

PENERAPAN METODE SUPPORT VECTOR MACHINE UNTUK PENGENALAN POLA AKSARA BATAK TOBA

Efdi Sarjono Panjaitan✉, Humuntal Rumapea, Indra Kelana Jaya

Fakultas Ilmu Komputer, Universitas Methodist Indonesia, Medan, Indonesia

Email: efdyspanjaitan@gmail.com

ABSTRACT

The usage of Batak Toba script has declined, and its complex forms pose challenges in pattern recognition. This study employs the Support Vector Machine (SVM) method to classify Batak Toba script patterns, utilizing Histogram of Oriented Gradients (HOG) as a feature extraction technique. The data used comes from various sources, totaling 285 script images. After preprocessing, SVM was applied to separate characters into two main classes, which were further subdivided into subclasses until final classification was achieved. The results show that the combination of HOG and SVM can classify Batak Toba script characters with an accuracy of 89,47%. This research makes a significant contribution to the preservation of Batak Toba script and has broader potential applications in pattern recognition and image classification.

Keyword: *Batak Toba Script, Support Vector Machine, Histogram of Oriented Gradients, Pattern Recognition.*

ABSTRAK

Penggunaan aksara Batak Toba semakin berkurang, dan kompleksitas bentuknya menimbulkan tantangan dalam pengenalan pola. Penelitian ini menggunakan metode Support Vector Machine (SVM) untuk mengklasifikasikan pola aksara Batak Toba dengan bantuan Histogram of Oriented Gradients (HOG) sebagai teknik ekstraksi fitur. Data yang digunakan berasal dari berbagai sumber, dengan total 285 citra aksara. Setelah melalui tahap preprocessing, SVM diterapkan untuk memisahkan karakter menjadi dua kelas utama dan kemudian dibagi lagi menjadi subkelas hingga mencapai klasifikasi akhir. Hasil penelitian menunjukkan bahwa kombinasi HOG dan SVM mampu mengklasifikasikan karakter aksara Batak Toba dengan akurasi 89,47%. Penelitian ini memberikan kontribusi penting dalam upaya pelestarian aksara Batak Toba dan memiliki potensi aplikasi lebih luas dalam pengenalan pola dan klasifikasi citra.

Kata Kunci: *Aksara Batak Toba, Support Vector Machine, Histogram of Oriented Gradients, Pengenalan Pola.*

PENDAHULUAN

Aksara Batak Toba merupakan salah satu warisan budaya Nusantara yang memiliki nilai historis dan kultural yang tinggi. Namun, seiring dengan perkembangan zaman, penggunaan aksara ini semakin langka. Kompleksitas bentuk aksara Batak Toba yang beragam, serta variasi gaya tulisan yang dihasilkan dari berbagai sumber, menimbulkan tantangan dalam pengembangan algoritma pengenalan pola yang efektif. Tantangan ini diperparah dengan adanya noise pada citra serta kemiripan antar karakter yang dapat menyebabkan kesalahan dalam proses pengenalan pola. Dalam konteks pelestarian budaya, pengembangan teknologi yang mampu mengenali dan mengklasifikasikan aksara Batak Toba menjadi sangat penting untuk mendokumentasikan dan menjaga keberlangsungan penggunaan aksara ini di masa mendatang.

Dengan memanfaatkan SVM, sebuah algoritma klasifikasi yang telah terbukti efektif dalam berbagai

konteks, penelitian ini diharapkan dapat membuka pintu baru untuk pemahaman, dokumentasi, dan pemeliharaan aksara Batak Toba. SVM adalah algoritma yang digunakan untuk klasifikasi atau pengenalan pola dan regresi. SVM bekerja dengan cara memisahkan data ke dalam kelas-kelas yang berbeda berdasarkan pada karakteristik dan fitur-fitur yang ada dalam data tersebut (Yulianti et al., 2019).

