

Machine Learning untuk Identifikasi Tanda Tangan Menggunakan GLCM dan *Euclidean Distance*

Gresiva Devi Angel¹, Resty Wulanningrum²

^{1,2}Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri

E-mail: *¹angelcharles46@gmail.com, ²resty0601@gmail.com

Abstrak – Tanda tangan (*signature*) merupakan salah satu bukti identitas seseorang dalam melakukan pengesahan sebuah dokumen atau presensi yang sering digunakan. Tanda tangan juga menjadi fitur pembeda untuk identifikasi seseorang. Keberadaan tanda tangan dalam sebuah dokumen menyatakan bahwa pihak yang menandatangani, mengetahui dan menyetujui seluruh isi dokumen. Mengenal bentuk tanda tangan seseorang sangatlah penting dalam melakukan pencocokan tanda tangan terhadap presensi atau dokumen untuk mengetahui apakah benar yang bertanda tangan adalah orang yang bersangkutan atau tidak. Oleh karena itu, dibuat sebuah sistem identifikasi tanda tangan menggunakan Gray Level Co-Occurrence Matrix (GLCM) dan *Euclidean Distance*. Pengujian menunjukkan hasil terbaik pada skenario 1 dengan tingkat akurasi 67,5%, skenario 2 dengan akurasi 67%, dan yang paling rendah skenario 3 dengan akurasi 57,5%. Penggunaan dimensi berbeda antara citra training dan citra testing menghasilkan akurasi tertinggi pada skenario 1 sebesar 65%, skenario 2 dengan akurasi 14%, dan yang paling rendah pada skenario 3 dengan akurasi hanya 10%.

Kata Kunci — *Euclidean Distance*, GLCM, Tanda Tangan

1. PENDAHULUAN

Sistem biometrika merupakan teknologi pengenalan diri dengan menggunakan bagian tubuh atau perilaku manusia. Tanda tangan (*signature*) merupakan salah satu contoh biometrika berdasarkan tingkah laku manusia [1].

Sistem pengenalan diri adalah sistem untuk mengenali identitas seseorang secara otomatis dengan menggunakan teknologi komputer. Contohnya, sistem pengenalan tulisan menggunakan tanda tangan yang berfungsi sebagai pembuktian [2].

Tanda tangan telah menjadi fitur pembeda untuk identifikasi seseorang [3]. Tanda tangan banyak digunakan pada berbagai macam dokumen kegiatan yang memerlukan tanda tangan sebagai identifikasi dan kebenaran dokumen. Salah satu contoh kegiatan yang memerlukan tanda tangan pada dokumen adalah kegiatan administrasi [4]. Keberadaan tanda tangan dalam sebuah dokumen menyatakan bahwa pihak yang menandatangani, mengetahui dan menyetujui seluruh isi dokumen [5]. Tidak ada tanda tangan asli yang sama persis. Sehingga biometrika tanda tangan masih menjadi topik riset yang menantang [6].

Penelitian pertama dari Ismi Amalia tahun 2014 [3] dengan judul “Pengenalan citra tanda tangan menggunakan GLCM dan PNN” memiliki hasil fitur yang diperoleh adalah perbedaan *dissimilarity*, *entropy* dan *homogeneity*. Pengenalan dan verifikasi dilakukan menggunakan *probabilistic neural network*. Algoritma ini telah diuji pada 100 tanda tangan. Penelitian ini menunjukkan akurasi

rata-rata PNN dalam melakukan pengenalan tanda tangan adalah sebesar 71%.

Penelitian kedua dari Resty Wulanningrum tahun 2014 [5] dengan judul “Implementasi principal component analysis untuk identifikasi citra tanda tangan” memiliki hasil dari penggunaan partisi citra dan nilai *threshold* akan menghasilkan hasil identifikasi citra tanda tangan yang bervariasi untuk proses analisa yang akan digunakan. Berdasarkan hasil 4 skenario, didapatkan hasil terbaik pada skenario 2 dan 4 dengan nilai *threshold* sebesar 0-255 menggunakan reduksi dimensi 20 yaitu sebesar 98%.

