



SKRIPSI

**Pengenalan Huruf Hiragana
Menggunakan Metode *Histogram of
Oriented Gradients (HOG)* dan *Support
Vector Machine (SVM)***

TSABITA SAFANA MUSTOFA
NPM 21081010282

DOSEN PEMBIMBING

Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT.
Hendra Maulana, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**



SKRIPSI

**Pengenalan Huruf Hiragana
Menggunakan Metode *Histogram of
Oriented Gradients (HOG)* dan *Support
Vector Machine (SVM)***

TSABITA SAFANA MUSTOFA
NPM 21081010282

DOSEN PEMBIMBING

Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT.
Hendra Maulana, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

Pengenalan Huruf Hiragana Menggunakan Metode Histogram of Oriented Gradients (HOG) dan Support Vector Machine (SVM)

Oleh :

TSABITA SAFANA MUSTOFA

NPM. 21081010282

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 4 Juni 2025.

Menyetujui,

Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT.
NPT. 222198 60 816400


.....

(Pembimbing I)

Hendra Maulana, S.Kom., M.Kom.
NPT. 201198 31 223248


.....

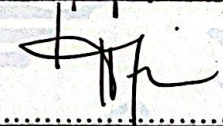
(Pembimbing II)

Yisti Vita Via, S.ST., M.Kom.
NIP. 19860425 202121 2 001


.....

(Ketua Penguji)

Afina Lina Nurlaili, S.Kom., M.Kom.
NIP. 1993121 3202203 2 010


.....

(Anggota Penguji)

Mengetahui,

Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**Pengenalan Huruf Hiragana Menggunakan Metode
Histogram of Oriented Gradients (HOG) dan Support
Vector Machine (SVM)**

Oleh :

TSABITA SAFANA MUSTOFA

NPM. 21081010282



Menyetujui,

Koordinator Program Studi Informatika

Fakultas Ilmu Komputer

Fetty Tri Anggraeny, S.Kom., M.Kom.

NIP. 19820211 2021212 005

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Yang bertandatangan di bawah ini:

Nama Mahasiswa : Tsabita Safana Mustofa
NPM : 21081010282
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu Lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 12 Juni 2025
Yang Membuat Pernyataan,



TSABITA SAFANA MUSTOFA
NPM. 21081010282

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Tsabita Safana Mustofa / 21081010282
Judul Skripsi : Pengenalan Huruf Hiragana Menggunakan Metode *Histogram of Oriented Gradients* (HOG) dan *Support Vector Machine* (SVM)
Dosen Pembimbing : 1. Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT.
2. Hendra Maulana, S.Kom., M.Kom.

Hiragana adalah sistem penulisan dasar yang paling sering digunakan untuk menulis kata-kata asli bahasa Jepang. Huruf Hiragana memiliki tantangan tersendiri dalam proses pengenalan dikarenakan huruf-huruf tersebut memiliki bentuk unik. Penelitian ini bertujuan mengembangkan sistem pengenalan huruf Hiragana secara otomatis dengan menggunakan metode HOG sebagai ekstraksi fitur dan SVM sebagai metode klasifikasi. Dataset yang digunakan terdiri dari 4.600 citra huruf Hiragana dengan jumlah masing-masing 100 citra untuk setiap kelas huruf. Pada tahap *preprocessing*, seluruh citra diubah terlebih dahulu ke format *grayscale*, kemudian diubah ukurannya menjadi 64x64 piksel. Ekstraksi fitur dilakukan menggunakan HOG dengan parameter ukuran sel 8x8 piksel, blok 2x2 sel, dan 9 bin orientasi. Model SVM dilatih menggunakan tiga kernel yang berbeda dengan pengujian terhadap beberapa kombinasi parameter C dan gamma dengan variasi pembagian proporsi data yang berbeda untuk mendapatkan performa terbaik. Hasil pengujian menunjukkan bahwa model mencapai akurasi sebesar 97,50%, *precision* 97,59%, *recall* 97,50%, dan *f1-score* 97,49% pada model HOG-SVM dengan kernel RBF, nilai C = 10, dan gamma = 0.01 pada proporsi dataset 80:20. Performa ini menunjukkan bahwa metode HOG dan SVM sangat efektif dalam mengenali huruf Hiragana, bahkan ketika huruf tersebut memiliki bentuk yang mirip. Model terbaik ini kemudian diimplementasikan dalam antarmuka *web* menggunakan Flask untuk memudahkan pengguna dalam mencoba pengenalan huruf secara interaktif. Penelitian ini diharapkan dapat menjadi dasar pengembangan sistem pembelajaran bahasa Jepang berbasis teknologi.