Citra digital adalah kumpulan pixel dalam sejumlah baris dan sejumlah kolom. Dengan perkataan lain, citra adalah matrix pixel. Adapun pixel adalah elemen terkecil yang menyusun gambar. Setiap pixel mempunyai alamat, yang dinyatakan dalam indeks baris dan indeks kolom. Sistem koordinat untuk pixel-pixel yang menyusun suatu citra dengan ukuran $M \times N$ pixel. Indeks kolom maupun baris dimulai dari 0. Dengan menggunakan notasi (x,y) untuk mewakili suatu pixel, maka pixel pada pojok kiri atas mempunyai koordinat $(0,0)$ (Kadir, 2019).

Dalam proses klasifikasi pola citra aksara Batak Toba menggunakan algoritma Support Vector Machine dibutuhkan data citra yang telah melalui tahap preprocessing. Teknik preprocessing yang umum digunakan adalah metode Histogram of Oriented Gradients (HOG). HOG beroperasi dengan menggunakan distribusi lokal intensitas cahaya dari gradien, yang cocok untuk menggambarkan garis dan bentuk objek. Citra dibagi menjadi sel-sel yang merepresentasikan daerah-daerah yang akan dibandingkan dan dibagi secara merata. Daerah-daerah ini kemudian dikelompokkan menjadi blok-blok yang selanjutnya dinormalisasi untuk memastikan konsistensi terhadap perubahan pencahayaan atau efek fotometrik. Fitur-fitur yang diekstraksi dihitung di atas blok-blok ini, dan blok-blok ini dikenal sebagai Histogram of Oriented Gradient (Leidiyana & Warta, 2022).

Penelitian ini bertujuan untuk menerapkan metode Histogram of Oriented Gradient dalam ekstraksi fitur citra serta metode Support Vector Machine untuk melakukan klasifikasi terhadap citra aksara Batak Toba.

TINJAUAN PUSTAKA

Citra Digital

Gambar digital dapat dilihat sebagai array dua dimensi yang terdiri dari piksel yang disusun dalam kolom dan baris. Secara umum, gambar mewakili kumpulan piksel, dan piksel adalah elemen terkecil dari sebuah gambar. Setiap piksel memiliki alamat yang dibentuk oleh indeks baris dan kolom. Untuk gambar berukuran $M \times N$, indeks piksel dimulai dari nol, yaitu piksel di sudut kiri atas berada pada koordinat (0,0) dalam notasi (x,y) (Rumapea et al., 2023).

Representasi gambar digital adalah metode untuk menggambarkan gambar dalam bentuk yang dapat diproses oleh komputer. Ini mencakup penggunaan matriks atau array angka untuk menggambarkan tingkat kecerahan atau warna setiap piksel. Dalam pengolahan gambar digital, ada beberapa bentuk representasi yang penting dipahami (Dijaya & Setiawan, 2023):

a. Citra Grayscale

Gambar grayscale hanya memuat informasi tentang tingkat kecerahan. Representasi digital setiap piksel pada gambar ini biasanya ditunjukkan dengan angka tunggal antara 0 dan 255. Angka 0 menunjukkan warna hitam, sementara 255 menunjukkan warna putih.

b. Citra Warna

Gambar berwarna memuat informasi tentang kecerahan dan warna. Representasi yang paling

umum adalah mode RGB (Merah, Hijau, Biru), di mana setiap piksel diwakili oleh tiga nilai yang menunjukkan intensitas masing-masing warna dalam rentang 0 hingga 255. Semakin banyak bit warna yang digunakan, semakin halus transisi warnanya.

c. Citra Biner

Gambar biner hanya memiliki dua nilai piksel: hitam dan putih. Ini dihasilkan setelah proses segmentasi, di mana piksel dibagi menjadi dua berdasarkan ambang batas tertentu. Representasi gambar biner menggunakan nilai 0 untuk hitam dan 1 untuk putih.

