

IDENTIFIKASI TULISAN TANGAN HURUF HIJAIYAH

Ardi Sanjaya¹⁾, Danang Wahyu Widodo²⁾

^{1), 2)} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri
Kampus 2 Jl KH Ahmad Dahlan Gg1 No 6 Mojoroto Kediri Jawa Timur
Email : dersky@gmail.com¹⁾, danayudo@yahoo.com²⁾

Abstrak

Penelitian ini mencoba memberikan alternatif untuk identifikasi huruf hijaiyah menggunakan partisi citra. Citra biner huruf hijaiyah di crop dengan batasan terluar berupa tepi huruf dan kemudian dilakukan partisi. Pada masing-masing bagian atau partisi, diambil nilai berupa prosentase jumlah piksel. Nilai prosentase jumlah piksel tiap partisi menjadi ciri dari citra tersebut. Selanjutnya nilai prosentase masing-masing partisi disimpan sebagai data training. Untuk proses pengenalan, citra uji juga dipartisi dan dihitung nilai prosentase piksel tiap partisi kemudian dicari selisih jarak terpendek terhadap data training menggunakan jarak Euclidean. Data training yang digunakan berupa tulisan tangan 28 huruf hijaiyah dengan partisi 3x3, 4x4, 5x5 dan 6x6. Data testing untuk pengujian pertama menggunakan data training itu sendiri. Rata-rata akurasi pengujian 100%. Pengujian kedua dan ketiga menggunakan data baru tulisan tangan 28 huruf hijaiyah dari sumber yang sama. Berdasarkan rata-rata hasil pengujian per partisi, semakin tinggi nilai pembagian partisi, maka akan menurunkan akurasi. Rata-rata pengujian pada partisi 3x3 sebesar 84.52%, pada partisi sebesar 77.38%, pada partisi 5x5 sebesar 72.62% dan pada partisi 6x6 sebesar 67.86%. Hal tersebut dikarenakan semakin tinggi nilai pembagian partisi, berdampak pada semakin banyak/detailnya ciri suatu huruf. Pada citra dengan partisi 3x3 dapat diartikan memiliki 9 ciri, partisi 4x4 memiliki 16 ciri, partisi 5x5 memiliki 25 ciri dan partisi 6x6 memiliki 36 ciri. Sehingga semakin banyak ciri menyebabkan toleransi kemiripan yang rendah.

Kata kunci : Identifikasi, huruf hijaiyah, partisi citra.

Abstract

This research is provide an alternative for identification of hijaiyah letters using image partitions. Binary image of hijaiyah letters are cropped outermost boundary of edge letters then partitions. In each part or partition, value is percentage of pixels and partition is a feature of image. Then the percentage value of each partition is stored as training data. For the introduction process, the test image is also partitioned and the percentage of pixels per partition is calculated, then the shortest distance to training data is determined using Euclidean distance. The training data used are 28 letters hijaiyah handwriting with 3x3, 4x4, 5x5 and 6x6 partitions. Data testing for the first test uses the training data itself. Average testing accuracy is 100%. The second and third tests used new handwritten data of 28 hijaiyah letters from the same source. Based on average test results per partition, the higher partitioning value, the lower accuracy. The average test on 3x3 partition is 84.52%, partition is 77.38%, partition 5x5 is 72.62% and partition 6x6 is 67.86%. This is due to the higher value of partition partitioning, the more characteristic / detailing of a letter will be. On images with 3x3 partitions can be interpreted as having 9 features, 4x4 partitions have 16 features, 5x5 partitions have 25 features and 6x6 partitions have 36 features. So that more characteristics cause low similarity tolerance.

Keywords : Identification, hijaiyah letter, image partition.

1. Pendahuluan

Teknologi informasi sudah berkembang sangat pesat dan penggunaannya sangat membantu manusia dalam menyelesaikan pekerjaannya. Misalnya membantu dalam hal memberikan koreksi dan penilaian secara otomatis pada ujian berbasis komputer. Peserta ujian cukup

memberi jawaban dari pertanyaan-pertanyaan yang diajukan oleh komputer melalui interaksi dengan layar yang representatif. Namun terdapat kendala ketika yang dinilai adalah hasil pekerjaan tulisan tangan misalnya pada obyek tulisan tangan huruf hijaiyah hasil karya/pekerjaan dari anak usia dini. Secara subyektif manusia mampu memberikan penilaian terhadap hasil pekerjaan tersebut namun ada kalanya kurang konsisten dan sulit menentukan secara pasti tingkat kemiripan tulisan tangan huruf hijaiyah terhadap huruf hijaiyah yang dijadikan acuan. Salah satu solusi untuk menyelesaikan permasalahan tersebut adalah dengan memanfaatkan pengolahan citra. Penelitian ini mencoba mengembangkan pengenalan atau identifikasi pola tulisan tangan huruf hijaiyah anak usia dini dengan menggunakan partisi citra dan *Euclidean distance*. Citra tulisan tangan huruf hijaiyah dipartisi atau dibagi 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.

