



UNIVERSITI PUTRA MALAYSIA

**PENGECAMAN TULISAN TANGAN TEKSJAWI MENGGUNAKAN
PENGKELAS MULTIKLAS**

KHAIRUDDIN BIN OMAR

FSKTM 2000 1

**PENGECAMAN TULISAN TANGAN TEKS JAWI MENGGUNAKAN
PENGKELAS MULTITARAS**

Oleh

KHAIRUDDIN BIN OMAR

**Tesis ini dikemukakan sebagai memenuhi keperluan bagi mendapatkan Ijazah
Doktor Falsafah di Fakulti Sains Komputer dan Teknologi Maklumat
Universiti Putra Malaysia
April 2000**



DEDIKASI

Dengan nama Allah yang Maha Pemurah lagi Maha Mengasihani

Penulis ingin merakamkan jutaan terima kasih di atas pengorbanan serta jasa kedua ibu-bapa yang telah bersusah payah membesarkan penulis sehingga apa yang penulis kecapai hari ini. Kedua-dua ibu bapa penulis telahpun kembali ke RahmatuLlah. Bapa penulis telah kembali ke RahmatuLlah pada 28 Julai 1988 di Tanah Suci Mekah manakala ibu penulis pada 27 Mac 2000 di Johor Bahru sebelum sempat penulis membentangkan tesis ini. Semoga roh mereka berdua dicucuri rahmat dan di tempatkan di kalangan orang-orang yang soleh. Jasa mu tetap dikenang.

Dedikasi ini ditujukan kepada isteri, Nor Hashimah binti Abu Bakar, dan anak-anak; Khairussyahidah, Khairunnajwa, Khairul Aiman, dan Khairunnur Aqilah, adik-adik di Johor Bahru; serta ibu dan bapa mertua di Mentakab, Pahang di atas pengorbanan, dorongan, sokongan, dan kesabaran yang diberikan sepanjang pengajian ini. Semoga Allah memberikan ganjaran yang setimpal di atas pengorbanan tersebut.



Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah.

**PENGECAMAN TULISAN TANGAN TEKS JAWI MENGGUNAKAN
PENGKELAS MULTIKLAS**

Oleh

KHAIRUDDIN BIN OMAR

April 2000

Pengerusi: Ramlan bin Mahmud, Ph.D.

Fakulti: Sains Komputer dan Teknologi Maklumat

Pengecaman tulisan tangan teks Jawi adalah satu tugas yang sangat mencabar di dalam bidang Pengecaman Aksara Optik (PAO) disebabkan Jawi adalah satu tulisan jenis bersambung. Tesis ini mengenegahkan teknik untuk memperbaiki kadar pengecaman teks Jawi tulisan tangan. Skema baru yang lebih cekap untuk prapemprosesan, penemberengan, penyarian fitur dan pormalan aksara, dan pengelasan telah direka untuk memenuhi objektif tersebut. Untuk prapemprosesan, kaedah pembetulan pencongan dan erotan menggunakan kaedah *histogram orientasi cerunan* (HOC) yang asalnya digunakan untuk dokumen Latin telah dimasukkan sebagai satu daripada modul prapemprosesan. Satu skema baru untuk penemberengan telah diperkenalkan. Ia berasaskan kepada gabungan kaedah unjuran profail histogram dan penentuan titik tembereng ubah suai (PTT) membentuk kaedah *penentuan titik tembereng* (PTT). Fitur-fitur disarikan daripada aksara yang telah ditemberengkan menggunakan tiga jenis fitur. Fitur-fitur ini ialah struktur, fitur *Momen Tak-berubah* (MTB) dan *Taburan Pikel Hitam* (TPH). Algoritma penyingkiran bahagian sekunder aksara Jawi (seperti titik-titik, “Á” “ dan maddah) juga telah diperkenalkan supaya dapat mengelakkan daripada salah cam sekunder ini.