Kata kunci: Hiragana, HOG, SVM, Pengolahan Citra, Pengenalan Karakter

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Tsabita Safana Mustofa / 21081010282
Thesis Title : Recognition of Hiragana Characters Using Histogram of Oriented Gradients (HOG) and Support Vector Machine (SVM)
Advisor : 1. Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT.
2. Hendra Maulana, S.Kom., M.Kom.

Hiragana is the most commonly used basic writing system for writing native Japanese words. Hiragana characters pose unique challenges in the recognition process due to their distinctive shapes. This study aims to develop an automatic Hiragana character recognition system using the HOG method for feature extraction and SVM as the classification method. The dataset used consists of 4,600 Hiragana character images, with 100 images for each character class. During the preprocessing stage, all images are first converted to grayscale format and then resized to 64x64 pixels. Feature extraction was performed using HOG with parameters of 8x8 pixel cell size, 2x2 cell block, and 9 orientation bins. The SVM model was trained using three different kernels with testing of several combinations of C and gamma parameters with varying data proportion divisions to obtain the best performance. The test results show that the model achieves an accuracy of 97.50%, precision of 97.59%, recall of 97.50%, and an F1-score of 97.49% on the HOG-SVM model with an RBF kernel, C = 10, and gamma = 0.01 on an 80:20 dataset split. This performance demonstrates that the HOG and SVM methods are highly effective in recognizing Hiragana characters, even when the characters have similar shapes. The best model was then implemented in a web interface using Flask to facilitate users in trying interactive character recognition. This research is expected to serve as a foundation for the development of Japanese language learning systems based on technology.

Keywords: Hiragana, HOG, SVM, Image Processing, Character Recognition

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya yang tak terhingga, sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Pengenalan Huruf Hiragana Menggunakan Metode *Histogram of Oriented Gradients* (HOG) dan *Support Vector Machine* (SVM)”** ini dengan baik.

Penyusunan skripsi ini merupakan salah satu tahapan penting dalam menempuh jenjang pendidikan Sarjana Komputer di Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur. Dalam kesempatan ini, penulis ingin menyampaikan rasa hormat dan terima kasih yang sebesar-besarnya kepada semua pihak yang telah berkontribusi, memberikan dukungan, serta doa dan motivasi selama proses penyusunan skripsi ini. Secara khusus, penulis menyampaikan ucapan terima kasih kepada:

1. Keluarga penulis, terutama kedua orang tua penulis, Bapak Mustofa dan Ibu Retno Palupi, sosok luar biasa yang selalu menjadi sumber kekuatan dan inspirasi, yang dengan penuh kesabaran, kasih sayang, serta doa yang tiada henti, senantiasa mendampingi langkah penulis. Kedua adik penulis, Azkiya Ramadhani Mustofa dan Saniya Isyfalana Mustofa, yang selalu mendoakan dan memberikan semangat selama proses penyusunan skripsi ini.
2. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT., selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur, atas dukungan dan kebijakan yang diberikan.
3. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT., selaku Dekan Fakultas Ilmu Komputer, atas arahan dan dukungan dalam pengembangan akademik.
4. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom., selaku Ketua Program Studi Informatika, atas bimbingan dan arahan akademik selama masa studi.
5. Ibu Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT. selaku Dosen Pembimbing I, dan Bapak Hendra Maulana, S.Kom., M.Kom. selaku Dosen Pembimbing II, yang dengan penuh kesabaran, ketelitian, serta dedikasi tinggi telah membimbing, mengarahkan, dan memberikan masukan yang sangat berarti dalam penyusunan skripsi ini.

