



### **Hak cipta dan penggunaan kembali:**

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

### **Copyright and reuse:**

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan ilmu pengetahuan dan teknologi telah memberikan banyak perubahan dalam kehidupan manusia. Perubahan tersebut menuntut manusia untuk menjalankan tugasnya secara efisien dan efektif. Salah satu perubahan yang terjadi adalah diterapkannya *paperless office* pada suatu perusahaan atau instansi. *Paperless office* merupakan pengurangan pemakaian kertas dengan mendigitalkan dokumen yang dimiliki suatu perusahaan atau instansi sehingga mengurangi pemakaian kertas secara fisik. Menurut Susilawati (2012), konsep *paperless office* dapat menunjang kinerja suatu lembaga pada masa kini dan yang akan datang, karena konsep digitalisasinya sangat menunjang dalam penyampaian arus informasi secara cepat, tepat, lengkap, dan yang paling penting efisien.

Untuk dapat mendigitalkan dokumen tercetak, suatu perusahaan atau instansi dapat menggunakan alat pemindai (*scanner*). Akan tetapi, dokumen digital yang dihasilkan dari alat pemindai tidak dapat diedit. Sehingga, untuk mengubah atau memperbaiki dokumen digital yang dihasilkan dari alat pemindai, suatu perusahaan atau instansi membutuhkan teknologi *Optical Character Recognition* (OCR) untuk dapat mengedit dokumen digital tersebut.

Menurut Mithe dkk. (2013), OCR adalah teknologi yang memungkinkan berbagai jenis dokumen dikonversi menjadi data yang dapat diedit dan dicari. OCR dapat digunakan untuk mengenali teks yang ada pada dokumen tercetak.

Namun, menurut Mithe dkk. (2013), kinerja OCR secara langsung bergantung pada kualitas dokumen input.

OCR telah muncul sebagai area penelitian utama sejak tahun 1950 (Bairagi, 2018). Sampai saat ini, OCR telah banyak diteliti dengan berbagai metode. Li dkk. (2016) melakukan penelitian mengenai OCR untuk pengenalan karakter pada naskah bahasa tradisional yang dinamai New Tai Lue menggunakan *backpropagation neural network* dengan akurasi akhir sebesar 98,04%. Selain penelitian yang dilakukan Li dkk. (2016), Radhiah dkk. (2018) melakukan penelitian untuk membandingkan tingkat akurasi pengenalan karakter pada huruf arab tercetak dari dua metode yang berbeda yaitu *backpropagation neural network* dengan *hidden markov model*. Hasil akhir penelitian tersebut menunjukkan bahwa tingkat akurasi dari pengenalan karakter arab terisolasi dengan metode *backpropagation neural network* lebih tinggi yaitu sebesar 100% untuk *font* Arial Unicode Ms, Tahoma, dan Times New Roman. Sementara, tingkat akurasi dari pengenalan karakter arab terisolasi dengan metode *hidden markov model* adalah sebesar 74% untuk *font* Arial Unicode Ms, 61% untuk *font* Tahoma, dan 77% untuk *font* Times New Roman.

Berdasarkan latar belakang tersebut, akan diimplementasikan jaringan saraf tiruan *backpropagation* untuk pengenalan karakter dalam penelitian ini. Pengenalan karakter akan dilakukan pada dokumen tercetak yang sudah berupa citra digital. Citra digital akan diproses dengan operasi morfologi dilasi dan erosi, teknik segmentasi *thresholding*, deteksi tepi *canny*, dan penghalusan citra dengan filter *gaussian*. Penelitian ini dilakukan dengan tujuan untuk mengimplementasikan jaringan saraf tiruan *backpropagation* pada dokumen

tercetak dan untuk mengetahui bagaimana tingkat akurasi dari implementasi jaringan saraf tiruan *backpropagation* pada dokumen tercetak tersebut.

## 1.2 Rumusan Masalah

Rumusan masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut :

1. Bagaimana mengimplementasikan jaringan saraf tiruan *backpropagation* untuk mengenali karakter pada dokumen tercetak ?
2. Bagaimana tingkat akurasi dari implementasi jaringan saraf tiruan *backpropagation* untuk pengenalan karakter pada dokumen tercetak ?

## 1.3 Batasan Masalah

Batasan masalah yang terdapat dalam penelitian ini adalah sebagai berikut :

1. Dokumen tercetak yang dimaksud adalah dokumen yang sudah dicetak dengan tulisan yang diketik menggunakan komputer.
2. Tulisan dalam dokumen tercetak hanya terdiri dari huruf abjad dengan *style regular* atau *bold*.
3. Jenis *font* yang digunakan dalam dokumen tercetak adalah Arial Unicode Ms, Tahoma, atau Times New Roman dengan minimal spasi baris 1,5.
4. Dokumen tercetak dicetak dengan tulisan berwarna hitam dan kertas berwarna putih.
5. Dokumen tercetak yang dijadikan sebagai *input* sudah berupa citra digital yang diambil menggunakan *scanner* dengan format .png atau .jpg.

6. Citra digital yang digunakan sebagai input tidak blur dengan ukuran minimal  $1600 \times 1200$  piksel.

#### 1.4 Tujuan Penelitian

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

1. Mengimplementasikan jaringan saraf tiruan *backpropagation* untuk pengenalan karakter pada dokumen tercetak.
2. Mengetahui tingkat akurasi dari implementasi jaringan saraf tiruan *backpropagation* untuk pengenalan karakter pada dokumen tercetak.

#### 1.5 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah sebagai berikut :

1. Bagi masyarakat, akan sangat membantu dalam mendigitalkan dokumen tercetak sehingga masyarakat dapat mengedit dokumen tercetak dalam format digital.
2. Bagi ilmu pengetahuan, dapat mengetahui tingkat akurasi dari implementasi jaringan saraf tiruan *backpropagation* untuk pengenalan karakter pada dokumen tercetak.
3. Bagi penulis, dapat mengimplementasikan jaringan saraf tiruan *backpropagation* untuk pengenalan karakter pada dokumen tercetak serta menerapkan ilmu yang sudah didapat selama perkuliahan.

## 1.6 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam laporan skripsi ini adalah sebagai berikut.

### BAB I PENDAHULUAN

Bab ini berisi latar belakang dilakukannya penelitian, rumusan masalah, batasan masalah, tujuan masalah, manfaat penelitian, dan sistematika penulisan.

### BAB II LANDASAN TEORI

Bab ini berisi teori-teori yang menjadi landasan dari penelitian yang dilakukan, seperti *Optical Character Recognition* (OCR), pengolahan citra digital, jaringan saraf tiruan, arsitektur jaringan saraf tiruan, *backpropagation*, dan fungsi aktivasi.

### BAB III METODOLOGI DAN PERANCANGAN SISTEM

Bab ini berisi metodologi penelitian, perancangan jaringan saraf tiruan, *flowchart*, dan rancangan antarmuka aplikasi.

### BAB IV IMPLEMENTASI DAN UJI COBA

Bab ini berisi spesifikasi sistem yang digunakan dalam penelitian ini, implementasi tampilan aplikasi, implementasi algoritma, uji coba jaringan saraf tiruan, dan evaluasi hasil.

### BAB V SIMPULAN DAN SARAN

Bab ini berisi simpulan dari hasil implementasi jaringan saraf tiruan *backpropagation* untuk pengenalan karakter pada dokumen tercetak dan saran untuk pengembangan penelitian selanjutnya.