Ia perlu dipisahkan terlebih dahulu sebelum melalui proses pengecaman. Hal ini dapat mengurangkan bilangan kelas aksara Jawi daripada 124 kepada 60. Sebanyak 200 sampel setiap kelas aksara Jawi telah diujikan untuk tujuan pengkelasan. Dua aras sistem pengkelasan terdiri daripada *Pengkelas Kumpulan berasaskan Ukuran Keserupaan* (PKUK) dan *Pengkelas berganda Genetik-Perambat-balik* (PGPB). Di aras pertama, PKUK menggunakan fitur struktur dan MTB untuk mengelompokkan kesemua aksara. Tujuh jenis primitif diperoleh menggunakan fitur struktur, dan proses pengelompokan berdasarkan kepada jenis primitif ini. Fitur MTB digunakan untuk mengirakan ukuran keserupaan dan kemudian menentukan kadar pengkelasan untuk setiap kumpulan. Setelah kesemua sampel aksara telah dikelompokkan, PGPB digunakan untuk mengelaskan setiap aksara dalam kumpulan masing-masing dan dilarikan secara berasingan. Kelas aksara yang terbanyak ialah 14 aksara. Di aras kedua, PGPB dilaksanakan dalam dua peringkat iaitu peringkat pembelajaran, dan peringkat ujian. Di peringkat pembelajaran, pengkelasan ini menggunakan fitur MTB dan TPH, manakala di peringkat ujian pengkelas ini menggunakan maklumat tambahan iaitu maklumat yang diperoleh ketika menyingkirkan jujuk sekunder, dan di samping fitur MTB serta TPH. Pemecahan masalah ini kepada dua aras telah mengurangkan masa pembelajaran yang diambil oleh pengkelas dan berjaya menambah kadar pengecaman. Tesis ini membicarakan secara terperinci setiap algoritma dan prestasinya terhadap sampel yang digunakan didalam ujikaji. Perbandingan juga dibuat terhadap kaedah pengawalan pemberat PB menggunakan pendekatan Sifar, Rawak, serta Rawak Nguyen-Widrow, di samping pendekatan ubah suai AG. Prestasi menggunakan AG (ubah suai) memberikan hasil pengkelasan yang dijanjikan.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirements for the degree of Doctor of Philosophy.

HANDWRITTEN JAWI TEXT RECOGNITION USING MULTILEVEL CLASSIFIER

By

KHAIRUDDIN BIN OMAR

April 2000

Chairman : Ramlan bin Mahmod, Ph. D.

Faculty : Computer Science and Information Technology

Recognizing Jawi handwritten text is a difficult task in the area of optical character recognition (OCR) since Jawi is a cursive type language. This thesis deals with techniques for improving the recognition rate of a cursive script Jawi text recognition system. A new efficient scheme for preprocessing, segmentation, features extraction and normalization of characters, and classification has been designed and implemented to achieve this objective. For preprocessing, skew and slant correction methods using the gradient orientation histogram (GOH) that are originally solved for Latin documents are included as one of the preprocessing modules. A new scheme for segmentation is proposed and its based on combinations of histogram profile projection and modified Determination of Segmentation Points methods to form Determination of Segmentation Points method (DSP). Features are extracted from segmented characters using three type of features, these are structure, Moment Invariant (MI), and Black Pixel Distribution (BPD). Removal of secondaries for Jawi character (such as dots as well as the other secondaries alif and maddah) are also introduced to avoid misclassifications due to these secondaries. They are separated before the characters are passed to the recognition process. This

removal is also reduced the number of classes to be processed from 124 to 60. Two hundred samples for each Jawi character classes are tested for classification purposes. The two level classification system consist of a group classifier based on similarity measure (GCSM) and multi classifier of Genetic Backpropagation (GBPC). At the first level, the GCSM used the structured and MI features to cluster all the characters. Seven types of primitives are obtained by using structured features, and the clustering process is based on these type of primitives. MI features are also used to obtain similarity measurement and to determine the classification rate for each group. After clustering all the sampels, the GBPC is used to classify all the characters within its group and implementation is done separately. The largest groups include only 14 characters of classes. At the second level, the GBPC is implemented in two steps, these include learning step, and testing step. In learning step, the classifier uses the MI and BPD features, while at the testing step, it uses the additional information previously recorded about the secondaries, as well as the MI and BPD features. By breaking this type of application into two levels of classification, the system has shown that the learning time taken was reduced and the recognition rate was increased. This thesis dicussed the details of each algorithm and its performance on the samples used in the experiment. We also present a comparison of several type of weight initializing for training BP such as zero value, random, and Nguyen-Widrow, as well as modified GA. The Performance of GA (modified) shows the promising results for classifications.



PENGHARGAAN

Pertamanya penulis ingin mengucapkan syukur kehadiran Allah s.w.t. yang telah memberikan kekuatan, kesabaran, dorongan, dan haluan untuk menyempurnakan penulisan tesis ini.

Seterusnya, penulis ingin merakamkan setinggi-tinggi penghargaan kepada Jawatan Kuasa Penyeliaan yang dipengerusikan oleh Dr. Ramlan bin Mahmod di atas bimbingan, tunjuk ajar, motivasi dan nasihat yang amat berguna sepanjang penyelidikan ini dijalankan. Penulis juga ingin merakamkan jutaan terima kasih kepada ahli Jawatan Kuasa Penyeliaan yang terdiri daripada Dr. Abd Rahman bin Ramli dan Dr. Md. Nasir bin Sulaiman di atas segala nasihat, bimbingan, komen dan sokongan mereka.

