

**PENGENALAN POLA HURUF JEPANG
MENGUNAKAN *TESSERACT ENGINE***

Skripsi

untuk memenuhi sebagian persyaratan

mencapai derajat Sarjana S-1



Disusun oleh :

Akhmad Imam Fahrizal

07650011

**Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Sunan Kalijaga
Yogyakarta
2013**



PENGESAHAN SKRIPSI/TUGAS AKHIR

Nomor : UIN.02/D.ST/PP.01.1/2694/2013

Skripsi/Tugas Akhir dengan judul : Pengenalan Pola Huruf Jepang Menggunakan *Tesseract Engine*

Yang dipersiapkan dan disusun oleh :
Nama : Akhmad Imam Fahrizal
NIM : 07650011
Telah dimunaqasyahkan pada : Rabu, 28 Agustus 2013
Nilai Munaqasyah : B -
Dan dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga

TIM MUNAQASYAH :

Ketua Sidang

Shofwatul 'Uyun, M.Kom
NIP. 19820511 200604 2 002

Penguji I

Sumarsono, M.Kom
NIP.19710209 200501 1 003

Penguji II

Nurochman, M.Kom
NIP. 19801223 200901 1 007

Yogyakarta, 10 September 2013
UIN Sunan Kalijaga
Fakultas Sains dan Teknologi
Dekan



Prof. Drs. H. Akh. Minhaji, M.A, Ph.D
NIP. 19580919 198603 1 002



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal :
Lamp :

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Akhmad Imam Fahrizal
NIM : 07650011
Judul Skripsi : Aplikasi Pengenalan Pola Huruf Jepang dengan Tesseract Engine

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam 6 tahun Bulan

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 20 Agustus 2013

Pembimbing

Shofwatul Uyun, M.Kom

NIP. 19820511-200604-02-001

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini:

Nama : Akhmad Imam Fahrizal

NIM : 07650011

Program Studi : Teknik Informatika

Fakultas : Sains dan Teknologi

Menyatakan bahwa skripsi dengan judul “**APLIKASI PENGENALAN POLA HURUF JEPANG DENGAN *TESSERACT ENGINE***” tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 20 Agustus 2013

Yang menyatakan,



Akhmad Imam Fahrizal

NIM. 07650011

KATA PENGANTAR

Alhamdulillah, segala puji bagi Allah *Subhanahu wa ta'ala* atas limpahan rahmat, hidayah, serta bimbingan-Nya. Shalawat serta salam semoga tercurah kepada Nabi Muhammad *Shallallohu 'alaihi wa sallam*. Akhirnya penulis dapat menyelesaikan penelitian Tugas Akhir yang berjudul **Aplikasi Pengenalan Pola Huruf Jepang dengan Tesseract Engine**. Oleh karena itu, dengan segala kerendahan hati pada kesempatan ini penulis mengucapkan banyak terima kasih kepada:

1. Prof. Drs. H. Akh. Minhaji, M.A., Ph.D selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga.
2. Bapak Agus Mulyanto, S.Si, M.Si. selaku Ketua Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga.
3. Ibu Shofwatul 'Uyun, M.Kom, selaku pembimbing yang selalu sabar membimbing, mengarahkan, memberikan nasehat dan saran selama penyusunan skripsi.
4. Keluarga Penulis yang telah memberikan dukungan selama penyusunan skripsi.
5. Seluruh teman-teman keluarga besar Program Studi Teknik Informatika, khususnya angkatan 2007 yang telah banyak sekali memberikan masukan, saran dan diskusi yang begitu berharga.

6. Serta semua rekan-rekan penulis di berbagai kegiatan maupun organisasi yang juga telah memberikan banyak sekali masukan dan kontribusi yang sangat berarti bagi penulis

Penulis merasa masih banyak sekali kekurangan dan kelemahan dalam penelitian ini, oleh karena itu segala kritik dan saran senantiasa penulis harapkan dari para pembaca. Akhir kata, semoga penelitian ini dapat menjadi panduan serta referensi yang sangat berguna bagi pembaca dan dapat dimanfaatkan sebaik-baiknya.

Yogyakarta, 20 Agustus 2013

Penulis

HALAMAN PERSEMBAHAN

Teriring ucapan syukur yang mampu ku ucapkan kepada-Mu ya Allah, semoga shalawat dan salam senantiasa tercurah kepada Baginda Nabi Muhammad SAW. Aku bersyukur kepadamu alhamdulillah, dengan bimbingan, karunia dan petunjuk-Mu, telah berhasil kuselesaikan tugas akhir ini. Ya Allah aku hanyalah manusia biasa, yang tanpa bimbinganmu aku bukanlah apa-apa. Aku hanyalah manusia biasa, yang tanpa orang-orang disekitarku aku juga bukanlah siapa-siapa... Untuk itu Ya Allah, ijinlanlah aku untuk menyampaikan rasa terimakasihku pada-Mu dan pada mereka... dan sudilah kiranya Engkau senantiasa menjaga dan menjaga mereka dijalan-Mu Ya Allah.. dari relung hati yang terdalam kusampaikan rasa terimakasihku :

- *Ibunda Rahmawati dan ayahanda Sri Istahdi tercinta yang sejak lahir sampai sekarang dan kelak akan senantiasa memberikan dukungan, do'a dan pengorbanannya untuk-ku. Terima kasih banyak karena selalu menyelipkan namaku disetiap do'a sepanjang sholat malam kalian. Ya Allah hanya seuntai doa yang aku panjatkan, Ampunilah dosa-doa mereka, sayangilah mereka seperti mereka menyayangiku diwaktu kecil dan Masukkanlah mereka kedalam surga-Mu.... Amiin.*
- *Kedua Adik Ku Tercinta Miftahur Rizky dan Mahfuzatun Ni'mah Sona*
- *Prof. Drs. H. Akh. Minhaji, M.A.,Ph.D selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga.*
- *Bapak Agus Mulyanto, M. Kom, selaku Ketua Program Studi Teknik Informatika. Keberadaan beliau dalam menuntut ilmu sudah seperti Bapak kedua bagi penulis yang selalu*

sabar membimbing, mengarahkan, memberikan nasehat dan saran yang tak ternilai harganya. Semoga Allah selalu melindungi Pak Agus dan keluarga :).