6. Ibu Yisti Vita Via, S.ST., M.Kom. dan Ibu Afina Lina Nurlaili, S.Kom., M.Kom. selaku dosen penguji, yang telah memberikan masukan, saran, dan koreksi yang sangat bermanfaat demi kesempurnaan skripsi ini.
7. Teman-teman seperjuangan penulis selama perkuliahan, Lintang jajan, Erica timnas, dan Delia mie ayam, yang sejak awal perkuliahan hingga tahap akhir masa perkuliahan ini selalu memberikan dukungan, semangat, serta menjadi bagian penting dalam setiap proses yang dijalani.
8. Raisah dracin dan Lativa informan, yang senantiasa memberikan dukungan dan menemani selama menjalani berbagai proses di semester akhir, menjadi bagian penting dalam menyelesaikan tahap akhir perkuliahan ini.
9. Teman baik penulis, Devita Aaliya Sahesti, yang selalu hadir memberikan dukungan moral, semangat, dan nasihat berharga di setiap proses, baik di saat senang maupun saat menghadapi berbagai tantangan.
10. Manuel Neuer, sebagai idola yang secara tidak langsung menjadi inspirasi bagi penulis. Ketekunan, konsistensi, dan dedikasinya dalam bidang yang digelutinya selalu menjadi motivasi untuk terus berjuang dan pantang menyerah.
11. Diri saya sendiri, atas usaha, konsistensi, dan komitmen yang telah dijalani dalam menyelesaikan setiap tahapan hingga skripsi ini dapat diselesaikan dengan sebaik mungkin.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi kesempurnaan penulisan. Semoga laporan ini dapat memberikan manfaat, baik bagi pembaca maupun bagi penulis sendiri.

Surabaya, 12 Juni 2025

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN	v
SURAT PERNYATAAN BEBAS PLAGIASI.....	vii
ABSTRAK	ix
ABSTRACT	xi
KATA PENGANTAR.....	xiii
DAFTAR ISI.....	xv
DAFTAR GAMBAR.....	xvii
DAFTAR TABEL	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Penelitian Terdahulu	5
2.2 Huruf Hiragana.....	6
2.3 Citra <i>Digital</i>	8
2.4 Pengolahan Citra <i>Digital</i>	8
2.5 <i>Machine Learning</i>	10
2.6 Python	11
2.7 <i>Histogram of Oriented Gradients (HOG)</i>	12
2.8 <i>Support Vector Machine (SVM)</i>	15
2.9 <i>Confusion Matrix</i>	18
2.10 Flask	20
BAB III METODOLOGI	21

3.1	Tahapan Penelitian	21
3.2	Studi Literatur	22
3.3	Pengumpulan Data	22
3.4	<i>Preprocessing</i>	25
3.5	Ekstraksi Fitur HOG	26
3.6	Perancangan Model SVM	33
3.7	Skenario Pengujian.....	36
3.8	Evaluasi Model.....	38
3.9	Implementasi <i>Website</i>	39
BAB IV HASIL DAN PEMBAHASAN		41
4.1	<i>Preprocessing</i>	41
4.2	Ekstraksi Fitur HOG	44
4.3	Pelatihan Model SVM.....	48
4.4	Pengujian Model	54
4.4.1	Pengujian Model Kernel Linear.....	54
4.4.2	Pengujian Model Kernel RBF.....	62
4.4.3	Pengujian Model Kernel Sigmoid.....	73
4.5	Evaluasi Model.....	84
4.5.1	Model HOG-SVM	84
4.5.2	Model SVM	89
4.5.3	Analisis Performa Kedua Model.....	95
4.6	Implementasi <i>Website</i>	97
BAB V KESIMPULAN DAN SARAN		113
5.1	Kesimpulan	113
5.2	Saran.....	114
DAFTAR PUSTAKA.....		115

DAFTAR GAMBAR

Gambar 2. 1 Huruf Hiragana.....	7
Gambar 2. 2 Akuisisi Citra	9
Gambar 2. 3 Macam-macam <i>Machine Learning</i>	10
Gambar 2. 4 Ilustrasi Metode SVM	15
Gambar 3. 1 Tahapan Penelitian	21
Gambar 3. 2 Tampilan dataset primer orang Jepang (a) citra huruf A, (b) citra huruf I, (c) citra huruf U, (d) citra huruf E, dan (e) citra huruf O.....	24
Gambar 3. 3 Tampilan dataset primer orang Indonesia (a) citra huruf A, (b) citra huruf I, (c) citra huruf U, (d) citra huruf E, dan (e) citra huruf O	25
Gambar 3. 4 Tampilan dataset sekunder (a) citra huruf A, (b) citra huruf I, (c) citra huruf U, (d) citra huruf E, dan (e) citra huruf O	25
Gambar 3. 5 Tahapan <i>Preprocessing</i>	25
Gambar 3. 6 Diagram Alir Ekstraksi Fitur HOG	26
Gambar 3. 7 Operator Sobel.....	27
Gambar 3. 8 Citra Huruf HA.....	27
Gambar 3. 9 Matriks Citra Huruf HA	28
Gambar 3. 10 Perhitungan Gradien Terhadap Sumbu x	28
Gambar 3. 11 Perhitungan Gradien Terhadap Sumbu y	29
Gambar 3. 12 Pergeseran Blok Secara Horizontal.....	31
Gambar 3. 13 Pergeseran Blok Secara Vertikal.....	31
Gambar 3. 14 Diagram Alir Perancangan Model SVM.....	33
Gambar 3. 15 <i>Wireframe Web Flask</i>	40
Gambar 4. 1 Hasil Citra <i>Grayscale</i> dan <i>Resize</i>	43
Gambar 4. 2 Hasil Pembagian Data	44
Gambar 4. 3 Hasil Ekstraksi Fitur HOG	47
Gambar 4. 4 Hasil Pengujian Kernel Linear Proporsi Dataset 80:20	55
Gambar 4. 5 <i>Confusion Matrix</i> Model Terbaik Kernel Linear Proporsi Dataset 80:20.....	56
Gambar 4. 6 Hasil Pengujian Kernel Linear Proporsi Dataset 70:30	57

