

UNIVERSITAS BINA NUSANTARA

Jurusan Teknik Informatika

Skripsi Sarjana Komputer

Semester Ganjil tahun 2001

PENGENALAN KARAKTER KANJI MANDARIN BERDASARKAN STRUKTUR KARAKTER DENGAN JARINGAN SYARAF TIRUAN

Susanna Widjaya 0331970938

Budi 0331970946

Heru Wintara O. 0331970947

Abstrak

Pengenalan pola karakter adalah bidang penelitian yang pada tiga dekade terakhir ini sangat menarik minat para peneliti. Salah satu obyek penelitian yang menarik untuk diteliti adalah karakter kanji Mandarin. Karena jumlah karakter kanji Mandarin sangat banyak dan kadang kala memiliki bentuk karakter yang hampir sama, sehingga dalam mengembangkan model pengenalan untuk karakter bahasa Mandarin adalah hal yang sulit.

Pada penelitian ini menawarkan alternatif karakteristik sebagai dasar pengenalan karakter kanji Mandarin. Penelitian ini menggunakan proses *thinning* dan *post-thinning* untuk pengolahan citra karakter kanji Mandarin, lalu proses *feature extraction* untuk mendapatkan karakteristik dari karakter.

Untuk proses pengklasifikasian dan pengenalan karakter, penelitian ini menggunakan *Backpropagation* dengan satu *hidden layer* yang mempunyai 10,15,20 dan 23 *node* pada *hidden layer* tersebut. Jumlah karakter yang mewakili penelitian ini adalah sebanyak 23 karakter sedangkan citra yang akan diuji sebanyak 138 citra. Hasil yang terbaik dicapai adalah pada konfigurasi 7:23:23 dengan menggunakan 23 *node* pada *hidden layer* tersebut.

Perlu untuk diteliti lebih lanjut mengenai banyaknya *node* yang terdapat dalam *hidden layer*, penggunaan karakteristik yang lain untuk menambah akurasi pengenalan dan banyaknya *node* pada masing-masing layer yang optimal untuk mendapatkan konfigurasi yang terbaik dan hasil yang lebih baik.

Kata kunci :

Thinning, Bckpropagation, Neural Network, Post-Thinning, Feature Extraction.

PRAKATA

Puji syukur kepada Tuhan Yang Maha Esa yang telah memberikan anugerahNya, sehingga skripsi ini dapat selesai dengan baik dan tepat pada waktunya. Skripsi ini diharapkan dapat bermanfaat bagi para mahasiswa-mahasiswi yang lainnya yang mengambil topik skripsi yang sama. Penulis juga menyadari bahwa skripsi ini masih jauh dari sempurna sehingga penulis juga mengharapkan saran yang bermanfaat bagi penelitian lebih lanjut.

Penulis juga ingin mengucapkan terima kasih kepada :

- Ibu Ir. Th. Widia S., M.M. selaku Rektor Universitas Bina Nusantara yang secara tidak langsung memberikan kesempatan kepada penulis untuk dapat menyelesaikan skripsi ini.
- Bapak Ir. Sablin Yusuf, M.Sc, M.CompSc. selaku Ketua Jurusan Teknik Informatika yang telah memberikan kesempatan kepada penulis untuk menyelesaikan skripsi ini.
- Ibu Anny Tandyo, M.Sc. selaku pembimbing yang telah memberikan pengetahuan yang bermanfaat sehingga penulisan skripsi ini dapat berjalan dengan baik.
- Orang tua dan kakak serta keluarga yang lainnya yang telah memberikan doa dan dorongan semangat sehingga penulisan skripsi ini dapat berjalan dengan baik.
- Teman-teman lainnya yang tanpa pamrih memberikan bantuan moril, tenaga dan doa.

- Para dosen yang telah memberikan saran-saran dan wawasan yang bermanfaat bagi penulisan skripsi ini.
- Serta pihak-pihak lainnya yang tidak mungkin disebutkan satu persatu.

Yang telah memberikan dorongan moril, tenaga dan waktu baik secara langsung maupun tidak langsung telah mendukung penulis sehingga skripsi ini dapat terselesaikan dengan baik dan tepat pada waktunya.