- *Ibu Shofwatul 'Uyun, M.Kom selaku Dosen Pembimbing yang telah banyak sekali memberikan ilmu-ilmu dan diskusi yang sangat mengakselerasi bagi pertambahan pengetahuan penulis, memberikan saran serta masukan kepada penulis dalam penyusunan skripsi. Semoga Allah senantiasa memberikan kemudahan dan petunjuk-Nya untuk bu 'Uyun dan keluarga :).*
- *Para dosen Teknik Informatika, Pak Mustaqim, Pak Sumarsono, Pak Nurrochman, Pak Bambang, Pak Nasirudin, Bu Ade, Bu Uyun, Bu Ulfah, Pak Taufik, Pak Anshari dan pak Bambang Robiin., Pak Aulia. terimakasih untuk semua ilmu yang telah dibagikan kepadaku... semoga Allah senantiasa memberikan kemudahan dan petunjuk-Nya untuk Bapak/Ibu dosen sekalian...*
- *Sahabat-sahabat seperjuangan Skripsi Super informatika, Mohammad Shobirin, Putera Aditya Dasopang, Andry Rachmadi, Fathur Rahman, Ismail Sembiring, Alfian Hidayat, Setya budi, M. Husna Mubarak, Mardiana Emi Nucifera, Annisa Affida, Fatma, dan banyak lagi yang lainnya yang tidak bisa aku sebutkan satu persatu. Banyak hal yang aku pelajari dari kalian. Semoga Allah selalu memberikan kemudahan untuk kalian sahabat ku.*
- *Sahabat Sahabat Informatika yang sudah lebih dahulu wisuda.*
- *Teman-Teman Dari Asrama Mahasiswa Kalimantan Selatan (AMKS) – HSU.*
- *Semua pihak yang telah memberikan bantuan dan dukungan kepada penulis dalam penyusunan skripsi yang tidak bisa disebutkan satu per satu.*

HALAMAN MOTTO

これは、ピースのケーキです

It's A Piece Of Cake

(Nanba Mutta – Uchuu Kyoudai)

I alone know that I'm okay with what I got

(Nobunaga Oda)

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN SKRIPSI / TUGAS AKHIR.....	ii
HALAMAN PERSETUJUAN SKRIPSI / TUGAS AKHIR.....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
HALAMAN PERSEMBAHAN	vii
HALAMAN MOTTO	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xiv
DAFTAR LAMPIRAN.....	xv
INTISARI.....	xvi
ABSTRACT	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Keaslian Penelitian	4
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI	5
2.1 Tinjauan Pustaka	5

2.2	Landasan Teori.....	9
2.2.1	Citra Digital.....	9
2.2.1.1	Definisi Citra Digital.....	9
2.2.1.2	Format Citra Digital.....	14
2.2.1.2.1	JPEG.....	14
2.2.1.2.2	TIFF.....	15
2.2.1.2.3	PNG.....	16
2.2.2	Pengolahan Citra Digital.....	16
2.2.3	Pengenalan Pola.....	17
2.2.3.1	Tesseract Engine.....	18
2.2.3.2	Pelatihan Tesseract.....	20
2.2.4	Bahasa Jepang.....	22
2.2.4.1	Lafal Vokal.....	23
2.2.4.2	Tulisan Bahasa Jepang.....	23
2.2.4.3	Kana.....	24
2.2.4.4	Hiragana.....	25
2.2.4.4.1	Definisi Hiragana.....	25
2.2.4.4.2	Kegunaan Hiragana.....	25
2.2.4.4.3	Huruf Hiragana.....	26
BAB III METODE PENELITIAN.....		29
3.1	Studi Pendahuluan.....	29
3.2	Pengumpulan Data.....	29
3.3	Kebutuhan Sistem.....	29

3.3.1	Perangkat Keras	29
2.2.2	Perangkat Lunak.....	30
3.4	Metode Pengenalan Pola	30
BAB IV HASIL DAN PEMBAHASAN		32
4.1	Tahap Pembuatan Gambar	32
4.1	Tahap Pelatihan	32
4.2.1	Pembuatan Box Files	33
4.2.2	Tahap Pelatihan Box	34
4.2.3	Unicharset	35
4.2.4	Font Properties	36
4.2.5	Clustering	36
4.2.6	Dictionary Data	38
4.2.6	Mengganti Nama.....	39
4.2.8	Penggabungan	40
4.3	Tahap Pengenalan Huruf.....	40
4.4	Pengujian Sistem	42
BAB V KESIMPULAN.....		46
5.1	Kesimpulan.....	46
5.2	Saran.....	46
DAFTAR PUSTAKA		48
LAMPIRAN.....		50

