

SISTEM IDENTIFIKASI TULISAN TANGAN HURUF HIJAKIYAH MENGUNAKAN PARTISI CITRA

Ardi Sanjaya¹⁾, Danang Wahyu Widodo²⁾

^{1), 2)} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri
Email : dersky@gmail.com¹⁾, danayudo@yahoo.com²⁾

ABSTRAK

Penelitian ini mencoba mengimplementasikan sistem identifikasi tulisan tangan huruf hijakiyah menggunakan partisi citra. Citra dalam format biner di crop untuk diambil area efektifnya dengan batsan titik terluar dari piksel pada citra huruf hijakiyah. Kemudian dilakukan partisi citra dengan tujuan untuk mengekstrak ciri. Ciri yang didapat yaitu prosentase jumlah pikse pada setiap partisi. Kemudian disimpan sebagai data training. Partisi yang digunakan pada penelitian ini yaitu 3x3, 4x4, 5x5, 6x6, 7x7, 8x8, 9x9 dan 10x10. Terdapat 3 skenario pengujian yaitu skenario pertamadata training menggunakan 28 citra tulisan tangan huruf hijakiyah dengan 1 data pada tiap jenis huruf dan data testing menggunakan data training skenario pertama. Skenario pengujian kedua data training menggunakan 28 citra tulisan tangan huruf hijakiyah dengan 3 data pada tiap jenis huruf dan data testing menggunakan data yang berbeda yaitu 28 citra tulisan tangan huruf hijakiyah dengan 3 data pada tiap jenis huruf. Skenario pengujian ketiga data training menggunakan 28 citra tulisan tangan huruf hijakiyah dengan 5 data pada tiap jenis huruf dan data testing menggunakan 28 citra tulisan tangan huruf hijakiyah dengan 3 data pada tiap jenis huruf. Hasil yang diperoleh yaitu penerapan partisi citra untuk ekstraksi ciri pada identifikasi tulisan tangan huruf hijakiyah memiliki akurasi yang rendah. Nilai akurasi tertinggi hanya apabila data training diujikan menjadi data testing. Pada data training yang jumlahnya hanya 3 tiap hurufnya memiliki akurasi rendah. Penambahan data training menjadi 5 tiap hurufnya hanya mampu menaikkan rata-rata akurasi 8.33% saja.

Kata Kunci : Identifikasi, huruf hijakiyah, partisi citra.

PENDAHULUAN

Perkembangan teknologi informasi saat ini sudah sangat pesat dimana sudah banyak pekerjaan atau proses kegiatan manusia mengimplementasikan teknologi informasi. Misalnya pengolahan data keuangan, pemberian keputusan dan analisa data. Disisi lain manusia terus berupaya menyederhanakan pekerjaan atau proses dengan memanfaatkan teknologi informasi. Salah satunya adalah proses menilai hasil tulisan tangan misalnya pada obyek tulisan tangan huruf hijakiyah. Apabila penilaian hasil tulisan tangan huruf hijakiyah dilakukan secara manual, secara manusiawi masih terdapat kelemahan berupa kurangnya konsistensi pada penilaian yang dilakukan. Manusia sulit memberi kepastian tentang hasil akhir dari tulisan tangan tersebut apakah sudah sesuai (sama) ataukah belum.

Penelitian ini mencoba mengimplementasikan pengenalan atau identifikasi pola tulisan tangan berupa huruf hijaiyah dengan menggunakan partisi citra dan *Euclidean distance*. Data berupa citra tulisan tangan huruf hijaiyah dibagi menggunakan partisi menjadi beberapa bagian yang sama dan diambil data ciri berupa jumlah piksel pada setiap partisi dan diubah menjadi nilai prosentase piksel. Selanjutnya dilakukan pencocokan dengan mencari selisih jarak yang paling kecil terhadap data latih (*training*) yang juga dipartisi dengan jumlah yang sama.