Histogram of Oriented Gradient

Ekstraksi fitur yang digunakan adalah Histogram of Oriented Gradient (HOG). HOG merupakan representasi lokal dari objek, memanfaatkan nilai intensitas gradien untuk mengekstraksi fitur yang sangat berguna dalam bidang computer vision dan pengolahan citra. Kelebihan HOG mencakup kemampuannya untuk mengidentifikasi tepi dan struktur gradien yang mencerminkan karakteristik yang kuat dari bentuk sebenarnya. Setiap citra memiliki karakteristik yang dapat diidentifikasi melalui distribusi gradiennya, yang diperoleh dengan membagi citra ke dalam area kecil yang disebut cell. Proses HOG memiliki tahapan sebagai berikut (Yohannes et al., 2021):

a. Konversi Citra RGB ke Grayscale

Untuk mengubah citra RGB menjadi citra grayscale, secara umum digunakan persamaan sebagai berikut:

$$G = (0,2125 \times R) + (0,7154 \times G) + (0,0721 \times B) \dots (1)$$

Dimana:

R = komponen warna merah.

G = komponen warna hijau.

B = komponen warna biru.

Persamaan ini digunakan untuk menghitung nilai intensitas grayscale berdasarkan kontribusi masing-masing komponen warna.

b. Menghitung nilai gradient

Untuk menghitung nilai gradien pada setiap piksel dalam citra grayscale, langkah pertama adalah membagi gambar menjadi wilayah-wilayah kecil yang terhubung yang disebut cell. Setiap cell tersebut dikelompokkan ke dalam bentuk block. Nilai gradien pada setiap piksel dihitung menggunakan persamaan sebagai berikut:

$$I_x(r,c) = I(r,c+1) - I(r,c-1) \dots \dots \dots (2)$$

$$I_y(r,c) = I(r+1,c) - I(r-1,c) \dots \dots \dots (3)$$

Di mana I adalah citra (image), dan r serta c masing-masing adalah baris dan kolom dari citra tersebut.

Setelah mendapatkan nilai I_x dan I_y , besar gradien μ dapat dihitung menggunakan persamaan:

$$\mu = \sqrt{I_x^2 + I_y^2} \dots \dots \dots (4)$$

Setelah nilai μ diperoleh, langkah berikutnya adalah menentukan nilai orientasi θ dengan menggunakan persamaan:

$$\theta_{(x,y)} = \text{mod} \left(\frac{180}{\pi} \tan^{-1} \left(\frac{G_x(x,y)}{G_y(x,y)} \right) \right) 180 \dots (5)$$

c. Orientasi Binning

Setelah menghitung nilai μ (besar gradien) dan orientasi untuk setiap blok, langkah berikutnya adalah melakukan orientasi binning. Ini dilakukan dengan mengelompokkan orientasi gradien ke dalam beberapa bin dalam histogram orientasi. Setiap bin mewakili rentang orientasi tertentu, tergantung pada representasi yang digunakan. Tujuan dari binning ini adalah untuk membuat representasi distribusi arah gradien yang lebih mudah dianalisis.

d. Menormalisasikan nilai tiap bin orientasi

Setelah nilai-nilai bin orientasi diperoleh, langkah selanjutnya adalah normalisasi blok. Normalisasi ini bertujuan untuk mengurangi sensitivitas terhadap variasi pencahayaan atau kontras dalam citra. Dengan menormalisasikan nilai orientasi pada tiap bin, kita memastikan bahwa perbedaan pencahayaan antar blok tidak mempengaruhi hasil pengenalan pola.

Support Vector Machine

Support Vector Machine (SVM) adalah metode klasifikasi. Pada intinya, ini adalah jenis algoritma pembelajaran terarah (supervised learning) yang menggabungkan kumpulan ide-ide terbaik dari komputasi yang telah muncul selama beberapa dekade. Salah satu teorinya yang paling menonjol adalah margin hyperplane kernel, yang awalnya diusulkan oleh Aronszajn pada tahun 1950, bersama dengan sejumlah konsep lain yang melengkapi kekurangannya. SVM memerlukan dataset pelatihan yang terdiri dari kelas positif dan negatif. Dataset ini digunakan oleh SVM untuk menghasilkan keputusan terbaik dalam memisahkan data dari dua kelas yang terpisah di ruang n-dimensi menggunakan hyperplane. Secara umum, SVM adalah pencarian untuk garis pemisah terbaik, atau yang disebut hyperplane, dengan keinginan maksimal untuk menjaga jarak atau margin maksimum antara dua kelompok kelas (Cholissodin et al., 2019).

