

BAB I

PENDAHULUAN

1.1 Latar Belakang

Warisan budaya yang kaya seperti di Jawa, Indonesia, tercermin dalam aksara yang unik, yaitu aksara Jawa, yang secara lokal dikenal dengan nama "Aksara Jawa." Aksara tradisional yang telah digunakan selama berabad-abad dalam teks-teks sastra dan sejarah ini merupakan bagian integral dari pelestarian budaya dan sejarah Jawa. Terlepas dari nilai historisnya, aksara Jawa menghadapi tantangan dalam representasi dan pelestarian digital modern [1], terutama karena struktur karakternya yang kompleks dan terbatasnya ketersediaan sumber daya digital. Kemajuan terbaru dalam *Machine Learning (ML)*, terutama *Deep Learning (DL)*, menawarkan solusi yang menjanjikan untuk pengenalan dan digitalisasi aksara tradisional secara otomatis. Teknik *DL* telah menunjukkan keberhasilan yang signifikan dalam berbagai bidang *image recognition* dan *Natural Language Processing (NLP)* [2], menjadikannya kandidat ideal untuk menangani kompleksitas aksara Jawa. Teknik-teknik ini dapat mempelajari representasi hirarkis dari data, sehingga efektif dalam mengenali pola dan struktur rumit yang melekat pada aksara Jawa.

Dalam beberapa tahun terakhir, *Convolutional Neural Networks (CNN)* telah muncul sebagai alat yang ampuh untuk tugas-tugas *image recognition*. *CNN*, sebuah kelas model *DL*, telah menunjukkan kinerja yang luar biasa dalam mengidentifikasi pola dan fitur di dalam gambar. Kemampuan *CNN* untuk secara otomatis dan adaptif mempelajari hierarki spasial fitur dari gambar input membuat *CNN* sangat cocok untuk tugas-tugas yang melibatkan data visual yang kompleks, seperti pengenalan karakter [3]. Model *CNNs* merupakan salah satu jenis jaringan syaraf tiruan *feed forward* yang memiliki kemampuan untuk mengekstraksi fitur dari data dengan struktur konvolusi dan merupakan salah satu metode yang digunakan oleh *DL* untuk menangani masalah klasifikasi citra [4]. Penerapan *CNNs* pada klasifikasi aksara Jawa sangat menjanjikan. *CNN* dapat secara efektif menangani variasi pola goresan, bentuk, dan ukuran aksara yang merupakan ciri khas dari aksara Jawa. Dengan melatih dataset besar aksara Jawa yang berlabel,

CNN dapat belajar mengidentifikasi dan mengklasifikasikan setiap aksara secara akurat, sehingga memungkinkan pengembangan sistem yang dapat diandalkan untuk digitalisasi aksara Jawa.

Beberapa karya telah menerapkan *CNN* untuk klasifikasi aksara Jawa dan menunjukkan hasil yang menjanjikan. Salah satu karya penting oleh [5] menggunakan model *deep CNN* untuk mengenali karakter aksara Jawa, mencapai akurasi 92,5% pada dataset khusus, para peneliti menyoroti pentingnya teknik augmentasi data untuk meningkatkan keragaman sampel pelatihan dan meningkatkan kemampuan generalisasi model. Karya penting lainnya oleh [6] mengeksplorasi penggunaan *transfer learning* dengan model *CNN* yang telah dilatih sebelumnya, seperti *VGG-16* dan *ResNet50*, yang disesuaikan dengan dataset karakter Jawa, pendekatan ini menghasilkan kinerja yang lebih baik dibandingkan dengan pelatihan dari awal, terutama mengingat terbatasnya ketersediaan data berlabel. Algoritma *back-propagation* untuk mengklasifikasikan karakter aksara Jawa yang berfokus pada *preprocessing* melibatkan beberapa tahap, dalam karya [7] gambar dikonversi ke *grayscale* kemudian *Canny edge detection* diterapkan untuk mendeteksi tepi pada foto, yang sangat penting untuk mengenali bentuk dan struktur aksara. Gambar-gambar tersebut berukuran 30x30 piksel untuk memastikan ukuran input yang konsisten untuk model jaringan syaraf tiruan, pendekatan ini menghasilkan sistem dengan akurasi 74% dalam mengenali karakter aksara Jawa tulisan tangan. Kemudian karya penting lainnya oleh [8], implementasi *framework KERAS* memperoleh 86,68% untuk tingkat akurasi yang menerapkan dataset 1000 gambar dan membutuhkan waktu 639,85 detik untuk mengklasifikasikan karakter aksara Jawa, efisiensi model ini membuatnya sangat cocok untuk aplikasi *mobile* dan *IoT*. Dalam karya [9] Proses *preprocessing* melibatkan pemotongan gambar menjadi ukuran 1275x877 piksel dan kemudian meresize gambar menjadi 64x64 piksel untuk mempercepat proses *CNN*, kemudian mengeksplorasi pengaruh penggunaan *dropout* dengan berbagai nilai (0, 25%, 50%, dan 75%) terhadap kinerja model. Hasil menunjukkan bahwa penggunaan *dropout* sebesar 0.75 menghasilkan akurasi testing tertinggi sebesar 73.60%.

