

Pengenalan Pola Aksara Jawa Tulisan Tangan dengan Jaringan Syaraf Tiruan Perambatan-Balik

HANDWRITTER JAWANESSE CHARACTER RECOGNATION SYSTEM USING ARTIFICIAL NETWORK BACK PROPAGATION

Farida Asriani dan Azis Wisnu Widhi Nugraha

*Prodi Teknik Elektro, Universitas Jenderal Soedirman
Jl. Kampus no.1, Grendeng, Purwokerto, Indonesia*

¹ida_elin@yahoo.com

²azis.wwn@unsoed.ac.id

Abstract— Automatic reading techniques of handwriting being developed. Computerization of Java script has started to be developed. However, there is no automated reader system for documents written in handwriting Java script. This research applies backpropagation neural network for recognition of handwritten java script. Preprocessing patterns done with Java script normalize the image size to be 40x40 pixels then feature extraction. Performance network training is 99.8% and performance examination is 95.81%.

Keywords— Java script, normalization, artificial neural networks, backpropagation.

PENDAHULUAN

Pada jaman kerajaan penulisan karya-karya sastra Jawa dilakukan dengan huruf Jawa tulisan tangan. Peninggalan karya-karya sastra Jawa yang ada sarat dengan ilmu pengetahuan. Ironisnya sebagian besar masyarakat terutama di pulau Jawa sudah tidak mengenal aksara Jawa. Agar ilmu pengetahuan yang tersurat dalam kitab Jawa terus dapat dikembangkan maka perlu adanya sistem pembaca dokumen aksara Jawa secara otomatis. Sistem ini akan dilandasi oleh sistem pengenalan aksara Jawa tulisan tangan yang dikembangkan dalam penelitian ini.

Perumusan masalah dalam penelitian ini adalah:

- bagaimana arsitektur yang tepat bagi jaringan syaraf tiruan perambatan-balik untuk pengenalan pola aksara Jawa tulisan tangan?,
- berapa besar parameter laju pelatihan dan momentum yang akan memberikan hasil pelatihan dan pengujian jaringan yang terbaik?, dan
- berapakah prosentase keberhasilan pengenalan pola yang dilakukan?.

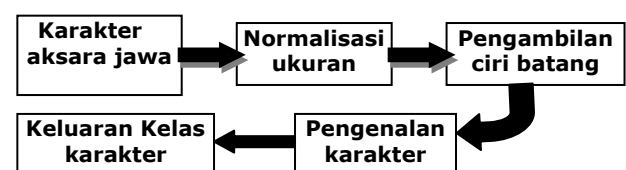
Berbagai sistem pengenalan tulisan tangan melibatkan fase scanning, digitalisasi, preprocessing, segmentasi, ekstraksi ciri dan klasifikasi (Tal et al, 1999). Untuk keperluan pengenalan maka perlu dilakukan pengolahan citra tulisan tangan untuk mendapatkan informasi yang berbeda untuk masing-masing kelas pola (Devijver and Kittler, 1982). Ekstraksi ciri batang dan ciri transisi sangat baik diterapkan pada pola tulisan tangan (Gardner et al, 1995; dan Mujiharjo, 2001).

Jaringan syaraf tiruan (artificial neural network) terkait erat dengan pengenalan pola. Jaringan syaraf tiruan dapat

mengatur dirinya untuk menghasilkan suatu respons yang konsisten dengan serangkaian masukan. Jaringan syaraf tiruan perambatan-balik (backpropagation) telah dikembangkan untuk melatih jaringan syaraf tiruan yang sampai pada suatu tingkat tertentu dapat melakukan generalisasi. Kemampuan untuk mengabaikan derau dan distorsi adalah hal yang sangat penting dalam pengenalan pola (Fausett, 1994). Jaringan syaraf tiruan perambatan balik telah banyak diterapkan untuk tujuan pengenalan pola (Yan et al, 1998; Mujiharjo, 2001; Petersen et al, 2002; Asriani, 2003).

METODE PENELITIAN

Model sistem yang dikembangkan ditunjukkan pada Gambar 1.



Gambar 1 Sistem yang akan dipakai untuk pengenalan pola aksara Jawa.

