

BAB IV

HASIL DAN PEMBAHASAN

4.1. Input yang Digunakan

Terdapat beberapa citra digital yang digunakan sebagai input dalam penelitian ini. Secara garis besar, keseluruhan input tersebut dapat dibedakan atas dua bagian: citra berupa setiap huruf Arab Melayu, dan citra berupa tulisan Arab Melayu.

Sebagian besar citra input digunakan sebagai data *training* (latihan), sekaligus sebagai data pada tahap pengenalan. Beberapa citra yang digunakan itu dapat dilihat pada Tabel 4.1 dan Gambar 4.1.

Tabel 4.1. Input berupa huruf Arab Melayu

Huruf Latin	Bentuk Sambung Awal	Bentuk Sambung Tengah	Bentuk Sambung Akhir	Huruf Arab Melayu
A		-	ا	alif
B	ب	ب	ب	ba
T	ت	ت	ت	ta
S	س	س	س	sa
J	ج	ج	ج	jim
C	چ	چ	چ	ce
H	ه	ه	ه	ha
Kh	خ	خ	خ	kho
D	د	-	د	dal
Z	ز	-	ز	zal
R	ر	-	ر	ro
Z	ڙ	-	ڙ	zai
S	ـ	ـ	ـ	sin
Sy	ـ	ـ	ـ	syin
S	ـ	ـ	ـ	sod

Huruf Latin	Bentuk Sambung Awal	Bentuk Sambung Tengah	Bentuk Sambung Akhir	Huruf Arab Melayu
d	ض	ضـ	ضـ	dod
t	ط	ـطـ	ـطـ	to
Z	ظـ	ـظـ	ـظـ	zo
G	غـ	ـغـ	ـغـ	gain
Ng	ـڠـ	ـڠـ	ـڠـ	nga
P	ـپـ	ـپـ	ـپـ	pa
k/q	ـقـ	ـقـ	ـقـ	qaf
K	ـڪـ	ـڪـ	ـڪـ	ka
G	ـڱـ	ـڱـ	ـڱـ	ga
L	ـڶـ	ـڶـ	ـڶـ	lam
M	ـڻـ	ـڻـ	ـڻـ	mim
N	ـڻـ	ـڻـ	ـڻـ	nun
W	ـڻـ	-	ـڻـ	waw
H	ـڻـ	ـڻـ	ـڻـ	ha
Y	ـڻـ	ـڻـ	ـڻـ	ya
Ny	ـڻـ	ـڻـ	ـڻـ	nya

كمبع
هنجر
كمبع مک

(a)

سکرا داتع
دو فاسخ
کلوار کمفع

(b)

Gambar 4.1. Input berupa tulisan Arab Melayu

بُوْيِهْ فُوَّنَهْ سُفَرَتِيْ كُوَاه
ادق دان كاڪق دودق دكت بافق

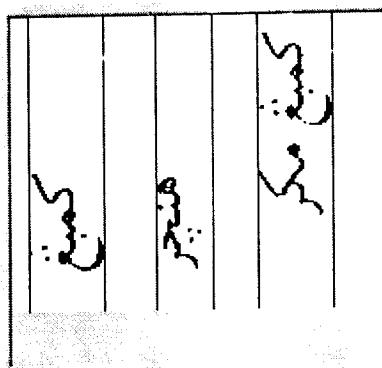
(c)

Gambar 4.1. Input berupa tulisan Arab Melayu

4.2. Segmentasi

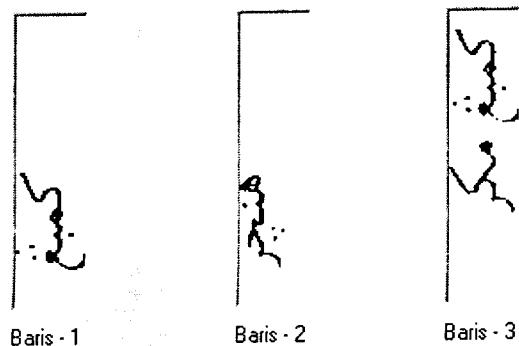
Proses segmentasi citra pertama kali dilakukan terhadap setiap baris tulisan Arab Melayu dengan berpatokan pada piksel yang teratas dan nilai piksel terendah. Sebelum proses ini dilakukan, citra terlebih dahulu diubah nilai intensitasnya menggunakan nilai *threshold* 200 sampai dengan 210 untuk membedakan nilai intensitas piksel antara 0 dan 255. Hal ini dilakukan agar penentuan piksel teratas dan terendah pada setiap baris dapat dilakukan dengan lebih mudah.