KAJIAN TEORI

Definisi Huruf Hijaiyah

Rupa huruf atau biasa juga dikenal dengan istilah *typeface* adalah salah satu elemen terpenting dalam desain grafis karena huruf merupakan sebuah bentuk universal untuk menghantarkan bentuk visual menjadi sebuah bahasa. Huruf hijaiyah merupakan huruf penyusun kata dalam Al-Qur'an. Seperti halnya di Indonesia yang memiliki huruf alfabet dalam menyusun sebuah kata menjadi kalimat, huruf hijaiyah juga memiliki peran yang sama (Andrian,2012). Adapun huruf hijaiyah terdiri atas beberapa karakter seperti gambar 1 di bawah :

خ	ح	ج	ث	ت	ب	ا
Kho	Kha	Jim	Tsa	Ta	Ba	Alif
ص	ش	س	ز	ر	ذ	د
Shod	Syin	Sin	Za	Ra	Dzal	Dal
ق	ف	غ	ع	ظ	ط	ض
Qof	Fa	Ghoin	Ain	Dhlo	Tho	Dhod
ي	ه	و	ن	م	ل	ك
Ya	Ha	Waw	Nun	Mim	Lam	Kaf

Gambar 1. Huruf Hijakiyah

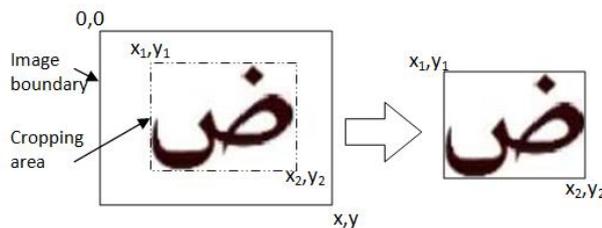
Citra

Citra adalah suatu representasi (gambaran), kemiripan atau imitasi dari suatu obyek. Citra sebagai keluaran suatu sistem perekaman data dapat bersifat optik berupa foto, bersifat analog berupa sinyal-sinyal video seperti gambar pada monitor televisi atau bersifat digital yang dpaat langsung disimpan pada suatu media penyimpanan. Citra digital adalah citra yang dapat diolah oleh komputer. Citra digital sebenarnya berupa angka-angka yang menunjukkan nilai intensitas pada masing-masing piksel (T. Sutoyo, 2009). Dari citra skala abu-abu diubah menjadi citra biner dengan operasi pengambangan (*thresholding*) atau biasa disebut juga dengan binerisasi citra. Citra biner adalah citra yang hanya memiliki nilai 0 dan 255 pada setiap pikselnya. Pengubahan nilai masing-masing piksel dari skala keabuan yang sudah dideteksi tepi menjadi nilai 0 atau 255 berdasarkan nilai *threshold* yang ditentukan (Balza, 2005). Persamaan yang digunakan untuk proses binerisasi adalah sebagai berikut :

$$\begin{aligned}
 K_o = & \begin{cases} 0, & \text{jika } K_i < \text{ambang} \\ 1, & \text{jika } K_i \geq \text{ambang} \end{cases} \\
 \text{Atau} & \\
 K_o = & \begin{cases} 0, & \text{jika } K_i \geq \text{ambang} \\ 1, & \text{jika } K_i < \text{ambang} \end{cases}
 \end{aligned}$$

Crop Citra

Pemotongan citra (*cropping citra*) merupakan cara mendapatkan area warna tertentu yang diamati (*area of interest*), yang bertujuan untuk mempermudah penganalisaan citra dan memperkecil ukuran penyimpangan citra (Angga et al, 2017). *Cropping Citra* yang dimaksud adalah membuang area yang tidak diperlukan dan menjadikan piksel terluar obyek sebagai nilai batas tepi area. *Cropping citra* ditunjukkan pada gambar 2 berikut :



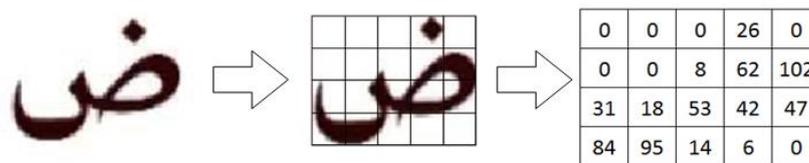
Gambar 2. Ilustrasi Crop Citra

Untuk mendapatkan piksel terluar dari obyek menggunakan metode scanning. Dimulai dari koordinat 0,0 sampai dengan lebar citra-1, tinggi citra-1. Variabel x1 dan y1 diisi lebih tinggi dari

lebar dan tinggi citra. Nilai x2 dan y2 diisi nol. Saat proses *scanning* menemukan piksel hitam di koordinat x,y maka apabila $x < x_1$ maka nilai x1 adalah x. Apabila $x > x_2$ maka nilai x2 adalah x. Apabila $y < y_1$ maka nilai y1 adalah y. Apabila $y > y_2$ maka nilai y2 adalah y.