A. Tahap-Tahap Penelitian

- 1) Studi literatur.
- 2) Pengumpulan data-data aksara Jawa tulisan tangan. *Data aksara Jawa yang akan dikenali adalah aksara dasar sebagai berikut.*

TABEL 1
AKSARA JAWA YANG AKAN DIKENALI

latin	jawa	latin	jawa	latin	jawa	latin	jawa
Ha	ꦲ	Da	ꦢ	Pa	ꦥ	Ma	ꦩ
Na	ꦤ	Ta	ꦠ	Dha	ꦢꦲ	Ga	ꦒ
Ca	ꦕ	Sa	ꦱ	Ja	ꦗ	Ba	ꦧ
Ra	ꦫ	Wa	ꦮ	Ya	ꦪ	Tha	ꦠꦲ
Ka	ꦏ	La	ꦭ	Nya	ꦤꦪ	Nga	ꦤꦒ

3) Preprocessing data. Pada tahap ini dilakukan ekstraksi ciri batang dari data-data yang sudah diperoleh. Ukuran citra ditetapkan pada 40x40 pixel. Ukuran penjadwalan citra divariasi.

4) Perancangan jaringan syaraf tiruan perambatan-balik.

5) Pelatihan jaringan syaraf tiruan. Melakukan pelatihan jaringan syaraf tiruan dengan beberapa nilai variabel jumlah neuron pada lapisan tersembunyi, laju pelatihan dan momentum. Variabel yang dipilih adalah variabel yang memberikan unjuk kerja jaringan yang terbaik.

6) Pengujian Jaringan. Pengujian jaringan dilakukan dengan menggunakan pola yang berbeda dengan pola yang digunakan untuk pelatihan.

7) Analisis. Pada analisis data akan ditentukan arsitektur, laju pelatihan, dan momentum yang paling cocok untuk pengenalan pola aksara jawa tulisan tangan. Kemudian akan dianalisa faktor-faktor yang mempengaruhi tinggi rendahnya tingkat pengenalan pola aksara jawa tulisan tangan.

HASIL DAN PEMBAHASAN

A. Preprocessing

Preprocessing awal diperlukan untuk mengisi keterbatasan pada langkah penguraian ciri, seperti adanya derau pada citra masukan. Dilakukan normalisasi sebanyak mungkin citra masukan dalam upaya untuk mengurangi variasi jenis tulisan tangan. Hasil yang telah dilakukan diuraikan pada bagian berikut.

B. Resize citra Aksara Jawa

Resize citra aksara jawa dilakukan karena tidak samanya ukuran kroping antara aksara satu dengan yang lainnya. Ukuran yang dipilih adalah 40x40 pixel karena dengan ukuran yang lebih kecil sulit diperoleh informasi detail tentang bentuk citra aksara jawa.

C. Normalisasi Ukuran Citra Aksara Jawa

Normalisasi ukuran karakter dilakukan karena tidak samanya ukuran karakter penulis yang satu dengan lainnya. Ukuran karakter memegang peranan penting dalam pengambilan ciri batang. Pada penelitian ini semua karakter dinormalisasi 40x40 piksel. Normalisasi dilakukan dengan fungsi Matlab:

Normalisasi = `imresize(CitraAsli,[40 40],'nearest')`

D. Ciri Batang

Telah dilakukan ekstraksi citra kedalam ciri batang dengan berbagai variasi penjadwalan yang berbeda-beda dan masing-masing menghasilkan vector input dengan ukuran masing-masing:

1. 168x1
2. 280x1
3. 504x1
4. 840x1
5. 1800x1

E. Menentukan laju pelatihan dan momentum

Pada tahap ini diambil jaringan dengan arsitektur tetap yaitu 504 neuron pada lapisan input, 60 neuron pada lapisan tersembunyi dan 20 neuron pada lapisan output. Jumlah pola pelatihan adalah 1000 pola (50 pola untuk tiap kelas). Dengan beberapa variasi parameter laju pelatihan dan momentum diperoleh hasil seperti ditunjukkan pada Tabel 2.

TABEL 2
UNJUK KERJA PELATIHAN TERHADAP VARIASI LAJU PELATIHAN DAN MOMENTUM

Laju Pelatihan	Momentum	Jumlah iterasi	Unjukkerja pelatihan
0,005	0,1	10000	95,4%
0,008	0,08	10000	95,8%
0,003	0,03	10000	96%
0,003	0,005	10000	95%
0,003	0,05	10000	96,6%

Mengacu pada Tabel 2, untuk pelatihan jaringan memberikan unjukkerja tertinggi saat diterapkan laju pelatihan sebesar 0,003 dan momentum sebesar 0,05.

