

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Aksara Sunda merupakan aksara yang digunakan oleh Suku Sunda dahulu untuk berkomunikasi, aksara tradisional dari Indonesia ini merupakan hasil karya ortografi melalui perjalanan sejarahnya sejak 5 abad yang lalu [1]. Kerumitan dan keunikan dari karakter Aksara Sunda membuat aksara ini sulit dalam pengenalannya. Pengenalan karakter dari tulisan tangan Aksara Sunda diperlukan pengetahuan khusus untuk dapat memahami arti dari setiap karakter. Sehingga untuk memudahkan dalam pengenalan Aksara Sunda dapat dikembangkan kegiatan yang berkaitan dengan komputerisasi sesuai dengan perkembangan zaman, salah satunya dengan melakukan penelitian terkait pengenalan pola dari citra tulisan tangan Aksara Sunda.

Penelitian sebelumnya mengenai pengenalan pola dari citra tulisan tangan Aksara Sunda telah dilakukan, diantaranya adalah Alif Kirana, dkk. (2020) dengan judul “Pengenalan Pola Aksara Sunda dengan Metode *Convolutional Neural Network*” untuk citra tulisan tangan responden yang diambil menggunakan pemindai didapatkan akurasi 85.5% [2], Nisa Amalia, dkk. (2020) dengan judul “Pengenalan Aksara Sunda Menggunakan Metode Jaringan Saraf Tiruan Backpropagation dan Deteksi Tepi Canny” didapatkan akurasi sebesar 76.19% [3], Imam Ramadhan, dkk. (2018) dengan judul “Pengenalan Pola Citra Tulisan Tangan Aksara Sunda dengan Metode *Convolutional Neural Network*” didapatkan akurasi sebesar 62% [4], Rizki Rahmat Riansyah, dkk. (2017) dengan judul “Sistem Pengenalan Aksara Sunda Menggunakan Metode *Modified Direction Feature* dan *Learning Vector Quantization*” didapatkan akurasi sebesar 78.67% [5].

Dari beberapa penelitian yang telah dilakukan sebelumnya, mengenai penerapan beberapa metode untuk pengenalan pola citra tulisan tangan Aksara Sunda didapatkan hasil yang belum sempurna atau masih ada beberapa aksara yang belum terklasifikasi, Namun, pada penelitian-penelitian sebelumnya khususnya metode *Convolutional Neural Network* (CNN) mendapatkan akurasi terbaik dibandingkan beberapa metode sebelumnya seperti Jaringan Saraf Tiruan (JST), *Modified Direction Feature* (MDF) dan *Learning Vector Quantization* (LVQ) hal ini diperkuat dengan adanya hasil survey [6] bahwa metode CNN memberikan hasil yang baik terhadap pengenalan dari pola tulisan tangan. Tetapi, pada penelitian sebelumnya metode CNN juga memiliki kekurangan ketika dilatih menggunakan sampel data yang sedikit dan tidak bervariasi sehingga akurasi yang didapatkan kurang maksimal.

Kurangnya sampel data dan data yang tidak bervariasi dapat mempengaruhi kinerja dari metode CNN, oleh karena itu penggunaan augmentasi data umum dilakukan untuk meningkatkan variasi citra dengan melakukan manipulasi transformasi dari citra. Augmentasi data merupakan proses memanipulasi atau memodifikasi data sedemikian rupa sehingga komputer akan mendeteksi bahwa data yang dimanipulasi adalah data yang berbeda, namun manusia masih dapat mengetahui bahwa data tersebut merupakan data yang sama [7]. Beberapa penelitian telah berhasil menerapkan augmentasi data guna meningkatkan kinerja dari metode CNN [8]–[10]. Oleh karena itu, dengan memberikan variasi citra yang banyak menggunakan augmentasi data dapat membantu metode CNN dalam mengenali fitur atau karakteristik dari citra secara lebih baik.

