

Verifikasi Citra Tanda Tangan Berdasarkan Ciri Pyramid Histogram of Oriented Gradient (PHOG) Menggunakan Metode Klasifikasi K-Nearest Neighbor

Latifa Nabila Harfiya¹, Agus Wahyu Widodo², Randy Cahya Wihandika³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹latifanaharfiya@gmail.com, ²a_wahyu_w@gmail.com, ³rendicahya@ub.ac.id

Abstrak

Tanda tangan secara luas telah diterima oleh masyarakat menjadi salah satu wujud alat verifikasi identitas seseorang. Untuk menghindari pemalsuan, penyalahgunaan, maupun kejahatan lainnya yang terkait maka diperlukan adanya verifikasi tanda tangan. Proses verifikasi tanda tangan menggunakan metode ekstraksi ciri berdasarkan *Pyramid Histogram of Oriented Gradient* (PHOG) untuk mendapatkan ciri dari area global hingga area lokal citra yang digunakan untuk proses setelahnya yaitu klasifikasi menggunakan metode *K-Nearest Neighbor* (K-NN). Terdapat berbagai parameter yang dapat mempengaruhi ekstraksi ciri PHOG dan metode klasifikasi K-NN seperti jumlah bin, level, rentang sudut, dan K. Adapun parameter tambahan yaitu jumlah data latih yang secara keseluruhan mempengaruhi hasil klasifikasi yang digunakan. Ekstraksi ciri dan metode klasifikasi dengan menggunakan parameter dan jumlah data latih terbaik berhasil menghasilkan akurasi tertinggi sebesar 99,5% pada data tanda tangan Indonesia asli dan 98,5% pada data tanda tangan Persia asli. Sedangkan pada data tanda tangan palsu akurasi yang dihasilkan sebesar 56% pada data tanda tangan Indonesia dan 35,5% pada data tanda tangan Persia. Hasil pengujian menunjukkan bahwa walaupun algoritme bekerja dengan baik dalam melakukan pengenalan, tetapi masih belum cukup baik membedakan tanda tangan asli dengan tanda tangan palsu yang memiliki kemiripan yang tinggi dengan tanda tangan asli.

Kata kunci: verifikasi tanda tangan, *Pyramid Histogram of Oriented Gradient*, *K-Nearest Neighbor*

Abstract

Signature have been widely accepted by people as one of many tools that used to verify a person's identity. Signature verification is highly needed in order to avoid crimes regarding signature's validity. The process of signature image verification uses feature extraction based on Pyramid Histogram of Oriented Gradient (PHOG) for extracting feature from global to local image area that used for the next process, classification using K-Nearest Neighbor method. There are some parameters that can affect the feature extraction of PHOG and K-NN as classification method such as number of bin, level, range of angles, and K. As for the additional parameters, namely the amount of training data that affect the overall results of the classification used. Feature extraction and classification by the method with the best parameter values and training data used produces the highest accuracy of 99.5% on Indonesian original signature data and 98.5% on the data of the Persian original signatures. While the forgery signatures data produces accuray only as much as 56% on data from Indonesia and 35,5% on data from Persian. Results from tests show that the algorithm is not good enough for distinguishing forgery signature that has high similarity with genuine signature even it is works well for recognizing genuine signature.

Keywords: signature verification, *Pyramid Histogram of Oriented Gradient*, *K-Nearest Neighbor*

1. PENDAHULUAN

Tanda tangan dikenal sebagai suatu atribut unik yang dimiliki secara personal yang

kemudian secara luas telah diterima oleh masyarakat menjadi salah satu wujud alat verifikasi identitas seseorang. Hingga kini, tanda tangan terus digunakan sebagai alat yang dapat mengotorisasi berbagai keperluan legalitas

untuk berbagai macam dokumen. Berdasarkan fakta tersebut, tanda tangan telah dinyatakan berperan sangat krusial dalam fungsi pemberi ciri suatu akta secara hukum (Widodo & Harjoko, 2015). Akan tetapi, untuk kepentingan tertentu tanda tangan menjadi rentan akan pemalsuan dan penyalahgunaan yang menyebabkan perlu dilakukannya verifikasi tanda tangan.

Verifikasi tanda tangan merupakan proses penentuan apakah tanda tangan tersebut asli atau tidak berdasarkan perbedaan antara variasi pada penandatanganan dengan pemalsunya. Tahapan pasti yang dilakukan pada pengembangan verifikasi tanda tangan ialah melakukan ekstraksi ciri yang cocok dengan citra tanda tangan yang dapat mengategorikan perbedaan antara tanda tangan setiap individu. Ciri yang baik harus mengandung informasi dengan tingkat ambiguitas paling rendah (Dasgupta, et al., 2016). Sejauh ini sejumlah ciri diklasifikasikan menjadi dua jenis yaitu ciri global dan ciri lokal. Ciri global menggambarkan tanda tangan sebagai keseluruhan sedangkan ciri lokal hanya menggambarkan segmen-segmen tertentu. Penentuan ciri yang akan digunakan untuk menjadi representasi yang efisien dari tanda tangan merupakan salah satu permasalahan mendasar dalam verifikasi tanda tangan.

Pyramid Histogram of Oriented Gradient (PHOG) merupakan ciri yang dapat membuat proses ekstraksi dan seleksi lebih efisien dan meningkatkan tingkat pengenalan. Ciri ini mengambil sifat spasial dari segmen lokal ke dalam satuan sementara citra direpresentasikan dengan Histogram of Oriented Gradient (HOG). Informasi spasial direpresentasikan dengan melakukan segmentasi citra menjadi beberapa area pada beberapa resolusi berdasarkan levelnya. Distribusi spasial ini yang menjadi keunggulan dari PHOG dalam melakukan verifikasi (Saidani & Kacem, 2014).

