

PENGENALAN HURUF KATAKANA DENGAN ADAPTIVE THRESHOLD DAN FITUR ZONING

Skripsi



Disusun Oleh

Ryan Ariestillman Chandra

71160006

PROGRAM STUDI INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2020

HALAMAN JUDUL
Pengenalan Huruf Katakana Dengan
Adaptive Threshold dan Fitur Zoning

Skripsi



Diajukan kepada Program Studi Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana
Sebagai Salah Satu Syarat dalam Memperoleh Gelar
Sarjana Komputer

Disusun Oleh

Ryan Ariestillman Chandra

71160006

PROGRAM STUDI INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2020

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
SKRIPSI/TESIS/DISERTASI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Kristen Duta Wacana, saya yang bertanda tangan di bawah ini:

Nama : Ryan Ariestillman Chandra
NIM : 71160006
Program studi : Informatika
Fakultas : Fakultas Teknologi Informasi (FTI)
Jenis Karya : Skripsi

demikian pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Kristen Duta Wacana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

**“PENGENALAN HURUF KATAKANA DENGAN ADAPTIVE THRESHOLD
DAN FITUR ZONING”**

berserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Kristen Duta Wacana berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama kami sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Yogyakarta
Pada Tanggal : 06 April 2020

Yang menyatakan



Ryan Ariestillman Chandra
NIM.71160006

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:
Pengenalan Huruf Katakana Dengan Adaptive Threshold Dan Fitur Zoning

Yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan sarja Program Studi Informatika, Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi kesarjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar kesarjanaan saya.

Yogyakarta, 6 April 2020



Ryan Ariestillman Chandra

71160006

HALAMAN PERSETUJUAN

Judul Skripsi : Pengenalan Huruf Katakana Dengan Adaptive Threshold
Dan Fitur Zoning

Nama Mahasiswa : Ryan Ariestillman Chandra

NIM : 71160006

Mata Kuliah : Skripsi (Tugas Akhir)

Kode : [TI0366]

Semester : Genap

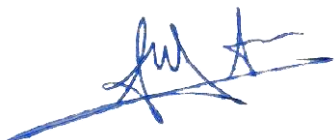
Tahun Akademik : 2019/2020

©UKDW

Telah diperiksa dan disetujui di
Yogyakarta,

Pada Tanggal

Dosen Pembimbing I



Aditya Wikan Mahastama, S.Kom.,
M.Cs.

Dosen Pembimbing II



Dra. Widi Hapsari, M.T

HALAMAN PENGESAHAN

SKRIPSI

Pengenalan Huruf Katakana Dengan Adaptive Threshold Dan Fitur Zoning

Oleh: 71160006 / Ryan Ariestillman Chandra

Dipertahankan di depan Dewan Penguji Skripsi

Program Studi Informatika Fakultas Teknologi Informasi

Universitas Kristen Duta Wacana – Yogyakarta

Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar

Sarjana Komputer

Pada tanggal 6 April 2020

Yogyakarta, 06 April 2020
Mengesahkan,

Dewan Penguji:

1. Aditya Wikan Mahastama, S.Kom.,
M.Cs.
2. Dra. Widi Hapsari, M.T
3. Danny Sebastian, S.Kom., M.M., M.T



Dekan



(Restyandito, S.Kom., MSIS, Ph.D.)

Ketua Program Studi



(Gloria Virginia, S.Kom., MAI, Ph.D)

UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas berkat dan karunia-Nya, penulis dapat menyelesaikan Tugas Akhir (TA) dengan judul “Pengenalan Huruf Katakana Dengan Adaptive Threshold Dan Fitur Zoning” untuk memenuhi persyaratan kelulusan di Universitas Kristen Duta Wacana.