DAFTAR GAMBAR

Gambar 2.1 Cara kerja tesseract.....	20
Gambar 2.2 Diagram Alir Pelatihan.....	21
Gambar 2.3 Hasil File	22
Gambar 2.4 Huruf Vokal dan Konsonan.....	26
Gambar 2.5 Gabungan Huruf Hiragana	27
Gambar 3.1 Diagram Alir Sistem	31
Gambar 4.1 Gambar Untuk Data Pelatihan.....	32
Gambar 4.2 Tampilan BoxMaker	33
Gambar 4.3 Tampilan JTessBoxEditor	34
Gambar 4.4 Pelatihan Box	35
Gambar 4.5 Proses Unicharset.....	36
Gambar 4.6 Tahap Shapeclustering	37
Gambar 4.7 Tahap mftraining.....	38
Gambar 4.8 Tahap cntraining.....	38
Gambar 4.9 Tampilan Antar Muka	41
Gambar 4.10 Gambar dengan 73 Huruf dan ukuran huruf 11	42
Gambar 4.11 Gambar dengan 73 Huruf dan ukuran huruf 18	43
Gambar 4.12 Gambar dengan 279 Huruf dan ukuran huruf 11	43
Gambar 4.13 Gambar dengan 279 Huruf dan ukuran huruf 18	44
Gambar 4.14 Contoh Gambar Tulisan Tangan Huruf Hiragana	44
Gambar 6.1 Gambar Original Pertama dengan ukuran huruf 11	51
Gambar 6.2 Gambar Original Kedua dengan ukuran huruf 11	51
Gambar 6.3 Contoh Gambar Hasil Scan Tulisan Tangan	51

DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian.....	8
Tabel 4.1 Dictionary Data.....	39
Tabel 4.2 Hasil Terbaik 56 Kali Pengujian.....	45
Tabel 6.1 Hasil Pengujian.....	52



DAFTAR LAMPIRAN

Lampiran A Perbandingan hasil pengujian.....	51
Lampiran B Pengujian Tulisan tangan.....	63
Curriculum Vitae.....	73



Aplikasi Pengenalan Pola Huruf Jepang dengan *Tesseract Engine*

Akhmad Imam Fahrizal

NIM. 07650011

INTISARI

Akhir-akhir ini pengolahan citra digital di banyak negara maju menjadi bidang yang digeluti oleh banyak peneliti karena menarik untuk diterapkan pada berbagai kegiatan, baik kegiatan analisis maupun produksi. Salah satu cabang dalam dari citra digital adalah pengenalan pola. Penelitian ini menggunakan *Tesseract* sebagai alat untuk mengenali pola dari huruf Hiragana. Penelitian ini dilakukan untuk mengetahui seberapa besar *Tesseract* mampu mengenali sebuah teks Jepang dan juga teks tulisan tangan.

Penelitian ini menggunakan 1 citra sebagai data latih yang berisi 74 huruf hiragana yang diproses melalui sebuah pelatihan dan menghasilkan data pelatihan untuk masing-masing huruf. Pada penelitian mempunyai beberapa kriteria pengujian berdasarkan ukuran huruf dan juga resolusi untuk mencari hasil terbaik dalam pengenalan pola

Sistem ini mampu mengenali 74 Huruf Hiragana dengan memakai *Tesseract Engine*. Sistem pengenalan pola ini juga mampu melakukan pelatihan data menggunakan *Tesseract Engine*. Sistem juga dapat mengenali citra dengan prosentase keberhasilan terbaik 98,24 % dengan resolusi gambar 200dpi dan ukuran huruf 18. Sistem ini juga bisa mengenali citra tulisan tangan dengan prosentase keberhasilan terbaik 90 % dengan resolusi gambar 200dpi.

Kata Kunci : sistem pengenalan pola, *Tesseract Engine*, Huruf Hiragana.

Pattern Recognition Letters Japanese applications with Tesseract Engine

Akhmad Imam Fahrizal

NIM. 07650011

ABSTRACT

Lately, digital image processing in many developed countries into fields cultivated by many researchers as attractive to apply to various activities, both analysis and production activities. One of the branches in the digital image is pattern recognition. This study uses Tesseract as a tool to recognize patterns of letters Hiragana. Penelitian was conducted to determine how much Tesseract is able to recognize a Japanese text and handwritten text.

This study uses one image as the training data which contains 74 letters hiragana are processed through a training and generate training data for each letter. In the study has several testing criteria based on font size and resolution for the best match in pattern recognition.

The system is able to recognize 74 letters Hiragana using Tesseract Engine. Pattern recognition system is also capable of doing the training data using Tesseract Engine. The system can also recognize the image with the best percentage of success 98,24 % with image resolutions 200dpi dan font size 18. This system can also recognize handwriting image with the best percentage of success 90 % With an image resolution 200dpi

Keywords: pattern recognition system, *Tesseract Engine*, Hiragana Letter.

BAB I

PENDAHULUAN

1.1. Latar Belakang

Sekarang ini masyarakat sudah mulai tertarik untuk mempelajari bahasa-bahasa asing diluar bahasa mereka sendiri, mulai dari bahasa jepang, inggris, perancis, jerman, madarin dan lain lain. mereka mulai membeli buku-buku ataupun kamus untuk mendukung pembelajaran mereka. pengembangan untuk pembelajaran itu sendiri sudah mulai berkembang pesat, dimulai dari kamus elektronik sampai *e-learning*.