Penjelasan singkat tentang Support Vector Machine (SVM). SVM pada dasarnya adalah klasifikasi linear, meskipun juga dapat digunakan untuk masalah non-linear. Pada dataset yang digunakan, terdapat berbagai kategori. SVM linear

biasanya diterapkan pada data yang dapat dipisahkan secara linear. Contoh yang diberikan adalah, dengan memperhatikan $x_i = \{x_1, x_2, \dots, x_n\} \in R^n$ dan $y_i \in \{-1, +1\}$, di mana x_i adalah titik data dan y_i adalah kelas data dari titik data x_i tersebut. Langkah pertama adalah menemukan hyperplane atau garis pemisah antara dua kelas. Ini menggunakan fungsi linear yang dapat diformulasikan secara matematis, seperti dalam persamaan.

$$f(x) = w \cdot x_i + b \dots \dots \dots (6)$$

Dalam SVM, bobot yang diberikan kepada support vector atau vektor yang tegak lurus terhadap hyperplane dilambangkan dengan w , dan b adalah variabel bias (threshold). Data x_+ adalah support vector untuk kelas positif, sedangkan x_- adalah support vector untuk kelas negatif dengan nilai alpha tertinggi. Fungsi keputusan klasifikasi menggunakan tanda $\text{sign}(f(x))$, yang menunjukkan apakah data dapat diklasifikasikan ke dalam kelas positif atau negatif. Aturan yang digunakan dalam pengklasifikasian data adalah:

- $\text{sign}(f(x)) = 1$ untuk kelas positif
- $\text{sign}(f(x)) = -1$ untuk kelas negatif

Ida dasar SVM adalah menemukan hyperplane pemisah yang optimal, yaitu dengan memaksimalkan margin antara dua kelas. Hyperplane optimal adalah yang berada di tengah-tengah antara dua kelompok objek di masing-masing kelas.

Confusion Matrix

Confusion matrix adalah metrik lain yang digunakan untuk menghitung akurasi dalam klasifikasi. Terdapat empat istilah utama yang digunakan untuk mengklasifikasikan hasil evaluasi saat menggunakan confusion matrix: True Positive, True Negative, False Positive, dan False Negative. True Positive (TP) dan True Negative (TN) adalah prediksi yang benar. TP adalah jumlah data positif yang diprediksi dengan benar sebagai positif, dan TN adalah jumlah data negatif yang diprediksi dengan benar sebagai negatif. False Positive (FP) terjadi ketika hasil positif muncul, padahal seharusnya negatif, dan False Negative (FN) terjadi ketika hasil negatif muncul, padahal seharusnya positif. Nilai akurasi, presisi, recall, dan f1-score untuk model klasifikasi akan dihitung berdasarkan confusion matrix ini. Visualisasi Confusion Matrix dapat dilihat pada Gambar 1.

		Actual Class	
		Yes	No
Predicted Class	Yes	TP	FN
	No	FP	TN

Gambar 1 Confusion Matrix

Berdasarkan nilai dari confusion matrix tersebut, dapat dihasilkan nilai akurasi, presisi, recall, dan f1-score sebagai berikut :

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

METODE PENELITIAN

Aksara Batak Toba

Surat Batak sering diklasifikasikan sebagai sebuah silabogram, namun ini jelas keliru karena aksara Batak sebagaimana juga aksara-aksara lainnya di Nusantara merupakan bagian dari rumpun tulisan Brahmi (India) yang lebih tepat dapat diklasifikasikan sebagai abugida (paduan antara silabogram dan abjad). Sebuah abugida terdiri dari aksara yang melambangkan sebuah konsonan sementara vokal dipasang pada aksara sebagai diakritik. Abugida adalah jenis tulisan yang bersifat fonetis dalam arti bahwa setiap bunyi bahasanya dapat dilambangkan secara akurat (Kozok, 2009).