Menggunakan input berupa citra pada gambar 4.1.(a), diperoleh hasil segmentasi seperti tampak pada gambar 4.2. Garis biru menandakan lokasi piksel tertinggi dan terendah pada setiap baris tulisan Arab Melayu. Posisi citra sengaja dirotasi -90 derajat untuk mempermudah proses selanjutnya.



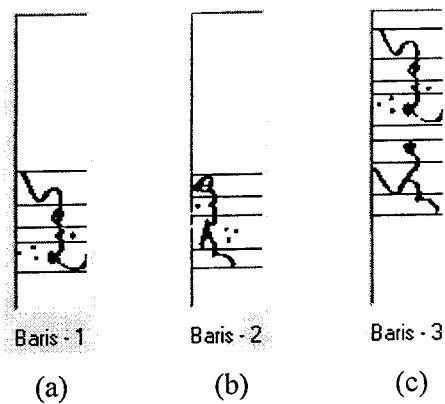
Gambar 4.2. Hasil segmentasi baris

Berdasarkan patokan garis tersebut, dengan melakukan pemotongan (*cropping*) didapatkan citra tulisan Arab Melayu untuk setiap baris seperti terlihat pada gambar 4.3.



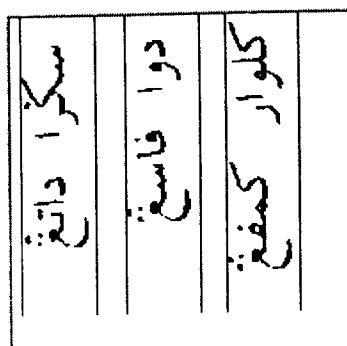
Gambar 4.3. Hasil pemotongan baris

Menggunakan hasil pemotongan baris pada gambar 4.3, selanjutnya dilakukan penandaan kolom untuk memisahkan setiap huruf Arab Melayu. Hasil penandaan ini dapat dilihat pada gambar 4.4.



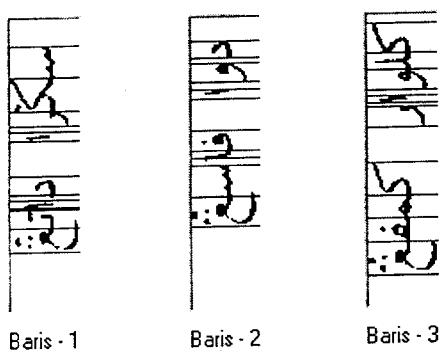
Gambar 4.4. Hasil segmentasi kolom (huruf)

Selanjutnya menggunakan input citra pada gambar 4.1.(b), dengan cara yang sama sebelumnya, diperoleh hasil segmentasi baris seperti terlihat pada gambar 4.5.



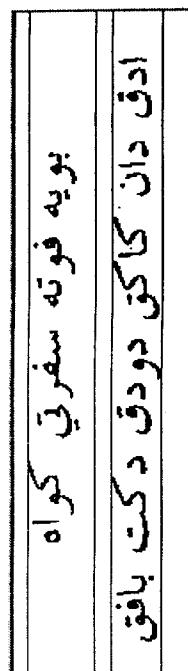
Gambar 4.5. Hasil segmentasi baris citra gambar 4.1(b)

Terakhir, hasil segmentasi kolomnya dapat dilihat pada gambar 4.6.

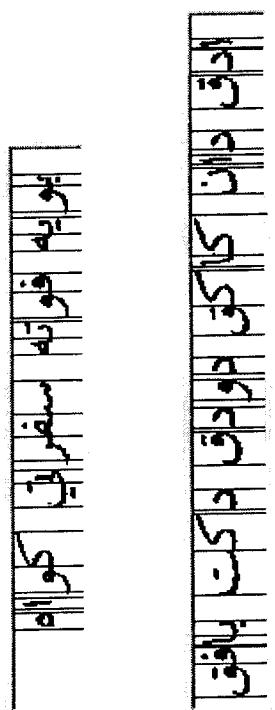


Gambar 4.6. Hasil segmentasi kolom citra gambar 4.1(b)

Hasil segmentasi baris dan kolom untuk citra gambar 4.1.(c) masing-masing dapat dilihat pada gambar 4.7 dan gambar 4.8.