Salah satu cabang dalam Kecerdasan buatan, Jaringan syaraf tiruan memiliki banyak keunggulan, diantaranya memungkinkan komputer untuk melakukan training (menerima set *input* dan menetapkan target yang dicapai). Dengan evolusi ilmu komputer yang terus berkembang, muncul kebutuhan untuk membangun suatu aplikasi yang mempunyai kemampuan untuk belajar.

Seperti yang diketahui, *input* untuk huruf atau kata dapat dilakukan dengan berbagai macam cara seperti keyboard. Akan tetapi dengan berkembangnya kesibukan manusia. Maka diperlukan suatu metode yang lebih praktis dalam melakukan *input*.

Input akan menjadi lebih mudah apabila komputer dapat mengenali karakter-karakter atau huruf yang di-*input*, baik itu secara cetak maupun tulisan tangan, dan telah diketahui bahwa huruf atau karakter itu tidak terbatas pada huruf latin atau abjad yang dikenal. Terdapat huruf atau karakter yang penulisannya

bersifat gambar atau lambang yang lebih dikenal dengan kaligrafi. Bergabai macam karakter atau huruf yang termasuk kaligrafi, diantaranya Mandarin, Jepang, Arab, dan masih banyak lagi.

Bahasa Jepang Sendiri punya tiga jenis huruf yaitu Hiragana, katagana, dan Kanji. Akan tetapi yang paling umum di temui dan paling mudah untuk dipelajari adalah huruf Hiragana, sedangkan Katagana biasanya dipakai untuk kata serapan asing, dan kanji biasanya lumayan sulit untuk dipelajari.

Metode untuk mengenali huruf itu sendiri ada bermacam-macam. Yang paling sering ditemui adalah dengan metode *back propagation*, akan tetapi peneliti mencoba untuk memakai *Tesseract Engine* karena keakuratan untuk mengenali huruf dengan menggunakan *Tesseract Engine* melebihi 90% (Ager, 2007).

Namun persentase keberhasilan tersebut adalah persentase keberhasilan untuk mengenali huruf Alfabet. Oleh karena itu peneliti mengajukan penelitian ini untuk mengetahui seberapa besar *Tesseract* bisa mengenali tulisan huruf Hiragana.

1.2. Rumusan Masalah

Berdasarkan Latar Belakang yang telah disebutkan diatas maka dapat disusun perumusan masalah yang perlu diselesaikan dalam penelitian ini yaitu :

1. Seberapa besar tingkat akurasi *Tesseract* dalam mengenali tulisan tangan huruf jepang dan teks jepang ?
2. Berapa prosentase terbaik *Tesseract* sehingga bisa mengenali tulisan

tangan huruf jepang dan teks jepang dengan baik dengan parameter yang telah ditentukan ?

1.3. Batasan Masalah

Batasan Masalah penelitian ini adalah sebagai berikut :

1. Aplikasi ini hanya menerima masukan Huruf Hiragana
2. Aplikasi ini hanya bisa membaca masukan dari pengguna per gambar.
3. Aplikasi ini hanya menerima input 3 tipe gambar (JPG, TIFF, PNG)
4. Parameter yang dipakai adalah besarnya huruf (11 dan 18) dan resolusi (100dpi, 200dpi, 300dpi, dan 600dpi).
5. Dalam pengujian, tipe file yang digunakan hanya TIFF.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Mengetahui seberapa akurat *tesseract* bisa mengenali tulisan tangan dan juga teks hiragana.
2. Mengetahui prosentase terbaik untuk mengenali tulisan tangan dan juga teks hiragana dengan parameter yang telah ditentukan.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

1. Menambahkan variasi baru dalam hal pengenalan pola huruf jepang.

2. Sebagai bahan referensi untuk penelitian lain
3. Pemahaman tentang *Tesseract Engine*

1.6.1 Keaslian Penelitian

Penelitian mengenai sistem pengenalan huruf sudah pernah dilakukan, akan tetapi tidak banyak yang membahas tentang sistem pengenalan huruf jepang yang memakai *Tesseract Engine*. Penelitian ini fokus terhadap konsep *Tesseract Engine*.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasar pada penelitian yang telah dilakukan oleh penulis mengenai sistem pengenalan pola huruf jepang menggunakan *Tesseract Engine*, maka dapat diambil kesimpulan sebagai berikut:

1. Pengenalan pola tulisan tangan huruf Hiragana dengan menggunakan *Tesseract Engine* dapat mengenali tulisan tangan dengan prosentase keberhasilan 52,90 % dan Prosentase keberhasilan untuk mengenali teks Hiragana sebesar 91,75 %
2. *Tesseract* bisa mengenali tulisan tangan dengan baik dalam keadaan resolusi 200dpi dengan prosentase keberhasilan 90 % , sedangkan untuk teks hiragana dalam ukuran huruf 18 dan resolusi 200dpi dengan prosentase keberhasilan 98,24 %

5.2 Saran

Penelitian yang telah dilakukan tentunya tidak lepas dari kekurangan pada setiap eksperimen. Beberapa saran dari peneliti untuk penelitian berikutnya adalah :

1. Dalam data pelatihan untuk huruf hiragana, peneliti tidak memakai data kamus, sehingga kemampuan sistem ini untuk membaca masukan dari pengguna masih kurang maksimal.

2. Dalam hal memasukkan hasil dari pelatihan data, pengguna juga harus memasukkan secara manual ke dalam folder Tesseract-OCR yang ada dalam komputer pengguna.