Di kesempatan ini juga penulis ingin mengucapkan terima kasih kepada semua rakan-rakan seperjuangan di atas dorongan yang telah diberi, terutamanya saudara Zurni, saudari Malathi, abang Ali, abang Rohamat, Shahidan, pensyarah-pensyarah di Fakulti ini, adik-adik di bilik siswazah seperti Ehsan, Kamaruzzaman, Azmi, Kamal, dan ramai lagi seperti Ummu Salmah yang banyak membantu persekitaran dan corak pembelajaran di Fakulti ini.

Tidak lupa juga kepada pihak Penaja iaitu Universiti Kebangsaan Malaysia yang telah membiayai segala perbelanjaan sepanjang pengajian ini. Ini termasuklah pembiayaan ketika penulis membuat lawatan sambil menjalankan penyelidikan pendik di Universiti Loughborough, UK selama 3 bulan 2 minggu pada 23hb.

Februari hingga 6 Jun 1997. Lawatan tersebut telah banyak mempengaruhi hasil penyelidikan yang dihasilkan ini. Juga kepada Fakulti Sains Komputer dan Teknologi Maklumat yang telah mengizinkan penggunaan segala kemudahan peralatan komputer termasuklah Internet.

Saya mengesahkan bahawa Jawatankuasa Pemeriksa bagi Khairuddin bin Omar telah mengadakan pemeriksaan akhir pada 11 April 2000 untuk menilai tesis Doktor Falsafah beliau yang bertajuk “Pengecaman Tulisan Tangan Teks Jawi Menggunakan Pengkelas Multiaras” mengikut Akta Universiti Pertanian Malaysia (Ijazah Lanjutan) 1980 dan Peraturan-Peraturan Universiti Pertanian Malaysia (Ijazah Lanjutan) 1981. Jawatankuasa Pemeriksa memperakukan bahawa calon ini layak dianugerahkan ijazah tersebut. Anggota Jawatankuasa Pemeriksa adalah seperti berikut:

Hamidah binti Ibrahim, Ph.D.
Wakil Dekan Pusat Pengajian Siswazah
Universiti Putra Malaysia
(Pengerusi)

Ramlan bin Mahmud, Ph.D.
Fakulti Sains Komputer dan Teknologi Maklumat
Universiti Putra Malaysia
(Ahli)

Abd. Rahman bin Ramli, Ph.D.
Fakulti Kejuruteraan
Universiti Putra Malaysia
(Ahli)

Md. Nasir bin Sulaiman, Ph.D.
Fakulti Sains Komputer dan Teknologi Maklumat
Universiti Putra Malaysia
(Ahli)



MOHD. GHAZALI MOHAYIDIN, Ph.D.
Professor
Timbalan Dekan Pusat Pengajian Siswazah
Universiti Putra Malaysia

Tarikh: **01 JUN 2000**

Tesis ini telah diserahkan kepada Senat Universiti Putra Malaysia dan telah diterima sebagai memenuhi keperluan untuk ijazah Doktor Falsafah.

Kamis Awang

KAMIS AWANG, Ph.D.

Profesor Madya

Dekan Pusat Pengajian Siswazah

Universiti Putra Malaysia

Tarikh: 13 JUL 2000



Saya mengaku bahawa tesis ini adalah hasil kerja saya yang asli melainkan petikan dan sedutan yang telah diberi penghargaan di dalam tesis. Saya juga mengaku bahawa tesis ini tidak dimajukan untuk ijazah-ijazah lain di Universiti Putra Malaysia atau di institusi-institusi lain.



(Khairuddin bin Omar)

Tarikh: 31 MEI 2000

KANDUNGAN

Muka Surat

DEDIKASI	ii
ABSTRAK	iii
ABSTRACT	v
PENGHARGAAN	vii
LEMBARAN PENGESAHAN	ix
PERNYATAAN KEASLIAN	xi
SENARAI JADUAL	xvii
SENARAI RAJAH	xix
SENARAI NAMA SINGKATAN	xxvi