Pada proses verifikasi, tanda tangan yang diuji akan dibandingkan dengan seluruh tanda tangan yang telah dilatih dan nilai *similarity* nya akan dianggap sebagai jarak di antara tanda tangan yang dibandingkan. Terdapat beberapa metode yang dapat digunakan untuk menentukan apakah jarak antara tanda tangan yang diuji dengan tanda tangan yang dibandingkan dapat diterima atau tidak. Metode yang dimaksud biasanya menggunakan algoritme untuk klasifikasi seperti *Multi-Layer Perceptron* (MLP), *K-nearest neighbor* (K-NN), dan

Support Vector Machine (SVM) (Zhang, et al., 2010).

Pada penelitian sebelumnya yang berkaitan dengan verifikasi tanda tangan telah dilakukan oleh Zhang, dkk. (2010) dengan menggunakan metode Pyramid Histogram of Oriented Gradient (PHOG) untuk membandingkan metode klasifikasi *Generalized Linear Model* (GLM), K-NN, *Fisher Linear Discriminant* (FLD), MLP, dan SVM. Penelitian terkait juga dilakukan oleh Widodo dan Harjoko (2015) dengan menggunakan ekstraksi ciri HOG, Histogram of Curvature (HoC), dan SVM sebagai metode pengklasifikasi dalam proses verifikasinya.

Berdasarkan uraian dan referensi di atas, penelitian ini menggunakan ekstraksi ciri PHOG dan metode klasifikasi K-NN dalam melakukan proses verifikasi citra tanda tangan untuk mendapatkan nilai akurasi yang akan diukur dengan menggunakan parameter FRR dan FAR. Dengan penelitian ini diharapkan hasil dapat menjadi solusi yang lebih baik dan menghasilkan tingkat keakuratan yang lebih tinggi dalam verifikasi tanda tangan.

2. HISTOGRAM OF ORIENTED GRADIENT

Ide dasar dari deskriptor *Histogram of Oriented Gradient* (HOG) adalah bahwa penampilan obyek dan bentuk lokal dalam sebuah citra dapat digambarkan oleh distribusi intensitas arah gradien tepi (Widodo & Harjoko, 2015). Penerapan deskriptor ini dapat dicapai dengan membagi citra ke dalam daerah-daerah kecil yang saling terhubung, yang disebut sel dan untuk setiap sel disusun histogram arah gradien atau orientasi tepi untuk tiap-tiap piksel yang berada di dalam sel. Kombinasi histogram-histogram ini kemudian yang menghasilkan suatu deskriptor. Sedangkan gradien pada masing-masing piksel didapatkan dengan menerapkan operator Sobel yang menggunakan filter di mana titik tengahnya bernilai 2 dan tetangganya bernilai 1. Dengan nilai seperti itu diasumsikan bahwa semakin jauh piksel tetangga dari titik perhitungan maka kontribusi intensitasnya akan semakin kecil sehingga diharapkan intensitas tepi yang dihasilkan akan meningkat (Madenda, 2015).

Dasar pembentukan ciri HOG di masing-masing sel adalah proses akumulasi piksel yang memiliki orientasi yang sama. Di mana arah gradien akan dikelompokkan menjadi beberapa

bagian yang disebut bin. Menurut penelitian yang dilakukan oleh Dalal dan Triggs (2005), untuk mendapatkan hasil deteksi terbaik jika jumlah bin yang digunakan untuk pembagian arah minimal berjumlah 8 bin. Namun pada penelitian lain disebutkan bahwa jumlah bin sebesar 9 memberikan hasil deteksi yang lebih optimal (Kobayashi, et al., 2008).

Orientasi direpresentasikan sebagai derajat tunggal dan derajat ganda dengan rentang sudut yang disesuaikan apakah gradien tergolong signed atau unsigned. Pada penelitian yang dilakukan oleh Shu, et al. (2011) dinyatakan bahwa penggunaan derajat tunggal dapat menghasilkan hasil yang lebih baik karena derajat tunggal menganggap tepi pada daerah yang terbalik memiliki orientasi yang berlawanan. Sehingga dengan derajat tunggal semakin banyak pola yang dapat dibedakan.

3. PYRAMID HISTOGRAM OF ORIENTED GRADIENT

PHOG awalnya diusulkan oleh Bosch, et al. (2007) untuk mendefinisikan representasi padat dari citra pada resolusi tertentu. Ciri ini mengambil sifat spasial dari segmen lokal ke dalam satuan sementara citra direpresentasikan dengan ciri *Histogram of Oriented Gradient* (HOG). Informasi spasial direpresentasikan dengan melakukan segmentasi citra menjadi beberapa area pada beberapa resolusi berdasarkan levelnya. Masing-masing citra kemudian dibagi ke dalam urutan grid spasial yang bertambah dengan menguadratkan jumlah segmen pada setiap arah sumbu. Jumlah titik pada sebuah sel pada suatu level adalah jumlah dari keempat sel yang dibagi pada level berikutnya, hal ini yang membentuk representasi piramid (Saidani & Kacem, 2014). Ilustrasi segmentasi citra yang dipiramidkan digambarkan pada Gambar 1.



