



**KLASIFIKASI CITRA TEKS UNTUK PENERJEMAHAN
BAHASA DAERAH MENGGUNAKAN *ARTIFICIAL NEURAL
NETWORK***

SKRIPSI

FAREL FATHURRAHMAN

1610511083

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN

JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

2020



**KLASIFIKASI CITRA TEKS UNTUK PENERJEMAHAN
BAHASA DAERAH MENGGUNAKAN *ARTIFICIAL NEURAL
NETWORK***

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Komputer**

FAREL FATHURRAHMAN

1610511083

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN

JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

2020

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Farel Fathurrahman

NIM : 1610511083

Tanggal : 18 Juni 2020

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 18 Juni 2020

Yang Menyatakan,



(Farel Fathurrahman)

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI
UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Farel Fathurrahman

NIM : 1610511083

Fakultas : Ilmu Komputer

Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

**Klasifikasi Citra Teks Untuk Penerjemahan Bahasa Daerah Menggunakan
*Artificial Neural Network***

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 18 Juni 2020

Yang Menyatakan,



(Farel Fathurrahman)

PENGESAHAN

Dengan ini dinyatakan bahwa Tugas Akhir berikut:

Nama : Farel Fathurrahman

NIM : 1610511083

Program Studi : Informatika

Judul Tugas Akhir : KLASIFIKASI CITRA TEKS UNTUK
PENERJEMAHAN BAHASA DAERAH
MENGUNAKAN *ARTIFICIAL NEURAL NETWORK*

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Henki Bayu Seta, S.Kom, MTI.
Ketua Penguji



Ing. Artambo B. Pangaribuan, BSc.
Anggota Penguji



Anita Muliawati, S.Kom., MTI.
Pembimbing I



Mayanda Mega Santoni, S.Kom,
M.Kom.
Pembimbing II



Dr. Ermanita, M.Kom.
Dekan



Anita Muliawati, S.Kom., MTI.
Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 18 Juni 2020



**KLASIFIKASI CITRA TEKS UNTUK PENERJEMAHAN
BAHASA DAERAH MENGGUNAKAN *ARTIFICIAL NEURAL
NETWORK***

FAREL FATHURRAHMAN

ABSTRAK

Bahasa daerah adalah bahasa yang digunakan untuk saling berkomunikasi di daerah-daerah tertentu. Banyak faktor yang membuat lemahnya kesadaran generasi sekarang ini untuk melestarikan bahasa daerahnya. Salah satunya adalah masih kurangnya sarana yang dapat digunakan untuk mengakses informasi dari bahasa daerah itu sendiri, sehingga ini menjadi salah satu kendala yang terjadi. Penelitian ini akan merancang sistem penerjemahan sebuah gambar/citra berisi teks bahasa Indonesia menjadi sebuah teks dalam bahasa daerah. Penelitian ini dimulai dari tahap pra proses, teknik segmentasi karakter pada citra menggunakan pelabelan *Connected Component Analysis*, lalu citra di lakukan ekstraksi kemudian citra karakter diklasifikasikan menggunakan metode *Artificial Neural Network*. Tahap selanjutnya, dilakukan penggabungan karakter menjadi sebuah teks. Setelah itu, proses penerjemahan menggunakan algoritma *Levensthein* untuk mencocokkan hasil klasifikasi teks dengan bahasa daerah. Penelitian ini diharapkan mampu untuk menerjemahkan citra teks berbahasa Indonesia menjadi teks bahasa daerah, untuk membantu melestarikan bahasa daerah di Indonesia.