BAB

I	PENDAHULUAN	1
	Pengenalan.....	1
	Latar Belakang Masalah.....	4
	Objektif Kajian.....	9
	Sumbangan Kajian.....	10
	Skop Kajian.....	11
	Struktur Organisasi Tesis.....	13
II	ULASAN KARYA	16
	Pengenalan.....	16
	Sejarah Perkembangan Pengecam Teks Optik Arab/Jawi.....	16
	Ciri-ciri Teks Jawi/Arab.....	23
	Keupayaan Satu Sistem PTOA.....	27
	Model Umum PTOA.....	30
	Prapemprosesan.....	31
	Latar Belakang Prapemprosesan.....	35
	Proses Perduaan (Ambang).....	35
	Langkah Bersyarat/Penyaman-Penipisan dan Kelicinan... ..	36
	Penipisan.....	37
	Pernormalan.....	40
	Garis Tapak.....	40
	Pengesanan Sudut Pencong.....	41
	Penemberengan.....	41
	Beberapa Konsep Penting dalam Penemberengan.....	43
	Pemecahan Halaman.....	44
	Penemberengan Perkataan.....	46
	Penyarian Fitur.....	47
	Kaedah-kaedah Pengkelasan.....	49
	Kaedah Sintaktik.....	49

	Kaedah Statistik.....	50
	Rangkaian Neural.....	52
	Pendekatan Hibrid.....	54
	Pembelajaran.....	56
	Pemprosesan Akhir.....	57
	Contoh Kaedah Penemberengan Aksara Arab.....	58
	Pendekatan Histogram.....	58
	Penemberengan Baris Teks.....	58
	Penemberengan Perkataan.....	60
	Penemberengan Aksara.....	61
	Berasaskan Jumlah Lajur.....	62
	Berasaskan Pertimbangan Jarak di antara Dua Puncak Lampau.....	62
	Pembaikan Lanjutan.....	63
	Pendekatan Penentuan Titik Tembereng Utama.....	64
	Pendekatan Pewarnaan Komponen Terkait.....	68
	Pendekatan Penentuan Titik Tembereng.....	71
	Ciri-ciri dan Keupayaan Sistem PTOA.....	74
	Kesimpulan.....	74
III	METODOLOGI DAN REKABENTUK PTOJ	78
	Pengenalan	78
	Struktur Senibina PTOJ.....	81
	Modul Prapemprosesan.....	82
	Sub-Modul Pengimbas Imej.....	83
	Sub-Modul Penapisan Hingar.....	83
	Sub-Modul Proses Ambang.....	83
	Sub-Modul Pengesan dan Pembetulan Pencongan.....	83
	Sub-Modul Penipisan.....	83
	Sub-Modul Penemberengan.....	83
	Modul Pernormalan dan Penyarian Fitur Aksara Jawi.....	84
	Sub-Modul Translasi	84
	Sub-Modul Penyingkir Juzuk Sekunder.....	84
	Sub-Modul Penskilan.....	84
	Sub-Modul Kelicinan.....	85
	Sub-Modul Penyari Fitur Struktur.....	85
	Sub-Modul Penyari Fitur MTB.....	85
	Sub-Modul Penyari Fitur TPH.....	85
	Modul Pengkelasan.....	85
	Sub-Modul PKUK.....	86
	Sub-Modul PGPB.....	86
	Sub-Modul Aplikasi.....	86
	Ringkasan.....	86
IV	PRAPEMPROSESAN IMEJ JAWI TULISAN TANGAN	87
	Pengenalan.....	87
	Keterangan Ujikaji.....	88
	Pemerolehan Data.....	89
	Penyingkiran Hingar - Penapisan Imej Jawi.....	89
	Proses Ambang.....	90



	Pengesan dan Penormalan Pencongan dan Erotan Teks Jawi.....	92
	Pembetulan Erotan.....	97
	Penipisan.....	100
	Penemberengan atau Pisahan Teks Jawi.....	104
	Penemberengan Baris Teks.....	106
	Penemberengan Perkataan.....	107
	Penemberengan Aksara.....	112
	Titik Tembereng Berpotensi.....	114
	Sifat-sifat Titik Tembereng Data Ujian.....	117
	Penemberengan Kasus Ke I.....	120
	Penemberengan Kasus Ke II.....	122
	Hasil Ujikaji Penemberengan.....	124
	Ringkasan.....	127
V	PERNORMALAN DAN PENYARIAN FITUR AKSARA.....	129
	Pengenalan.....	129
	Skema Pernormalan Aksara Jawi.....	130
	Translasi.....	131
	Penyingkiran Bahagian Sekunder Aksara Jawi.....	132
	Penskilan.....	140
	Kelicinan.....	141
	Hasil Uji Kaji Pernormalan Aksara Jawi.....	141
	Penyarian Fitur	145
	Penyarian Fitur Struktur.....	147
	Penyusoran Kontur.....	150
	Pengahampiran Poligon.....	151
	Kecembungan Poligon.....	155
	Hasil Uji Kaji Penyarian Fitur Struktur.....	156
	Penyarian Fitur Momen Tak-Berubah.....	160
	Penyarian Fitur Taburan Pikel Hitam.....	162
	Ringkasan.....	168
VI	KAEDAH PENGECAMAN.....	168
	Pengenalan.....	168
	Pengkelasan Mengikut Kumpulan.....	169
	Ukuran Keserupaan.....	171
	Pokok Kata Putus.....	173
	RN Perambat-balik (PB) – Pengenalan.....	175
	Seni Bina RKMA.....	176
	Fungsi Pengaktifan.....	177
	Model Latihan PB.....	178
	Pilihan Pengawalan Pemberat dan Pincang.....	182
	Pengawalan Pemberat Rawak.....	183
	Pengawalan Pemberat Rawak Nguyen-Widrow.....	184
	Algoritma PB.....	185
	Algoritma Latihan.....	185
	Algoritma Aplikasi.....	188
	Faktor-faktor yang Mempengaruhi Penumpuan Rangkaian.....	189
	Pengkelasan Berganda bagi RKMA.....	193
	Algoritma Genetik: Prinsip dan Fitur Asas.....	196