Penulis menyadari bahwa penyusunan Tugas Akhir (TA) ini tentunya tidak lepas dari berbagai kesulitan dan hambatan, sehingga penulis sangat berterima kasih untuk bimbingan dan bantuan dalam penyusunan Tugas Akhir (TA). Pada kesempatan ini, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Tuhan Yang Maha Esa, atas karunia dan berkat-Nya lah, skripsi ini dapat selesai dengan baik.
2. Kedua orang tua saya, Tonni Chandra (ayah) dan Harjati (ibu), yang selalu mendukung dan menyemangati saya, serta memotivasi saya selama penyusunan skripsi ini.
3. Saudara saya, Julia Elfreda Chandra, yang selalu mendukung dan menyemangati saya, serta memotivasi saya selama penyusunan skripsi ini.
4. Restyandito, S.Kom., MSIS, Ph.D., selaku dekan Fakultas Teknologi Informasi

5. Gloria Virginia, S.Kom., MAI, Ph.D, selaku Ketua Program Studi Teknologi Informasi
6. Aditya Wikan Mahastama, S.Kom., M.Cs., selaku dosen pembimbing I yang selalu memberikan bimbingan berupa saran dan kritik selama penyusunan skripsi ini.
7. Dra. Widi Hapsari, M.T, selaku dosen pembimbing II yang selalu memberikan bimbingan berupa saran dan kritik selama penyusunan skripsi ini.
8. Teman-teman dekat saya yang memberi dukungan dan mendorong saya untuk menyelesaikan Tugas Akhir (TA)
9. Seluruh Dosen dan Staff Fakultas Teknologi Informasi Kristen Duta Wacana untuk bantuan, pembelajaran, dan bimbingan dalam penyusunan skripsi ini.

INTISARI

Pengenalan Huruf Katakana Dengan Adaptive Threshold Dan Fitur Zoning

Bahasa Jepang merupakan bahasa yang cukup banyak digunakan. Bahasa Jepang sendiri terbagi menjadi tiga, yaitu aksara Kanji, aksara Hiragana, dan aksara Katakana. Salah satu aksara bahasa Jepang yaitu aksara Katakana, biasanya digunakan untuk menuliskan bahasa asing seperti bahasa Inggris ke bahasa Jepang. Aksara Katakana sendiri memiliki keunikan seperti memiliki sudut yang tajam dan garis lurus.

Bagi orang Indonesia, mempelajari bahasa Jepang cukup sulit karena menggunakan lebih terbiasa dengan alfabet. Penulis akan membuat suatu program untuk mengenali tulisan aksara Jepang menggunakan *adaptive threshold* tetapi hanya sebatas aksara Katakana. *Adaptive threshold* akan dilanjutkan dengan *zoning* yang akan digunakan untuk mengambil nilai yang akan digunakan untuk meengidentifikasi karakter. Program ini diharapkan dapat membantu orang dalam mempelajari bahasa Jepang.

Hasil akurasi terbaik dari program yang dibangun adalah 51,53% dengan *zone* yang digunakan adalah 4x4 dan menggunakan 8 data latih. Hasil akurasi terbaik setelah *zone 4x4* adalah *zone 8x8* dengan akurasi 56,1% untuk karakter katakana dasar, dan lebih kecil daripada akurasi terbaik sebanyak 1,25% bila disertai karakter disertai dengan *tenten* dan *maru*. Hasil akurasi terburuk terletak pada *zone 2x2* dengan tingkat akurasi 16,39% hingga 19,3%.

Kata Kunci : bahasa Jepang, katakana, *adaptive threshold*, *zoning*, *zone*, *euclidean distance*

Daftar Isi

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN	iv
HALAMAN PENGESAHAN	v
Daftar Isi.....	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah	2
1.4. Tujuan Penelitian	2
1.5. Metodologi Penelitian.....	2
1.6. Sistematika Penulisan	3
BAB 2 TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	4
2.2 Landasan Teori	5
2.2.1 Aksara Katakana.....	5
2.2.2 Optical Character Recognition	7
2.2.3 Preprocessing.....	7
2.2.4 Feature Extraction.....	9
2.2.5 Klasifikasi	9
BAB 3 METODOLOGI PERANCANGAN	11
3.1 Analisis Kebutuhan Program.....	11
3.1.1 Spesifikasi Perangkat Keras (<i>Hardware</i>)	11