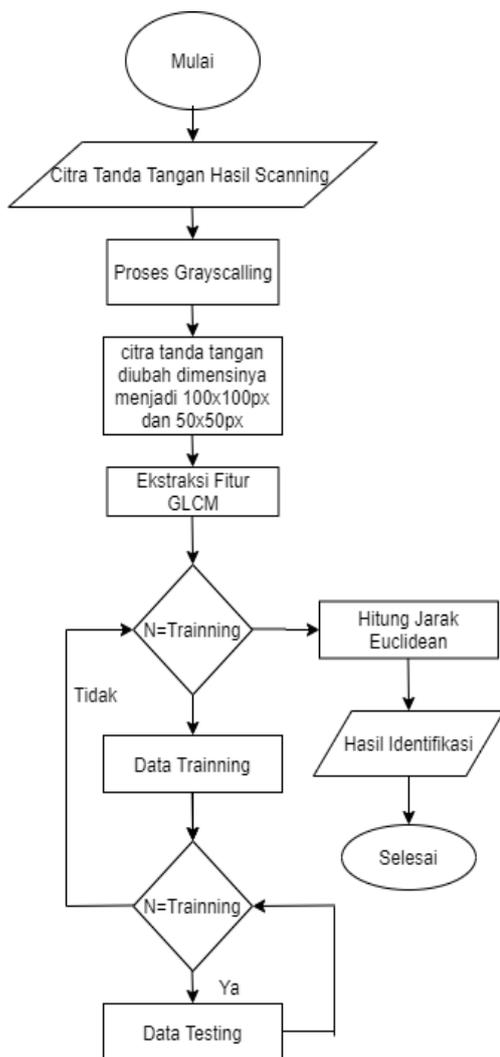
Penelitian ketiga dari Roni Salambue tahun 2013 [7] dengan judul “Pengenalan pola tanda tangan dengan metode *moment invariant* dan *euclidean distance*”. Pada penelitian ini vektor hasil *moment invariant* menjadi data input dalam proses pengenalan citra digital. Metode pengenalan menggunakan *Euclidean Distance* yang mengukur selisih antara vektor. Ukuran kemiripan citra digital berdasarkan pada selisih minimum yang terjadi antara citra pengujian dengan citra yang tersimpan dalam basis data.

Identifikasi tanda tangan seseorang pada umumnya masih dilakukan secara manual yaitu dengan mencocokkan tanda tangan pada waktu transaksi dengan tanda tangan asli. Contohnya dalam proses pembelajaran di kampus, dimana mahasiswa sering menitipkan tanda tangan untuk presensi yang membuat para dosen harus lebih teliti dalam membedakan apakah tanda tangan tersebut sudah sesuai dengan tanda tangan mahasiswa tersebut atau

tidak. Oleh karena itu, pencocokan karakteristik tanda tangan seseorang berdasarkan penggunaan dimensi yang berbeda yaitu untuk mengidentifikasi dan mengukur kemiripan tanda tangan seseorang menggunakan *Euclidean Distance* untuk pengenalan tanda tangan dan *Gray Level Co-Occurrence Matrix* (GLCM) untuk proses ekstraksi fitur.

2. METODE PENELITIAN

Tahapan yang dilakukan dalam pencocokan citra tanda tangan berdasarkan ekstraksi fitur menggunakan GLCM dan *Euclidean distance* ditunjukkan pada Gambar 1.



Gambar 1. Flowchart Sistem

2.1 Dataset

Data yang digunakan adalah citra tanda tangan dari sepuluh mahasiswa. Setiap mahasiswa mempunyai 10 citra tanda tangan, sehingga total citra yang digunakan dalam penelitian ini adalah 100 citra tanda tangan. Hasil tanda tangan sebanyak 100, dari 10 mahasiswa menunjukkan bahwa seseorang tidak bisa melakukan tanda tangan yang sama persis.

Data penelitian yang digunakan diambil dengan cara dikumpulkan dalam satu kertas yang kemudian dilakukan proses digitalisasi dengan menggunakan *scanner*. Contoh citra tanda tangan mahasiswa yang digunakan ditunjukkan pada Gambar 2.