Penilaian Keupayaan dalam AG.....	199
Prinsip dan Fitur Asas.....	200
Pengawalan Pemberat melalui Pengoptimuman AG.....	201
Awalan- Merawakkan VP.....	203
Pengiraan Nilai Keupayaan.....	203
Pilihan - Memilih VP untuk Terus Hidup.....	206
Pengeluaran Semula- Membina VP Baru dari VP Lama....	207
Membentuk Set Latihan untuk Aras berikutnya.....	208
Ringkasan.....	208
VII PENGECAMAN AKSARA JAWI TULISAN TANGAN.....	210
Pengenalan.....	210
Reka Bentuk Pengkelas Aksara Jawi Tulisan Tangan.....	212
Pengkelas Peringkat Pertama: Pengkelas Kumpulan.....	216
Pengkelas Peringkat Kedua: PB.....	219
Pengkelas Peringkat Ketiga: Algoritma Aplikasi PB (Ubah Suai)..	220
Spesifikasi Uji Kaji.....	230
Skema Rangkaian.....	230
Parameter-parameter PB.....	232
Parameter-parameter AG.....	232
Skema Baru AG.....	235
Pengeluaran Semula – Membina VP Baru dari VP Lama...	236
Implementasi.....	239
Hasil Ujikaji.....	247
Pengkelas Kumpulan.....	247
Pengoptimuman Pemberat AG.....	250
Prestasi Rangkaian PB (Fasa Latihan).....	254
Ringkasan.....	273
VIII KESIMPULAN DAN PERBINCANGAN.....	277
Pengenalan.....	277
Ringkasan Hasil dan Perbincangan.....	278
Modul Prapemprosesan Teks Jawi.....	278
SubModul Penemberengan Teks Jawi.....	279
Modul Penyarian Fitur dan Pernormalan Aksara Jawi.....	280
Modul Pengkelasan.....	281
Kekangan dan Limitasi.....	284
Cadangan dan Pembaikan.....	285
BIBLIOGRAFI.....	287
LAMPIRAN.....	301
A Contoh-contoh Teks Jawi.....	305
B-1 Contoh Data Latihan.....	306
B-2 Contoh Data Ujian	308
C-1 Matrik Keliru Kumpulan K1.....	310
C-2 Matrik Keliru Kumpulan K2.....	312
C-3 Matrik Keliru Kumpulan K3.....	313
C-4 Matrik Keliru Kumpulan K4.....	316
C-5 Matrik Keliru Kumpulan K5.....	317
C-6 Matrik Keliru Kumpulan K6.....	318

