

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Seiring perkembangan teknologi yang maju dan semakin canggih bersamaan dengan meningkatnya kebutuhan akan digitalisasi dokumen saat ini. Maka dibutuhkan sebuah *tools* untuk membantu mengolah suatu dokumen menjadi informasi dalam bentuk digital. Pengenalan objek (*object recognition*) adalah salah satu bidang yang banyak digunakan dalam lingkungan pengolahan citra. Aplikasi yang biasa dijumpai dalam melakukan identifikasi citra karakter untuk kemudian menjadi teks tulisan yaitu OCR (*Optical Character Recognition*) [1].

Penelitian OCR sudah banyak yang melakukannya salah satunya dalam mengenal karakter pada surat masuk. OCR tersebut dibantu dengan menggunakan metode *Backpropagation Neural Network (BPNN)*. Pada penelitian tersebut pengenalan karakter pada surat masuk memiliki nilai keberhasilan sebesar 84,56% dari 790 karakter pada 15 data citra dokumen surat masuk [1]. Pada penelitian tersebut sudah mencapai nilai keberhasilan 84,56% dari 790 karakter, artinya algoritma *Backpropagation Neural Network* yang digunakan sudah cukup baik. Dan penulis tersebut menyatakan bahwa tingkat keberhasilan pengenalan karakter hasilnya lebih baik ketika polanya telah dilatih terlebih dahulu dibandingkan dengan yang tidak dilatih. Pola yang dilatih menggunakan *font Arial* dengan persentase 71,95%, sedangkan yang tidak dilatih yaitu *font Times New Roman* dan *Courier New* adalah 40,24%. Dan peneliti tersebut menyatakan bahwa *pre-processing* perlu diperhatikan karena akan mempengaruhi tingkat akurasi yang diharapkan [1].

Penelitian lainnya yaitu penggunaan BPNN untuk mengenali pola karakter atau huruf Jawa. Rata-rata keakuratan BPNN dalam mengenali pola karakter huruf Jawa yaitu sebesar 99.563% untuk data sampel berupa data pelatihan, 61.359% untuk data sampel diluar data pelatihan, dan 75% untuk data sampel data pelatihan dan di luar data pelatihan. Pada penelitiannya membuktikan bahwa semakin banyak data pelatihan, semakin baik nilai keakuratan BPNN dalam mengenali seluruh pola karakter huruf Jawa, baik data pelatihan yang ikut diuji maupun yang tidak ikut

diuji. Ketika jumlah sampel berjumlah tiga sampel, persentase keakuratan yang didapatkan yaitu 65.667% sedangkan jumlah sampel sebanyak 15 sampel persentase keakuratan mencapai 99% [2].

Sertifikat adalah sebuah surat pernyataan atau tanda tertulis tercetak dari orang yang berwenang berupa bukti kepemilikan atau suatu kejadian. Contohnya sertifikat tanah, pelatihan, seminar, hasil ujian, dan sertifikat lainnya yang banyak orang pasti pernah melihat salah satu contohnya [3]. Sertifikat merupakan dokumen yang sangat penting contohnya pada bidang administrasi atau yang berhubungan dengan manajemen manusia. Karakteristik yang dimiliki oleh dokumen sertifikat sangat khusus karena memiliki banyak jenis variasi gambar [4]. Maka dari itu kemajuan OCR pada dokumen sertifikat ini sangat dibutuhkan mengingat banyaknya sertifikat yang harus diolah. Teks yang terdapat pada suatu sertifikat biasanya tidak hanya menggunakan satu jenis teks saja. Suatu sertifikat tentu sangat memungkinkan ukuran, *font*, warna, tipe teks (seperti: *bold*, *italic*, ataupun *underline*) akan berbeda-beda.

Dari uraian tersebut, peneliti termotivasi untuk melakukan analisis mengenai karakter dengan menggunakan algoritma *Backpropagation Neural Network* pada kasus dokumen sertifikat.

1.2 Rumusan Masalah

Berdasarkan latar belakang pada uraian diatas, masalah utama yang dapat diangkat adalah menganalisis bagaimana cara sistem dapat mengenali citra menjadi teks digital/ketik. Dan pada kasus ini masalah tersebut berfokus terhadap analisis pengenalan karakter menggunakan *Backpropagation Neural Network* pada kasus dokumen sertifikat.

1.3 Maksud dan Tujuan

Maksud dan tujuan penelitian ini adalah untuk menganalisa tingkat akurasi keberhasilan algoritma BPNN (*Backpropagation Neural Network*) dalam mengenali karakter pada citra atau gambar dokumen sertifikat.