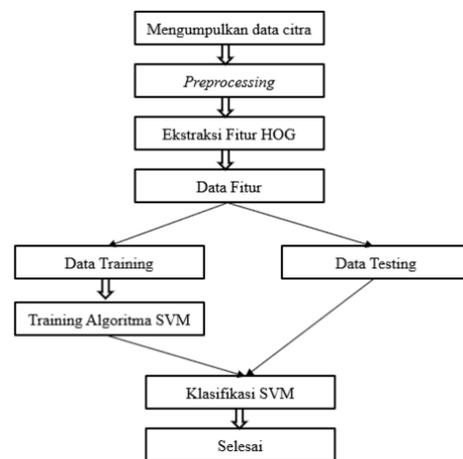
Aksara Batak terdiri dari dua kelompok huruf utama: ina ni surat dan anak ni surat. Ina ni surat terdiri dari 19 karakter, sedangkan kelompok lainnya, yaitu anak ni surat, terdiri dari 6 karakter. Penelitian ini akan mengklasifikasikan citra yang berasal dari ina ni surat. Karakter-karakter pada ina ni surat dapat dilihat pada Tabel 1.

Tabel 1. Ina ni Surat

Huruf Latin	Aksara	Huruf Latin	Aksara
a	ᐃ	ma	ᐃᐅ
ha	ᐃᐇ	ta	ᐃᐇᐅ
ba	ᐃᐇᐅ	sa	ᐃᐇᐅᐅ
pa	ᐃᐇᐅᐅ	ya	ᐃᐇᐅᐅᐅ
na	ᐃᐇᐅᐅᐅ	nga	ᐃᐇᐅᐅᐅᐅ
wa	ᐃᐇᐅᐅᐅᐅ	la	ᐃᐇᐅᐅᐅᐅᐅ
ga	ᐃᐇᐅᐅᐅᐅᐅ	nya	ᐃᐇᐅᐅᐅᐅᐅᐅ
ja	ᐃᐇᐅᐅᐅᐅᐅᐅ	i	ᐃᐇᐅᐅᐅᐅᐅᐅᐅ
da	ᐃᐇᐅᐅᐅᐅᐅᐅᐅ	u	ᐃᐇᐅᐅᐅᐅᐅᐅᐅᐅ
ra	ᐃᐇᐅᐅᐅᐅᐅᐅᐅᐅ		

Tahapan Penelitian

Penelitian ini memiliki beberapa langkah-langkah yang harus dilalui yang dirangkum dalam Framework pada Gambar 2.



Gambar 2. Framework Penelitian

Tahapan penelitian dapat dijelaskan sebagai berikut:

- Input dataset: Menyiapkan dataset yang telah dikumpulkan untuk diolah menggunakan program python. Dataset yang digunakan pada penelitian ini adalah 228 citra sebagai data latih dan 57 citra sebagai data uji.
- Preprocessing data: Peneliti mengolah data sebelum menggunakan SVM. Preprocessing meliputi pengumpulan data dan pengolahan data.
- Ekstraksi fitur: Peneliti mengambil informasi dari citra ina ni surat yang diperlukan untuk mengklasifikasikannya. Metode ekstraksi fitur yang digunakan adalah Histogram of Oriented Gradient. Pada proses ini citra akan diolah untuk mendapatkan fitur citra.
- Pembagian kelas ina ni surat: Tahapan ini merupakan proses pembagian 19 karakter ina ni surat ke dalam 2 kelas berdasarkan kemiripan bentuknya. 10 karakter ina ni surat yang memiliki kemiripan bentuk akan berada dalam kelas 1 dan sisanya masuk ke dalam kelas 2. Kemudian kelas 1 akan dibagi kembali menjadi 2 kelas dan demikian juga kelas 2. Begitu seterusnya hingga menyisakan 1 kelas untuk 1 karakter ina ni surat. Dengan metode tersebut dihasilkan 18 kelas baru untuk setiap ina ni surat.
- Peneliti menggunakan Support Vector Machine untuk membangun model klasifikasi dengan konsep multilevel classification. Proses klasifikasi dilakukan secara bertahap. Pada level pertama, model akan mengklasifikasikan 2 kelas ina ni surat yang telah dibagi sebelumnya. Jika hasil klasifikasi bernilai negatif (-), maka data tersebut masuk ke dalam kelas 1, sedangkan hasil positif (+) masuk ke kelas 2, kemudian dilanjutkan ke level berikutnya. Proses ini berlanjut dengan klasifikasi pada setiap level hingga hanya tersisa 2 kelas akhir yang akan diklasifikasikan.