DAFTAR PUSTAKA

- Ager, Simon.1998. *Japanese (Nihongo)* Url :
<http://www.omniglot.com/writing/japanese.htm> diakses 29 Januari 2013.
- Astamar, Muhammad Fachrul. 2012. *Aplikasi pengenalan bahasa jepang berbasis android*. Jakarta : BINUS.
- Daniel, Michael Witanto. 2006. *Pengenalan Karakter Mandarin dengan Metode Back Propagation*. Jakarta : BINUS.
- Danoedoro, Projo. 1996. *Pengolahan Citra Digital*. Fakultas Geografi. Universitas Gadjah Mada : Yogyakarta
- Eriawan, Yoddy. 2007. *Aplikasi pengenalan aksara-aksara dasar huruf jepang dengan menggunakan macromedia flash 8*. Depok : Universitas Gunadharma.
- Ginting, Priskilla Br.2010. *Kajian Steganografi Dengan Metode Bit-Plane Complexity Segmentation (BPCS) Pada Dokumen Citra Terkompresi*. Universitas Sumatra Utara : Sumatra Utara.
- Handariningsih, Riski Purwo. 2012 *Application of Recognition and analysis of handwriting character using Freeman Chain methode*. Tesis. Jakarta.
- Hasnan, Ahmad. 2009. *Pengantar Bahasa Jepang dan Indonesia*. Url :
<http://bando07.files.wordpress.com/2009/10/bahasa-jepang-dan-indonesia1.pdf>
diakses : 18 Agustus 2013.
- Kardian, Aqwam Rosadi. 2012. *Pengolahan Citra Digital* Url :
<http://aqwam.staff.jak-stik.ac.id/files/30.-pengolahan-citra%5B7%5D.pdf>
diakses : 18 Agustus 2013.
- Lioyod, Colin, Tsujita, Etsuko.2005.*Mahir Berbahasa Jepang dalam sepekan*. Kesaint Blanc. Yogyakarta.
- Nurhidayati, Ajeng.2007.*Aplikasi Penerjemah semantik bahasa jepang ke bahasa indonesia dengan algoritma pencocokan pola*. FMIPA UGM. Yogyakarta.
- ShinjukuCenter.2006. *Hiragana*. Url : <http://www.shinjukucenter.com/hiragana.php>
diakses : 18 Agustus 2013.
- Sholahudin, Asep. 2012. *Penerapan Neural Network Tentang Metode Propagation pada pengenalan Pola Huruf*. Makalah Proyek Akhir. Bandung.

- Smith, Ray. 2007. *Tesseract OCR engine What it is, where it came from, where it is going*. Url : tesseract-ocr.googlecode.com/files/TesseractOSCON.pdf diakses : 29 Januari 2013.
- Starostin, Sergei A., Anna V. Dybo, and Oleg A. Mudrak. 2003. *Etymological Dictionary of the Altaic Languages, 3 volumes*. Leiden: Brill Academic Publishers.
- Surajaya, Ketut. 1999. *Bahasa Jepang Modern*. Erlangga. Jakarta.
- Sutoyo, T. Dkk. 2009. *Teori Pengolahan Citra Digital*. Yogyakarta: Penerbit Andi.
- Tania, Ricky. 2008. *Aplikasi Pembelajaran Bahasa Jepang Mengenal Huruf Kanji Berbasis computer*. Bandung: Unikom.
- Tyagita, Benediktus Aditya Citta. 2010. *Perancangan Program Aplikasi Pengenalan Pola Katakana Menggunakan Metode Transformasi Wavelet dan Back Propagation*. Jakarta: BINUS.
- Utari, Widya Ayu. 2012. *Pengenalan pola dengan menggunakan metode Back Propagation menggunakan Matlab*. Depok: Universitas Gunadharma.
- Verstraeten, Cédric. . *How to Train Tesseract 3.01*. Url : <http://www.cedricve.me/2013/04/12/how-to-train-tesseract/> Diakses : 12 Juni 2013.
- White, Nick. 2012 *Training Tesseract for Ancient Greek OCR*. Department of Clasics and Ancient History, Durham, UK.



LAMPIRAN

LAMPIRAN A

Perbandingan Hasil Pengujian dengan gambar yang dipakai untuk menguji

Ada 56 tipe gambar yang dipakai oleh penguji, dengan hanya 2 gambar original tetapi dibagi-bagi berdasarkan resolusi dan besar huruf dan 10 gambar hasil *scan* tulisan tangan huruf hiragana . 2 gambar original tersebut bisa dilihat pada gambar 6.1 dan gambar 6.2 sedangkan contoh gambar tulisan tangan bisa dilihat pada gambar 6.3

ゆめものがたり ゆめ たぶ んなさけ よろしくたのみます よしき せんしょ せつとく
 なかなか ひとちがい げいしゃ ふだどめになる でんせつ あくまで いじましい へん
 しょく

Gambar 6.1 gambar original pertama dengan ukuran huruf 11

とまくめちゃくちゃにあばれる よみもの とつぷ からだ しんたい にんたいてき ぼど
 みんとん どけ りやくず すけっち かもつ かしや にだい とらんく とぶんにわかる
 ひはんひひよ とぎ げんご ことばおもちいる きけん きけんである きけんおもたらず
 それどころか しっしょにする きゅりよくする しかるべく せいじつである しんじつで
 ある じむしょ ふせん みなと ひかくする はんこする おこしてやる たてもの けん
 せつ かいはつ こずい こずいにやれる ぎんこ ぎんこいんぎんこぎよむ ぎんこか こ
 ろん ころんする とさつする とさつぎよむ まくら やれる ゆるぐ たすける ぞかす
 る しなもの せいおかぶれの れつおつくれる ならぶ れつ あたらしい かくしん