Gambar 2. Citra tanda tangan yang digunakan

2.2 Pra-pemrosesan

Sebelum masuk ke tahap ekstraksi fitur terlebih dahulu dilakukan proses *cropping* pada citra tanda tangan hasil *scan* seperti pada Gambar 2 dengan menggunakan *snipping tool*. Contoh hasil pemotongan ditunjukkan pada Gambar 3.

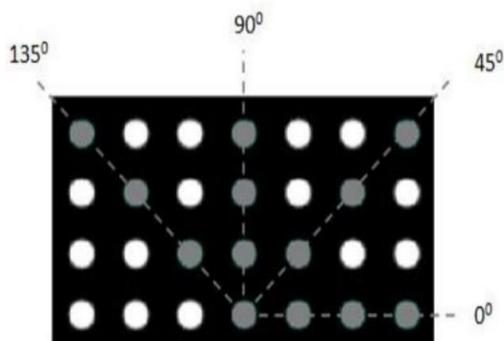


Gambar 3. Contoh hasil *cropping*

2.3 Ekstraksi Fitur dengan GLCM

Gray Level Co-Occurrence (GLCM) pertama kali diusulkan oleh Haralick pada tahun 1973 dengan fitur untuk menjelaskan pola spasial[8]. GLCM menggunakan perhitungan tekstur pada orde kedua. Pengukuran tekstur pada orde pertama menggunakan perhitungan statistik didasarkan pada nilai piksel. Pada orde kedua, hubungan antarpasangan dua piksel citra asli diperhitungkan. Orientasi dibentuk dalam empat arah sudut dan/atau jarak. Sebagai contoh, gambar berikut memperlihatkan empat arah untuk GLCM[9]. *Gray Level Co-Occurrence Matrix* (GLCM) juga

merupakan matrik yang menggambarkan frekuensi munculnya pasangan piksel pada jarak d dan orientasi arah dengan sudut θ dalam citra yang digunakan untuk menghitung fitur-fitur GLCM. Jarak d yang digunakan adalah 1 yang dinyatakan dalam piksel, sementara untuk orientasi sudut dinyatakan dalam derajat dengan sudut 0° , 45° , 90° , dan 135° [10].



Gambar 4. Contoh arah GLCM

Fitur tekstur yang diekstraksi dengan GLCM dalam penelitian ini adalah *contrast*, *correlation*, *energy*, dan *homogeneity*, seperti ditunjukkan pada rumus (1) sampai (4) berikut :

$$\text{Contrast} \sum_{n=1}^L n^2 \{ \sum_{|i-j|=n} \text{GLCM}(i, j) \} \dots\dots\dots(1)$$

$$\text{Correlation} \sum_{n=1}^L \sum_{j=1}^L (ij) (\text{GLCM}(i, j) - \mu_i \mu_j) \dots\dots\dots(2)$$

$$\text{Energy} \sum_{n=1}^L \sum_{j=1}^L (\text{GLCM}(i, j) \log(\text{GLCM}(i, j) \dots\dots\dots(3)$$

$$\text{Homogeneity} \sum_i \sum_j \frac{1}{1+(i-j)^2} P(i, j) \dots\dots\dots(4)$$

Setiap nilai dari *contrast*, *correlation*, *energy*, dan *homogeneity* akan dihitung dan hasil perhitungannya dijadikan sebuah informasi fitur untuk proses identifikasi dengan *Euclidean distance*.

2.4 Klasifikasi dengan *Euclidean Distance*

Euclidean Distance adalah metode yang menghitung akar dari kuadrat perbedaan 2 vektor (*root of square difference between 2 vector*)[9] atau juga dikenal sebagai metode klasifikasi tetangga terdekatnya dengan menghitung jarak antara dua buah obyek[5]. Rumus perhitungan *euclidean* ditunjukkan pada (1):

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \dots\dots\dots(5)$$

Dimana :

d_{ij} = Jarak objek antara objek i dan j
 n = Dimensi Data

x_{ik} = Koordinat dari obyek i pada dimensi k
 x_{jk} = Koordinat dari obyek j pada dimensi k

2.5 Perangkat Keras dan Perangkat Lunak

Perangkat keras yang digunakan dalam penelitian ini adalah *Processor Intel Celeron N2830* memori RAM 2 GB DDR3 dan *harddisk* 500 GB. Perangkat lunak yang digunakan adalah Sistem Operasi Windows 8 dan Matlab R2013a.