3.1.2 Spesifikasi Perangkat Lunak (<i>Software</i>)	11
3.1.3 Kebutuhan <i>Library</i>	12
3.2 Blok Diagram Program	12
3.3 Perancangan Desain Antarmuka Program	14
3.4 Perancangan Pengujian Program	15
BAB 4 IMPLEMENTASI DAN ANALISIS PROGRAM	16
4.1 Implementasi Program.....	16
4.1.1 Tampilan Interface	16
4.1.2 Hasil Uji Program.....	18
4.2 Analisis Hasil Program.....	28
4.2.1 Persentase Pengenalan Karakter	28
4.2.2 Metode Global Adaptive Thresholding.....	31
4.2.3 Metode Segmentasi.....	31
BAB 5 KESIMPULAN DAN SARAN	33
5.1 Kesimpulan.....	33
5.2 Saran	33
Daftar Pustaka.....	34
LAMPIRAN	36

DAFTAR TABEL

Tabel 2.1 Daftar aksara Katakana dasar yang akan digunakan.....	6
Tabel 2.2 Daftar aksara Katakana <i>tenten</i> dan <i>maru</i> yang akan digunakan	7
Tabel 4.1 Hasil uji program zone 4x4 dengan 8 data latih untuk katakana dasar (1).....	18
Tabel 4.1 Hasil uji program zone 4x4 dengan 8 data latih untuk katakana dasar (2).....	19
Tabel 4.2 Hasil uji program zone 4x4 dengan 8 data latih untuk katakana <i>tenten</i> dan <i>maru</i>	20
Table 4.3 Hasil uji program zone 8x8 dengan 8 data latih untuk katakana dasar (1).....	21
Tabel 4.3 Hasil uji program zone 8x8 dengan 8 data latih untuk katakana dasar (2).....	22
Tabel 4.4 Hasil uji program zone 8x8 dengan 8 data latih untuk katakana <i>tenten</i> dan <i>maru</i>	23
Tabel 4.5 Hasil uji program zone 2x2 dengan 4 data latih untuk katakana dasar (1).....	24
Tabel 4.5 Hasil uji program zone 2x2 dengan 4 data latih untuk katakana dasar (2).....	25
Tabel 4.6 Hasil uji program zone 2x2 dengan 4 data latih untuk katakana <i>tenten</i> dan <i>maru</i>	26
Tabel 4.7 Rangkuman keseluruhan data uji.....	27

DAFTAR GAMBAR

Gambar 2.1 Adaptive Threshold	8
Gambar 3.1 Blok Diagram Program	13
Gambar 3.2 Rancangan Desain Antarmuka Web.....	14
Gambar 3.3 Rancangan Desain Antarmuka Web Halaman Hasil.....	14
Gambar 4.1 Tampilan antarmuka halaman utama	16
Gambar 4.2 Tampilan antarmuka halaman hasil.....	16
Gambar 4.3 Tampilan antarmuka daftar karakter	17
Gambar 4.4 Trimming	32

©UKDW

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Bahasa Jepang merupakan bahasa yang cukup banyak digunakan. Bahasa Jepang sendiri terbagi menjadi tiga, yaitu aksara Kanji, aksara Hiragana, dan aksara Katakana. Biasanya, di negara Jepang sendiri aksara Katakana dan aksara Hiragana sudah dipelajari semenjak berada pada taman kanak-kanak atau disebut TK. Bahasa Jepang semakin populer dengan adanya manga dan anime. Namun, bahasa Jepang sendiri cukup rumit untuk dipelajari bagi masyarakat Indonesia. Hal ini bisa diketahui dari pola karakter aksara yang unik yang berbeda dari huruf bahasa Indonesia (Mawitjere, 2017).