Gambar 4.7. Hasil segmentasi baris citra gambar 4.1(c)



Gambar 4.8. Hasil segmentasi kolom citra gambar 4.1(c)

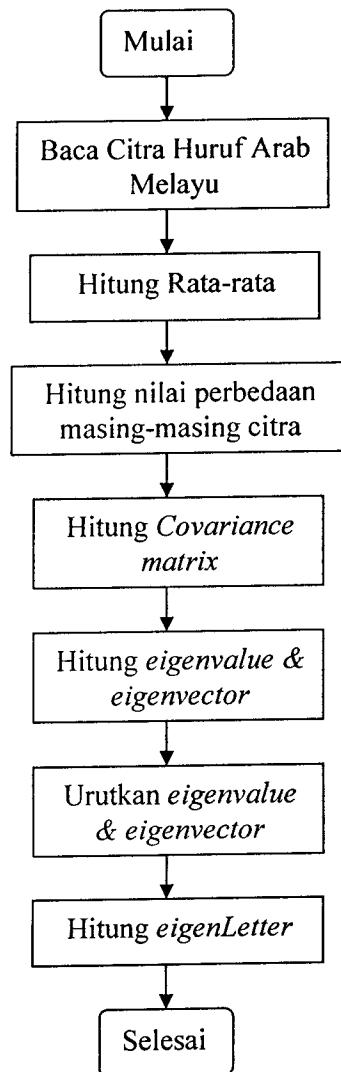
4.3. Pengenalan Huruf Arab Melayu

Hasil proses segmentasi yang telah didapatkan digunakan sebagai masukan untuk proses pengenalan huruf Arab Melayu. Dalam proses pengenalan ini pada prinsipnya terlebih dahulu dilakukan ekstraksi ciri dan klasifikasi terhadap huruf-huruf yang akan dikenali. Keseluruhan proses ini terangkum dalam penggunaan Metode *Principal Component Analysis* (PCA). Pada bagian terakhir akan dilanjutkan dengan penghitungan *Eucledian distance*, sebagai upaya dalam mengenali suatu huruf yang diinginkan berdasarkan klasifikasi yang telah didapatkan.

4.3.1. Penghitungan *EigenLetter*

Gambar 4.9 adalah diagram alir yang menggambarkan algoritma yang akan digunakan dalam menghitung *eigenLetter*. Dari gambar 4.9 terlihat jelas bahwa untuk menghitung *eigenLetter* prosesnya diawali dengan pembacaan citra yang akan digunakan sebagai data input pada tahap training. Setelah itu dihitung nilai rata-rata piksel pada masing-masing citra dan dilanjutkan dengan menghitung nilai perbedaan dari masing-masing citra (lihat persamaan (2.4)). Hasil perbedaan ini digunakan untuk menghitung matriks *covariance* menggunakan persamaan (2.7).

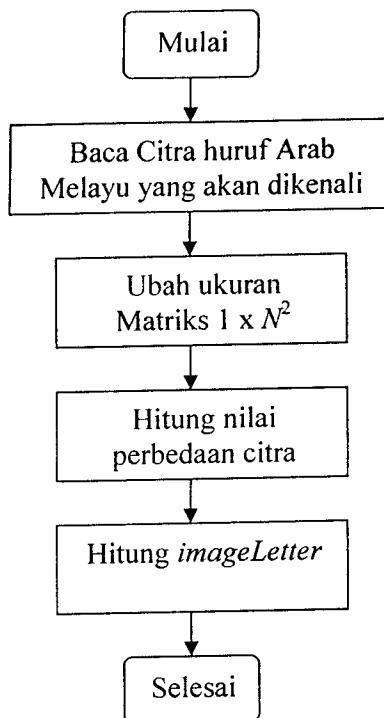
Langkah berikutnya adalah menghitung nilai eigen dan vektor eigen. Pada penghitungan nilai ini digunakan perangkat lunak MATLAB. Hasil vektor eigen yang didapatkan kemudian diurutkan untuk mendapatkan ciri yang terbesar. Nilai vektor eigen ini kemudian dikalikan dengan data keseluruhan citra untuk mendapatkan nilai *eigenLetter* menggunakan persamaan (2.11).