C-7 Matrik Keliru Kumpulan K7..... 319

BIODATA DIRI..... 321



SENARAI JADUAL

Jadual		Muka Surat
1	Aksara Jawi dalam semua bentuk (bentuk hujung BH, mula BM, tengah BT, dan tunggal BTG).....	24
2	Ciri-ciri dan Keupayaan beberapa Sistem PTOA.....	76
3	Hasil Penemberengan Algoritma Cadangan.....	124
4	Saiz Keseluruhan Imej Teks Jawi setelah melalui Proses Prapemprosesan.....	127
5	Kumpulan Aksara Jawi: Bilangan Primitif.....	158
6	Kumpulan Aksara Jawi: Penyusunan Semula.....	217
7	Kumpulan Aksara Jawi: Pengkelasan.....	221
8	Saiz Sub-pengkelas PB.....	231
9	Nilai Jangkaan Kumpulan $M^{(j)}$.	248
10	Hasil Pengkelasan Kumpulan.....	249
11	Generasi Penamat bagi Keseluruhan Kumpulan Primitif.....	251
12	Set Pemberat AG Kumpulan 1 (aras input-tersembunyi).....	252
13	Set Pemberat AG Kumpulan 1 (aras tersembunyi-output).....	253
14	Hasil PB untuk Kumpulan 1.....	255
15	Hasil PB untuk Kumpulan 2.....	257
16	Hasil PB untuk Kumpulan 3.....	259
17	Hasil PB untuk Kumpulan 4.....	261
18	Hasil PB untuk Kumpulan 5.....	263
19	Hasil PB untuk Kumpulan 6.....	265
20	Hasil PB untuk Kumpulan 7.....	267

21	Kadar Pengecaman setiap kumpulan untuk setiap Kaedah Pengawalan Pemberat.....	272
----	---	-----



SENARAI RAJAH

Rajah		Muka Surat
1	Bentuk Lazim Ciri-ciri Aksara Arab/Jawi (a) Bentuk 1: Tindanan Memugak. (b) Bentuk 2: Ligatur. (c) Bentuk 3: Berbaris. (d) Bentuk 4: Bersambung pada satu baris.....	6
2	Penulisan Jawi adalah berbentuk sambung. Satu perkataan Jawi dengan (a) bersambung sepenuhnya; (b) separa; (c) tidak bersambung sepenuhnya.....	23
3	Terdapat lebih kurang 21 aksara Jawi yang mempunyai titik, atau lengkok.....	25
4	Juzuk-juzuk sekunder dalam penulisan Jawi (a) Dua jenis aksara yang sama gelung tetapi mempunyai bilangan titik yang berbeza; (b) Dua aksara yang sama bilangan puncaknya tetapi berbeza bilangan titiknya; (c) garis memugak yang tidak menyentuh gelung; (d) garis memugak yang menyentuh gelung; (e) aksara ‘ا’ menyentuh aksara ‘ا’ menyentuh aksara ‘ج’; (g) garis condong tidak menyentuh juzuk utama aksara ك dan (h) garis condong menyentuh juzuk utama dalam aksara ك.....	25
5	Aksara bertindan. (a) Perkataan ‘amal’, dibina dari aksara “ع”, “م”, dan “ل” untuk membentuk ligatur; (b) perkataan ‘dan’, dibina dari aksara “د”, “ا”, “و”, dan “ن” membentuk pertindanan memugak dengan “ا” di atas “د”; (c) pertindanan memugak dengan “ا” di tengah-tengah “د”.....	26
6	Sifat Saiz dan Ketebalan aksara Jawi. Berbeza-beza dari segi panjang dan tebalnya. (a) perkataan ‘hutang’, dibina dari aksara “ه”, “ر”, “ت”, dan “غ”; (b) perkataan ‘hutang’ membentuk lurah senyap, seperti pada tembereng yang menghubungkan aksara “ه” dan “ر”.....	26
7	Garis Tapak.....	26
8	Satu Contoh Sistem PTOA dengan Lima Peringkat Operasi.....	30

9	Penemberengan Baris Teks menggunakan Pendekatan Khella(1992).....	59
10	Satu contoh pemisahan baris oleh kaedah Khella(1992). (a) Hasil daripada empat langkah pertama. (b) dan (c) hasil daripada langkah 6.....	59
11	Satu Contoh Penemberengan Melalui Pendekatan Khella(1992) yang pertama.....	63
12	Penemberengan Baris-baris Teks oleh pendekatan Romeo-Pekker <i>et al.</i> (1995). (a) Mengesan baris tulisan. (b) Kontur bagi perkataan-perkataan yang bertindan	65
13	Mengesan Garis Tapak.....	66
14	Penemberengan Aksara oleh pendekatan Romeo-Pekker <i>et al.</i> (1995). (a) Sebelum fasa pembetulan. (b) Selepas fasa pembetulan.....	66
15	Penemberengan Aksara oleh pendekatan Romeo-Pekker <i>et al.</i> (1995) menggunakan kaedah kontur sebelah atas. (a) Kontur sebelah atas perkataan. (b) Kontur bertapis bagi perkataan yang sama. (c) Penemberengan perkataan.....	67
16	Kejadian sudut pada aksara yang bersambung. (a) Empat bentuk bagi aksara “ح “ yang berbeza iaitu tunggal, hujung, tengah, dan mula. (b) Contoh dua perkataan Arab, arah anak panah menunjukkan sudut terbentuk daripada cantuman aksara-aksara.....	71
17	Syarat-syarat penemberengan oleh pendekatan Bushofa dan Spann(1997). (a) Tetingkap 7 X 7 digunakan untuk diletakkan pada titik di antara dua aksara yang hendak dipisahkan; (b) Tetingkap 3 X 3 digunakan untuk memisahkan aksara “ح “ (x: tidak kira apa nilai sekalipun, y:’1’ untuk aksara “ع “ atau aksara “ـ “ atau “0” selainnya).....	72
18	Contoh hasil penemberengan pendekatan Bushofa dan Spann(1997). (a) Satu contoh teks Arab. (b) Penemberengan ke atas teks yang sama, arah anak panah menunjukkan kegagalan algoritma ini.....	73
19	Struktur Pengecaman Teks Jawi Tulisan Tangan.....	81
20	Prapemprosesan Imej Jawi.....	88
21	Satu Contoh teks Jawi. (a) Teks Asal. (b) Imej Mentah. (c)	

