

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Aksara Sunda merupakan salah satu peninggalan sejarah masyarakat Sunda yang berperan sebagai sistem tulisan tradisional sekaligus mencerminkan identitas budaya masyarakat Sunda. Penggunaan aksara Sunda telah ditemukan sejak abad ke-15, dengan daun lontar sebagai media penulisan aksara Sunda kuno pada masa itu. Aksara ini digunakan untuk berbagai keperluan, seperti mencatat sejarah dinasti kerajaan Sunda, informasi medis, pertanian, dan lainnya (Suryani, Paulus, Hadi, Darsa, Burie, 2017). Sebagai warisan budaya yang kaya, aksara Sunda memiliki nilai-nilai penting yang perlu dilestarikan agar tetap menjadi simbol identitas masyarakat Sunda bagi generasi mendatang. Menurut Hayat (2023), pelestarian bahasa daerah merupakan bagian dari amanat kebudayaan nasional yang sejalan dengan Pasal 32 ayat (2) Undang-Undang Dasar Negara Republik Indonesia Tahun 1945.

Kondisi aksara Sunda kuno yang semakin mengalami degradasi kualitas akibat penuaan alami, bencana alam, dan faktor lainnya mengancam kelestarian warisan budaya ini. Oleh karena itu, digitalisasi aksara Sunda menjadi langkah mendesak untuk pelestarian (Suryani, dkk., 2017). Salah satu metode digitalisasi yang populer adalah *Optical Character Recognition* (OCR), teknologi yang mengubah teks dari gambar atau tulisan tangan menjadi format digital yang dapat diolah oleh mesin (Wei, Chuluunbandi, Tuyatsetseg, Altangerel, 2024). Teknologi OCR terus berkembang dan kini mampu menawarkan akurasi tinggi untuk berbagai jenis teks (AlKendi, Gechter, Heyberger, & Guyeux, 2024). Implementasi OCR dalam banyak bidang menunjukkan manfaat signifikan, seperti efisiensi waktu, kemudahan akses, dan pengelolaan dokumen digital (Pitura, 2023). Namun, sebagian besar penelitian OCR lebih banyak berfokus pada pengenalan teks latin (Robby, Tandra, Susanto, Harefa & Chowanda, 2019), sehingga aplikasi untuk aksara lokal seperti aksara Sunda masih menjadi tantangan yang belum banyak dieksplorasi.

Berbagai penelitian sebelumnya telah mengusulkan metode yang berbeda untuk pengenalan aksara lokal. Misalnya, *K-Nearest Neighbour* (K-NN) digunakan dalam pengenalan aksara Kaganga Lampung (Istiqphara, Faida & Darajat, 2023) dan aksara Bali (Alfarisi & Subandi, 2022), sedangkan transliterasi aksara Bali juga dilakukan dengan metode SVM dan CNN (Widiarti & Suparwito, 2023). Untuk aksara Sunda, model *Convolutional Neural Network* (CNN) banyak diterapkan dalam penelitian, menghasilkan akurasi tinggi dalam rentang 85–97% (Purnama, A. dkk., 2022; Rahmawati, Shelvi, Hidayat & Mubarak, 2021; Nathanael & Wasito, 2023; Ripera, 2023; Kirana, 2020). Metode lain juga telah diusulkan, seperti *Modified Direction Feature* dan *Learning Vector Quantization* dengan akurasi 78,67% (Riansyah, Nurhasanah, Dewi, 2017), model HOG dan SVM dengan akurasi 90,83% (Taufiqurahman, M., 2019), hingga YOLO v8 dengan akurasi 76,2% (Arifadilah, dkk., 2024). Berdasarkan tinjauan tersebut, penelitian berbasis CNN masih menjadi pilihan utama dengan hasil akurasi yang kompetitif. Namun, sebagian besar penelitian sebelumnya hanya berfokus pada pengenalan citra aksara pada tingkat karakter dan aksara modern, sementara tantangan pengenalan tulisan tangan aksara kuno, seperti aksara Sunda kuno, belum banyak dieksplorasi.

Kesiman, dkk., (2018) menyatakan bahwa pelatihan pada level kata cenderung memberikan performa yang lebih optimal dalam pengenalan kata dibandingkan dengan pelatihan pada tingkat teks baris. Oleh karena itu, pendekatan berbasis kata yang digunakan dalam penelitian ini menjadi relevan dalam konteks transliterasi dan pengenalan aksara Sunda kuno, terutama karena fokus utamanya adalah pada struktur kata. Lebih lanjut, sifat aksara Sunda kuno yang kompleks, seperti tulisan yang pudar, adanya kebisingan acak, perubahan warna, variasi gaya penulisan yang tidak terstandarisasi, serta jarak antar karakter yang beragam, menjadikannya sulit untuk dikenali hanya dengan metode CNN konvensional. Kondisi ini mengindikasikan perlunya kombinasi arsitektur ekstraksi dan klasifikasi yang lebih efektif untuk mengatasi tantangan tersebut (Suryani, dkk., 2017).