f. Hasil keluaran: Diperoleh hasil keluaran klasifikasi ina ni surat berupa citra ina ni surat dan huruf latinnya.

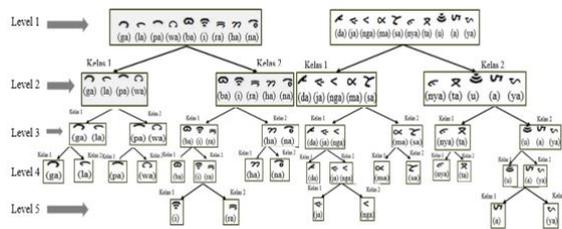
HASIL DAN PEMBAHASAN

Ekstraksi Fitur

Citra yang digunakan pada penelitian ini adalah citra RGB berukuran 28x28 piksel. Citra ini kemudian melalui tahap preprocessing untuk menghitung nilai fiturnya. Dengan menggunakan nilai parameter pixel per cell=14x14, cell per block= 2x2, dan orientasi bin=6, maka dihasilkan 24 nilai fitur untuk setiap citra.

Pembagian Kelas Ina Ni Surat

Dalam proses klasifikasi ina ni surat menggunakan SVM Linear, seluruh karakter tersebut tidak dapat diklasifikasikan sekaligus. Oleh karena itu, 19 ina ni surat akan diklasifikasikan secara bertahap dengan cara membagi mereka ke dalam beberapa kelas. Pembagian kelas ina ni surat dapat dilihat pada Gambar 4.



Gambar 3. Hasil Pembagian Kelas

Pembagian kelas 1 pada level 1 dapat dilihat dari pola dan bentuk yang memiliki lengkungan yang sama sedang karakter-karakter pada kelas 2 memiliki sudut pada setiap karakternya kecuali karakter u yang memiliki lengkungan kebawah sehingga tidak dimasukkan kedalam kelas 1. Pembagian kelas pada level 2 dan selanjutnya dilakukan dengan konsep yang sama yaitu dengan membandingkan pola setiap karakter untuk dikelompokkan menjadi kelas-kelas baru. Contoh pada karakter ga, la, pa dan wa, karakter ga dan la memiliki kesamaan pola garis pendek pada ujung karakter sedangkan pa dan wa memiliki kesamaan hanya pada lengkungannya.

Klasifikasi Support Vector Machine

Pada proses klasifikasi SVM, digunakan konsep multilevel classification atau klasifikasi secara bertahap. Pada level pertama dilakukan pengklasifikasian kelas 1 (9 karakter ina ni surat) dengan kelas 2 (10 karakter ina ni surat). Jika hasil menunjukkan menunjukkan citra kelas 1 maka proses klasifikasi dilanjutkan ke SVM selanjutnya. Perilaku serupa dilakukan jika hasil menunjukkan citra adalah

kelas 2 maka proses klasifikasi dilanjutkan pada SVM selanjutnya. Misalnya akan mengklasifikasikan citra ina ni surat karakter Ga dan La. Maka proses klasifikasi akan melewati 4 tahap atau level pengklasifikasian.