Gambar 1. Ilustrasi segmentasi citra yang dipiramidkan

Untuk melakukan ekstraksi ciri PHOG dari citra tanda tangan maka langkah yang dilakukan adalah sebagai berikut:

1. Menghitung gradien: metode paling umum adalah dengan mengaplikasikan titik pusat turunan pada kedua arah horizontal maupun vertikal. Secara spesifik metode ini membutuhkan filtrasi citra *grayscale* atau citra biner dengan suatu kernel filter. Kernel filter yang digunakan pada penelitian ini merupakan operator Sobel seperti pada Persamaan 1:

$$D_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad D_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad (1) \quad (2.9)$$

Dengan besarnya gradien dihitung pada Persamaan 2:

$$G = \sqrt{G_x^2 + G_y^2} \quad (2)$$

Orientasi gradien ditunjukkan pada Persamaan 3:

$$R_o = \arctan \frac{G_y}{G_x} \quad (3)$$

2. Membuat sel histogram (bin orientasi): langkah kedua perhitungan melibatkan pembuatan histogram sel. Setiap piksel dalam sel memiliki bobot untuk kanal histogram berdasarkan orientasi gradien. Kanal histogram tersebar menggunakan rentang sudut menggunakan gradien *signed* atau *unsigned* yaitu dari 0° hingga 180° atau 0° hingga 360° .
3. Menghitung matriks orientasi dan gradien: mengisi matriks pertama dengan banyaknya gradien yang berkoresponden dengan nilai bin histogram. Dengan cara yang sama, matriks kedua akan diisi dengan besar gradien berkoresponden.
4. Menggabungkan HOG pada setiap area dan mendapatkan deskriptor PHOG akhir

4. K-NEAREST NEIGHBOR

Dalam mesin pembelajaran, terdapat banyak sekali metode klasifikasi terkenal yang salah satunya adalah metode K-NN. Metode ini didasarkan pada fungsi jarak, contohnya jarak Euclidean untuk membandingkan pasangan sampel data. Metode ini pertama kali diusulkan

oleh T.M. Cover dan P.E. Hart san telah menjadi metode yang paling banyak digunakan dalam memecahkan kasus-kasus klasifikasi (Aouidate, 2011). Algoritme K-NN memiliki konsistensi yang sangat kuat dengan hal yang paling mempengaruhi ketepatan algoritme K-NN adalah ada atau tidaknya fitur-fitur yang tidak relevan atau bobot fitur dimaksud tidak setara dengan relevansinya terhadap klasifikasi. Algoritme K-NN dijelaskan pada tahap-tahap sebagai berikut (Priambodo, 2015):

1. Definisikan nilai K.
2. Lakukan penentuan jarak antara input dengan semua data uji.
3. Urutkan data berdasarkan hasil penentuan jarak.
4. Bentuk kelompok berdasarkan nilai dengan ketetanggaan terdekat.
5. Pilih nilai dengan kemunculan paling sering dari tetangga terdekat sebagai parameter prediksi untuk data selanjutnya.

Jarak dalam penelitian ini ditentukan dengan cara menggunakan *Euclidean distance* (2-distance) (Medjahed, et al., 2013). Perhitungan jarak *Euclidean* ditunjukkan pada Persamaan 4:

$$d(X_i, X_j) = \sqrt{\sum_{r=1}^n (X_{ir} - X_{jr})^2} \quad (4)$$

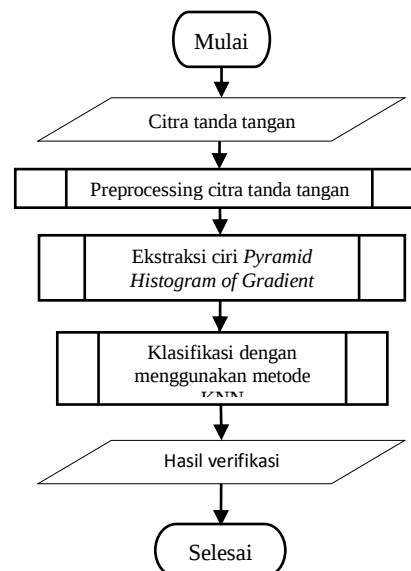
Keterangan:

- d : Jarak *Euclidean* antara X_i dengan X_j
 X_i : Data ke-i yang akan dilakukan perhitungan
 X_j : Data ke-j yang akan dilakukan perhitungan

5. METODE PENELITIAN

Secara garis besar terdapat tiga proses utama yang diimplementasikan ke dalam program untuk penelitian ini yaitu input, pemrosesan yang terdiri dari pemrosesan data latih dan pengujian, dan output. Pada pemrosesan data latih, citra tanda tangan harus diperbaiki kualitasnya dengan cara melakukan *preprocessing* yang terdiri dari binerisasi, *image denoising*, deteksi dan koreksi kemiringan, dan normalisasi. Setelah kualitas diperbaiki maka cita akan diambil cirinya dengan menggunakan ekstraksi ciri PHOG nya. Ciri dari citra yang

telah diambil kemudian disimpan ke dalam sebuah matriks yang akan digunakan dalam proses pengujian. Pada proses pengujian, salah satu citra tanda tangan akan dipilih kemudian dilakukan proses yang sama dengan pemrosesan data latih yang dilanjutkan dengan pencarian kelas citra yang diuji dengan melakukan perhitungan untuk klasifikasi menggunakan metode K-NN. Metode ini memiliki tahap pelatihan yang datanya didapatkan dari matriks yang digunakan untuk menyimpan nilai ciri dari proses pelatihan yang sebelumnya telah dilakukan. Dengan input berupa citra tanda tangan maka keluaran atau output yang dihasilkan nantinya berupa kelas dan hasil verifikasi dari klasifikasi data uji. Alur dari proses implementasi ditunjukkan pada Gambar 2.



Gambar 2. Diagram alir metode penelitian

6. PENGUKURAN KINERJA

Proses pengukuran kinerja atau pengujian program hasil dari implementasi menggunakan skenario di mana citra tanda tangan dengan identitas yang diklaim diinputkan ke dalam program hasil implementasi penelitian yang kemudian akan diputuskan penerimaan atau penolakan atas klaim yang diajukan. Terdapat beberapa skenario dalam proses verifikasi, yaitu (Zhang, et al., 2010):

- False rejection (FR)

Sistem dikenai pertanyaan apakah tanda tangan yang diinputkan merupakan milik pemilik yang benar dan respon dari sistem adalah negatif.