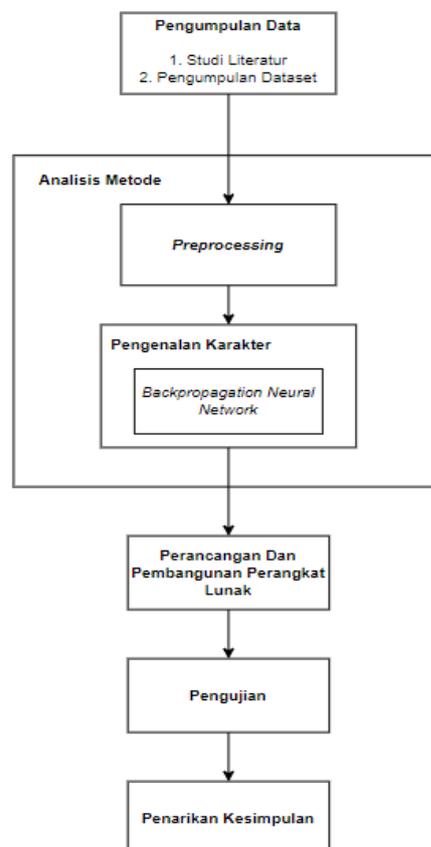
1.4 Batasan Masalah

Berikut batasan-batasan yang ditentukan agar permasalahan dapat diselesaikan sesuai dengan tujuan yang diharapkan:

- a. Objek yang digunakan adalah file gambar yang didapatkan hasil foto dari sertifikat aslinya.
- b. Citra yang digunakan adalah citra berwarna dengan format JPG/JPEG (*Join Photographics Expert Group*).
- c. Karakter yang akan dikenali berupa abjad huruf kapital (A-Z), huruf kecil (a-z) dan angka (0-9).
- d. Pengambilan gambar atau citra dilakukan dengan sebaik mungkin dengan pencahayaan yang cukup agar hasil gambar sertifikat terlihat dengan jelas.
- e. Menggunakan algoritma BPNN (*Backpropagation Neural Network*).
- f. Penelitian ini dibantu dengan aplikasi berbahasa Python.

1.5 Metodologi Penelitian

Metodologi penelitian yang digunakan pada penelitian ini yaitu *Case Studies Research* yang merupakan metode yang berfokus kepada sebuah kasus dengan mengumpulkan dan penggalan data yang lebih pada objek tersebut [5]. Gambaran pada metode *Case Studies Research* sebagai berikut:



Gambar 1.1 Metodologi Penelitian

1.5.1 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan pada penelitian ini melalui dua proses yang dibutuhkan, yaitu:

a. Studi Literatur

Studi literatur yaitu melakukan proses pengumpulan literasi dari penelitian-penelitian berupa jurnal ilmiah dan buku-buku dari dalam atau luar negeri.

b. Pengumpulan Dataset

Pengumpulan dataset yaitu melakukan proses pengumpulan data yang akan digunakan sebagai data masukan dari aplikasi yang akan dibangun.

1.5.2 Analisis Metode

Analisis metode merupakan melakukan analisa terhadap metode-metode yang akan digunakan dalam penelitian ini, terdapat dua bagian metode yang digunakan, yaitu:

a. *Preprocessing*

Preprocessing merupakan langkah awal yang dilakukan sebelum melakukan proses utama pada penelitian ini. *Preprocessing* dilakukan untuk dapat membantu melancarkan proses utama yaitu dalam pengenalan karakter. Tahap *preprocessing* akan menggunakan metode yang berbeda pada tahap pelatihan dan pengujian. Tahapan *preprocessing* yang ada pada penelitian ini yaitu:

1. *Grayscale*, yaitu proses perubahan citra berwarna menjadi keabuan.
2. *Maximally Stable Extremal Regions*, yaitu salah satu metode mendeteksi daerah-daerah objek pada citra.
3. *Binarization*, yaitu proses mengubah citra keabuan menjadi citra biner.
4. *Segmentation*, yaitu penyederhanaan citra suatu karakter.
5. *Resize*, yaitu proses perubahan ukuran citra.