Gambar 4.9. Algoritma penghitungan *eigenLetter*

4.3.2. Penghitungan *ImageLetter*

ImageLetter dihitung ketika fase pengenalan dilakukan terhadap suatu huruf Arab Melayu. Algoritma untuk menghitung nilai *imageLetter* ini dapat dilihat pada gambar 4.10.



Gambar 4.10. Algoritma penghitungan *imageLetter*

Pada gambar 4.10, *imageLetter* dihitung menggunakan persamaan (2.12).

4.3.3. Penghitungan *Eucledian Distance*

Eucledian distance baru dapat dihitung ketika nilai *eigenLetter* dan *imageLetter* telah diperoleh. Penghitungannya menggunakan persamaan (2.10). Nilai terkecil dari *Eucledian distance* ini akan memberikan citra yang bersesuaian dengan citra yang akan dikenali.

Berikut algoritma untuk menghitung *Eucledian distance* ini :

```

for i=1 to N2
    for j=1 to M
        Hasil(i,j) ← eigenLetter(i,j)-imageLetter(i,1);
    end
end
  
```

```

for h=1:M
    for i=1 to N2
        Jarak(1,h) ← Jarak(1,h)+(Hasil(i,h)*Hasil(i,h))
    end
    Jarak2(1,h) ← sqrt(Jarak(1,h));
end

```

4.3.4. Hasil Pengenalan

Hasil pengenalan terhadap citra yang ada pada Tabel 4.1 dapat dilihat pada Tabel 4.2.

Tabel 4.2. Hasil pengenalan terhadap huruf Tabel 4.1

No	Huruf Arab Melayu yang dikenali	Hasil pengenalan seharusnya	Hasil pengenalan terjadi	Deskripsi	<i>Euclidian distance</i>
1	ا	A	A	Sukses	0
2	ل	A	A	Sukses	8.8×10^{-13}
3	م	B	B	Sukses	7.5×10^{-13}
4	ن	B	B	Sukses	0
5	ه	B	B	Sukses	0
6	ت	T	T	Sukses	0
7	ز	T	T	Sukses	0
8	ب	T	T	Sukses	0
9	س	S	S	Sukses	0
10	د	S	S	Sukses	0
11	ف	S	S	Sukses	0
12	ج	J	J	Sukses	0
13	خ	J	J	Sukses	0
14	چ	J	J	Sukses	0
15	ڦ	C	C	Sukses	0
16	ڻ	C	C	Sukses	7.4×10^{-13}
17	ڦ	C	C	Sukses	0
18	ڦ	H	H	Sukses	0

No	Huruf Arab Melayu yang dikenali	Hasil pengenalan seharusnya	Hasil pengenalan terjadi	Deskripsi	<i>Eucledian distance</i>
50	ع	G	G	Sukses	0
51	خ	G	G	Sukses	0
52	ص	G	G	Sukses	0
53	ظ	Ng	Ng	Sukses	0
54	ڙ	Ng	Ng	Sukses	0
55	ڦ	Ng	Ng	Sukses	0
56	ڦ	P	P	Sukses	0
57	ڦ	P	P	Sukses	0
58	ڦ	P	P	Sukses	0
59	ڦ	Q	Q	Sukses	0
60	ڦ	Q	Q	Sukses	0
61	ڦ	Q	Q	Sukses	0
62	ڦ	K	K	Sukses	0
63	ڦ	K	K	Sukses	0
64	ڦ	K	K	Sukses	0
65	ڦ	G	G	Sukses	0
66	ڦ	G	G	Sukses	0
67	ڦ	G	G	Sukses	0
68	ڦ	L	L	Sukses	0
69	ڦ	L	L	Sukses	0
70	ڦ	L	L	Sukses	0
71	ڦ	M	M	Sukses	0
72	ڦ	M	M	Sukses	0
73	ڦ	M	M	Sukses	0
74	ڦ	N	N	Sukses	0
75	ڦ	N	N	Sukses	0
76	ڦ	N	N	Sukses	0
77	ڦ	W	W	Sukses	0
78	ڦ	W	W	Sukses	0