Gambar 6.2 gambar original Kedua dengan ukuran huruf 11

んむつめ
 んむつめ
 んむつめ
 んむつめ
 んむつめ

Gambar 6.3 Contoh gambar hasil scan tulisan tangan

Tabel 6.1 berikut adalah hasil dari 56 kali pengujian

Tabel 6.1 Hasil Pengujian

Nama Gambar	Ukuran Huruf	Resolusi	Hasil Pengujian	Kesalahan	Persen berhasil
Img_0001	11	100dpi	ゆぬちねつねくたゆ ゆめた ふ、んなさけよろしくたの みさすよしきせんしよせつ とく なかなかひとちがいげいし やふせごどぬよこなる でん せつあくまでいしましいへん しよく	10 huruf	86,31 %
Img_0002	11	200dpi	ゆめものねけこり ゆめたぶ んえにさぬすよろしくたの みますよしきせんしよせつ とく なねけよまぬひとちがいげ いしやおすどどめをこげよ る でんせつあくまでいしま しいへん しよく	11 huruf	84,94 %
Img_0003	11	300dpi	ゆめものねきそこり ゆめた ぶんそよさねすよろしくた のみますよしきせんしよせ つとく なえつけよまぬひとちがい をずしへしやふずどどめを こずよる でんせつあくまで いしましいへん しよく	13 Huruf	82,19 %
Img_0004	11	600dpi	ゆめものがたね ゆめたぶん なさけよろしくたのみます よしきせんしよせつとく なまぬげよまぬひとちねき いげしへしやふほどどめね こそよる でんせつあくまで	8 Huruf	89,04 %

			いじましいへん しよく		
Img_0005	18	100dpi	ゆめものぬぢたゆ ゆめたぶ んすよさるすよろしくたの み ますよしきせんしよせつと くなぬへげよぬへひとちが いげいしやふだどめになる でんせつあくまでい じましいへんしよく	6 Huruf	91,79 %
Img_0006	18	200dpi	ゆめものがたゆゆめたぶ んさけよろしくたのみ ますよしきせんしよせつと くなかなかひとちが いげいしぬふだどめになる でんせつあくまでい じましいへんしよく	2 Huruf	97,27 %
Img_0007	18	300dpi	ゆめものがたゆゆめたぶ んさけよろしくたのみ ますよしきせぬふはこえむ つとくなかなかひとちが いげいしやふだどめになる でんせつあくまでい じましいへんしよく	5 Huruf	93,15 %
Img_0008	18	600dpi	ゆめものがたねゆめたぶ んさけよろしくたのみ ますよしきせんしよせつと くなかなかひとちが いげいしぬふだどめになる でんせつあくまでい じましいへんしよく	2 Huruf	97,27 %
Img_0009	11	100dpi	とまくめちやくちやてこあら ますするよみものつつか らだんたいにんたいてき ぱど みんとんどけりやくずけ つちかもつかしやにだいと らんくとふんをこわひする ひはまんひひよとぎげんご ことばおもちいるきけんき てずんであるきよよんぎ、	37 Huruf	86.74

			<p>きおたらばず それろかしつしよにするき ゆゆにくするしかるべくせ いじつであるしんじつで あるしむけしよふせんみな とひかくするはずるしてや るたてもものけん せつかいはつこずいこずね べちこそあれろぎんこぎこ いんぎんはよむつんこかこ ろんころんするとさつする とさつきよむまくらやれね ろゆるぐたすけるそげいす るしならのせいねかぶすし のれつねつくれろならぶれ つあたらしいかくしん</p>		
Img_0010	11	200dpi	<p>とまくめちやくちやるこあを まれるよみちのとつぷから ずごしんたいにんたいてき ぼど みんとんどけゆやくずすけ つちかもつかしやにずごし ぺとらんくとぷんにわろす る ひはんひひよとぎげんごこ とばおもちいるきけんきけ んであるきをずんおもたら す それどころかしつしよにす るきゆゆよくするしかるべ くせしへじつであるしんじ つで あるじむしよふせんみえこ とひかくするはんこするお こしてやるたてもものけん せつかいはつこずいこずし けこやれるぎんこぎんこし ぺんぎ、んこぎよむぎんこ かこ ろんころんするとさつする とさつき、よむあまくらや</p>	28 Huruf	89.96 %

			れる ゆるぐたすける ぞまつ へす る しなもの せしべおねあぶ れの れつおつくれる ならぶ れつ あたらしい かくしん		
Img_0011	11	300dpi	とまくめちやくちやにあばれ る よみちの とつぷ からだし んたい にんたい てき ぼど みんとん どけ ゆやく ずすけ つち かつ かしや にだい と らんく とぶんに わける ひはん ひひよ とぎげんご こ とばおちちいる きけん きけ んである きけんおもたらず それどころか しつしよにす る きゆねよくする しかるべ く せいじつである しんじつ で ある じむしよ ふせん みずよ と ひねまくする はんこする おこしてやる たても のけん せつ かい はつ こずい こずい にやれる ぎんこ ぎんこ いん ぎんこ ぎよむ ぎんこ かこ ろん ころんする とさつする とさつぎよむ まくら やれる ゆるぐたすける ぞかす る しなもの せいおかぶれの れつおつくれる ならぶれつ あたらしい かくしん	7 Huruf	97,49 %
Img_0012	11	600dpi	とまくめちやくちやにあばれ る よみもの とつぷ からだし んたい にんたい てき ぼど みんとん ぼとけま かやく ず すけつを あねけぬつまぬしや にだい とらぬれこくとおん に わける ひはん ひひよ とぎげんご こ とばおもちいる きけん きけ んである きけんおもたらず それどころか しつしよにす	16 Huruf	94,27 %