Jakarta, 29 Januari 2001

Penulis

DAFTAR ISI

	Halaman
Halaman Judul Luar	I
Halaman Judul Dalam	ii
Halaman Persetujuan <i>Hardcover</i>	iii
Halaman Pernyataan Dewan Penguji	iv
Abstrak	v
Prakata	vi
Daftar Isi	viii
Daftar Tabel	xii
Daftar Gambar	xiii
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Ruang Lingkup	2
1.3 Tujuan dan Manfaat	3
1.4 Metodologi	3
1.5 Sistematika Penulisan	4
BAB 2. LANDASAN TEORI	6
2.1 Citra	6
2.1.1 Pengertian Citra	6

2.1 Pengenalan Pola (Pattern Recognition)	10
2.2.1 Pengertian Pengenalan Pola	10
2.2.2 Pendekatan dalam Pattern Recognition	11
2.3 Thinning	11
2.4 Format File	12
2.4.1 Pengenalan Format File Bitmap (BMP)	12
2.4.2 Struktur File Bitmap (BMP)	13
2.5 Backpropagation	14
2.5.1 Sejarah Backpropagation	14
2.5.2 Algoritma Backpropagation	16
2.5.3 Normalisasi	20
2.6 Artificial Neural Network	21
2.6.1 Pengertian Artificial Neural Network	21
2.6.2 Biological Neural System	23
2.6.3 Sejarah Perkembangan Artificial Neural Network	24
2.6.4 Bagian Artificial Neural Network	24
2.7 Karakter Kanji Mandarin	26
2.7.1 Sejarah Karakter Kanji Mandarin	26
2.7.2 Jenis-Jenis Goresan Karakter Kanji Mandarin	26
BAB 3. ANALISIS DAN PERANCANGAN	28
3.1 Pemrosesan Citra Biner	29
3.1.1 Thinning	29
3.1.2 Algoritma Thinning	34

3.1.3 Post Thinning	36
3.1.4 Algoritma Teknik Maximum Circle	38
3.2 Ekstrasi Karakteristik	39
3.2.1 Pendeteksian Part	40
3.2.2 Algoritma Pendeteksian Part	42
3.2.3 Pendeteksian Closed Loop	43
3.2.4 Algoritma Pendeteksian Closed Loop	45
3.2.5 Pendeteksian Ending	45
3.2.6 Algoritma Pendeteksian Ending	46
3.2.7 Pendeteksian 4 Macam Cabang	47
3.2.8 Algoritma Pendeteksian Cabang	53
3.3 Pelatihan	55
3.4 Pengujian	57

BAB 4. IMPLEMENTASI DAN EVALUASI	59
4.1 Perangkat Keras dan Piranti Lunak yang diperlukan	59
4.2 Parameter Jaringan Syaraf Tiruan	59
4.3 Pelatihan dengan Menggunakan 23 Node pada Hidden Layer	67
4.4 Pelatihan dengan Menggunakan 20 Node pada Hidden Layer	68
4.5 Pelatihan dengan Menggunakan 15 Node pada Hidden Layer	70
4.6 Pelatihan dengan Menggunakan 10 Node pada Hidden Layer	72
4.7 Perbandingan Seluruh Konfigurasi JST	74
4.8 Tabel Error Tolerance pada Konfigurasi yang ada	76

BAB 5. KESIMPULAN DAN SARAN	77
5.1 Kesimpulan	77
5.2 Saran	77

DAFTAR PUSTAKA

RIWAYAT HIDUP

LAMPIRAN

DAFTAR TABEL

	Halaman
Tabel 4.1 Tabel Daftar Karakter Kanji Mandarin	60
Tabel 4.2 Tabel Karakteristik Karakter Kanji Mandarin	62
Tabel 4.3 Tabel Hasil Pengenalan Karakter pada Konfigurasi 7:23:23	67
Tabel 4.4 Tabel Hasil Pengenalan Karakter pada Konfigurasi 7:20:23	69
Tabel 4.5 Tabel Hasil Pengenalan Karakter pada Konfigurasi 7:15:23	71
Tabel 4.6 Tabel Hasil Pengenalan Karakter pada Konfigurasi 7:10:23	73
Tabel 4.7 Tabel Error Tolerance pada Konfigurasi 7:23:23	76
Tabel 4.8 Tabel Error Tolerance pada Konfigurasi 7:20:23	76
Tabel 4.9 Tabel Error Tolerance pada Konfigurasi 7:15:23	76
Tabel 4.10 Tabel Error Tolerance pada Konfigurasi 7:10:23	77