No	Huruf Arab Melayu yang dikenali	Hasil pengenalan seharusnya	Hasil pengenalan terjadi	Deskripsi	<i>Eucledian distance</i>
79	ا	H	H	Sukses	0
80	هـ	H	H	Sukses	0
81	كـ	H	H	Sukses	0
82	لـ	Y	Y	Sukses	0
83	مـ	Y	Y	Sukses	0
84	نـ	Y	Y	Sukses	0
85	ـ	Ny	Ny	Sukses	0
86	ـ	Ny	Ny	Sukses	0
87	ـ	Ny	Ny	Sukses	0

Selanjutnya hasil pengenalan terhadap huruf Arab Melayu hasil segmentasi untuk input gambar 4.1(a) (lihat gambar 4.4) dapat dilihat pada tabel 4.3, dan hasil pengenalan untuk input gambar 4.1(b) dapat dilihat pada tabel 4.4. Sedangkan hasil pengenalan terhadap input gambar 4.1(c) dituangkan pada tabel 4.5.

Keseluruhan proses pengenalan sukses dilakukan, namun terdapat perbedaan nilai *Eucledian distance* yang cukup signifikan. Nilai *Eucledian distance* yang relatif paling kecil terdapat pada tabel 4.2 karena semua huruf yang dikenali merupakan citra-citra yang digunakan pada fase *training*.

Tabel 4.3. Hasil pengenalan terhadap huruf gambar 4.1(a)

Baris ke-	Huruf Arab Melayu yang dikenali	Hasil pengenalan seharusnya	Hasil pengenalan terjadi	Deskripsi	<i>Eucledian distance</i>
1	سـ	K	K	Sukses	922.35
	ـ	M	M	Sukses	0
	ـ	B	B	Sukses	312.66
	ـ	NG	NG	Sukses	226.03

Baris ke-	Huruf Arab Melayu yang dikenali	Hasil pengenalan seharusnya	Hasil pengenalan terjadi	Deskripsi	Eucledian distance
2	ا	H	H	Sukses	256.17
	ن	N	N	Sukses	516.61
	س	C	C	Sukses	426.93
	ر	R	R	Sukses	762.88
3	ك	K	K	Sukses	1.8×10^{-2}
	م	M	M	Sukses	457.98
	ب	B	B	Sukses	500.62
	گ	NG	NG	Sukses	1.0×10^3
	Spasi	Spasi	Spasi	Sukses	0
	ـ	M	M	Sukses	1.1×10^3
	ڪ	K	K	Sukses	0
	,	R	R	Sukses	703.17

Tabel 4.4. Hasil pengenalan terhadap huruf gambar 4.1(b)

Baris ke-	Huruf Arab Melayu yang dikenali	Hasil pengenalan seharusnya	Hasil pengenalan terjadi	Deskripsi	Eucledian distance
1	ـ	S	S	Sukses	1.2×10^{-12}
	ـ	G	G	Sukses	0
	,	R	R	Sukses	876.34
	Spasi	Spasi	Spasi	Sukses	0
	ا	A	A	Sukses	0
	ـ	Spasi	Spasi	Sukses	0
	ـ	D	D	Sukses	0
	ـ	Spasi	Spasi	Sukses	211.45
	ـ	A	A	Sukses	0
	ـ	Spasi	Spasi	Sukses	0
	ـ	T	T	Sukses	0
	ـ	NG	NG	Sukses	0

	5	D	D	Sukses	0
2	Spasi	Spasi	Spasi	Sukses	0
	W	W		Sukses	0
	Spasi	Spasi	Spasi	Sukses	0
	A	A		Sukses	676.12
	Spasi	Spasi	Spasi	Sukses	0
	P	P		Sukses	0
	A	A		Sukses	742.86
	Spasi	Spasi	Spasi	Sukses	102.49
	S	S		Sukses	0
	NG	NG		Sukses	0
3	5	K	K	Sukses	1.8×10^3
	L	L		Sukses	0
	W	W		Sukses	0
	Spasi	Spasi	Spasi	Sukses	0
	A	A		Sukses	1.0×10^3
	Spasi	Spasi	Spasi	Sukses	0
	R	R		Sukses	0
	Spasi	Spasi	Spasi	Sukses	0
	5	K	K	Sukses	1.8×10^3
	M	M		Sukses	0
	P	P		Sukses	0
	NG	NG		Sukses	1.6×10^3