Partisi Citra

Segmentasi citra merupakan suatu proses pengelompokkan citra menjadi beberapa *region* berdasarkan kriteria tertentu. Berdasarkan pengertiannya, segmentasi memiliki tujuan menemukan karakteristik khusus yang dimiliki suatu citra (Sa'adah, 2009). Partisi citra adalah membagi citra menjadi beberapa blok dimana masing-masing blok atau bagian memiliki ukuran yang sama besar (Sanjaya, 2014). Salah satu tujuan dilakukan partisi terhadap citra adalah menghitung jumlah piksel masing-masing blok pada data *training* dan data *testing* dan mengubahnya menjadi nilai prosentase piksel yang mewakili nilai masing-masing blok partisi. Nilai prosentase piksel didapat dari jumlah piksel keseluruhan dibagi jumlah piksel pada partisi. Nilai prosentase pada data *training* dan *testing* kemudian dicari selisih yang terpendek menggunakan *Euclidean* (Sanjaya, 2014)..



Gambar 3. Ilustrasi partisi citra

Dari ilustrasi diatas, dapat diartikan bahwa huruf Dhod memiliki ciri pada partisi 4 terdapat 26 piksel, pada partisi 9 terdapat 8 piksel dan seterusnya.

Pencocokan

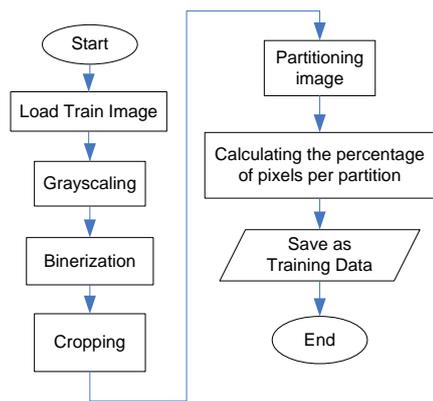
Untuk mencari nilai kemiripan data yang diuji terhadap data latih dari nilai masing-masing ciri menggunakan jarak euclidean. Jarak *Euclidean* (*Euclidean distance*) yaitu metrika yang digunakan untuk menghitung kesamaan 2 vector. Jarak *Euclidean* menghitung akar dari kuadrat perbedaan 2 vector (*root of square difference between 2 vectors*) (Putra, 2009). Rumus atau persamaan penghitungan jarak ditulis sebagai berikut :

$$d_e = \sqrt{\sum_{k=1}^m (f_{d_{i,k}} - k_j)^2}$$

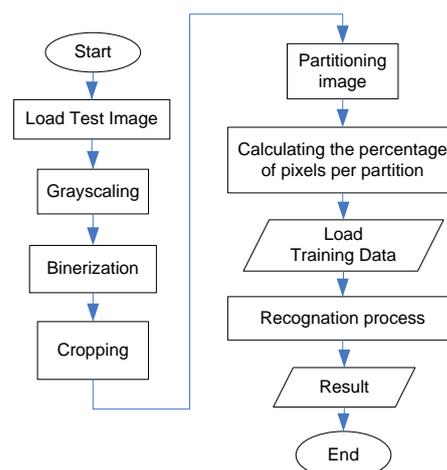
METODE PENELITIAN

Pada penelitian ini menggunakan 3 skenario pengujian. Skenario pertama yaitu menggunakan 28 jenis citra tulisan tangan huruf hijakiyah dengan 1 data citra untuk masing-masing huruf dan total untuk data training yaitu 28. Untuk pengujiannya (*testing*) menggunakan 28 citra pada data training. Skenario kedua menggunakan 28 jenis citra tulisan tangan huruf hijakiyah dengan 3 data citra untuk masing-masing huruf dan total untuk data training yaitu 84. Untuk *testing*nya menggunakan 28 jenis citra tulisan tangan huruf hijakiyah dengan 3 data citra untuk masing-masing huruf yang berbeda dengan data training pada skenario 2 dan total untuk data *testing*nya adalah 84. Untuk skenario ketiga, yaitu menggunakan 28 jenis citra tulisan tangan huruf hijakiyah dengan 5 data citra untuk masing-masing huruf dan total untuk data training yaitu 140. Untuk *testing*nya menggunakan 28 jenis citra tulisan tangan huruf hijakiyah dengan 3 data citra untuk masing-masing huruf dan total untuk data *testing* yaitu 84.