Pada penelitian ini berfokus pada bagaimana penggunaan augmentasi data dapat membantu untuk meningkatkan akurasi dalam melakukan pengenalan pola citra tulisan tangan Aksara Sunda menggunakan metode CNN yang belum dilakukan sebelumnya, dengan membandingkan hasil akurasi antara model *baseline* dan model yang telah dilatih menggunakan augmentasi data. Serta menambah jumlah *layer* atau jaringan pada metode CNN dengan menerapkan arsitektur LeNet-5 berdasarkan kesimpulan dari penelitian sebelumnya. Metode CNN dipilih karena memiliki akurasi yang baik dari penelitian-penelitian sebelumnya untuk pengenalan pola citra tulisan tangan Aksara Sunda.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan yang telah diuraikan sebelumnya, maka rumusan permasalahan penelitian ini adalah rendahnya akurasi yang diperoleh dari penelitian dengan metode CNN pada jumlah sampel data yang sedikit dan data yang tidak bervariasi.

1.3 Maksud dan Tujuan

Maksud dari penelitian ini adalah menerapkan augmentasi data untuk pengenalan pola citra tulisan tangan Aksara Sunda menggunakan metode CNN. Adapun tujuan yang ingin dicapai dari penelitian ini adalah untuk mengukur perbandingan akurasi yang didapatkan dengan membandingkan antara *baseline model* dan model yang telah dilatih menggunakan augmentasi data.

1.4 Batasan Masalah

Agar penelitian ini lebih terarah sesuai dengan tujuan yang diinginkan, maka batasan masalah dari penelitian ini adalah sebagai berikut:

1. Data masukan berupa citra hasil *scanning* dari tulisan tangan Aksara Sunda dengan format citra JPG/JPEG
2. Aksara yang digunakan untuk pengenalan sebanyak 25 Aksara Sunda (ngalagena) yaitu: ba, ca, da, fa, ga, ha, ja, ka, kha, la, ma, na, nga, nya, pa, qa, ra, sa, sya, ta, va, wa, xa, ya, dan za
3. Aksara berposisi tegak dan tidak bersambung
4. Kertas yang digunakan adalah kertas jenis HVS beralatar belakang putih dan ditulis menggunakan spidol
5. Metode *preprocessing* yang digunakan adalah *grayscale (averaging)*, *thresholding (NICK thresholding)*, dan *resize (nearest neighbour interpolation)*
6. Metode augmentasi data yang digunakan adalah *flipping (horizontal flip)*, *rotation*, dan *translation*
7. Hasil akurasi didapatkan dengan melakukan evaluasi dari data uji menggunakan *confusion matrix*

1.5 Metodologi Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah menggunakan metode penelitian kuantitatif. Metode penelitian kuantitatif adalah

metode penelitian yang berguna untuk mendapatkan informasi berupa data yang dapat diukur. Adapun alur dari penelitian ini adalah seperti yang tergambar pada Gambar 1.1.



Gambar 1.1 Alur Penelitian

1.5.1 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Studi Literatur

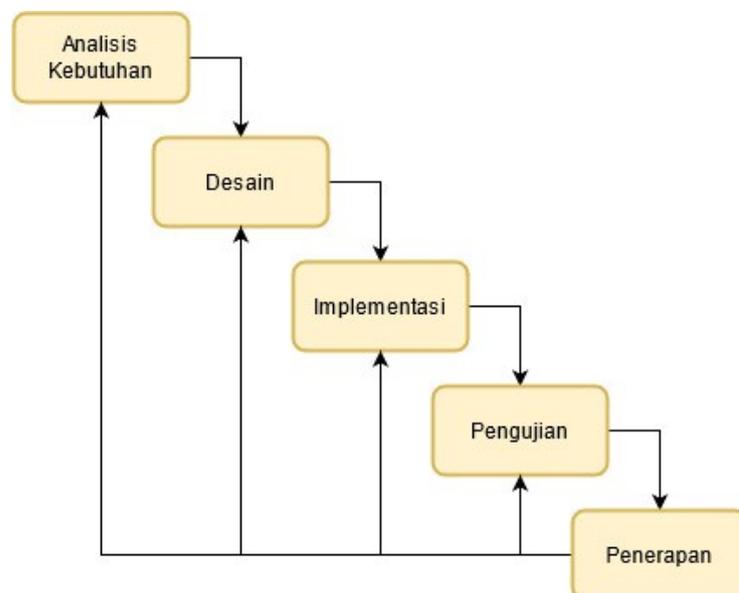
Studi literatur dilakukan dengan melakukan pencarian dan pengumpulan sumber-sumber literatur yang menunjang penelitian ini. Sumber-sumber literatur tersebut didapat dengan melakukan studi terhadap sumber bacaan seperti jurnal, *paper*, buku, serta laporan tugas akhir yang berkaitan dengan penelitian.