Tabel 4.5. Hasil pengenalan terhadap huruf gambar 4.1(c)

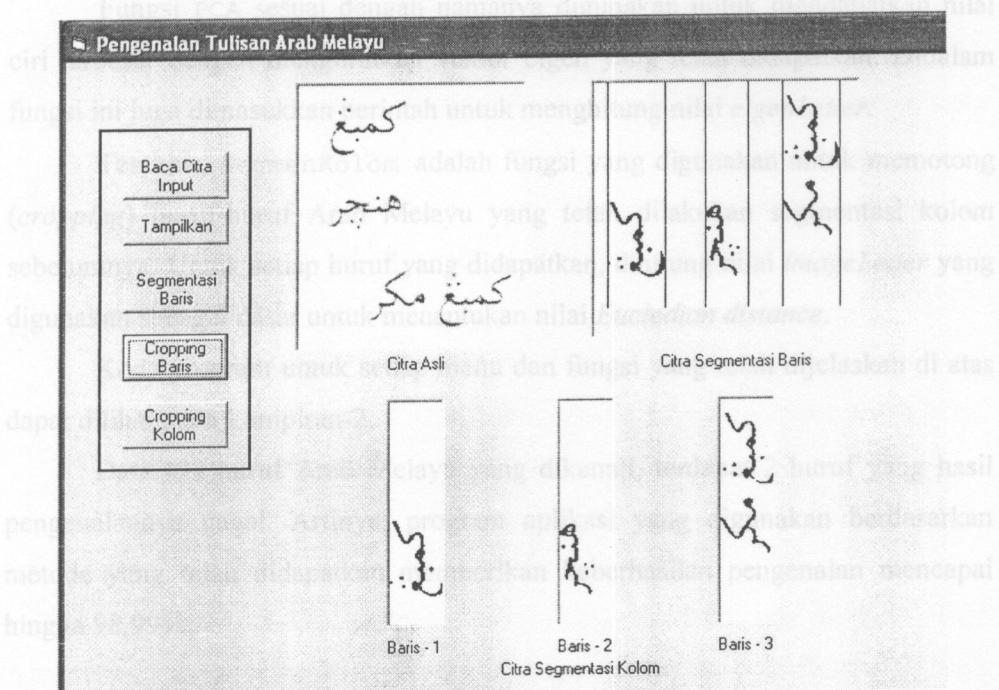
Baris ke-	Huruf Arab Melayu yang dikenali	Hasil pengenalan seharusnya	Hasil pengenalan terjadi	Deskripsi	<i>Euclidian distance</i>
1	:	B	98	Sukses	1.2×10^{-12}
	,	W	105	Sukses	927.16
	Spasi	Spasi	Spasi	Sukses	255
	ؒ	Y	99	Sukses	0
	ؑ	H	100	Sukses	0
	Spasi	Spasi	Spasi	Sukses	0
	ؔ	P	101	Sukses	0
	,	W	78	Sukses	1.2×10^3
	Spasi	Spasi	Spasi	Sukses	0
	ؓ	T	102	Sukses	0
	ؑ	H	100	Sukses	403.36
	Spasi	Spasi	Spasi	Sukses	0
	ؚ	S	91	Sukses	1.2×10^3
	ؔ	P	53	Sukses	1.5×10^3
	,	R	103	Sukses	0
	Spasi	Spasi	Spasi	Sukses	275.29
	ؒ	T	A	Gagal	1.3×10^3
	,	Y	R	Gagal	1.5×10^3
2	Spasi	Spasi	Spasi	Sukses	0
	ؕ	K	104	Sukses	0
	,	W	105	Sukses	0
	Spasi	Spasi	Spasi	Sukses	255
	ؑ	A	89	Sukses	775.89
	Spasi	Spasi	Spasi	Sukses	0
	ؑ	H	100	Sukses	1.2×10^3
	ؑ	A	A	Sukses	768.12
	Spasi	Spasi	Spasi	Sukses	0
	ؔ	D	D	Sukses	0
	Spasi	Spasi	Spasi	Sukses	0