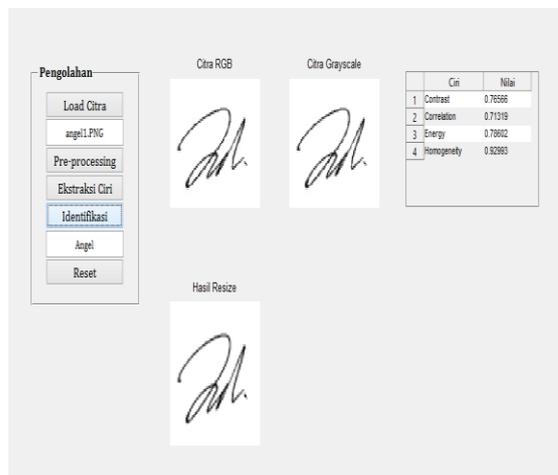
2.6 Akurasi

Evaluasi penentuan hasil identifikasi atau pencocokan citra tanda tangan menggunakan GLCM dan *Euclidean Distance* ditentukan dengan akurasi yang dicapai dalam mengidentifikasi data *testing* setelah dilakukan *training*. Tingkat akurasi diperoleh dengan perhitungan :

$$\text{Akurasi} = \frac{\text{data testing benar}}{\text{total data testing}} \times 100\% \dots\dots\dots(6)$$

2.7 Desain *Interface*

Pada penelitian ini memiliki desain *interface* yang ditunjukkan pada Gambar 5. Dimana terlihat bahwa citra tanda tangan angel dikenali dengan angel.



Gambar 5. Form desain *interface*

3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan metode *Gray Level Co-Occurrence Matrix* untuk proses ekstraksi fitur citra tanda tangan. Citra hasil dari ekstraksi fitur yang diperoleh digunakan sebagai input untuk pencocokan tanda tangan dengan menggunakan metode *Euclidean Distance*.

Gray Level Co-Occurrence Matrix (GLCM) digunakan untuk mengekstraksi fitur dari sebuah citra. 4 fitur yang digunakan *contrast*, *correlation*, *energy*, dan *homogeneity*. Dilihat dari Gambar 5 bahwa pada setiap fitur masing-masing citra tanda tangan menunjukkan nilai-nilai yang berbeda. Oleh

karena itu, pencocokan tanda tangan dengan metode ini sangat baik.

Uji coba dilakukan untuk mengidentifikasi citra tanda tangan ke dalam 10 kelas, yaitu tanda tangan dari 10 orang yang berbeda. Hasil ekstraksi fitur-fitur *contrast*, *correlation*, *energy*, dan *homogeneity* digunakan untuk identifikasi dengan *euclidean distance*.

No	Citra Tanda Tangan	Contrast	Correlation	Energy	Homogeneity
1		0.76566	0.71319	0.78602	0.92993
2		1.7752	0.51384	0.72397	0.90111
3		0.40909	0.70758	0.8692	0.95906
4		0.83091	0.66154	0.66969	0.89511
5		1.1677	0.58772	0.62282	0.87207
6		0.4001	0.81514	0.74853	0.93157
7		0.60202	0.65175	0.72041	0.91598
8		1.0174	0.76686	0.6483	0.88439
9		0.97859	0.68839	0.77744	0.92577
10		1.0531	0.6942	0.76508	0.92115

Gambar 6. Hasil ekstraksi fitur citra tanda tangan

Berdasarkan hasil uji coba implementasi identifikasi tanda tangan menggunakan GLCM dan *Euclidean distance* dengan total 100 citra tanda tangan dari 10 mahasiswa masing-masing mahasiswa mempunyai 10 citra tanda tangan dengan ukuran 100 piksel x 100 piksel yang akan digunakan sebagai data *training* dan data *testing*, maka hasil yang didapatkan dari skenario 1-3 sebagai berikut :