	Imej sesudah ditapis. (d) Imej sesudah proses ambang.....	91
22	Topeng Penjejak Sisi Sobel. (a) operator q dan (b) operator p.....	94
23	Histogram Orientasi Cerunan.....	95
24	Satu contoh imej teks Jawi (a) Imej terpencong sebanyak 14° . (b) Histogram orientasi cerunan imej Jawi dalam (a).....	96
25	Tranformasi Ricih Imej Jawi.....	98
26	Satu contoh perkataan Jawi (a) Imej dengan keadaan ricih. (b)Histogram orientasi cerunan. (c) Pembetulan ricih perkataan (sudut ricih ialah 15°).....	99
27	Imej Jawi setelah diperbetulkan sebanyak 14°	100
28	Penipisan. (a) Satu contoh Jawi asal sebelum dinipiskan. (b) Imej dinipiskan menggunakan algoritma Naccache dan Shinghal (lelaran = 7).....	103
29	Imej dinipiskan menggunakan algoritma Naccache dan Shinghal (lelaran = 15).....	104
30	Satu contoh imej Jawi. (a) Imej Teks Jawi yang telah dinipiskan. (b) Sesudah baris ditemberengkan. (c) Sesudah perkataan ditemberengkan.....	108
31	Kegagalan Pendekatan Khella. (a) aksara “ج” bertindan di bawah aksara “ب”. (b) aksara “ج”, “ر”, dan “س” bertindan di antara satu sama lain.....	109
32	Proses pemisahan subtembereng berdasarkan algoritma PKT.....	111
33	Perkataan Jawi Asal. (a) Perkataan “شَمْفُو” yang bermaksud syampu. (b) Perkataan “هَوْتَع” yang bermaksud hutang.....	112
34	Kegagalan dalam pendekatan Khella. (a) aksara bergelung (aksara ب dan ر) dan bertitik (aksara ث). (b) aksara gelung (aksara هـ).....	113
35	Satu contoh perkataan bahasa Malaysia yang ditulis dalam tulisan Jawi (disebut “sampu” atau سَمْفُو). Titik tembereng tidak sering berlaku pada garis tapak malah kadang kala boleh	

	berlaku di atas atau di bawah garis tapak yang dikirakan. Garis tapak yang diperoleh daripada kaedah histogram berlaku pada kedudukan piksel ke 31, manakala titik tembereng berlaku pada kedudukan 83, 142, dan 198 yang diperoleh daripada algoritma cadangan yang akan dijelaskan dalam bahagian berikut.....	115
36	Satu imbasan imej Jawi secara mengufuk dibuat di sepanjang garis tapak dan kemudian satu lagi imbasan secara memugak dibuat untuk mencari titik pusat piksel P sama ada ke atas atau ke bawah sehingga menemui piksel yang paling berpotensi iaitu TTB seperti yang dipaparkan dalam templet bersaiz 7x7. (a) Titik TTB berlaku pada pusat P. (b) Titik TTB berlaku bukan pada titik P.....	116
37	Jenis corak cantuman utama bagi aksara Jawi tulisan tangan.....	117
38	Contoh Konfigurasi Cantuman Piksel. (a) Cantuman dalam Rajah 37 (a) dan (b) . (b) Rajah 37(c) . (c) Rajah 37(d) . (d) Rajah 37(e) dan (f)	118
39	Tetingkap Uji. p menunjukkan piksel pusat.....	118
40	Hasil penemberengan imej dalam Rajah 33 menggunakan algoritma cadangan.(a) Tanpa anjakan ke kanan. (b) Dengan anjakan.....	121
41	Bentuk Ligatur. (a) Perkataan Jawi Asal. (b) Hasil Algoritma II.....	123
42	Contoh Aksara Cantuman.....	126
43	Skema Baru Algoritma Pernormalan Aksara Jawi.....	130
44	Aksara Jawi dengan Juzuk Utama Asas dan Juzuk Sekunder....	132
45	Perubahan tanda. (a) Dalam arah mengufuk dan memugak (x, y). (b) Jarak di antara dua titik pada perubahan tanda.....	136
46	Aksara “ ﷲ ” dengan tiga kontur.....	137
47	Aksara “ ﷲ ” dengan tiga kontur.....	138
48	Aksara “ ﷲ ” selepas penyingkiran juzuk sekundernya.....	139
49	Aksara “ ﷲ ” selepas penyingkiran juzuk sekundernya.....	139