- False acceptance (FA)

Sistem dikenai pertanyaan apakah tanda tangan yang diinputkan merupakan milik pemilik yang salah dan respon dari sistem adalah positif.

Dari kedua skenario tersebut pengujian kinerja dari program hasil implementasi penelitian verifikasi citra tanda tangan diukur berdasarkan *False Acceptance Rate* (FAR) yang didefinisikan pada Persamaan 5 dan *False Rejection Rate* (FRR) yang didefinisikan pada Persamaan 6.

$$FAR = \frac{\text{Jumlah FA}}{\text{Jumlah tanda tangan palsu}} \times 100\% \quad (5)$$

$$FRR = \frac{\text{Jumlah FR}}{\text{Jumlah tanda tangan asli}} \times 100 \quad (6)$$

Pengujian dengan FAR dilakukan dengan menginputkan sejumlah citra tanda tangan yang merupakan tanda tangan palsu ke dalam program kemudian untuk setiap citra tanda tangan palsu yang diterima atau diverifikasi sebagai tanda tangan asli akan meningkatkan nilai FAR. Sebaliknya, untuk meningkatkan nilai FRR maka dalam pengujiannya dilakukan dengan menginputkan sejumlah citra tanda tangan yang merupakan tanda tangan asli ke dalam program yang kemudian citra tanda tangan asli diverifikasi sebagai tanda tangan palsu (Widodo dan Harjoko, 2015).

7. HASIL DAN PEMBAHASAN

Pengujian dilakukan terhadap citra tanda tangan yang berasal dari dua sumber data yaitu dari Indonesia dan Persia yang didptkn dri penelitian oleh Widodo & Harjoko (2015). Masing-masing sumber data terdiri dari contoh-contoh tanda tangan oleh 20 penandatangan. Terdapat 24 citra tanda tangan asli dan 10 citra tanda tangan palsu dari sumber data Indonesia serta 20 citra tanda tangan asli dan 10 citra tanda tangan palsu dari sumber dat Persia. Citra yang digunakan sebagai data latih merupakan citra tanda tangan asli dan citra yang digunakan sebagai data latih berupa citra tanda tangan asli dan palsu dari 20 orang tersebut dari setiap sumber data. Data citra disimpan dalam format file yang beragaram yaitu *.jpg*, *.png*, dan *.gif* dengan banyak data yang digunakan sebagai data latih maupun data uji disesuaikan dengan skenario pengujian.

7.1 Pengujian nilai bin, level, dan rentang sudut terhadap akurasi

Pengujian ini bertujuan untuk mencari nilai inialisasi bin, level, dan rentang sudut ekstraksi

ciri PHOG terbaik dengan membandingkan tingkat akurasi. Sama seperti pengujian sebelumnya, data tanda tangan yang digunakan memiliki 3 macam yaitu data tanda tangan yang berasal dari Indonesia dan Persia. Data latih yang digunakan ialah 100 data citra tanda tangan dari setiap negara asal yang terdiri dari 5 tanda tangan asli dari 20 penandatangan. Data uji yang digunakan ialah 200 data citra tanda tangan yang terdiri dari 10 tanda tangan asli selain data latih yang dipilih secara acak oleh program dari 20 penandatangan tersebut. Nilai K yang digunakan pada proses klasifikasi pada pengujian ini bernilai 3.

Tabel 1. Hasil FRR (%) pengujian data tanda tangan Indonesia

Derajat	180				360			
Level Bin	0	1	2	3	0	1	2	3
7	32	12,5	5,5	12,5	30	11	4,5	12,5
8	31,5	9	4,5	11	35	16	7,5	11,5
9	31	11	6	11	32,5	15,5	8	13
10	25	10,5	5	12	31	15,5	6,5	11,5

Banyaknya nilai FRR yang berbeda-beda sebagai hasil dari pengujian untuk nilai bin, level, dan rentang sudut ditunjukkan pada Tabel 1 untuk data tanda tangan Indonesia dan Tabel 2 untuk data tanda tangan Persia. Dari nilai FRR tersebut kemudian didapatkan nilai akurasi untuk dihitung rata-rata pada setiap nilai bin, level, maupun rentang sudut.

Tabel 2. Hasil FRR (%) pengujian data tanda tangan Persia

Derajat	180				360			
Level Bin	0	1	2	3	0	1	2	3
7	25,5	7,5	5,5	7	28,5	7,5	7	7
8	26	8	6,5	8	26	9,5	6	7,5
9	23	6,5	5,5	6,5	25	9	4,5	6,5
10	24,5	7,5	5,5	6,5	28	6	5	7

7.1.1 Hasil pengujian nilai bin dan analisis terhadap akurasi

Variasi nilai bin yang digunakan pada pengujian ini ada 4 yaitu 7, 8, 9, 10 dan nilai bin tersebut diimplementasikan pada pemrosesan data latih dan juga pada pengujian. Setiap nilai

bin dikombinasikan dengan nilai level dan rentang sudut yang berbeda-beda.