Salah satu aksara bahasa Jepang yaitu aksara Katakana, biasanya digunakan untuk menuliskan bahasa asing seperti bahasa Inggris ke bahasa Jepang. Berbeda dengan aksara Hiragana yang berfungsi untuk melengkapi bagian yang tak tercakup oleh *Kanji*. Aksara Katakana sendiri memiliki keunikan seperti memiliki sudut yang tajam dan garis lurus (*chokusenteki*) (Mawitjere, 2017).

Bagi orang Indonesia, mempelajari bahasa Jepang cukup menantang. Hal ini dikarenakan perbedaan karakter pada bahasa Indonesia dan bahasa Jepang. Karena orang Indonesia lebih terbiasa dengan alfabet, ketika dihadapkan dengan huruf-huruf dari bahasa Jepang akan kesulitan dan hal ini membuat proses belajar Bahasa Jepang menjadi lebih lambat dibanding belajar bahasa Inggris (Mawitjere, 2017).

Dengan berkembangnya bidang teknologi, hal ini dapat diatasi dengan cara yaitu membuat program pengenalan tulisan aksara Jepang. Untuk tugas akhir ini, penulis akan membuat suatu program untuk mengenali tulisan aksara Jepang menggunakan *adaptive threshold* tetapi hanya sebatas aksara Katakana. Program ini akan mengolah tulisan tangan menjadi dalam bentuk citra digital.

Program ini diharapkan dapat membantu orang dalam mempelajari bahasa Jepang, terutama aksara Katakana melalui pengolahan input tulisan tangan menjadi citra digital.

1.2. Rumusan Masalah

Berdasarkan latar belakang, masalah pun dirumuskan menjadi :

1. Bagaimana tingkat akurasi pengenalan apabila menggunakan *adaptive threshold* dan beberapa ukuran *zoning*?

1.3. Batasan Masalah

Batasan-batasan masalah untuk penelitian ini adalah sebagai berikut :

1. Pola tulisan tangan yang akan diolah hanya sebatas aksara Katakana
2. Fitur yang digunakan adalah fitur biner area
3. Input berupa tulisan tangan aksara Katakana dasar beserta Katakana dengan *tenten* dan *maru* ditambah 1 karakter tanda baca
4. Huruf Katakana yang akan dikenali berjumlah 46 huruf dasar, dari a sampai n. Ditambah 25 huruf Katakana dengan *tenten* dan *maru* dan 1 karakter tanda baca.
5. Input berasal dari hasil *scan* dengan posisi lurus/horizontal dan tidak miring/vertikal

1.4. Tujuan Penelitian

Tujuan dari penelitian ini yaitu untuk mengukur kinerja *adaptive threshold* dan *zoning* dalam mengenali tulisan tangan karakter Katakana menggunakan *adaptive thresholding*.

1.5. Metodologi Penelitian

Secara garis besar, metode ini terdiri dari 4 tahap, yaitu tahap pengambilan gambar, tahap pemrosesan gambar, lalu tahap *Feature Extraction*, dan terakhir klasifikasi. Metode yang digunakan pertama adalah pengambilan gambar terhadap teks yang memiliki huruf aksara Katakana, kemudian diproses menggunakan *adaptive threshold* yang kemudian diolah lagi menggunakan *zoning*, setelah data

diolah menggunakan *zoning* data akan diklasifikasikan menggunakan *euclidean distance*.

1.6. Sistematika Penulisan

Laporan penelitian dalam penulisan tugas akhir ini terdiri menjadi 5 bab, yaitu: Pendahuluan, Tinjauan Pustaka, Metodologi Perancangan, Pembahasan, dan Kesimpulan dan Saran. Penjelasan setiap bagian adalah sebagai berikut:

1. Pendahuluan

Terdiri dari Latar Belakang Masalah, Rumusan Masalah, Batasan Masalah, Tujuan Penulisan, Metodologi Penelitian, dan Sistematika Penulisan.