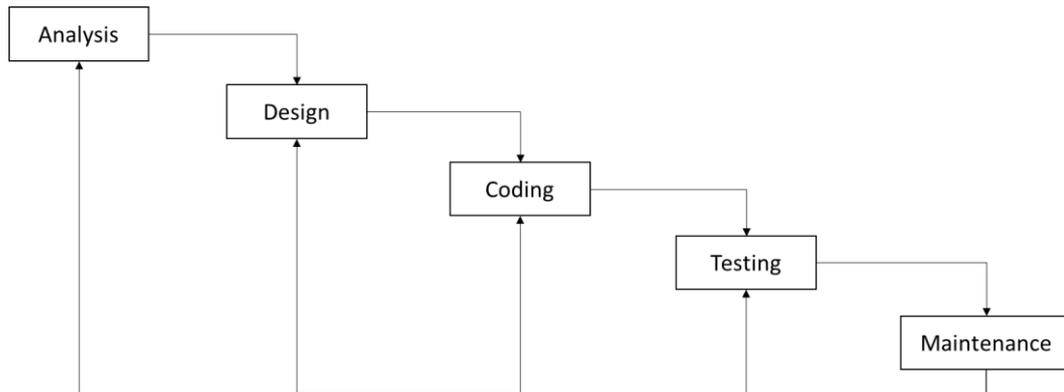
b. Pengenalan Karakter

Metode dalam proses pengenalan karakter menggunakan *Backpropagation Neural Network*, yaitu menggunakan konsep jaringan

saraf tiruan dan dibantu dengan metode *backpropagation* dalam mengenali pola-pola citra sehingga dikenali sebagai karakter.

1.5.3 Metode Pembangunan Perangkat Lunak

Tahapan proses pembangunan perangkat lunak menggunakan metode *The Classic Life Cycle (Paradigma Waterfall)*. Tahapan yang terjadi pada metode tersebut adalah sebagai berikut:



Gambar 1.2 Metode Pembangunan Perangkat Lunak

1. *Analysis*, yaitu tahap menganalisa kebutuhan yang diperlukan dalam pelaksanaan proyek pembangunan perangkat lunak.
2. *Design*, yaitu tahap membuat model umum dari hasil analisis kebutuhan yang telah ditetapkan dari tahap analisis.
3. *Coding*, yaitu melakukan penulisan kode menggunakan suatu bahasa pemrograman sesuai dengan desain yang sudah direncanakan.
4. *Testing*, yaitu tahap untuk menguji program yang telah dibangun untuk memastikan tidak ada kesalahan dan berjalan sesuai dengan perencanaan yang telah ditetapkan.
5. *Maintenance*, yaitu tahap untuk melakukan perubahan jika diperlukan sebagai perkembangan dan penyempurnaan pada perangkat lunak yang sedang dibangun.

1.5.4 Pengujian

Tahap pengujian meliputi proses pengujian sistem pengenalan karakter menggunakan metode *black box* dan *white box*, serta menghitung akurasi keakuratan dari metode *Backpropagation Neural Network* pada citra dokumen sertifikat.

1.5.5 Penarikan Kesimpulan

Tahap akhir yang dilakukan adalah dengan menarik kesimpulan berdasarkan hasil pengujian yang telah dilakukan pada proses pengenalan karakter dengan menggunakan metode *Backpropagation Neural Network*.

1.6 Sistematika Penulisan

Sistematika penulisan skripsi yang akan disusun terkait penelitian pembangunan aplikasi akan dikerjakan sebagai berikut:

BAB 1 PENDAHULUAN

Bab 1 berisi tentang gambaran dari penelitian yang akan dilakukan yang terdiri dari latar belakang masalah, rumusan masalah, identifikasi masalah, maksud dan tujuan, batasan masalah, metode penelitian, dan sistematika penulisan.

BAB 2 TINJAUAN PUSTAKA

Bab 2 menguraikan tentang teori dasar yang terkait dan pembahasan pendukung secara analisis pada permasalahan dengan penelitian yang sedang dilakukan.

BAB 3 ANALISIS DAN PERANCANGAN

Bab 3 membahas pembahasan analisa dan perancangan dari sistem yang akan dibangun diantaranya analisis sistem, analisis data masukan, analisis pada *processing* dan klasifikasi, analisis data keluaranm analisis non fungsional, perancangan basis data, perancangan antar muka, dan perancangan jaringan semantik.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab 4 membahas segala tahapan yang terjadi pada tahapan implementasi dari hasil analisis dan perancangan yang telah ditentukan serta hasil pengujian dari sistem yang dibangun menggunakan pengujian *white box*.

BAB 5 KESIMPULAN DAN SARAN

Isi dari bab 5 sebagai kesimpulan dari hasil penelitian yang telah dilakukan. Selain itu berisi saran kepada pembaca untuk dapat melakukan pengembangan yang lebih baik dari penelitian ini.