Alur pemrosesan *training* disajikan pada gambar 4 dan *testing* disajikan pada gambar 5 sebagai berikut :

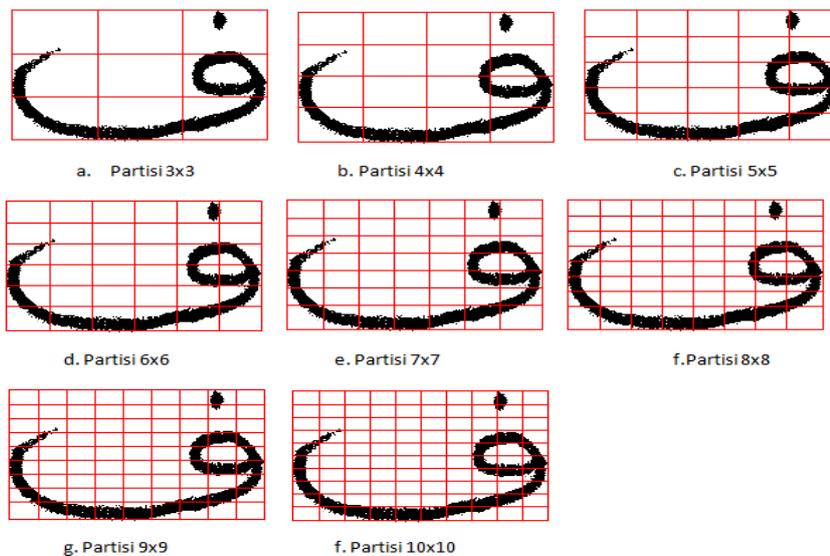


Gambar 4. Alur Data Training



Gambar 5. Alur Data Testing

Partisi yang digunakan pada penelitian ini yaitu partisi ukuran 3x3, 4x4, 5x5, 6x6, 7x7, 8x8, 9x9 dn 10x10. Berikut adalah ilustrasi dari partisi yang digunakan :



Gambar 6. Ilustrasi partisi yang digunakan

HASIL DAN PEMBAHASAN

Berdasarkan pengujian skenario 1,2 dan 3 didapat hasil sebagai berikut :

Tabel 1. Hasil Akurasi Pengujian (%)

Partisi	Skenario 1		Skenario 2			Skenario 3		
	Pengujian		Pengujian			Pengujian		
	1	1	2	3	1	2	3	
3x3	100	64.29	71.43	75.00	78.57	82.14	78.57	
4x4	100	67.86	60.71	67.86	71.43	78.57	75.00	
5x5	100	60.71	71.43	64.29	67.86	78.57	71.43	
6x6	100	57.14	67.86	64.29	64.29	78.57	75.00	
7x7	100	53.57	64.29	53.57	57.14	67.86	57.14	
8x8	100	39.29	46.43	50.00	53.57	64.29	60.71	
9x9	100	39.29	53.57	50.00	46.43	60.71	53.57	
10x10	100	32.14	53.57	42.86	42.86	60.71	46.43	

Tabel 2. Rata-rata Akurasi Pengujian(%)

Partisi	Skenario		
	1	2	3
3x3	100	70.24	79.76
4x4	100	65.48	75.00
5x5	100	65.48	72.62
6x6	100	63.10	72.62
7x7	100	57.14	60.71
8x8	100	45.24	59.52
9x9	100	47.62	53.57
10x10	100	42.86	50.00

Pada skenario 1, didapat hasil akurasi 100% di semua ukuran partisi. Hal tersebut menunjukkan sistem mampu mengenali sesuai data trainingnya. Pada skenario kedua, mengalami penurunan akurasi. Peneliti mengambil salah satu perhitungan pada pengujian ke 1 partisi 7x7. Didapat data sebagai berikut :

Tabel 3. Hasil Pengujian ke 1 partisi 7x7 Skenario ke 2

NO	CITRA	INITIAL	DIKENALI	EUCLIDEAN	KET
1	Ain.bmp	AIN	AIN	12.72641	Benar
2	Alif.bmp	ALIF	ALIF	17.65147	Benar
3	Ba.bmp	BA	BA	14.08788	Benar
4	Dal.bmp	DAL	DAL	17.32615	Benar
5	Dhlo.bmp	DHLO	DHLO	15.35483	Benar
6	Dhod.bmp	DHOD	DHOD	14.82141	Benar
7	Dza.bmp	DZA	DZAL	21.25338	Salah
8	Dzal.bmp	DZAL	DZAL	17.97838	Benar
9	Fa.bmp	FA	FA	10.0044	Benar
10	Ghoin.bmp	GHOIN	AIN	17.25845	Salah
11	Ha.bmp	HA	HA	13.18031	Benar
12	Jim.bmp	JIM	KHA	14.23834	Salah
13	Kha.bmp	KHA	KHA	17.128	Benar
14	Khaf.bmp	KHAF	KHAF	5.882163	Benar
15	Kho.bmp	KHO	KHO	15.21034	Benar
16	Lam.bmp	LAM	LAM	14.64366	Benar
17	Mim.bmp	MIM	KHO	19.56	Salah
18	Nun.bmp	NUN	NUN	12.08073	Benar
19	Qof.bmp	QOF	FA	9.913212	Salah
20	Ro.bmp	RO	DZAL	12.28081	Salah
21	Shod.bmp	SHOD	SHOD	14.00656	Benar
22	Sin.bmp	SIN	SYIN	14.355	Salah
23	Syin.bmp	SYIN	HA	19.97729	Salah
24	Ta.bmp	TA	NUN	8.05565	Salah
25	Tho.bmp	THO	DHLO	13.74045	Salah
26	Tsa.bmp	TSA	NUN	8.722603	Salah
27	Wawu.bmp	WAWU	DAL	18.37048	Salah
28	Ya.bmp	YA	BA	13.85617	Salah

