

## PENERAPAN JARINGAN SYARAF TIRUAN UNTUK MENDETEKSI POSISI WAJAH MANUSIA PADA CITRA DIGITAL

Setyo Nugroho<sup>1</sup>, Agus Harjoko<sup>2</sup>

<sup>1</sup>Jurusan Teknik Informatika, STMIK STIKOM Balikpapan  
Jln. Kapten P. Tendean 2A, Balikpapan, 76121

<sup>2</sup>Program Pascasarjana Ilmu Komputer, Universitas Gadjah Mada, Yogyakarta

### Abstrak

Salah satu bagian penting dalam proses pengenalan wajah adalah mendeteksi posisi wajah. Dalam makalah ini kami merancang dan mengimplementasikan sistem pendeteksi posisi wajah dengan menggunakan jaringan syaraf tiruan. Sistem ini dilatih dengan menggunakan contoh-contoh wajah yang diberikan. Algoritma Quickprop dan metode active learning digunakan untuk mempercepat proses pelatihan sistem. Dari hasil eksperimen dengan menggunakan 23 file citra berisi 149 wajah, sistem pendeteksi wajah ini memberikan hasil detection rate 71,14% dan false positive 62.

**Kata kunci:** deteksi wajah, jaringan syaraf tiruan, quickprop, active learning

### 1. Pendahuluan

Dengan semakin tingginya kemampuan komputer untuk memproses data dalam jumlah besar dengan kecepatan yang tinggi, sistem biometrik semakin banyak diaplikasikan dalam kehidupan manusia. Sistem biometrik adalah sistem untuk melakukan identifikasi dengan cara menggunakan ciri-ciri fisik atau anggota badan manusia, seperti sidik jari, retina mata, suara. Teknologi biometrik ini memiliki beberapa kelebihan seperti tidak mudah hilang, tidak dapat lupa, tidak mudah dipalsukan, dan memiliki keunikan yang berbeda antara manusia satu dengan yang lain.

Salah satu cara yang digunakan dalam sistem biometrik adalah pengenalan wajah. Sistem pengenalan wajah bertujuan untuk mengidentifikasi wajah seseorang dengan cara membandingkan wajah tersebut dengan database wajah yang sudah ada. Dalam sistem pengenalan wajah, pendeteksian posisi wajah merupakan salah satu tahap yang penting karena di dunia nyata wajah dapat muncul di dalam citra dengan berbagai ukuran dan posisi, dan dengan latar belakang yang bervariasi [Hjelmas, 2001].

Dalam makalah ini kami merancang dan mengimplementasikan sistem pendeteksi wajah manusia yang dapat memberikan output berupa jumlah, posisi, dan ukuran wajah manusia yang ditemukan di dalam suatu citra digital.

### 2. Jaringan Syaraf Tiruan

Jaringan syaraf tiruan adalah suatu sistem pemrosesan informasi yang cara kerjanya memiliki kesamaan tertentu dengan jaringan syaraf biologis [Fausett, 1994]. Jaringan syaraf tiruan dikembangkan sebagai model matematis dari syaraf biologis dengan berdasarkan asumsi bahwa:

- Pemrosesan terjadi pada elemen-elemen sederhana yang disebut neuron.

- Sinyal dilewatkan antar neuron melalui penghubung.
- Setiap penghubung memiliki bobot yang akan mengalikan sinyal yang lewat.
- Setiap neuron memiliki fungsi aktivasi yang akan menentukan nilai sinyal output.

Jaringan syaraf dapat digolongkan menjadi berbagai jenis berdasarkan pada arsitekturnya, yaitu pola hubungan antara neuron-neuron, dan algoritma trainingnya, yaitu cara penentuan nilai bobot pada penghubung.