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Proses Sampling	7
Gambar 2.2 Pendekatan 8-Tetangga	7
Gambar 2.3 Pendekatan 4-Tetangga	8
Gambar 2.4 Edge Pixel	9
Gambar 2.5 Diagram Hubungan Antara Tiga Bidang Studi	9
Gambar 2.6 Gambaran dari Pattern Recognition	11
Gambar 2.7 Multi-Layer Error Backpropagation Network	15
Gambar 2.8 Backpropagation	16
Gambar 2.9 Fungsi Sigmoid Bipolar	16
Gambar 2.10 Forward Propagation Step	18
Gambar 2.11 Backward Propagation Step	20
Gambar 2.12 Simulasi Model dari Neural Network	21
Gambar 2.13 Sinapsis	23
Gambar 2.14 Struktur Neuron	23
Gambar 2.15 Fungsi Aktivasi	25
Gambar 2.16 Contoh dari Jenis Karakter Kanji Mandarin	27
Gambar 3.1 Gambaran Umum Proses Pengenalan Karakter Kanji Mandarin	28
Gambar 3.2 Proses yang Terjadi dalam Tahap Pemrosesan Citra	29
Gambar 3.3 Citra Asal dan Skeleton yang Mengandung Cabang Palsu	30

Gambar 3.4 Gambaran 7 Jenis Edge Pixel Beserta Konfigurasi Tetangganya	31
Gambar 3.5 Edge Pixel-Edge Pixel yang tidak boleh dihapus	33
Gambar 3.6 Proses Thinning dari Citra Asal (1) Hingga didapat Skeletonnya (9)	35
Gambar 3.7 Contoh Cabang Palsu Pada Citra Skeleton	36
Gambar 3.8 Ilustrasi Teknik Maximum Circle	37
Gambar 3.9 Citra Asal dan Skeleton Hasil Proses Post-Thinning	39
Gambar 3.10 Kegiatan yang dilakukan Dalam Ekstrasi Karakter	39
Gambar 3.11 Contoh Gambar untuk Menunjukkan Part, Elips Menandakan Satu Part	41
Gambar 3.12 Part pada Citra Post-Thinned Skeleton	41
Gambar 3.13 Ilustrasi Pendeteksian Part	42
Gambar 3.14 Contoh Closed Loop pada Citra Asal	43
Gambar 3.15 Closed Loop pada Citra Post-Thinned Skeleton	44
Gambar 3.16 Ilustrasi Pendeteksian Closed Loop	44
Gambar 3.17 Ending pada Citra Post-Thinned Skeleton	46
Gambar 3.18 Contoh 24 Macam Pola Cabang 6	47
Gambar 3.19 Contoh 15 Macam Pola Cabang 5	48
Gambar 3.20 Contoh 4 Macam Pola Cabang 4	48
Gambar 3.21 Contoh Lain Cabang 6, 5 dan 4	49
Gambar 3.22 Contoh Penomoran Pixel	49
Gambar 3.23 Dua Macam Pixel Tetangga 4	50
Gambar 3.24 Tiga Macam Pixel Tetangga 8	50
Gambar 3.25 Bentuk Umum 24 Pola Cabang 6	51
Gambar 3.26 Bentuk Umum 15 Pola Cabang 5	52

Gambar 3.27 Bentuk Umum 4 Pola Cabang 4	53
Gambar 3.28 Gambaran Umum Proses Pelatihan	55
Gambar 3.29 Gambaran Umum Proses Pengujian	57
Gambar 4.1 Grafik Error Tolerance pada Konfigurasi 7:23:23	67
Gambar 4.2 Grafik Error Tolerance pada Konfigurasi 7:20:23	69
Gambar 4.3 Grafik Error Tolerance pada Konfigurasi 7:15:23	71
Gambar 4.4 Grafik Error Tolerance pada Konfigurasi 7:10:23	73
Gambar 4.5 Grafik Error Tolerance Seluruh Konfigurasi	75