

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pengenalan tulisan tangan atau *handwriting recognition* merupakan bentuk aplikasi dari pengenalan pola melalui gambar (S, M, P, & S, 2020), serta merupakan salah satu topik utama dalam bidang *computer vision*, hal ini disebabkan oleh beberapa tantangan yang diberikan seperti beragamnya gaya tulisan tangan dalam bahasa dan media yang berbeda serta tingkat kerusakan pada media tersebut.

Tulisan tangan secara umum terbagi menjadi dua, yaitu tulisan tangan *offline* dimana tulisan dibuat dari media non-digital seperti kertas dan manuskrip kuno, serta tulisan tangan *online* dimana tulisan dibuat pada media digital seperti penulisan menggunakan *touchscreen*, *electronic signature device*, serta alat pembantu lain yang menghasilkan tulisan tangan digital secara langsung. Manuskrip kuno merupakan salah satu media tulisan tangan *offline* yang memiliki tingkat pengenalan kata serta huruf yang cukup rendah oleh sistem pengenalan tulisan yang telah ada seperti BI-LSTM, hal ini disebabkan karena banyaknya media manuskrip mengalami kerusakan yang cukup berat dikarenakan oleh usianya.

Aksara Sunda diperkirakan sudah ada sejak abad ke-14 yang ditandai dengan ditemukannya aksara Sunda kuno pada prasasti-prasasti serta umumnya ditemukan pada daun lontar. Aksara Sunda pada daun lontar yang

tertua ditemukan dari abad ke-15, hal ini pula yang membuat aksara Sunda bisa dikatakan sebagai sebuah warisan suku Sunda.

Benchmark yang dilakukan dengan BI-LSTM (Kesiman, et al., 2018) mendapatkan *word error rate* (WER) untuk pengenalan kata manuskrip lontar Sunda kuno (Suryani, Paulus, Hadi, Darsa, & Burie, 2017) sebesar 75.52%, dimana angka *error* tersebut masih bisa dikatakan cukup tinggi, hal ini dikarenakan fitur ekstraksi pada input gambar yang kurang baik. Model lain dengan fitur ekstraksi yang lebih baik diperlukan untuk mendapatkan fitur penting dari input untuk kemudian diproses oleh *classifier*, salah satu ekstraksi fitur gambar yang efektif adalah *Convolutional Neural Network* (CNN). Penelitian lebih lanjut diperlukan untuk mengetahui hasil dari penerapan fitur ekstraksi CNN serta metode atau model lain yang dapat menurunkan WER dimana semakin rendah tingkat *error* maka tingkat kesalahan prediksi pada model semakin rendah seperti penggunaan *Attention* pada model berbasis *encoder-decoder* dengan *Recurrent Neural Network* (RNN), telah terbukti dapat meningkatkan akurasi pada tugas terjemahan suatu bahasa (Luong, Pham, & Manning, 2015). Berdasarkan permasalahan tersebut maka diusulkan penelitian yang berjudul **“Convolutional Recurrent Neural Network dengan Attention untuk Pengenalan Tulisan Tangan Manuskrip Daun Lontar Sunda Kuno”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah penelitian ini adalah:

1. Bagaimana menerapkan *Convolutional Recurrent Neural Network + Attention* (CRNN + *Attention*) untuk menurunkan WER pengenalan tulisan tangan manuskrip daun lontar Sunda kuno?
2. Bagaimana perbandingan WER CRNN + *Attention* dengan WER BI-LSTM?

1.3 Batasan Masalah

Beberapa hal yang menjadi batasan dalam penelitian ini adalah:

1. *Dataset* dibagi menjadi dua, dengan masing-masing merupakan gambar serta label per kata dari *dataset* berjumlah 1426 untuk data latih dan 317 untuk data tes serta *character mapping* sebanyak 80.
2. Acuan metrik yang diukur berupa WER.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dibuat, maka tujuan penelitian ini adalah:

1. Menerapkan model CRNN + *Attention* untuk menurunkan WER pengenalan tulisan tangan manuskrip daun lontar Sunda kuno.
2. Perbandingan nilai WER CRNN + *Attention* dengan WER BI-LSTM.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah:

1. Turunnya WER pengenalan manuskrip lontar Sunda kuno dengan metode CRNN + *Attention*.
2. Mengetahui perbandingan WER dengan metode BI-LSTM.

1.6 Metodologi Penelitian

Metodologi penelitian ini menjelaskan tentang jenis penelitian, jenis serta sumber data, tahapan penelitian, serta tahapan eksperimen. Tahapan penelitian terdiri dari:

1. Pendahuluan. Tahapan ini terdiri dari penentuan topik, identifikasi serta perumusan masalah, penentuan tujuan, penentuan batasan masalah serta metodologi penelitian.
2. Studi pustaka, Tahap ini dilakukan studi pustaka dari literatur yang berkaitan dengan penelitian.
3. Pengumpulan data. Tahap ini dilakukan pengumpulan data eksperimen yang akan digunakan dalam penelitian.
4. Eksperimen. Tahapan ini merupakan tahapan implementasi dari penelitian yang akan dilakukan, Eksperimen ini memiliki beberapa tahapan yang terdiri dari, perancangan model, *preprocessing* data, pengujian (*model training*, *model testing*), dan evaluasi.
5. Kesimpulan dan Saran, Tahapan ini melakukan penarikan kesimpulan serta saran berdasarkan hasil eksperimen.

1.7 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam penulisan tugas akhir adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas laporan secara garis besar yang meliputi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metodologi penelitian, serta sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini menjelaskan dasar-dasar teori atau konsep yang diperlukan dan berhubungan dalam penelitian yang dilakukan.

BAB III METODOLOGI

Bab ini membahas metode atau langkah-langkah yang akan dilakukan pada penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi hasil serta pembahasan implementasi dari metodologi yang telah dibuat, serta analisis terhadap penelitian yang dilakukan.

BAB V KESIMPULAN DAN SARAN

Bab ini membahas kesimpulan yang didapat dari hasil penelitian beserta saran-saran yang dapat menjadi masukan untuk penelitian lain.