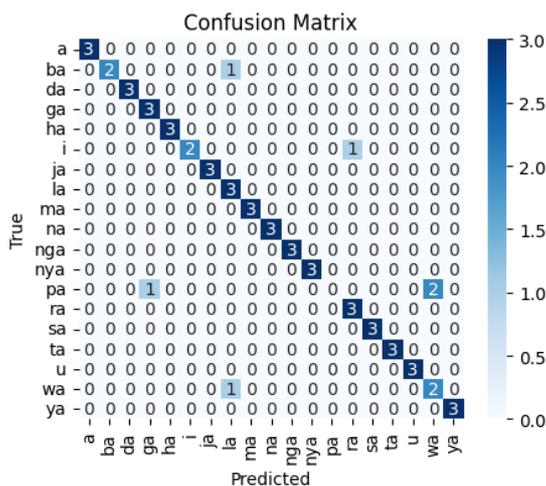
Setelah dilakukan pengujian dengan program, sebanyak 51 dari 57 citra dapat dikenali oleh algoritma pengklasifikasian. Hasil pengujian dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pengujian

No	Citra	Kelas Asli	Prediksi Kelas	Status Prediksi
1.	5	A	A	Benar
2.	5	A	A	Benar
3.	5	A	A	Benar
4.	5	Ba	La	Salah
5.	5	Ba	Ba	Benar
6.	5	Ba	Ba	Benar
7.	5	Da	Da	Benar
8.	5	Da	Da	Benar
9.	5	Da	Da	Benar
10.	5	Ga	Ga	Benar
11.	5	Ga	Ga	Benar
12.	5	Ga	Ga	Benar
13.	5	Ha	Ha	Benar
14.	5	Ha	Ha	Benar
15.	5	Ha	Ha	Benar
16.	5	I	Ra	Salah
17.	5	I	I	Benar
18.	5	I	I	Benar
19.	5	Ja	Ja	Benar
20.	5	Ja	Ja	Benar
21.	5	Ja	Ja	Benar
22.	5	La	La	Benar
23.	5	La	La	Benar
24.	5	La	La	Benar
25.	5	Ma	Ma	Benar
26.	5	Ma	Ma	Benar
27.	5	Ma	Ma	Benar
28.	5	Na	Na	Benar
29.	5	Na	Na	Benar
30.	5	Na	Na	Benar
31.	5	Nga	Nga	Benar
32.	5	Nga	Nga	Benar
33.	5	Nga	Nga	Benar
34.	5	Nya	Nya	Benar
35.	5	Nya	Nya	Benar
36.	5	Nya	Nya	Benar
37.	5	Pa	Wa	Salah
38.	5	Pa	Wa	Salah

39.	Pa	Ga	Salah
40.	Ra	Ra	Benar
41.	Ra	Ra	Benar
42.	Ra	Ra	Benar
43.	Sa	Sa	Benar
44.	Sa	Sa	Benar
45.	Sa	Sa	Benar
46.	Ta	Ta	Benar
47.	Ta	Ta	Benar
48.	Ta	Ta	Benar
49.	U	U	Benar
50.	U	U	Benar
51.	U	U	Benar
52.	Wa	Wa	Benar
53.	Wa	Wa	Benar
54.	Wa	La	Salah
55.	Ya	Ya	Benar
56.	Ya	Ya	Benar
57.	Ya	Ya	Benar

Berdasarkan hasil pengujian program klasifikasi dengan SVM berikut adalah akurasi menggunakan metode confusion matriks yang ditampilkan pada Gambar 5.