2. Tinjauan Pustaka

Memuat teori-teori yang berkaitan dengan pengenalan aksara Katakana menggunakan *adaptive thresholding* dengan *zoning*.

3. Metodologi Perancangan

Berisi metode-metode dan teori yang akan di implementasikan terkait perancangan program.

4. Pembahasan

Berisi pembahasan-pembahasan dari rancangan program yang telah dibuat pada bab 3.

5. Kesimpulan dan Saran

Terdiri dari kesimpulan akhir dari keseluruhan program dan saran-saran untuk pengembangan program

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari Penelitian yang telah dilakukan, maka dapat disimpulkan :

1. Hasil akurasi rata-rata dari katakana dasar dan tenten dan maru terbaik dari program yang dibangun adalah 51,53% dan zone yang digunakan adalah zone 4x4.
2. Zone terbaik untuk karakter katakana dasar ada pada zone 8x8 dengan akurasi sebesar 56,1%.
3. Zone terbaik untuk karakter katakana dengan tenten dan maru ada pada zone 4x4 dengan akurasi sebesar 43,9%
4. Semakin banyak data latih semakin tinggi akurasi pada zone 4x4.

5.2 Saran

Saran penelitian dengan metode lain, yaitu:

- Menggunakan metode local adaptive thresholding, sehingga kontras pada citra bisa lebih bebas.

Daftar Pustaka

- Abdullah, S., Omar, K., Zaini, A. S., Petrou, M., & Khalid, M. (2016). Determining adaptive thresholds for image segmentation for a license plate recognition system. *(IJACSA) International Journal of Advanced Computer Science and Applications*.
- Arifianto, T. (2017). Segmentasi Aksara Pada Tulisan Aksara Jawa Menggunakan Adaptive Threshold. *SMATIKA*.
- Gonçalves, D. N., Gonçalves, C. d., de Assis, T. F., & da Silva, M. A. (2014). Analysis of the difference between the Euclidean distance and the actual road distance in Brazil. *Transportation Research Procedia*.
- Jana, R., Chowdhury, A. R., & Islam, M. (2013). Optical Character Recognition from Text Image. *International journal of computer applications technology and research*.
- Liberti, L., Lavor, C., Maculan, N., & Mucherino, A. (2012). Euclidean distance geometry and applications.
- Lubis, I., Rahmat, R. F., Lubis, H., Ardianshar, A., & Syahputra, M. F. (2018). Image processing method in implementation of handwriting identification for Japanese katakana characters. *IOP*.
- Mawitjere, I. (2017). PENINGKATAN PENGUASAAN HURUF JEPANG (HIRAGANA DAN KATAKANA) MELALUI MEDIA BUKU BERGAMBAR. *Universitas Negeri Manado (UNIMA)*.
- Patel, D. K., Som, T., Yadav, S. K., & Singh, M. K. (2012). Handwritten Character Recognition Using Multiresolution Technique and Euclidean Distance Metric. *Journal of Signal and Information Processing*.
- Rani, N. S., Verma, S. K., & Joseph, A. (2016). A Zone Based Approach for Classification and Recognition of Telugu Handwritten Characters. *International Journal of Electrical and Computer Engineering (IJECE)*.
- Roy, P., Dutta, S., Dey, N., Dey, G., Chakraborty, S., & Ray, R. (2014). Adaptive Thresholding: A comparative study . *International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*.
- Singh, P., & Budhiraja, S. (2011). Feature Extraction and Classification Techniques in O.C.R. Systems for Handwritten Gurmukhi Script – A

Survey. *International Journal of Engineering Research and Applications (IJERA)*.

Sutramiani, N. P., Putra, I. K., & Sudarma, M. (2015). Local Adaptive Thresholding Pada Preprocessing Citra Lontar Aksara Bali. *JURNAL TEKNOLOGI ELEKTRO (JTE)*.

©UKDW