			<p>るきゆよくするしかるべくせいじつであるしんじつで</p> <p>あるじむしよふせんみなとひかくするはんこするおこしてやるたてものけん</p> <p>せつかいはつこずいこずいにやれるぎんこぎんこいんぎんこぎよむぎんこかこ</p> <p>ろんころんするとさつするとさつぎよむまくらやれるゆるぐたすけるぞかするしなものをせいおかぶれのれつおつくれるならぶれつあたらしねもわべしめえ</p>		
Img_0013	18	100dpi	<p>とまくめちやくちやにあばれるよみものつつぶからだしんたいにんたいてきばどみんとんどけゆやくずすけつちかもつかしやにだいとらんくと</p> <p>ぶんにわけるひはんひひよとぎげんごことばおちちいるきけんきけんであるきけんおもたらずそれどころかしつしよにするきゆりよくするしかるべくせいじつであるしんじつであるじむしよふせんみなとひかくするはんこするおこしてやるたてものけんせつかいはつこずいこずいにやれるぎんこぎんこいんぎんこぎよむぎんこかころんころんするとさつするとさつぎよむまくらやれるゆるぐたすけるぞかするしなものせ</p>	3 Huruf	98,93 %

			いおかぶれのれつおつ くれるならぶれつあたらし いかくしん		
Img_0014	18	200dpi	とまくめちやくちやにあばれ るよみものつつぶから だしんたいにんたいてきば どみんとんどけのや くずすけつちかちつかしぬ にだいとらんくと ぶんにわけるひはんひひよ とぎげんごことばおも ちいるきけんきけんである きけんおもたらずそれ どころかしつしよにするき ゆりよくするしかるべく せいじつであるしんじつで あるじむつしよふせんみ などひかくするはんこする おこしてやるたてもの けんせつかいはつこずいこ ずいにやれるぎんこ ぎんこいんぎんこぎよむぎ んこかころんころんする とさつするとさつぎよむま くらやれるゆるぐた すけるぞかするしなものせ いおかぶれのれつおつ くれるならぶれつあたらし いかくしん	5 Huruf	98.24 %
Img_0015	18	300dpi	とまくめちやくちねこあねま れるよみものつつぶから だしんえていにんた、しへ てきばどみんとんどけゆや くずすねまつちかもつかし やにたいとらんくと ぶんねこわねするひはんひ ひよとぎげんごごことをま おも ちいるきねすんきけんであ るきねすんおもすこらすそ れ	25 Huruf	91.03 %

			<p>どころかしつしよにする き ゆゆよくする しかるべく せしへじつである しんじつ である じむしよふせんみ など ひかくする はんこする おこしてやる たてもの けんせつかしへによつ ござ い ござしへねこやれる ぎん こ</p> <p>ぎんこいんぎんこぎよむ ぎ んこか ころん ころんする とさつする とさつき、よむ まくら やれる ゆるぐた すねせる そけぬする しぜよ もの せしへおえねぶれのれ つおつ くれる ならぶれつ あたらし いかくしん</p>		
Img_0016	18	600dpi	<p>とまくめちやくちやにあばれ る よみもの とつぷ から だ しんそこしへにんすこし へてき ぼどみんとん どけゆ や</p> <p>くずすねすつちかもつかし やにせきしへとらんくと ぶんをこわねする ひによん ひひよとき、げんご、こと ねまおも</p> <p>ちいる きけん きねせんであ る きねまんおもすこらすそ れ</p> <p>どころねこ しつしよねこす る きゆゆよくする しかるべ く</p> <p>せいじつである しんじつで ある じむしよふせんみ など ひかくする はんこする おこしてやる たてもの けんせつかしへをまつ ござ しへ ござしへねこやれる ぎ</p>	29 Huruf	89,61 %

			んこ ぎんこいんぎんこぎよむぎ んこかころんころんする とさつする とさつき、よむ まくらやれる ゆるぐた すける ぞかする しえにもの せしへおぬぬぶれのれつお つ くれる ならぶれつ あせこら ししへかく しん		
Img_0017	-	100dpi	まをむつぬ んむつぬ そしむ つめ ろまつめ よをきやつめ	5 huruf	75 %
Img_0018	-	200dpi	えしむつぬ をむにつぬ そし ちつめ ろまつきつかをきつ め	7 huruf	65 %
Img_0019	-	300dpi	きこむつめ んむしつぬ んむ やつめ みをまつめ えまけめ	4 huruf	80%
Img_0020	-	600dpi	えしむつぬ んむにつぬ そし むやつきくきてえしまつき つよをえやつぬ	8 Huruf	60 %
Img_0021	-	100dpi	らゆるれろ うゆるれちらつ るれろ やらつるれろ うゆるねろ	11 huruf	56 %
Img_0022	-	200dpi	らゆるれろ らゆうれちらつ るねろ からつろれう うゆるれろ	10 huruf	60 %
Img_0023	-	300dpi	らゆるれろ らゆうれちらつ るれろ あらつろれろ ろゆるれろ	11 huruf	56 %
Img_0024	-	600dpi	らゆるれろ らのうれちらつ	10 huruf	60 %