Tabel 3. Hasil akurasi (%) pengujian nilai bin pada data tanda tangan Indonesia

Derajat Level Bin	180				360				Rata-rata
	0	1	2	3	0	1	2	3	
7	68	87,5	94,5	87,5	70	89	95,5	87,5	84,938
8	68,5	91	95,5	89	65	84	92,5	88,5	84,25
9	69	89	94	89	67,5	84	92	87	84
10	75	90,5	95	88	69	94,5	93,5	88,5	86,75

Nilai rata-rata akurasi tertinggi pada pengujian data tanda tangan Indonesia didapatkan ketika nilai bin yang digunakan adalah 10 dengan nilai rata-rata akurasi sebesar 86,75%. Untuk pengujian data tanda tangan Persia, nilai rata-rata akurasi tertinggi didapatkan ketika nilai bin yang digunakan adalah 9 dengan nilai rata-rata akurasi sebesar 89,188%.

Tabel 4. Hasil akurasi (%) pengujian nilai bin pada data tanda tangan Persia

Derajat Level Bin	180				360				Rata-rata
	0	1	2	3	0	1	2	3	
7	74,5	92,5	94,5	93,5	71,5	92,5	93,5	93,5	88,063
8	74	92,5	93,5	92,5	74,5	90,5	94,5	92,5	87,813
9	77,5	93,5	94,5	93,5	75,5	91,5	95,5	93,5	89,188
10	75,5	92,5	94,5	93,5	72,5	94,5	95,5	93,5	88,75

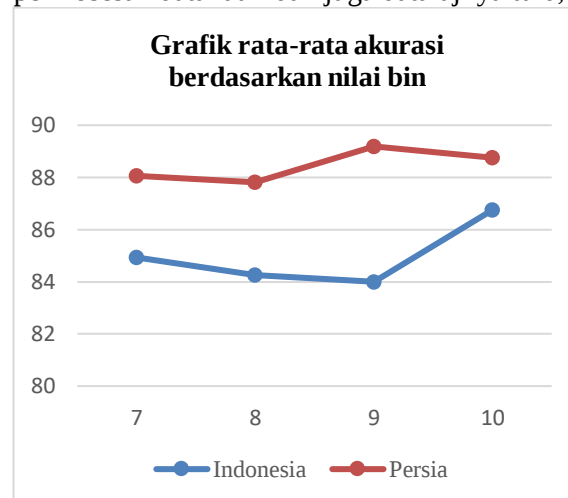
Rata-rata nilai akurasi yang didapatkan dari nilai bin yang berbeda seperti yang ditunjukkan pada Tabel 3 dan Tabel 4 direpresentasikan pada grafik pada Gambar 3. Dari grafik tersebut terlihat bahwa peningkatan nilai rata-rata akurasi tidak terjadi dengan pola yang sama pada setiap data tanda tangan. Nilai rata-rata akurasi tertinggi pada masing-masing data juga tidak diperoleh dari nilai bin yang sama dengan nilai rata-rata akurasi tertinggi secara keseluruhan didapatkan dari pengujian dengan menggunakan data tanda tangan Persia.

Nilai bin yang semakin besar mempengaruhi distribusi dari histogram menjadi semakin spesifik sehingga pada awalnya diperkirakan peningkatan nilai bin dapat meningkatkan akurasi. Pengujian ini menunjukkan bahwa peningkatan nilai bin

ternyata tidak selalu meningkatkan akurasi walaupun dari dua data yang diuji didapatkan nilai bin dengan akurasi terbaik dengan menggunakan nilai bin yang tertinggi. Hal ini dimungkinkan terjadi karena faktor pendukung lainnya seperti nilai level dan nilai rentang sudut yang digunakan cukup beragam.

7.1.2 Hasil pengujian nilai level dan analisis terhadap akurasi

Pada pengujian ini terdapat 4 variasi nilai level yang diimplementasikan dalam pemrosesan data latih dan juga data uji yaitu 0,



Gambar 3 Grafik rata-rata akurasi berdasarkan nilai bin

1, 2, dan 3. Level dengan nilai 0 merupakan level di mana citra belum disegmentasi, level dengan nilai 1 merupakan nilai terendah sebagai tingkatan piramida segmentasi citra, sedangkan nilai 3 merupakan nilai tertinggi sebagai tingkatan piramida segmentasi citra. Tingkatan piramida pada pengujian ini hanya mencapai level 3 karena data tanda tangan yang diujikan dilakukan normalisasi ukuran dengan ukuran citra keluaran adalah sebesar 200 x 200 piksel.

Tabel 5. Hasil akurasi (%) pengujian nilai level pada data tanda tangan Indonesia

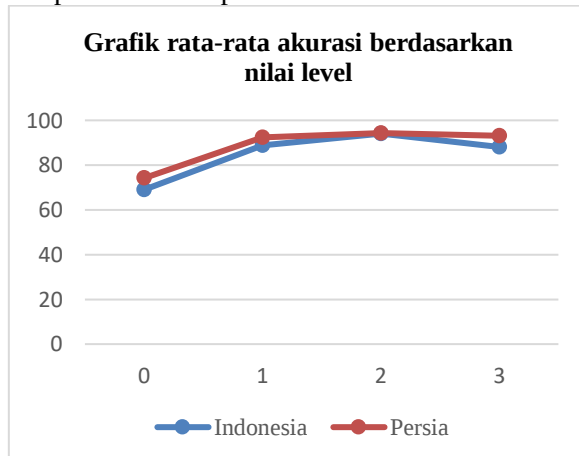
Derajat Level Bin	180				360			
	0	1	2	3	0	1	2	3
7	68	87,5	94,5	87,5	70	89	95,5	89,5
8	68,5	91	95,5	89	65	84	92,5	89
9	69	89	94	89	67,5	84,5	92	87
10	75	90,5	95	88	69	94,5	93,5	87
Rata-rata	Level = 0: 69	Level = 1: 88,75	Level = 2: 94,0625	Level = 3: 88,125				