2. Pengumpulan Sampel Data Tulisan Tangan Aksara Sunda

Pengumpulan sampel data tulisan tangan Aksara Sunda dalam penelitian ini dilakukan dengan mengumpulkan 75 orang responden, masing-masing responden menuliskan 25 karakter tulisan tangan Aksara Sunda (ngalagena) sehingga terkumpul sampel sebanyak 1875 citra. Contoh penulisan 25 karakter tulisan tangan dapat dilihat pada Error: Reference source not found, kemudian data akan dipotong per karakter dan dikelompokkan pada sebuah *folder* berdasarkan kelas dari data tersebut. Dari 1875 sampel data akan dibagi menjadi sebanyak 1500 citra untuk digunakan sebagai data latih, dan 375 citra untuk digunakan sebagai data uji.

1.5.2 Metode Pembangunan Perangkat Lunak

Dalam penelitian ini, metode pembangunan perangkat lunak yang digunakan adalah *Software Development Life Cycle (SDLC) Waterfall Model*. Berikut merupakan gambaran dari metode pembangunan perangkat lunak *Waterfall Model* pada Gambar 1.2.



Gambar 1.2 Waterfall Model

Berikut adalah langkah-langkah yang dilakukan dalam penelitian ini dengan menggunakan metode pembangunan perangkat lunak *Waterfall Model*:

1. Analisis Kebutuhan

Pada langkah ini dilakukan analisis untuk mendefinisikan kebutuhan fungsional dan non-fungsional dari perangkat lunak yang akan dibangun.

2. Desain

Pada langkah ini dilakukan perancangan dari perangkat lunak yang akan dibangun meliputi perancangan prosedural dan antarmuka pengguna.

3. Implementasi

Setelah melakukan perancangan pada langkah sebelumnya, dilakukan implementasi ke dalam bentuk kode program yang nantinya akan menghasilkan perangkat lunak.

4. Pengujian

Pada langkah ini dilakukan pengujian terhadap perangkat lunak yang dibangun, untuk memastikan perangkat lunak yang dibangun tidak terdapat kesalahan.

5. Penerapan

Pada langkah ini dilakukan penerapan perangkat lunak ke perangkat *server* sesuai dengan kebutuhan.

1.6 Sistematika Penulisan

Sistematika penulisan disusun untuk memberikan gambaran umum mengenai penulisan penelitian ini. Sistematika penulisan dari penelitian ini adalah sebagai berikut:

BAB 1 PENDAHULUAN

Pada bab ini diuraikan mengenai latar belakang permasalahan, rumusan masalah, maksud dan tujuan penelitian, batasan masalah penelitian, metodologi penelitian, serta sistematika penulisan penelitian.

BAB 2 TINJAUAN PUSTAKA

Pada bab ini dibahas mengenai konsep dasar dan teori-teori pendukung yang berkaitan dengan penelitian penelitian. Teori yang akan dibahas pada bab ini adalah mengenai Aksara Sunda, citra digital, pengolahan citra digital, metode CNN, dan *confusion matrix*.

BAB 3 ANALISIS DAN PERANCANGAN

Pada bab ini membahas mengenai analisis metode-metode yang digunakan sebelum dilakukan implementasi serta perancangan dari model klasifikasi dan *prototype system* yang akan dibangun. Bab ini memuat analisis mengenai analisis masalah, analisis sistem, analisis citra masukan, analisis metode *preprocessing*, analisis metode augmentasi data, analisis metode klasifikasi, serta perancangan *prototype system*.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Pada bab ini dilakukan pembahasan mengenai implementasi dari metode-metode yang sudah dilakukan analisis dan perancangan sebelumnya kedalam bentuk kode program, serta dilakukan proses pengujian dari model klasifikasi yang dibangun menggunakan data uji sehingga dapat diketahui hasil dari penelitian ini.

BAB 5 KESIMPULAN DAN SARAN

Pada bab ini diuraikan mengenai kesimpulan dari hasil yang didapatkan pada penelitian yang dilakukan, serta saran-saran terkait pengembangan penelitian berikutnya agar penelitian-penelitian terkait dapat lebih baik.