Gambar 4. Confusion Matriks

Untuk menghitung persentase akurasi digunakan rumus:

$$\begin{aligned}
 Acc &= (TP(A) + TP(BA) + TP(DA) + TP(GA) \dots \\
 &\quad + TP(YA)) / (TP(A) + TP(BA) \\
 &\quad + TP(DA) \dots + FP(A) + FP(BA) \dots \\
 &\quad + FN(A)) \times 100\% \\
 Acc &= \frac{2 + 2 + 3 + 3 + 2 + 0 + \dots + 3}{2 + 2 + 3 + \dots + 0 + \dots + 0} \times 100\% \\
 Acc &= \frac{51}{57} \times 100\% \\
 Acc &= 89,47\%
 \end{aligned}$$

KESIMPULAN

Berdasarkan hasil penelitian pada bab sebelumnya dapat disimpulkan bahwa kombinasi antara ekstraksi fitur HOG dan algoritma SVM dapat digunakan untuk mengklasifikasikan karakter ina ni surat dengan akurasi yang memadai. Beberapa point yang dirangkum dalam penelitian ini antara lain:

1. Penerapan pembagian kelas-kelas ina ni surat bekerja dengan efektif namun masih memiliki ruang untuk hasil yang lebih optimal.
2. HOG mampu mengekstraksi fitur penting dari gambar karakter sehingga memudahkan algoritma klasifikasi untuk mengklasifikasikan ina ni surat.
3. SVM memanfaatkan fitur tersebut untuk melakukan klasifikasi dengan baik. Tingkat akurasi sebesar 89,47% menunjukkan bahwa metode ini cukup andal, meskipun masih ada ruang untuk perbaikan lebih lanjut dalam meningkatkan akurasi klasifikasi.

Penelitian ini dapat menjadi dasar untuk penelitian lanjutan dalam pengembangan sistem klasifikasi karakter berbasis gambar, dan juga dapat diaplikasikan dalam bidang lain yang memerlukan teknik pengenalan pola dan klasifikasi yang serupa.

DISEMINASI

Artikel ini telah diseminasikan pada Seminar Nasional Teknologi Informasi dan Komunikasi (SEMNASTIK) APTIKOM Tahun 2024 yang diselenggarakan oleh Universitas Methodist Indonesia pada tanggal 24-26 Oktober 2024.

DAFTAR PUSTAKA

Cholissodin, I., Sutrisno, Soebroto, A. A., Hasanah, U., & Febiola, Y. I. (2019). *Buku Ajar AI, Machine Learning & Deep Learning*. UB FILKOM.
<https://www.researchgate.net/publication/348003841>

Dijaya, R., & Setiawan, H. (2023). Buku ajar Pengolahan Citra Digital. In M. Tanzil (Ed.), *Umsida Press*. UMSIDA Press.

Kadir, A. (2019). *Langkah Mudah Pemograman OpenCV & Python* (p. 158). PT Elex Media Komputindo.

Leidiyana, H., & Warta, J. (2022). Implementasi Metode SVM untuk Klasifikasi Bunga dengan Ekstraksi Fitur Histogram of Gradient (HOG). *Journal of Information and Information Security (JIFORTY)*, 3(1), 89.
<http://ejurnal.ubharajaya.ac.id/index.php/jiforty>

Rumapea, H., Sinambela, M., Jaya, I. K., & Sarkis, I. M. (2023). Prediction of Rainfall in North Sumatera Using Machine Learning. 2023 *International Conference of Computer Science and Information Technology (ICOSNIKOM)*, 1–4.

<https://doi.org/10.1109/ICoSNIKOM60230.2023.10364504>

- Yohannes, Y., Udjulawa, D., & Ivan Sariyo, T. (2021). Klasifikasi Jenis Jamur Menggunakan SVM dengan Fitur HSV dan HOG. *PETIR*, 15(1), 113–120.
<https://doi.org/10.33322/petir.v15i1.1101>
- Yulianti, R., Pasek, G., Wijaya, S., & Bimantoro, D. F. (2019). *Pengenalan Pola Tulisan Tangan Suku Kata Aksara Sasak Menggunakan Metode Moment Invariant dan Support Vector Machine (Handwritten Sasak Ancient Script Recognition using Moment Invariant and Support Vector Machine)*. <http://jcosine.if.unram.ac.id/>