Tabel 6. Hasil akurasi (%) pengujian nilai level pada data tanda tangan Persia

Derajat	180				360			
Level	0	1	2	3	0	1	2	3
1								
Bin								
7	74, 5	92, 5	94, 5	93	71, 5	92, 5	93	93
8	74	92	93, 5	92	74	90, 5	94	92, 5
9	77	93, 5	94, 5	93, 5	75	91	95, 5	93, 5
10	75, 5	92, 5	94, 5	93, 5	72	94	95	93
Rata-rata	Level = 0: 74,188		Level = 1: 92,3125		Level = 2: 94,3125		Level = 3: 93	

Nilai rata-rata akurasi tertinggi pada pengujian data tanda tangan Indonesia didapatkan ketika nilai level yang digunakan adalah 2 dengan nilai rata-rata akurasi sebesar 94,0625%. Untuk pengujian data tanda tangan Persia, nilai rata-rata akurasi tertinggi didapatkan ketika nilai level yang digunakan adalah 2 dengan nilai rata-rata akurasi sebesar 94,3125%.

Rata-rata nilai akurasi yang didapatkan dari nilai level yang berbeda seperti yang ditunjukkan pada Tabel 5 dan Tabel 6 direpresentasikan pada



Gambar 4. Grafik rata-rata akurasi berdasarkan nilai level

grafik pada Gambar 4. Dari grafik tersebut terlihat bahwa terjadi peningkatan nilai rata-rata akurasi hanya terjadi pada satu nilai level. Peningkatan nilai rata-rata akurasi yang cukup signifikan hanya terjadi ketika level bernilai 2. Sehingga dapat disimpulkan bahwa semakin banyak level atau tingkatan yang digunakan untuk segmentasi tidak selalu mempengaruhi besarnya nilai rata-rata akurasi. Pada percobaan ini didapatkan nilai level terbaik dengan nilai rata-rata akurasi tertinggi adalah 2.

7.1.3 Hasil pengujian nilai rentang sudut dan analisis terhadap akurasi

Rentang sudut yang digunakan pada pengujian ini adalah 180 dan 360 dengan tujuan membandingkan nilai akurasi terkait pembagian rentang sel histogram. Pada data tanda tangan Indonesia sel histogram yang terbagi dalam sejumlah bin dalam rentang 0 hingga 180 memiliki nilai rata-rata akurasi 90,875% seperti yang ditunjukkan pada Tabel 7. Pada pengujian data tanda tangan Persia dengan sel histogram yang terbagi dalam sejumlah bin dalam rentang 0 hingga 180 memiliki nilai rata-rata akurasi tertinggi dengan nilai 88,78% seperti yang ditunjukkan pada Tabel 8. Dapat disimpulkan dari pengujian ini bahwa rentang sudut terbaik untuk pembagian rentang sel histogram adalah 180.

Tabel 7. Hasil akurasi (%) pengujian nilai rentang sudut pada data tanda tangan Indonesia

Derajat	180				360			
Level	0	1	2	3	0	1	2	3
Bin								
7	68	87, 5	94, 5	87, 5	70	89	95, 5	87, 5
8	68, 5	91	95, 5	89	65	84	92, 5	88, 5
9	69	89	94	89	67, 5	84, 5	92	87
10	75	90, 5	95	88	69	94, 5	93, 5	88, 5
Rata-rata	85,668				84,281			

Tabel 8. Hasil akurasi (%) pengujian nilai rentang sudut pada data tanda tangan Persia

Derajat	180				360			
Level	0	1	2	3	0	1	2	3
1								
Bin								
7	74, 5	92, 5	94, 5	93	71, 5	92, 5	93	93
8	74	92	93, 5	92	74	90, 5	94	92, 5
9	77	93, 5	94, 5	93, 5	75	91	95, 5	93, 5
10	75, 5	92, 5	94, 5	93, 5	72	94	95	93
Rata-rata	88,78				88,125			

7.2 Pengujian Nilai K

Pengujian ini bertujuan untuk mencari nilai K terbaik yang memiliki nilai akurasi paling tinggi dengan nilai K yang diuji antara lain 1, 3, 5, dan 7. Pengujian ini dilakukan dengan menggunakan perhitungan jarak *Euclidean*,

dengan nilai bin, level, dan rentang sudut terbaik pada pengujian sebelumnya serta jumlah data latih yang sama dengan pengujian sebelumnya.

Pada pengujian sebelumnya untuk data tanda tangan Indonesia didapatkan nilai bin terbaik adalah 10, nilai level terbaik adalah 2, dan nilai rentang sudut terbaik adalah 180 dengan nilai K yang digunakan adalah 3. Hasil pengujian nilai K pada data tanda tangan Indonesia ditunjukkan pada Tabel 9 menunjukkan bahwa nilai K terbaik adalah 1 dengan akurasi sebesar 98%. Perbandingan hasil klasifikasi dari pengujian sebelumnya yang menggunakan nilai K sebesar 3 dengan penggunaan nilai K terbaik yaitu 1 tampak

Tabel 9. Hasil pengujian nilai K pada data tanda tangan Indonesia

K	FRR	Akurasi
1	2	98
3	5	95
5	12,5	87,5
7	23,5	76,5

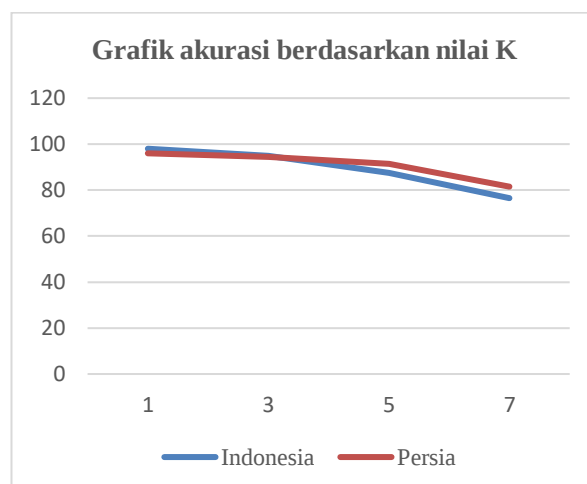
bahwa perbaikan nilai K dapat mengurangi kesalahan penolakan klaim.