2. Dasar teori

2.1 Deskripsi 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 [1]. 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 Hijaiyah

2.2 Citra Abu-abu

Citra skala keabuan (*grayscale*) adalah citra yang memiliki nilai 0 sampai 255 untuk setiap pikselnya. Untuk mengubah citra bitmap menjadi skala keabuan menggunakan nilai rata-rata dari ketiga elemen warna citra bitmap [2], menggunakan persamaan :

$$K_o = \frac{R_i + G_i + B_i}{3} \tag{1}$$

2.3 Citra Biner

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

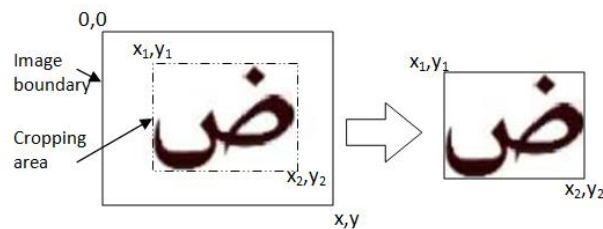
dari skala keabuan menjadi nilai 0 atau 255 berdasarkan nilai *threshold* yang ditentukan. Persamaan yang digunakan untuk proses binerisasi adalah [2] :

$$K_o = \begin{cases} 0, & \text{jika } K_i < \text{ambang} \\ 1, & \text{jika } K_i \geq \text{ambang} \end{cases} \quad (2)$$

$$K_o = \begin{cases} 0, & \text{jika } K_i \geq \text{ambang} \\ 1, & \text{jika } K_i < \text{ambang} \end{cases} \quad (3)$$

2.4 Crop Citra

Pemotongan citra (*cropping* citra) merupakan cara mendapatkan area tertentu yang diamati (*area of interest*), yang bertujuan untuk mempermudah penganalisaan citra dan memperkecil ukuran penyimpanan citra [3]. *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 $x < x1$ maka nilai x1 adalah x. Apabila $x > x2$ maka nilai x2 adalah x. Apabila $y < y1$ maka nilai y1 adalah y. Apabila $y > y2$ maka nilai y2 adalah y. *Pseudocode* diilustrasikan sebagai berikut :

```

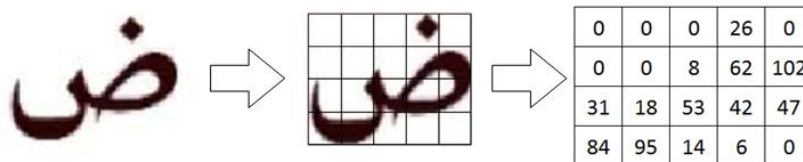
Var
  x, y, x1, y1, x2, y2: integer;
Begin
  x1 := 999999; x2 := 0;
  y1 := 999999; y2 := 0;
  for x:=0 to Image1.Width-1 do
    for y:=0 to Image1.Height-1 do
      if Image1.Canvas.Pixels[x,y]=clBlack then
        begin
          if x<x1 then x1:=x;
          if x>x2 then x2:=x;
          if y<y1 then y1:=y;
          if y>y2 then y2:=y;
        end;
      end;
    end;
  End;

```

2.5 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 [4]. Partisi citra adalah membagi citra menjadi beberapa blok dimana masing-masing blok atau bagian memiliki ukuran yang sama besar [5]. 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* [5].



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.

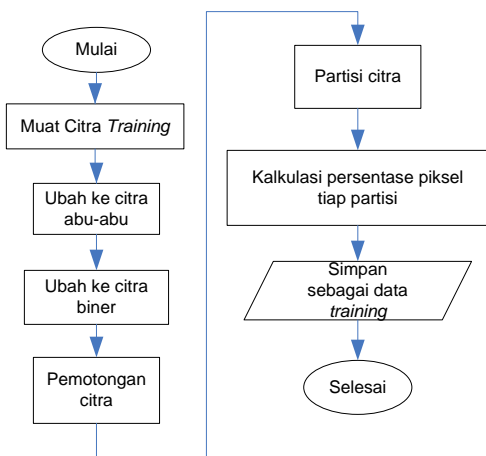
2.6 Euclidean Distance

Jarak *Euclidean* (*Euclidean distance*) yaitu rumus yang digunakan untuk menghitung kesamaan 2 vektor. Jarak *Euclidean* menghitung akar dari kuadrat perbedaan 2 vektor (*root of square difference between 2 vectors*) [6]. Rumus atau persamaan penghitungan jarak ditulis sebagai berikut (persamaan 4) :

