberapa kondisi tertentu. Pratalamariam dan Resi Subali memiliki precision yang rendah namun recall tinggi, yang berarti model sering mengklasifikasikan karakter lain sebagai mereka. Ini menunjukkan bahwa pola visual mereka mungkin terlalu mirip dengan karakter lain, sehingga membingungkan model dalam proses klasifikasi. Secara keseluruhan, model *CrossViT* mencapai akurasi 89% pada dataset pengujian yang terdiri dari 1.552 sampel. Nilai precision rata-rata sebesar 0.90, recall 0.88, dan f1-score 0.88 menunjukkan bahwa model ini memiliki kemampuan klasifikasi yang cukup baik. Namun, masih terdapat beberapa karakter yang perlu diperbaiki, terutama yang memiliki fitur visual mirip dengan karakter lain.

5.2. Saran

1. Penyempurnaan dataset dengan menambahkan lebih banyak variasi gambar untuk karakter yang sulit dikenali.
2. Optimasi hyperparameter lebih lanjut untuk meningkatkan akurasi model.
3. Penggunaan teknik augmentasi data untuk meningkatkan ketahanan model terhadap berbagai kondisi pencahayaan dan sudut pandang gambar.
4. Secara keseluruhan, model *Cross Vision Transformer* telah terbukti mampu mengenali sebagian besar karakter wayang dengan akurasi tinggi. Dengan beberapa penyempurnaan, model ini memiliki potensi besar untuk dikembangkan lebih lanjut agar lebih akurat dalam mengidentifikasi karakter-karakter yang lebih kompleks atau memiliki kemiripan visual tinggi.
5. Perbanyak variasi dari karakter wayang, jadi jangan hanya menggunakan 1 referensi jenis wayang pada setiap karakternya.