Tabel 1. Skenario 1-3 dengan ukuran 100 piksel x 100 piksel

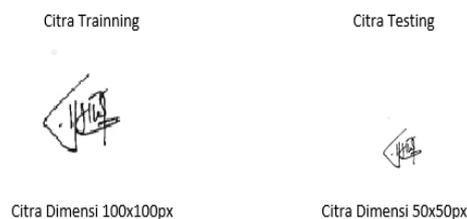
Uji Coba ke-	Data Training	Data Testing	Akurasi
Skenario 1	60	40	67,5%
Skenario 2	40	60	67%
Skenario 3	20	80	57,5%

Pada skenario 1 dalam pengujian sistem identifikasi citra tanda tangan, uji coba yang dilakukan pada data *training* sebanyak 60 data citra dengan ukuran 100 piksel x 100 piksel dan data *testing* sebanyak 40 data diluar data *training* didapatkan hasil akurasi yaitu sebesar 67,5%. Tebal dan tipisnya sebuah tanda tangan sangat mempengaruhi tingkat akurasi.

Pada skenario 2 dalam pengujian sistem identifikasi citra tanda tangan, uji coba yang dilakukan pada data *training* sebanyak 40 data citra dan yang berukuran 100 piksel x 100 piksel dengan data *testing* 60 data diluar data *training*. Hasil akurasi yang didapatkan pada skenario 2 sebesar 67%.

Pada skenario uji coba 3, menggunakan data *training* sebanyak 20 data citra yang berukuran 100 piksel x 100 piksel dengan data *testing* sebanyak 80 data citra diluar data *training*. Hasil akurasi yang didapatkan pada skenario 3 yaitu sebesar 57%. Nilai akurasi yang didapatkan lebih rendah dibandingkan pada skenario 1 dan 2, hal ini disebabkan data *testing* yang digunakan lebih besar jumlahnya dibandingkan dengan data *training*, sehingga mempengaruhi hasil akurasi.

Selanjutnya pada uji coba kedua dilakukan perubahan ukuran dimensi pada citra tanda tangan dengan ukuran 50 piksel x 50 piksel, maka hasil akurasi dari skenario 1-3 sebagai berikut :



Gambar 7. Perbedaan citra *training* dan citra *testing*

Dari hasil uji coba citra tanda tangan data *training* dengan ukuran 100 piksel x 100 piksel dan data *testing* dengan ukuran 50 piksel x 50 piksel dapat dilihat pada tabel 2, hasil akurasi tertinggi yang didapatkan dari 3 skenario sebesar 65%.

Tabel 2. Skenario 1-3 dengan ukuran 50 piksel x 50 piksel

Uji Coba ke-	Data Training	Data Testing	Akurasi
Skenario 1	60	40	65%
Skenario 2	40	60	14%
Skenario 3	20	80	10%

Dalam pengujian sistem identifikasi citra tanda tangan menggunakan skenario 1 ujicoba dilakukan pada data *training* sebanyak 60 data citra dengan ukuran 100 piksel x 100 piksel, dan data *testing* sebanyak 40 data diluar data *training* dengan

ukuran 50 piksel x 50 piksel, maka hasil akurasi yang didapatkan adalah 65%. Namun ada tanda tangan yang tetap dapat dikenali oleh sistem ada juga yang dikenali sebagai tanda tangan orang lain. Hal ini disebabkan oleh ukuran pixel yang diubah menjadi pecah.

Pada skenario 2 menggunakan data *training* sebanyak 40 data citra dengan ukuran 100 piksel x 100 piksel, data *testing* sebanyak 60 data diluar data *training* dengan ukuran 50 piksel x 50 piksel, dan hasil akurasi yang diperoleh sebesar 14%.

Dan pada skenario 3 menggunakan data *training* sebanyak 20 data citra dengan ukuran 100 piksel x 100 piksel, data data *testing* sebanyak 80 citra diluar data *training* dengan ukuran 50 piksel x 50 piksel dan akurasi yang didapatkan hanya 10% saja.