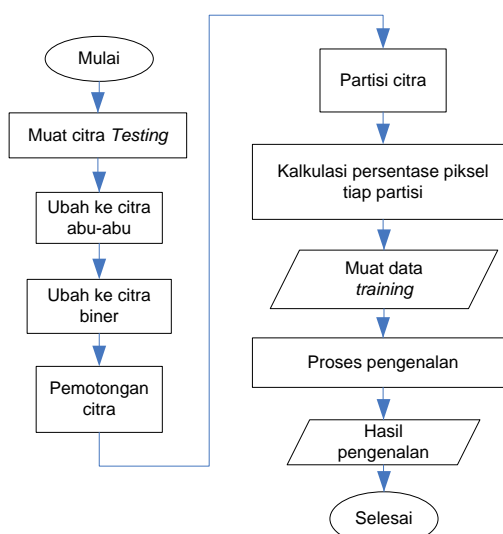
$$d_e = \sqrt{\sum_{k=1}^m (fd_{i,k} - k_j)^2} \tag{4}$$

3. Metodologi Penelitian

Data *training* yang digunakan adalah 28 citra alfabet huruf hijaiyah yang diambil dari scan tulisan tangan. Data *testing* mengambil dari data *training* dan scan tulisan tangan anak usia dini sejumlah 56 sampel huruf, sehingga total data *testing* 84. Alur pemrosesan *training* disajikan pada gambar 4 dan *testing* disajikan pada gambar 5.



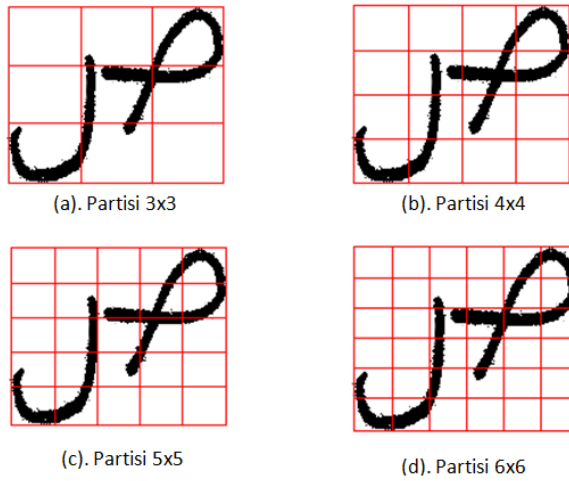
Gambar 4. Alur Data Training



Gambar 5. Alur Data Testing

4. Hasil dan Pembahasan

Partisi yang digunakan pada penelitian ini adalah 3x3, 4x4, 5x5 dan 6x6. Gambar 6 berikut merupakan ilustrasi partisi yang digunakan dalam penelitian :



Gambar 6. Ilustrasi partisi yang digunakan

Pengujian pertama dilakukan dengan menjadikan 28 data citra huruf hijaiyah yang sebelumnya dijadikan sebagai data *training* sebagai data *testing*. Pengujian kedua dan ketiga peneliti menggunakan 28 data citra huruf hijaiyah baru yang diambil dari sumber sama. Didapatkan hasil sebagai berikut :

Tabel 1. Hasil Pengujian

No	Partisi	Akurasi Pengujian (%)			Rata-Rata
		1	2	3	
1	3x3	100	75.00	78.51	84.52
2	4x4	100	67.85	64.28	77.38
3	5x5	100	60.71	57.14	72.62
4	6x6	100	50.00	53.57	67.86

Pada pengujian pertama, didapatkan akurasi 100% untuk semua ukuran partisi. Ini menunjukkan sistem mampu mengenali sesuai data *training*. Pengujian kedua dan ketiga mengalami penurunan akurasi. Peneliti mengambil satu sampel perhitungan untuk dianalisa yaitu pada pengujian kedua menggunakan partisi 3x3. Huruf Shod dari data *testing* yang diujikan terhadap data *testing* didapat hasil sebagai berikut :

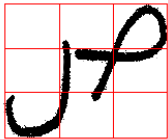
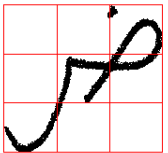
Tabel 2. Analisa Hasil Pengujian Kedua Pada Huruf Shod Terhadap Data *Training* Partisi 3x3