Terlepas dari kemajuan ini, masih ada tantangan untuk mencapai akurasi dan keandalan yang tinggi dalam klasifikasi aksara Jawa. Rangkaian aksara yang beragam dan rumit, ditambah dengan variasi kontekstual dan pengikat, membutuhkan model yang kuat dan canggih. Selain itu, kelangkaan set data yang komprehensif dan beranotasi untuk aksara Jawa menjadi rintangan yang signifikan untuk melatih model *DL* secara efektif. *Hyperparameter Tuning* adalah proses penting dalam mengoptimalkan *CNN* untuk klasifikasi karakter aksara Jawa, yang dikenal sebagai aksara Jawa. Struktur aksara Jawa yang unik dan kompleks membutuhkan pemilihan dan penyesuaian *hyperparameter* yang cermat untuk mencapai akurasi yang tinggi dan kinerja yang kuat. *Hyperparameter Tuning* adalah proses mengoptimalkan parameter dalam sebuah model untuk mencapai hasil terbaik [10]. Hal ini melibatkan penyesuaian parameter seperti *optimizer*, *learning rate*, *batch normalization*, *dropout rate*, *weight decay*, *momentum*, *gradient clipping* dan jumlah lapisan dalam jaringan untuk mendapatkan konfigurasi yang menghasilkan kinerja yang optimal [10] [11].

Dalam naskah ini penulis mengusulkan dengan pendekatan baru yang fokus pada penerapan teknik *Hyperparameter Tuning* dan *preprocessing* untuk mengoptimalkan klasifikasi aksara Jawa. Dengan memanfaatkan teknik *Hyperparameter Tuning* yang meliputi pengaturan parameter seperti *optimizer*, *batch normalization*, *dropout rate*, *learning rate*, *weight decay*, *momentum*, *batch size*, jumlah lapisan dalam jaringan, dan *epoch*, penulis bertujuan untuk menemukan konfigurasi terbaik yang dapat meningkatkan kinerja model *CNN* dalam mengenali aksara Jawa. Selain itu, penulis mengimplementasikan teknik *preprocessing* yang melibatkan *Inter-Area Interpolation*, *Grayscale Converter*, *Laplacian of Gaussian (LoG)*, *reshape*, dan *normalization* untuk meningkatkan kualitas data input yang digunakan dalam pelatihan model.

Pendekatan yang penulis usulkan mencakup eksperimen komprehensif yang mengevaluasi berbagai kombinasi *hyperparameter* dan teknik *preprocessing* untuk menentukan setelan optimal. Melalui optimasi ini, penulis berharap dapat mencapai tingkat akurasi dan keandalan yang lebih tinggi dalam klasifikasi aksara Jawa, sehingga bisa mengatasi tantangan kompleksitas karakter dan variasi kontekstual

yang ada. Hasil dari penelitian ini diharapkan tidak hanya meningkatkan performa klasifikasi aksara Jawa tetapi juga memberikan panduan yang bermanfaat bagi penelitian masa depan dalam digitalisasi aksara tradisional menggunakan metode *DL* yang lebih baik. Naskah ini dibagi menjadi 5 bagian. Bagian 1 mencakup pendahuluan, bagian 2 tinjauan pustaka, bagian 3 metode penelitian serta memberikan desain sistem yang diusulkan, bagian 4 merinci hasil dan evaluasi. Terakhir, kesimpulan disajikan di Bagian 5.

1.2 Rumusan Masalah

Rumusan masalah dari penelitian yang diusulkan adalah bagaimana mengatasi keragaman dan kompleksitas aksara Jawa yang rumit. Untuk itu, penelitian ini menggunakan pendekatan baru dengan memanfaatkan *Hyperparameter Tuning* serta mengeksplorasi teknik *preprocessing* untuk mencapai akurasi dan keandalan yang lebih tinggi dalam klasifikasi aksara Jawa.

1.3 Tujuan Penelitian

Penelitian yang penulis usulkan berfokus pada teknik praproses data dan optimasi model *CNN* untuk klasifikasi aksara Jawa dengan memanfaatkan *Hyperparameter Tuning*, yang bertujuan untuk menemukan kombinasi parameter terbaik yang dapat meningkatkan kinerja model dalam mengklasifikasikan aksara Jawa. Selain itu, penelitian ini akan mengeksplorasi teknik praproses data yang diusulkan untuk meningkatkan kualitas data input dan hasil klasifikasi.

1.4 Batasan Masalah

Penelitian dibatasi pada hal-hal berikut:

1. Dataset yang digunakan terbatas pada aksara Jawa dasar yang tersedia secara publik di platform *Kaggle* dan *Roboflow*.
2. Fokus pada teknik *Hyperparameter Tuning* yang meliputi *optimizer*, *batch normalization*, *dropout rate*, *learning rate*, *weight decay*, *momentum*, *batch size*, jumlah lapisan dalam jaringan, dan *epoch*.

3. Teknik *preprocessing* yang diterapkan meliputi *Inter-Area Interpolation*, *Grayscale Converter*, *Laplacian of Gaussian (LoG)*, *reshape*, dan *normalization*.
4. Evaluasi kinerja model *CNN* terbatas pada akurasi klasifikasi aksara Jawa dan tidak mencakup aspek lain seperti kecepatan komputasi.

1.5 Kontribusi

Penelitian ini diharapkan dapat memberikan kontribusi sebagai berikut:

1. Menghasilkan model *CNN* yang lebih akurat dan andal untuk klasifikasi aksara Jawa, sehingga mendukung pelestarian digital aksara tradisional.
2. Menyediakan wawasan baru tentang penggunaan teknik *Hyperparameter Tuning* dalam mengoptimalkan model *CNN* untuk tugas klasifikasi aksara.
3. Menyajikan teknik *preprocessing* yang efektif untuk meningkatkan kualitas data input dan hasil klasifikasi aksara Jawa.
4. Menyediakan referensi dan panduan yang bermanfaat bagi peneliti lain yang tertarik dalam bidang digitalisasi aksara tradisional dan aplikasi *DL* pada klasifikasi karakter.
5. Mendukung pelestarian budaya dan sejarah Jawa melalui digitalisasi aksara Jawa, serta meningkatkan kesadaran dan apresiasi terhadap warisan budaya yang kaya tersebut.