Kata Kunci : bahasa daerah, *cca*, *artificial neural network*, penerjemahan bahasa daerah

TEXT IMAGE CLASSIFICATION FOR LOCAL LANGUAGE TRANSLATION USING ARTIFICIAL NEURAL NETWORK

FAREL FATHURRAHMAN

ABSTRACT

Local language are used to communicate with each other in certain areas. Many factors make the young generation's awareness to preserve the local language. One of them is still lack of means to access the local languages, so this is one of the problems that occurs. In this study, would design a translation system of an image containing Indonesian text into a local language text. This research starts from the pre-process stage, character segmentation techniques in the image using Connected Component Analysis labeling. After that, the image is extracted then the character image is classifying using the Artificial Neural Network method. The next step is to merge characters into a text. After that, the translation process uses the Levensthein algorithm to match the results of the text classification with local language. This research is expected to be able to translate the image of Indonesian texts into local language text, to help preserve local language in Indonesia.

Keywords : *local language, cca, artificial neural network, local language translation*

KATA PENGANTAR

Puji Syukur penulis panjatkan kepada Allah Subhanahu Wa Ta'ala yang sudah melimpahkan rahmat dan hidayah-Nya, penulis mampu menyelesaikan skripsi yang berjudul “Klasifikasi Citra Teks Untuk Penerjemahan Bahasa Daerah Menggunakan *Artificial Neural Network*” dapat diselesaikan.

Pada kesempatan ini penulis menyampaikan penghargaan dan ucapan terima kasih kepada:

1. Kedua orang tua peneliti yang bernama Ambang Mulya, dan Ibu Yetrina beserta adik saya yang bernama Fathan Athallah Rahman dan Falby Naufal Rahman serta seluruh keluarga besar yang selalu mendukung dan mendoakan serta memotivasi.
2. Ibu Dr. Ermatita, M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jakarta.
3. Ibu Anita Muliawati, S.Kom., MTI. selaku Ketua Program Studi jurusan Informatika dan juga selaku pembimbing skripsi.
4. Ibu Mayanda Mega Santoni, S.Kom, M.Kom, selaku Dosen Pembimbing skripsi yang sudah meluangkan waktu dan bersedia untuk membimbing penulis dalam menyusun skripsi ini dari awal hingga akhir penulisan.
5. Teman-teman penulis Hadaya, Kevin, Jamal, Billy, Iqbal, Haris, dan Kubis.
6. Teman-teman seperbimbingan Donni, Aksen, Andre, Maul, Basith, Abi, dan Rafif.
7. Teman-teman Program Studi S1 Informatika Angkatan 2016 Fakultas Ilmu Komputer.
8. Semua pihak yang tidak dapat disebutkan satu per satu yang telah membantu penulis selama ini.

Akhir kata, penulis berharap Allah SWT berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga skripsi ini dapat bermanfaat bagi pihak – pihak yang membutuhkan.

Jakarta, 18 Juni 2020

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
PERNYATAAN ORISINALITAS	ii
LEMBAR PERSETUJUAN.....	iii
ABSTRAK	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiii
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan Penelitian.....	3
1.4. Manfaat Penelitian.....	3
1.5. Ruang Lingkup	3
1.6. Luaran yang di harapkan	4
1.7. Sistematika Penulisan.....	4
BAB 2 TINJAUAN PUSTAKA	6
2.1. Penelitian Terdahulu.....	6
2.2. Pengolahan Citra	6
2.1.1. Citra Biner.....	7
2.1.2. <i>Thresholding</i>	7
2.3. Ekstraksi Fitur	8

2.4.	<i>Artificial Neural Network</i>	8
2.4.1.	<i>Multi Layer Perceptron (MLP)</i>	8
2.5.	<i>Connected Component Analysis</i>	9
2.6.	<i>Edit Distance</i>	10
BAB 3 METODOLOGI PENELITIAN.....		12
3.1.	Tahapan Penelitian	12
3.1.1.	Perumusan Masalah	12
3.1.2.	Studi Pustaka.....	12
3.1.3.	Pengumpulan Data	13
3.1.4.	Perancangan Sistem	14
3.1.5.	Implementasi Sistem	15
3.1.6.	Pengujian Sistem.....	17
3.1.7.	Evaluasi Hasil.....	18
3.2.	Perangkat Penelitian	18
3.2.1.	Perangkat Keras (<i>Hardware</i>)	18
3.2.2.	Perangkat Lunak (<i>Software</i>).....	19
3.3.	Jadwal Penelitian	19
BAB 4 HASIL DAN PEMBAHASAN.....		20
4.1	Pengumpulan Data Citra	20
4.2	Perancangan Sistem.....	23
4.2.1.	Pra Proses Data Citra	23
4.2.2.	Ekstraksi Ciri.....	25
4.2.3.	Perancangan ANN.....	26
4.2.4.	Penentuan Parameter	29
4.2.5.	Implementasi Sistem	35
4.2.6.	Proses Pengujian	35