Pada pengujian sebelumnya untuk data tanda tangan Persia didapatkan nilai bin terbaik adalah 10, nilai level terbaik adalah 2, dan nilai rentang sudut terbaik adalah 180 dengan nilai K yang digunakan adalah 3. Hasil pengujian nilai K pada data tanda tangan Indonesia ditunjukkan pada Tabel 10 menunjukkan bahwa nilai K terbaik adalah 1 dengan akurasi sebesar 96%. Perbandingan hasil klasifikasi dari pengujian sebelumnya yang menggunakan nilai K sebesar 3 dengan penggunaan nilai K terbaik yaitu 1.

Tabel 10. Hasil pengujian nilai K pada data tanda tangan Persia

K	FRR	Akurasi
1	4	96
3	5,5	94,5
5	9	91,5
7	17,5	81,5

Hasil pengujian nilai K yang ditunjukkan pada Tabel 9 dan Tabel 10 direpresentasikan ke dalam bentuk grafik pada Gambar 5. Dari grafik tersebut dapat dilihat bahwa semakin besar nilai K maka semakin kecil akurasi yang didapatkan. Hal ini dapat terjadi karena semakin besar nilai K maka semakin banyak kemungkinan variasi kelas yang diambil sebagai tetangga terdekat.



Gambar 5. Grafik akurasi berdasarkan nilai K

7.3 Pengujian jumlah data latih

Pengujian ini bertujuan untuk menemukan jumlah data latih yang memiliki akurasi yang tertinggi. Terdapat 3 jumlah data latih yang diuji yaitu 5, 10, dan 14 untuk data tanda tangan Indonesia dan 2 jumlah data latih yang diuji untuk data tanda tangan Persia yaitu 5 dan 10. Pengujian terhadap jumlah data latih dilakukan dengan menggunakan nilai bin, level, dan rentang sudut terbaik untuk masing-masing data dan nilai K terbaik yaitu 1 sesuai dengan pengujian sebelumnya.

Tabel 11. Hasil pengujian jumlah data latih pada data tanda tangan Indonesia

Jumlah data latih	FRR	Akurasi
5	2	98
10	1,5	98,5
14	0,5	99,5

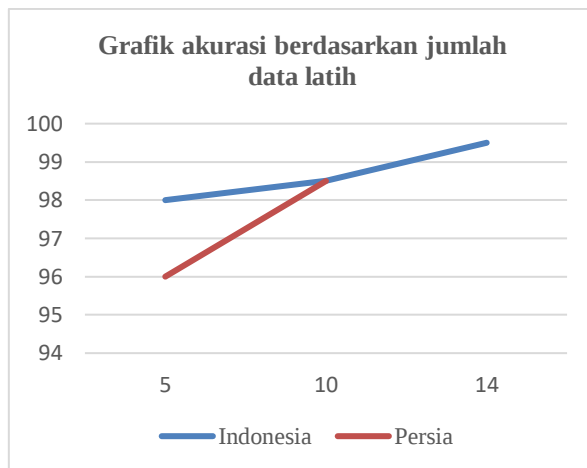
Tabel 12. Hasil pengujian jumlah data latih pada data tanda tangan Persia

Jumlah data latih	FRR	Akurasi
5	4	96
10	1,5	98,5

Tabel 11 menunjukkan bahwa dari 3 macam jumlah data latih pada pengujian untuk data tanda tangan Indonesia didapatkan akurasi mencapai 99,5% dengan menggunakan data latih terbanyak yaitu 14. Tabel 12 menunjukkan bahwa dari 2 macam jumlah data latih pada pengujian untuk data tanda tangan Persia didapatkan akurasi mencapai 98,5% dengan menggunakan data latih terbanyak yaitu 10. Terdapat beberapa tanda tangan yang masih salah dikenali sehingga mengalami kesalahan

penolakan oleh program walaupun nilai K yang digunakan sudah yang terbaik dan pengetahuan (data latih) yang dimiliki sistem cukup banyak. Kasus seperti ini dapat terjadi karena bentuk tanda tangan yang memang mirip antara kelas yang satu dengan lainnya terutama setelah citra tanda tangan dikenai *preprocessing* citra seperti rotasi pada koreksi kemiringan, *thinning*, dan normalisasi ukuran citra.

Secara keseluruhan hasil pengujian jumlah data latih direpresentasikan ke dalam grafik pada Gambar 6. Grafik tersebut menunjukkan bahwa semakin banyak data latih yang digunakan mempengaruhi akurasi menjadi semakin baik. Hal ini dikarenakan semakin banyak jumlah data latih maka variasi bentuk tanda tangan untuk setiap kelasnya juga bermacam-macam sehingga hasil klasifikasi dapat lebih baik.



Gambar 6. Grafik akurasi berdasarkan jumlah data latih

7.4 Pengujian dengan menggunakan data tanda tangan palsu

Pengujian ini bertujuan untuk mengetahui akurasi dalam membedakan citra tanda tangan yang asli dengan yang palsu. Data yang diuji berjumlah 200 yang terdiri dari 10 tanda tangan palsu dari 20 orang untuk setiap data asalnya yaitu Indonesia dan Persia. Pengujian menggunakan kelas yang sama dengan pengujian-pengujian sebelumnya. Data latih yang digunakan yaitu 14 untuk setiap kelasnya pada data tanda tangan Indonesia dan 10 untuk setiap kelas pada data tanda tangan Persia dengan nilai K, bin, level, dan rentang sudut terbaik sesuai dengan hasil pengujian sebelumnya. Hasil pengujian menggunakan data palsu ditunjukkan pada Tabel 13 dengan akurasi untuk data tanda tangan Indonesia mencapai

56% dan untuk data tanda tangan Persia sebesar 35,5%.