50	Aksara “د”. (a) Imej asal. (b) Penskilan 16 x 16 sebelum dilicinkan. (c) Sesudah dilicinkan.....	142
51	Aksara “—”. (a) Imej asal. (b) Penskilan 16 x 16 sebelum dilicinkan. (c) Sesudah dilicinkan.	143
52	Aksara “ث”. (a) Imej asal. (b) Penskilan 16 x 16 sebelum dilicinkan. (c) Sesudah dilicinkan	144
53	Kod Syarat Negatif.....	148
54	Kod-kod Syarat Positif.....	149
55	Satu contoh primitif aksara “—”. Titik hujung ditandakan dengan nilai negatif iaitu -4 dan -1. Titik cabang ditandakan dengan nilai -9. Nilai positif menandakan kerangka aksara.....	151
56	Penghampiran Poligon bagi aksara “ن” menggunakan algoritma PP dengan panjang minimum garis disetkan kepada dua.....	153
57	Pemecahan Aksara “د” dalam bentuk primitif.....	156
58	Pemecahan Aksara “ب” dalam bentuk primitif.....	156
59	Pemecahan Aksara “م” dalam bentuk primitif.....	157
60	Pemecahan aksara “ث” dalam bentuk primitif.....	157
61	Kumpulan Primitif oleh Bushofa dan Spann(1997).....	159
62	Pengezonan imej 16x16 kepada 16 segiempat bersaiz 4x4.....	164
63	Fitur-fitur Tersarikan bagi Aksara “ا”. (a) Aksara Tunggal. (b) Aksara Hujung.....	163
64	Fitur-fitur tersarikan bagi aksara “ح”	165
65	Fitur-fitur tersarikan bagi aksara “ف”.....	165
66	Fitur-fitur tersarikan bagi aksara “ك”.....	166



67	Fitur-fitur tersarikan bagi aksara “لـ “.....	166
68	Contoh Satu Pokok Kata Putus.....	174
69	Seni Bina Rangkaian Kehadapan Multiaras atau RKMA.....	176
70	Butiran terperinci satu nod. (a) Nod x'_2 . (b) Fungsi Pengaktifan Sigmoid Perduaan.....	178
71	Pengkelas multi RKMA dengan skema konsensus. N RN dilatih secara berasingan untuk mengelaskan satu corak input dengan menggunakan kaedah konsensus untuk membuat kata putus pengelasan secara kolektif.....	194
72	Rangkaian Neural Selari dengan Unit Pengawal.....	195
73	Algoritma Genetik.....	201
74	Skema baru Pengkelas Aksara Jawi (PB-1 hingga PB-7 menunjukkan Perambat-balik bagi K1 hingga K7, masing-masing).....	214
75	Mengelompok aksara mengikut bilangan primitif.....	219
76	Pengkelas Utama (aras tertinggi).....	222
77	Kumpulan 1 – Bilangan Primitifnya adalah satu.....	224
78	Kumpulan 2 – Bilangan Primitifnya adalah dua.....	225
79	Kumpulan 3 – Bilangan Primitifnya adalah tiga	226
80	Kumpulan 4 – Bilangan Primitifnya adalah empat	227
81	Kumpulan 5 – Bilangan Primitifnya adalah lima	227
82	Kumpulan 6 – Bilangan Primitifnya adalah enam	228
83	Kumpulan 7 – Bilangan Primitifnya adalah tujuh	229
84	RN PB dengan Satu Aras Tersembunyi.....	230
85	Pengawalan Pemberat ditentukan oleh AG melalui nilai Keupayaannya. Keupayaan pemberat yang dicari itu dapat dianggarkan dengan menguji set pemberat awalan pada RN yang sebenar, kemudian dilatih dengan set data latihan serta diujikan dengan data ujian.....	240
86	Struktur Hibrid AG+PB bagi satu aras RKMA.....	241