4.3	Pengujian Sistem	37
4.4	Evaluasi Sistem	39
BAB V PENUTUP.....		42
5.1	Kesimpulan.....	42
5.2	Saran.....	42
DAFTAR PUSTAKA		44
RIWAYAT HIDUP.....		1
LAMPIRAN.....		3

DAFTAR GAMBAR

Gambar 1 Huruf “B” dan representasi biner dari derajat keabuannya.....	7
Gambar 2 Citra <i>Thresholding</i>	7
Gambar 3 Arsitektur MLP sederhana (<i>credit: Kriesel</i>).....	9
Gambar 4 4-Konektivitas (<i>4-Connected Neighbors</i>)	10
Gambar 5 8-Konektivitas (<i>8-Connected Neighbors</i>)	10
Gambar 6 Tahapan Penelitian	12
Gambar 7 <i>Flowchart</i> Perancangan Sistem.....	14
Gambar 8 Proses <i>testing</i> dan <i>training</i> pada ANN.....	17
Gambar 9 Proses Pengambilan Data	21
Gambar 10 Proses <i>Cropping</i> Citra Alfabet dan Citra Teks.....	21
Gambar 11 Citra Data Huruf Kapital	22
Gambar 12 Citra Data Huruf Kecil	22
Gambar 13 Penamaan Citra Data Teks	22
Gambar 14 Data Citra .csv	24
Gambar 15 Pra <i>Processing</i> Pada Citra	24
Gambar 16 Hasil Segmentasi	24
Gambar 17 Ukuran citra baru 32 x 32 piksel	25
Gambar 18 Pelatihan dan pengujian ANN.....	26
Gambar 19 Rancangan Arsitektur ANN	28
Gambar 20 Model <i>Summary</i>	28
Gambar 21 <i>Classification Report</i>	35
Gambar 22 Hasil Klasifikasi Citra Teks	36
Gambar 23 Hasil Terjemahan	37
Gambar 24 Citra rusak	39
Gambar 25 Kesalahan proses segmentasi pada citra rusak.....	40
Gambar 26 Kesalahan Pelabelan Objek Bukan Huruf.....	40

DAFTAR TABEL

Tabel 1 Contoh Operasi pada <i>Edit Distance</i>	11
Tabel 2 Jadwal Penelitian.....	19
Tabel 3 Contoh Penamaan Citra Data Alfabet.....	22
Tabel 4 Contoh Penamaan Citra Data Teks	23
Tabel 5 Hasil Ekstraksi Ciri	25
Tabel 6 Percobaan <i>Batch Size</i>	29
Tabel 7 Percobaan <i>Epoch</i>	30
Tabel 8 Percobaan <i>Learning Rate</i>	32
Tabel 9 Akurasi model berdasarkan nilai <i>Batch Size</i>	33
Tabel 10 Akurasi model berdasarkan nilai <i>Epoch</i>	33
Tabel 11 Akurasi model berdasarkan nilai <i>Learning Rate</i>	34
Tabel 12 Akurasi Klasifikasi dari setiap <i>Device</i> dan <i>Font</i>	38
Tabel 13 Akurasi Terjemahan dari setiap <i>Device</i> dan <i>Font</i>	39
Tabel 14 Contoh Klasifikasi Benar dan Terjemahan Benar.....	40
Tabel 15 Contoh Klasifikasi Salah dan Terjemahan Benar	41
Tabel 16 Contoh Klasifikasi Salah dan Terjemahan Salah	41