			るねろ けらのろれう うゆるれろ		
Img_0025	-	100dpi	やゆよわを わゆよわそ そゆよおを やゆよるておを やゆよおを	6 huruf	76 %
Img_0026	-	200dpi	そやゆよわを めゆよわを すそゆよおを やゆよるてお を やゆよおを	5 huruf	80 %
Img_0027	-	300dpi	ふききつゆよわちこ ぬゆよわそ やゆまおを ゆやゆよらてお を やゆよおを	8 huruf	68 %
Img_0028	-	600dpi	ふやゆよわを ぬぬゆよわそ おそゆよおを やゆよきてお を やゆよおを	6 huruf	76 %
Img_0029	-	100dpi	まみめも まみめち ままめも まみめも まみあえ	4 huruf	80 %
Img_0030	-	200dpi	まみめも まみめも ままめも まみめも まみあえ	2 huruf	90 %
Img_0031	-	300dpi	まみめも まみめも ままめも まみめも まみあえ	3 huruf	85 %
Img_0032	-	600dpi	-	20 huruf	0 %
Img_0033	-	100dpi	はひふへもおけこひろつめ けまけこひさつの そおろま	17 huruf	32 %

			ひらきへしちち		
Img_0034	-	200dpi	はひふへいち はひろつへう まはひふのしちろけこにひ ろつへしま	8 huruf	68 %
Img_0035	-	300dpi	はひらへほほひろつへらま はひふのほはひふへしま	6 huruf	76 %
Img_0036	-	600dpi	ほひふへほけこひふへいこ ほひろつけまはひふのほい しひらつへしま	10 huruf	60 %
Img_0037	-	100dpi	-	20 huruf	0 %
Img_0038	-	200dpi	えのをしまてわねてまを てさつてののをけきてをけ め	13 huruf	35 %
Img_0039	-	300dpi	-	20 huruf	0 %
Img_0040	-	600dpi	ほきへほこへやへなくこを たは くこつてみやよさへけみめ くさえ へるほはとここまつくへり てとやゆてへひけへすわ けよここへあまう	20 huruf	0 %
Img_0041	-	100dpi	たちてと すきろゆてとすきちてとせ さをとたさこと	9 huruf	55 %
Img_0042	-	200dpi	てちてと すきろこてとたさてとせき さとたちてと	9 huruf	55 %
Img_0043	-	300dpi	てちてこと すきをてとたちてとせちさ とたちこと	9 huruf	55 %
Img_0044	-	600dpi	たちてへとすきをてとたち てとせさきことねちこと	7 huruf	65 %
Img_0045	-	100dpi	さしませそさしすせえるし すせそるしをてせそえとし すせを	7 huruf	72 %
Img_0046	-	200dpi	さしきてせそさしすせをさ	6 huruf	76 %

			しすせそ りしませそをしす せを		
Img_0047	-	300dpi	さしませそ さしすせる あし すせそもしをてせを りしす せを	8 huruf	68 %
Img_0048	-	600dpi	さしませを さしすせる えしすせそそえ しませを えしすせを	9 huruf	64 %
Img_0049		100dpi	えよ きくことかきくけこ ねこきこへしすこかきしも すせかさりうすう	13 huruf	48 %
Img_0050	-	200dpi	えふ きなとかきくけここ うつこきこおまでこかきし けきこういきくそうみこ	12 huruf	52 %
Img_0051	-	300dpi	えよ きなかきくけここ う つこきくほおまこ えいき してえまきこうさきさくの うよこ	12 huruf	52 %
Img_0052	-	600dpi	すえこ きな かきくけこ うつこきこぬく こてこかきなもうとういき くろ うけこ	13 huruf	48 %
Img_0053	-	100dpi	あいうえお あはうえをよ あはつえわや あしへうえお あしへうえつをよ	8 huruf	68 %
Img_0054	-	200dpi	-	20 huruf	0 %
Img_0055	-	300dpi	-	20 huruf	0 %
Img_0056	-	600dpi	はほへりゆらほぬけめから めゆうのへ	20 huruf	0 %

Lampiran B Pengujian Tulisan Tangan

んむつめ
んむつめ
んむつめ
んむつめ
んむつめ



らりるれろ

らりるれろ

らりるれろ

らりるれろ

らりるれろ



やゆよおを
やゆよおを
やゆよおを
やゆよおを
やゆよおを



まみぬ毛

まみぬ毛

まみぬ毛

まみぬ毛

まみぬ毛



はひふ〜は

はひふ〜は

はひふ〜は

はひふ〜は

はひふ〜は



たぐいおの
たぐいおの
たぐいおの
たぐいおの
たぐいおの



たちごと

たちごと

たちごと

たちごと

たちごと



신하세주
신하세주
신하세주
신하세주
신하세주



かま<けと

かま<けと

かま<けと

かま<けと

かま<けと



あいうえお

あいうえお

あいうえお

あいうえお

あいうえお

あいうえお



CURICULUM VITAE

Nama : Akhmad Imam Fahrizal

Tempat Tanggal Lahir : Amuntai, 11 September 1990

Alamat Asal : Jl. Patmaraga 24 RT 3 Kebun Sari, Amuntai,
Kalimantan Selatan

Alamat Tinggal : Jl. Babadan 505 Gedong Kuning, Bantul, Yogyakarta

Email : Sanosuke.Ichizen@yahoo.com
Sanosuke.Ichizen@gmail.com

Institusi : Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universita Islam Negeri Sunan Kalijaga Yogyakarta