Dari tabel 3 diatas, pada data ke 24 dimana citra testing huruf Ta dikenali sebagai huruf Nun dengan jarak Euclidean 8.60.

Tabel 4. Analisa pengujian 1 partisi 7x7 skenario kedua

NO	CITRA		CITRA TESTING	
	TRAINING	INITIAL		JARAK EUCLIDEAN
1	TA1.bmp	TA	TA_Uji_S2_1.bmp	13.44891074
2	TA2.bmp	TA	TA_Uji_S2_1.bmp	14.84697612
3	TA3.bmp	TA	TA_Uji_S2_1.bmp	11.95820639
4	NUN1.bmp	NUN	TA_Uji_S2_1.bmp	8.055650191
5	NUN2.bmp	NUN	TA_Uji_S2_1.bmp	14.21300109
6	NUN3.bmp	NUN	TA_Uji_S2_1.bmp	10.37847291

Berdasar tabel 4 diatas, dapat diketahui bahwa citra testing huruf TA pada file TA_uji_S2_1.bmp memiliki jarak Euclidean paling dekat dengan citra training NUN1.bmp sehingga dikenali sebagai huruf NUN. Hal tersebut menyebabkan akurasi menurun.

Pada skenario ketiga, data di tabel 2 menunjukkan terdapat peningkatan akurasi dikarenakan data training yang ditambah. Namun secara keseluruhan semakin tinggi nilai pembagian partisi maka nilai akurasi menurun. Hal tersebut disebabkan semakin tinggi nilai pembagian partisi berdampak pada ciri citra yang menjadi semakin detail.

KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan bahwa penerapan partisi citra untuk ekstraksi ciri pada identifikasi tulisan tangan huruf hijakiyah memiliki akurasi yang rendah. Nilai akurasi tertinggi hanya apabila data training diujikan menjadi data testing. Pada data training yang jumlahnya hanya 3 tiap hurufnya memiliki akurasi rendah. Penambahan data training menjadi 5 tiap hurufnya hanya mampu menaikkan rata-rata akurasi 8.33% saja.

Untuk penelitian selanjutnya, agar digabungkan dengan proses verifikasi dengan metode lain yang relevan agar didapat akurasi yang lebih tinggi.

DAFTAR PUSTAKA

- Andrian, W. 2012. Analisa Dan Implementasi Metode Kohonen Neural Network Untuk Pengenalan Karakter Huruf Arab. UIN Sultan Syarif Kasim Riau
- Balza , A. 2005. *Teknik Pengolahan Citra Digital menggunakan DELPHI*. Yogyakarta. Ardi Publishing.
- Angga, A.M. Ichwan. M,Budhi, D. U. 2017. *Pengenalan Tingkat Kematangan Tomat Berdasarkan Citra Warna Pada Studi Kasus Pembangunan Sistem Pemilihan Otomatis*. Jurnal Teknik Informatika dan Sistem Informasi, Vol 3 No 3, e-ISSN 2443-2229
- Sa'adah. S.2009. *Segmentasi Citra*.
http://digilib.itelkom.ac.id/index.php?option=com_content&view=article&id=575:segmentasi-citra&catid=20:informatika&Itemid=14. diakses 15 Mei 2017
- Sanjaya, Ardi. 2014. *Identifikasi Personal Berdasarkan Bentuk Tangan*. Prosiding Seminar Nasional Teknologi Informasi dan Multimedia 2014. ISSN 2302-3805
- Putra. Darma. 2009. *Sistem Biometrika*. Yogyakarta. Andi Publishing