Gambar 4. 7 <i>Confusion Matrix</i> Model Terbaik Kernel Linear Proporsi Dataset 70:30.....	58
Gambar 4. 8 Hasil Pengujian Kernel Linear Proporsi Dataset 60:40	59
Gambar 4. 9 <i>Confusion Matrix</i> Model Terbaik Kernel Linear Proporsi Dataset 60:40.....	60
Gambar 4. 10 Hasil Pengujian Kernel RBF Proporsi Dataset 80:20	63
Gambar 4. 11 <i>Confusion Matrix</i> Model Terbaik Kernel RBF Proporsi Dataset 80:20	64
Gambar 4. 12 Hasil Pengujian Kernel RBF Proporsi Dataset 70:30	66
Gambar 4. 13 <i>Confusion Matrix</i> Model Terbaik Kernel RBF Proporsi Dataset 70:30	67
Gambar 4. 14 Hasil Pengujian Kernel RBF Proporsi Dataset 60:40	69
Gambar 4. 15 <i>Confusion Matrix</i> Model Terbaik Kernel RBF Proporsi Dataset 60:40	70
Gambar 4. 16 Hasil Pengujian Kernel Sigmoid Proporsi Dataset 80:20	74
Gambar 4. 17 <i>Confusion Matrix</i> Model Terbaik Kernel Sigmoid Proporsi Dataset 80:20.....	75
Gambar 4. 18 Hasil Pengujian Kernel Sigmoid Proporsi Dataset 70:30	77
Gambar 4. 19 <i>Confusion Matrix</i> Model Terbaik Kernel Sigmoid Proporsi Dataset 70:30.....	78
Gambar 4. 20 Hasil Pengujian Kernel Sigmoid Proporsi Dataset 60:40	80
Gambar 4. 21 <i>Confusion Matrix</i> Model Terbaik Kernel Sigmoid Proporsi Dataset 60:40.....	81
Gambar 4. 22 <i>Classification Report</i> Model HOG-SVM.....	86
Gambar 4. 23 Hasil Prediksi Model HOG-SVM	87
Gambar 4. 24 Citra Huruf Salah Prediksi Model HOG-SVM	88
Gambar 4. 25 <i>Confusion Matrix</i> Model SVM	90
Gambar 4. 26 <i>Classification Report</i> Model SVM.....	91
Gambar 4. 27 Hasil Prediksi Model SVM	93
Gambar 4. 28 Citra Huruf Salah Prediksi Model SVM	94
Gambar 4. 29 Tampilan Halaman Beranda.....	101
Gambar 4. 30 Tampilan Halaman Prediksi	101

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	18
Tabel 3. 1 Dataset huruf Hiragana	23
Tabel 3. 2 Hasil Histogram Gradien	31
Tabel 3. 3 Skenario Pengujian	37
Tabel 3. 4 Evaluasi Performa Kedua Model	39
Tabel 4. 1 Performa Model Terbaik Kernel Linear.....	61
Tabel 4. 2 Performa Model Terbaik Kernel RBF	72
Tabel 4. 3 Performa Model Terbaik Kernel Sigmoid	83
Tabel 4. 4 Hasil Pengujian Model HOG-SVM	85
Tabel 4. 5 Hasil Evaluasi Performa Kedua Model.....	95
Tabel 4. 6 Pengujian Pada Prediksi Huruf Hiragana.....	102
Tabel 4. 7 Pengujian Pada Prediksi Huruf Non-Hiragana	109

Halaman ini sengaja dikosongkan