1		2		3		4	
Citra <i>Training</i>	Euclidean Distance	Citra <i>Training</i>	Euclidean Distance	Citra <i>Training</i>	Euclidean Distance	Citra <i>Training</i>	Euclidean Distance
AIN	48.07	DZAL	41.34	KHO	37.83	SIN	18.14
ALIF	42.13	FA	46.81	LAM	37.24	SYIN	21.27
BA	45.87	GHOIN	27.19	MIM	47.52	TA	33.21
DAL	39.28	HA	30.21	NUN	42.14	THO	44.06
DHLO	38.01	JIM	40.30	QOF	34.07	TSA	39.32

DHOD	7.67	KHA	55.52	RO	40.55	WAWU	27.04
DZA	42.61	KHAF	28.57	SHOD	14.54	YA	45.34

Berdasar tabel 2, dapat diketahui bahwa nilai jarak *Euclidean* yang terkecil adalah 7.67 terhadap citra *training* huruf Dhod atau dikenali sebagai Dhod. Nilai jarak *Euclidean* terkecil seharusnya terhadap huruf Shod. Berikut disajikan perbandingan data citra *testing* huruf Shod dan data *training* huruf Dhod.

Tabel 3. Perbandingan Prosentase Piksel Pada Skenario 3 Partisi 3x3

	Citra Huruf	Prosentase Piksel Tiap Partisi (%)									
Data <i>Testing</i> = Shod		<table border="1"> <tr> <td>0</td> <td>1.87</td> <td>23.98</td> </tr> <tr> <td>0</td> <td>30.8</td> <td>13.35</td> </tr> <tr> <td>23.45</td> <td>6.53</td> <td>0</td> </tr> </table>	0	1.87	23.98	0	30.8	13.35	23.45	6.53	0
0	1.87	23.98									
0	30.8	13.35									
23.45	6.53	0									
Data <i>Training</i> = Dhod		<table border="1"> <tr> <td>0</td> <td>0.79</td> <td>24.97</td> </tr> <tr> <td>0.14</td> <td>34.29</td> <td>17.06</td> </tr> <tr> <td>21.34</td> <td>1.4</td> <td>0</td> </tr> </table>	0	0.79	24.97	0.14	34.29	17.06	21.34	1.4	0
0	0.79	24.97									
0.14	34.29	17.06									
21.34	1.4	0									

Berdasarkan rata-rata hasil pengujian per partisi, semakin tinggi nilai pembagian partisi, maka akan menurunkan akurasi. Rata-rata pengujian pada partisi 3x3 sebesar 84.52%, pada partisi sebesar 77.38%, pada partisi 5x5 sebesar 72.62% dan pada partisi 6x6 sebesar 67.86%. Hal tersebut dikarenakan semakin tinggi nilai pembagian partisi, berdampak pada semakin banyak/detailnya ciri suatu huruf. Pada citra dengan partisi 3x3 dapat diartikan memiliki 9 ciri, partisi 4x4 memiliki 16 ciri, partisi 5x5 memiliki 25 ciri dan partisi 6x6 memiliki 36 ciri. Sehingga semakin banyak ciri menyebabkan toleransi kemiripan yang rendah.

5. Kesimpulan

Meski secara visual data citra *training* dan data *testing* memiliki kemiripan namun apabila diekstraksi menjadi ciri berupa prosentase jumlah piksel tiap partisi citra ternyata memiliki nilai yang berbeda. Semakin tinggi nilai partisi maka nilai toleransi kemiripan menjadi semakin rendah. Relatifitas dari penulisan huruf hijaiyah dengan tangan juga menjadi salah satu faktor penentu akurasi atau keberhasilan dalam identifikasi.

Daftar Pustaka

- [1] Andrian. W, “Analisa Dan Implementasi Metode Kohonen Neural Network Untuk Pengenalan Karakter Huruf Arab”, UIN Sultan Syarif Kasim Riau, 2012.
- [2] Achmad, Ir. Balza, Teknik Pengolahan Citra Digital menggunakan DELPHI, Yogyakarta, Ardi Publishing, 2005.
- [3] Angga. A.M. Ichwan. M. Budhi, D. U, , “Pengenalan Tingkat Kematangan Tomat Berdasarkan Citra Warna Pada Studi Kasus Pembangunan Sistem Pemilihan Otomatis”, Jurnal Teknik Informatika dan Sistem Informasi, Vol. 3, No. 3, pp 550-564, 2017.

- [4] Sa'adah. Siti, 2009, *Segmentasi Citra*, http://digilib.ittelkom.ac.id/index.php?option=com_content&view=article&id=575:segmentasi-citra&catid=20:informatika&Itemid=14, diakses 15 Mei 2017
- [5] Sanjaya. Ardi, 2014, "Identifikasi Personal Berdasarkan Bentuk Tangan", Prosiding Seminar Nasional Teknologi Informasi dan Multimedia 2014,
- [6] Putra.Darma, *Sistem Biometrika*, Yogyakarta, Andi Publishing, 2009.