F. Menentukan jumlah neuron pada lapisan tersembunyi

Dengan menggunakan arsitektur jaringan 504 neuron pada lapisan input, p (dengan variasi p: 50, 60 dan 70) neuron pada lapisan tersembunyi dan 20 neuron pada lapisan output serta digunakan 1000 pola aksara jawa untuk pelatihan. diperoleh hasil seperti ditunjukkan pada Tabel 3.

TABEL 3
PENGARUH JUMLAH LAPISAN TERSEMBUNYI TERHADAP UNJUK KERJA

Laju pelatihan	Momentum	p	Jumlah iterasi	Unjukkerja pelatihan	Unjukkerja pengujian
0,003	0,05	50	10000	92,80%	58,33%
0,003	0,05	60	10000	96%	63,33%
0,003	0,05	70	10000	96,60%	67%

G. Unjukkerja Jaringan dari Berbagai Arsitektur

Setelah didapatkan nilai-nilai parameter pelatihan jaringan yang meliputi laju pelatihan, momentum dan jumlah neuron pada lapisan tersembunyi, selanjutnya dilakukan pelatihan jaringan dengan berbagai arsitektur, dan hasilnya ditunjukkan pada Tabel 4. Pada tahap ini jaringan dilatih sampai 10000 epoch.

TABEL 4
PENGARUH JUMLAH NEURON MASUKAN TERHADAP UNJUK KERJA

Neuron input	p	Neuron output	Jumlah iterasi	Unjukkerja pelatihan	Unjukkerja pengujian
168	70	20	10000	81,80%	65,70%
280	70	20	10000	90,80%	64,10%
504	70	20	10000	96,60%	67,00%
840	70	20	10000	98,70%	67,84%

Dari table diatas dapat dilihat bahwa unjukkerja pengujian dari berbagai arsitektur jaringan masih rendah. Hal ini sangat dipengaruhi oleh :

- preprosesing* citra aksara jawa yang dilakukan belum memberikan informasi yang rinci tentang pola-pola aksara jawa yang akan diinputkan ke jaringan syaraf tiruan;
- pelatihan jaringan belum mencapai kondisi jenuh. (belum dicapai kondisi minimum dari jumlah kuadrat galatnya) sehingga perlu ditingkatkan jumlah epoch pelatiahannya;
- kurang beragamnya pola aksara jawa tulisan tangan yang digunakan sebagai pola pelatihan jaringan syaraf tiruan.

Upaya yang dilakukan untuk meningkatkan unjukkerja jaringan syaraf tiruan adalah:

- memperkecil ukuran jendela pada ekstraksi ciri batang citra aksara jawa menjadi 5x5 pixel sehingga diperoleh informasi yang lebih rinci tentang bentuk-bentuk vertical, horizontal dan diagonal edari citra aksara jawa yang akan dikenali. Pada ekstraksi ini dihasilkan vektor input dengan ukuran 1800x1.
- Dilakukan penambahan variasi pola pelatihan menjadi 1840 pola (92 pola untuk tiap kelas).
- Menambah jumlah epoch pelatihan sampai tidak terjadi penurunan yang signifikan dari jumlah kuadrat galat.

Hasil yang dicapai ditunjukkan pada Tabel 5.

TABEL 5 UNJUK KERJA SISTEM SETELAH DILAKUKAN UPAYA PERBAIKAN

Neuron input	p	Neuron output	Jumlah iterasi	Unjukkerja pelatihan	Unjukkerja pengujian
1800	70	20	50000	99,80%	95,81%

Mengacu pada Tabel 4 dan Tabel 5 maka jaringan dengan arsitektur 1800 neuron pada lapisan input, 70 neuron pada lapisan tersembunyi dan 20 neuron pada lapisan output yang dilatih dengan laju pelatihan 0,003 dan momentum 0,05 dengan jumlah iterasi 50000 kali memberikan unjukkerja yang terbaik yaitu 99,8% untuk pelatiahannya dan 95,81% untuk pengujiannya. Kesalahan pengenalan tiap kelas aksara ditunjukkan pada Tabel 6.