ج	K/Q	K/Q	Sukses	0
Spasi	SPasi	Spasi	Sukses	0
د	D	D	Sukses	769.85
Spasi	Spasi	Spasi	Sukses	218.20
ا	A	A	Sukses	772.41
Spasi	Spasi	Spasi	Sukses	0
ن	N	N	Sukses	0
Spasi	Spasi	Spasi	Sukses	0
ك	K	K	Sukses	1.1×10^3
ا	A	A	Sukses	1.1×10^3
Spasi	Spasi	Spasi	Sukses	0
ك	K	K	Sukses	744.14
ج	K/Q	K/Q	Sukses	1.5×10^3
Spasi	Spasi	Spasi	Sukses	0
د	D	D	Sukses	938.22
Spasi	Spasi	Spasi	Sukses	0
و	W	W	Sukses	996.87
Spasi	spasi	Spasi	Sukses	53.57
د	D	D	Sukses	743.56
Spasi	Spasi	Spasi	Sukses	0
ج	K/Q	K/Q	Sukses	787.12
Spasi	Spasi	Spasi	Sukses	0
د	D	D	Sukses	722.35
Spasi	Spasi	Spasi	Sukses	0
ك	K	K	Sukses	0
ت	T	T	Sukses	0
Spasi	Spasi	Spasi	Sukses	0
ب	B	B	Sukses	884.69
ا	A	A	Sukses	726.65
Spasi	Spasi	Spasi	Sukses	0
پ	P	P	Sukses	703.49
ج	K/Q	K/Q	Sukses	0

Arab Melayu, sebelum menghasilkan file-file yang akan digunakan pada

4.4. Program Aplikasi

Keseluruhan hasil yang telah didapatkan pada bagian-bagian sebelumnya, dihitung dan ditentukan menggunakan sebuah program aplikasi yang dibuat dengan tujuan untuk melakukan simulasi perhitungan. Antarmuka dari program aplikasi tersebut seperti terlihat pada gambar 4.9.



Gambar 4.9. Antarmuka program simulasi

Program pada gambar 4.9 digunakan untuk membaca nilai piksel citra, dan untuk keperluan validasi pembacaan nilai tersebut ditampilkan kembali. Kedua hal ini dilakukan oleh menu Baca Citra Input dan Tampilkan. Menu Segmentasi Baris digunakan untuk melakukan segmentasi baris terhadap citra asli. Hasilnya akan ditampilkan pada Citra Segmentasi Baris. Batas segmentasi untuk setiap baris ditunjukkan oleh garis vertikal berwarna biru.

Untuk memisahkan setiap baris yang telah didapatkan, pemotongannya dilakukan menggunakan menu Cropping Baris. Terakhir, menu Cropping Kolom digunakan untuk melakukan segmentasi kolom pada setiap baris tulisan

Arab Melayu, sekaligus menghasilkan *file-file* yang akan digunakan pada perangkat lunak MATLAB.

Beberapa prosedur dan fungsi yang digunakan di MATLAB adalah: BentukMatrix, PCA, dan SegmenKolom. BentukMatrix digunakan untuk membaca citra-citra yang digunakan sebagai input pada fase *training*. Adapun citra-citra input yang digunakan dapat dilihat pada Lampiran-1.

Fungsi PCA sesuai dengan namanya digunakan untuk mendapatkan nilai ciri terbesar dengan mengurutkan vektor eigen yang telah didapatkan. Didalam fungsi ini juga dimasukkan perintah untuk menghitung nilai *eigenLetter*.

Terakhir, SegmenKolom adalah fungsi yang digunakan untuk memotong (*cropping*) huruf-huruf Arab Melayu yang telah dilakukan segmentasi koiom sebelumnya. Untuk setiap huruf yang didapatkan, dihitung nilai *imageLetter* yang digunakan sebagai dasar untuk menentukan nilai *Eucledian distance*.

Kode program untuk setiap menu dan fungsi yang telah dijelaskan di atas dapat dilihat pada Lampiran-2.

Dari 199 huruf Arab Melayu yang dikenali, terdapat 2 huruf yang hasil pengenalannya gagal. Artinya, program aplikasi yang digunakan berdasarkan metode yang telah didapatkan memberikan keberhasilan pengenalan mencapai hingga 98,99%.