Pada skenario 2 dan 3 akurasi yang diperoleh lebih rendah daripada skenario 1. Hal ini dikarenakan data *testing* yang digunakan lebih besar dari data *training*.

4. SIMPULAN

Identifikasi citra tanda tangan berdasarkan hasil 3 skenario menggunakan metode ekstraksi fitur *Gray Level Co-Occurrence Matrix* (GLCM) dan *Euclidean Distance* untuk menghitung jarak kedekatan citra tanda tangan, didapatkan hasil terbaik pada skenario 1 dengan tingkat akurasi 67,5%, skenario 2 dengan akurasi 67%, dan yang paling rendah skenario 3 dengan akurasi 57,5%. Penggunaan dimensi yang berbeda antara citra *training* dan citra *testing* mempengaruhi hasil akurasi dari identifikasi tanda tangan. Penggunaan dimensi ini menghasilkan akurasi tertinggi pada skenario 1 sebesar 65%, skenario 2 dengan akurasi 14%, dan yang paling rendah pada skenario 3 dengan akurasi hanya 10%.

5. SARAN

Berdasarkan uji coba pada penelitian yang dilakukan, berikut saran-saran yang diberikan untuk penelitian lanjutan. 1. Menambahkan data tanda tangan palsu dengan menggunakan metode FRR & FAR, 2. Menggunakan Python agar lebih mudah, 3. Gunakan citra tanda tangan yang lebih tebal agar hasil akurasinya lebih bagus.

DAFTAR PUSTAKA

- [1] Putra, Darma. 2009. *Sistem Biometrika*. Yogyakarta. ANDI.
- [2] Suwignyo, Hadi. 2009. Keabsahan Cap Jempol sebagai Pengganti Tanda Tangan dalam Pembuatan Akta Otentik. *Journal Notarus*. Vol.1, No.1. Diakses 15 November 2013. Online. <https://ejournal.undip.ac.id/index.php/notarius/article/view/1126>
- [3] Amalia, Ismi. 2014. Pengenalan Citra Tanda Tangan Menggunakan Gray Level Co-Occurrence Matrix (GLCM) dan Probabilistic Neural Network (PNN). *Jurnal Teknologi*. Vol.14, No.1:29-34.
- [4] Novandra, Gagah, Muhammad Zidny Na'fan, Tri Ginanjar Laksana. 2018. Perancangan Aplikasi Android Identifikasi Tanda Tangan Menggunakan Multi Layer Perceptron. Vol.03, No.1:76-83.
- [5] Wulanningrum, Resty, Ema Utami, Armadyah Ambrowati. 2014. Implementasi Principal Component Analysis untuk Identifikasi Citra Tanda Tangan. *Seminar Nasional Teknologi Infomasi dan Multimedia*. Diakses 8 Februari 2014. Online. <https://www.ojs.amikom.ac.id/index.php/semnasteknomedia/article/view/1058>
- [6] Kumar, M. 2012. Signature Verification Using Neural Network. *IJCSE*. Vol.4, No.9: 1498-1504.
- [7] Salambue, Roni. 2013. Pengenalan Pola Tanda Tangan dengan Metode Moment Invariant dan Euclidean Distance. *Prosiding Semirata FMIPA*. Vol.1, No.1. Online. <http://jurnal.fmipa.unila.ac.id/semirata/article/view/918>
- [8] Haralick R.M, Shanmugan K, Dinsten I.H. 1973. *Textural Features For Image Classification*. IEEE Trans. Syst. Man. Cybern. 3(6), hal.610-621.
- [9] Kadir, Abdul, Adhi Susanto. 2013. *Teori dan Aplikasi Pengolahan Citra*. Yogyakarta: ANDI.
- [10] Anggoro, Wahyu. Implementasi Ekstraksi Fitur Tekstur Gray Level Co-occurrence Matrices (GLCM) untuk Pengelompokan Citra Tenun Menggunakan Algoritma K-Means. Online. http://eprints.dinus.ac.id/18245/2/jurnal_17791.pdf