Mempertimbangkan karakteristik kompleks aksara Sunda kuno, penelitian relevan mengusulkan penggabungan metode *Convolutional Neural Network* (CNN) dan *Recurrent Neural Network* (RNN) yang dikenal sebagai CRNN. Dalam arsitektur ini, CNN berfungsi untuk mengekstraksi fitur visual, sedangkan RNN memproses urutan dari fitur-fitur tersebut (Zhu, dkk., 2018). Meskipun CRNN efektif dalam mengolah data sekuensial seperti tulisan tangan, penelitian Yu, Ibrayim & Hamdulla, (2023) mengungkapkan bahwa model ini memiliki kelemahan dalam mengenali teks yang tidak beraturan, dengan panjang sekuensial yang bervariasi, atau informasi yang tidak lengkap. Untuk mengatasi keterbatasan ini, *Connectionist Temporal Classification* (CTC) diusulkan sebagai tambahan, yang terbukti mampu meningkatkan akurasi pengenalan teks secara signifikan dibandingkan penggunaan CRNN saja. Kombinasi CRNN-CTC telah diterapkan dalam berbagai studi relevan, seperti pengenalan aksara Nepal (Nakarmi, Sthapit, Shakya, Chulyadyo, Bal, 2024), teks Urdu (Nasir & Malik, 2024), aksara Pegon (Rudelviyani, dkk., 2024), dan aksara Amharic (Gondere, dkk., 2022). Penelitian-penelitian ini menunjukkan efektivitas model CRNN-CTC dalam mengatasi tantangan pengenalan tulisan tangan bahasa non-Latin dengan karakteristik kompleks yang beragam.

Berdasarkan berbagai penelitian sebelumnya, terlihat bahwa aksara Sunda kuno memiliki tantangan unik yang memerlukan pendekatan teknologi modern untuk memastikan kelestariannya. Digitalisasi aksara Sunda kuno menggunakan model yang lebih canggih seperti CRNN-CTC menjadi sangat relevan, mengingat sifat aksara yang kompleks serta hasil positif dari penelitian serupa pada bahasa non-Latin. Penelitian ini bertujuan untuk mengeksplorasi efektivitas model CRNN-CTC dalam pengenalan tulisan tangan aksara Sunda kuno berbasis kata, sebagai langkah untuk mendukung pelestarian budaya melalui digitalisasi yang lebih optimal. Dengan ini, diharapkan penelitian dapat memberikan kontribusi signifikan terhadap pengembangan teknologi OCR untuk aksara tradisional, sekaligus memperkaya upaya pelestarian budaya Indonesia.

1.2 Rumusan Masalah Penelitian

Berdasarkan uraian pada bagian sebelumnya, rumusan masalah pada penelitian ini adalah sebagai berikut :

1. Bagaimana evaluasi kinerja model CRNN-CTC dalam pengenalan tulisan tangan aksara Sunda kuno?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah pada penelitian ini, tujuan dari penelitian adalah sebagai berikut:

1. Mengevaluasi kinerja model CRNN-CTC dalam rekognisi tulisan tangan aksara Sunda kuno.

1.4 Manfaat Penelitian

Penelitian ini dilakukan dengan harapan akan bermanfaat bagi sebagian pihak, antara lain :

1.4.1. Manfaat Teoritis

Memberikan kontribusi dalam pengembangan ilmu pengetahuan, khususnya di bidang rekognisi tulisan tangan berbasis model deep learning, dengan menambah kajian tentang penerapan arsitektur CRNN-CTC pada aksara non-Latin seperti aksara Sunda kuno.

1.4.2. Manfaat Praktis

Hasil dari penelitian ini secara praktis diharapkan dapat menghasilkan sistem pengenalan aksara Sunda kuno yang dapat digunakan sebagai dasar pengembangan aplikasi transliterasi otomatis atau pendidikan berbasis teknologi.

1.4.3. Manfaat Akademis

Hasil dari penelitian ini secara akademis diharapkan dapat berkontribusi sebagai acuan bagi penelitian sejenis tentang model arsitektur CRNN-CTC dalam rekognisi teks non-Latin.

1.5 Batasan Penelitian

Berikut batasan penelitian yang telah ditentukan dalam penelitian:

1. Dataset yang digunakan merupakan jenis aksara Sunda manuskrip pada daun lontar berasal dari abad ke-15 dengan format level kata.
2. Fokus penelitian hanya sebatas transliterasi aksara Sunda ke bentuk latinnya dan tidak mendalami terjemahannya.
3. Evaluasi performa model dibatasi pada metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER) dan *Overall accuracy*.
4. Penelitian ini terbatas pada pengembangan model rekognisi aksara Sunda kuno dan tidak mencakup implementasi sistem OCR yang lengkap atau integrasi ke aplikasi yang siap digunakan secara langsung.

1.6 Sistematika Penulisan Skripsi

Sistematika penulisan skripsi berperan sebagai pedoman penulisan agar dalam penulisan ini lebih terarah, maka skripsi ini dibagi menjadi beberapa bab. Adapun struktur organisasi skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menggambarkan latar belakang penelitian, identifikasi masalah penelitian, rumusan masalah penelitian, tujuan penelitian, manfaat/signifikansi penelitian dan struktur organisasi skripsi.

BAB II KAJIAN PUSTAKA

Pada bab ini berisi teori-teori relevan beserta kajian penelitian-penelitian sebelumnya yang relevan dengan topik pembahasan.

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan metode dan desain penelitian, sumber data penelitian, instrumen penelitian, alat dan bahan yang digunakan pada penelitian.

BAB IV TEMUAN DAN PEMBAHASAN

Pada bab ini menjelaskan tentang hasil pengolahan data dan pembahasan yang akan menjelaskan temuan dan pembahasan data hasil penelitian mencakup evaluasi model dengan metrik yang digunakan.

BAB V SIMPULAN DAN REKOMENDASI

Pada bab ini menjelaskan kesimpulan dari temuan penelitian dan saran berdasarkan pembahasan yang telah diuraikan untuk penelitian selanjutnya.