Tabel 13. Hasil pengujian menggunakan data tanda tangan palsu

Data	FAR	Akurasi
Indonesia	44	56
Persia	64,5	35,5

Sulitnya program mengenali tanda tangan palsu dikarenakan miripnya tanda tangan palsu yang diuji dengan data tanda tangan asli. Seperti data tanda tangan Persia yang memiliki data palsu yang sangat menyerupai data aslinya sehingga kesalahan penerimaan klaim sangat tinggi.



Gambar 7. Perbandingan tanda tangan asli dan palsu pada data tanda tangan Persia dengan data tanda tangan Indonesia

Perbandingan kemiripan tanda tangan asli dan palsu pada dua sumber data ditunjukkan pada Gambar 7. Selain itu untuk data palsu jumlah data latih yang terlalu banyak akan mengurangi akurasi karena semakin banyak variasi bentuk tanda tangan dalam satu kelas yang dapat diterima sebagai tanda tangan asli.

8. KESIMPULAN

Ciri PHOG sangat mempengaruhi proses verifikasi citra tanda tangan di mana histogram-histogram yang digabungkan dari perhitungan pada area global hingga ke area lokal citra

memberikan kinerja yang lebih baik dibandingkan dengan histogram yang dihasilkan dari perhitungan pada area global saja. Level atau tingkatan segmentasi citra sangat mempengaruhi akurasi di mana kenaikan akurasi tertinggi didapatkan ketika level bernilai sebanyak 2.

Pada pengujian data tanda tangan asli metode K-NN bekerja dengan baik sebagai metode klasifikasi pada proses verifikasi citra dengan nilai K terbaik sebesar 1 dan penurunan akurasi didapatkan ketika nilai K diperbesar. Jumlah data latih juga memberikan pengaruh terhadap nilai akurasi di mana semakin banyak jumlah data latih perorang semakin meningkatkan nilai akurasi. Dengan menggunakan parameter terbaik dari hasil pengujian rata-rata akurasi tertinggi untuk data tanda tangan Indonesia asli mencapai 99,5% dan untuk data tanda tangan Persia asli sebesar 98,5%.

Sebaliknya, pengujian pada data tanda tangan palsu menghasilkan akurasi yang cukup kecil yang berarti walaupun algoritme yang diimplementasi telah berhasil melakukan pengenalan dengan baik, tetapi masih belum cukup baik membedakan antara yang asli dengan yang palsu. Penggunaan nilai K yang kecil dan jumlah data latih yang banyak justru memperbanyak variasi bentuk tanda tangan dalam satu kelas yang dapat diterima sebagai tanda tangan asli. Pengujian dengan menggunakan parameter yang terbaik yang telah didapatkan terhadap data tanda tangan Indonesia palsu hanya menghasilkan akurasi sebesar 56% dan untuk data tanda tangan Persia asli sebesar 35,5%.

9. DAFTAR PUSTAKA

- Aouidate, A. M. & Riadh, A. B. A. A. 2011. Survey of Nearest Neighbor Condensing Techniques. *International Journal of Advanced Computer Science and Application*, 2(11), 59-64.
- Bosch, A., Zisserman, A. & Munoz, X. 2007. Representing Shape with A Spatial Pyramid Kernel. Amsterdam. *Proceedings of the 6th ACM International conference on Image and Video Retrieval*, 401-408.
- Dalal, N. & Triggs, B. 2005. Histogram of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 886-893.
- Dasgupta, J., Bhattacharya, K. & Chandra, B. 2016. A holistic approach for Off-line handwritten cursive word recognition using directional feature based on Arnold transform. *Pattern Recognition Letters*.
- Kobayashi, T., Hidaka, A. & Kurita, T. 2008. Selection of Histograms of Oriented Gradients Features for Pedestrian Detection. In *ICONIP 2007. Berlin Heidelberg, 2008*. Springer-Verlag.
- Medjahed, S. A., Saadi, T. A. & Benyettou, A., 2013. Breast Cancer Diagnosis by using K-Nearest Neighbor with Different Distances and Classification Rules. *International Journal of Computer Applications* (0975 - 8887), 62(1), pp. 1-5.
- Priambodo, A., Dewi, C. & Triwiratno, A. 2015. Implementasi Metode K-Nearest Neighbor untuk Identifikasi Penyakit Tanaman Jeruk Keprok berdasarkan Citra Daun. *Thesis*. Universitas Brawijaya.
- Saidani, A. & Kacem, A. 2014. Pyramid Histogram of Oriented Gradient for Machine-printed/Handwritten and Arabic/Latin word discrimination. *International Conference of Soft Computing and Pattern Recognition*, 267-272.
- Shu, C., Ding, X. & Fang, C. 2011. Histogram of the Oriented Gradient for Face Recognition. *Tsinghua Science and Technology*, 16(2), 216-224.
- Widodo, A. W. & Harjoko, A. 2015. Sistem Verifikasi Tanda Tangan Off-line berdasar Ciri Histogram of Oriented Gradient (HOG) dan Histogram of Curvature (HoC). *Jurnal Teknologi Informasi dan Ilmu Komputer*, 2(1), 1-10.
- Zhang, B. 2010. Off-line signature verification and identification by pyramid histogram of oriented gradients. *International Journal of Intelligent Computing and Cybernetics*, 3(4), 611 – 630.