Adanya kesalahan pengenalan ini disebabkan karena :

- penulis data aksara jawa tidak terbiasa menulis aksara jawa sehingga terjadi beberapa kesalahan;
- aksara Jawa memiliki bentuk yang rumit dan antara aksara satu dan yang lainnya ada yang memiliki banyak kesamaan.

TABEL 6 KESALAHAN PENGENALAN PADA SETIAP KELAS

No	Aksara	Prosentase Kesalahan	Dikenali sebagai	No	Aksara	Prosentase Kesalahan	Dikenali sebagai
1	ᮊᮧ	4,41	ᮊᮧ ᮊᮧ ᮊᮧ	11	ᮊᮧ	4,41	ᮊᮧ ᮊᮧ ᮊᮧ
2	ᮊᮧ	7,35	ᮊᮧ ᮊᮧ	12	ᮊᮧ	1,47	ᮊᮧ
3	ᮊᮧ	4,41	ᮊᮧ	13	ᮊᮧ	0	-
4	ᮊᮧ	1,47	ᮊᮧ	14	ᮊᮧ	0	-
5	ᮊᮧ	1,47	ᮊᮧ	15	ᮊᮧ	4,41	ᮊᮧ
6	ᮊᮧ	0	-	16	ᮊᮧ	2,94	ᮊᮧ
7	ᮊᮧ	0	-	17	ᮊᮧ	1,47	ᮊᮧ
8	ᮊᮧ	10,29	ᮊᮧ	18	ᮊᮧ	5,88	ᮊᮧ ᮊᮧ
9	ᮊᮧ	11,76	ᮊᮧ ᮊᮧ ᮊᮧ	19	ᮊᮧ	4,41	ᮊᮧ
10	ᮊᮧ	5,88	ᮊᮧ ᮊᮧ	20	ᮊᮧ	4,41	ᮊᮧ ᮊᮧ

KESIMPULAN DAN SARAN

A. Kesimpulan

Beberapa hal yang sangat mempengaruhi unjuk kerja pengenalan pola aksara jawa antara lain.

- Preprosesing* pola, jumlah pola yang dilatih, dan parameter-parameter pelatihan pola sangat mempengaruhi unjukkerja pengenalan pola aksara jawa.
- Variasi bentuk tulisan yang dikenali semakin banyak variasinya maka semakin baik unjuk kerjanya
- Pemilihan arsitektur jaringan syaraf tiruan.
- Jumlah iterasi pada pelatihan jaringan.

B. Saran

Saran untuk penelitian selanjutnya.

- Dikembangkan ke objek dalam bentuk kata.
- Diterapkan metode klasifikasi yang lain seperti SVM untuk memperkecil tingkat komutasinya.

DAFTAR PUSTAKA

- Asriani, F. 2003, *Evaluasi Jaringan Syaraf Tiruan Perambatan Balik untuk Pengenalan Pola Digit Tulisan Tangan*, Majalah Ilmiah, Unsoed, Purwokerto.
- Devijver, P. A. and Kittler J. 1982, *Pattern Recognition: a Statistical Approach*, Prentice Hall International, London.
- Fausett L. 1994, *Fundamentals of Neural Networks*, Prentice Hall, Englewood Cliffs, New Jersey.
- Gader, P.D., Keller, J.M., Cai, J., 1995, "A Fuzzy Logic System for the Detection and Recognition of Handwritten Street Number", *IEEE Transactions on Fuzzy Systems*, Vol. 3. NO. 1, pp. 83-94.
- Mudjiharjo. P., 2001, *Penerapan Jaringan Syaraf Tiruan Perambatan-Balik untuk Pengenalan Kode Pos Tulisan Tangan*, Tesis S2, UGM, Yogyakarta.
- Petersen M. Egmont, , Ridder D. de, and Handels, H., 2002, *Image processing with neural networks—a review*, *Journal of Pattern Recognition*, pp 2279–2301
- Tal Steinhertz, Ehud Rivlin, and Nathan Intrator, 1999 "Offline Cursive Script Word Recognition -a Survey", *International Journal on Documents Analysis and Recognition (IJ DAR)*, pp.90–110, September 1999.
- Yann Lecun, I'Eon Bottou, Yoshua Bengio, and Patrick Haffner, 1998, *Gradient-Based Learning Applied to Document Recognition*, *IEEE Transactions on Neural Networks*.