#### 2.1 Multi-Layer Perceptron

Multi-Layer Perceptron adalah jaringan syaraf tiruan *feed-forward* yang terdiri dari sejumlah neuron yang dihubungkan oleh bobot-bobot penghubung. Neuron-neuron tersebut disusun dalam lapisan-lapisan yang terdiri dari satu lapisan input (*input layer*), satu atau lebih lapisan tersembunyi (*hidden layer*), dan satu lapisan output (*output layer*). Lapisan input menerima sinyal dari luar, kemudian melewatkannya ke lapisan tersembunyi pertama, yang akan diteruskan sehingga akhirnya mencapai lapisan output [Riedmiller, 1994].

Setiap neuron  $i$  di dalam jaringan adalah sebuah unit pemrosesan sederhana yang menghitung nilai aktivasinya yaitu  $s_i$  terhadap input eksitasi yang juga disebut net input  $net_i$ .

$$net_i = \sum_{j \in pred(i)} s_j w_{ij} - \theta_i \quad (1)$$

di mana  $pred(i)$  melambangkan himpunan predesesor dari unit  $i$ ,  $w_{ij}$  melambangkan bobot koneksi dari unit  $j$  ke unit  $i$ , dan  $\theta_i$  adalah nilai bias dari unit  $i$ . Untuk membuat representasi menjadi lebih mudah, seringkali bias digantikan dengan suatu bobot yang terhubung dengan unit bernilai 1. Dengan demikian bias dapat diperlakukan secara sama dengan bobot koneksi.

## 2.2 Supervised Learning

Tujuan pada pembelajaran *supervised learning* adalah untuk menentukan nilai bobot-bobot koneksi di dalam jaringan sehingga jaringan dapat melakukan pemetaan (*mapping*) dari input ke output sesuai dengan yang diinginkan. Pemetaan ini ditentukan melalui satu set pola contoh atau data pelatihan (*training data set*).

Setiap pasangan pola  $p$  terdiri dari vektor input  $x^p$  dan vektor target  $t^p$ . Setelah selesai pelatihan, jika diberikan masukan  $x^p$  seharusnya jaringan menghasilkan nilai output  $t^p$ . Besarnya perbedaan antara nilai vektor target dengan output aktual diukur dengan nilai error yang disebut juga dengan *cost function*:

$$E = \frac{1}{2} \sum_{p \in P} \sum_n (t_n^p - s_n^p)^2 \quad (2)$$

di mana  $n$  adalah banyaknya unit pada *output layer*. Tujuan dari training ini pada dasarnya sama dengan mencari suatu nilai minimum global dari  $E$ .

## 2.3 Algoritma Backpropagation

Salah satu algoritma pelatihan jaringan syaraf tiruan yang banyak dimanfaatkan dalam bidang pengenalan pola adalah backpropagation. Algoritma ini umumnya digunakan pada jaringan syaraf tiruan yang berjenis multi-layer feed-forward, yang tersusun dari beberapa lapisan dan sinyal dialirkan secara searah dari input menuju output.

Algoritma pelatihan backpropagation pada dasarnya terdiri dari tiga tahapan [Fausett, 1994], yaitu:

- Input nilai data pelatihan sehingga diperoleh nilai output
- Propagasi balik dari nilai error yang diperoleh
- Penyesuaian bobot koneksi untuk meminimalkan nilai error.

Ketiga tahapan tersebut diulangi terus-menerus sampai mendapatkan nilai error yang diinginkan. Setelah training selesai dilakukan, hanya tahap pertama yang diperlukan untuk memanfaatkan jaringan syaraf tiruan tersebut.

Secara matematis [Rumelhart, 1986], ide dasar dari algoritma backpropagation ini sesungguhnya adalah penerapan dari aturan rantai (*chain rule*) untuk menghitung pengaruh masing-masing bobot terhadap fungsi error:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial s_i} \frac{\partial s_i}{\partial w_{ij}} \quad (3)$$

dan

$$\frac{\partial s_i}{\partial w_{ij}} = \frac{\partial s_i}{\partial net_i} \frac{\partial net_i}{\partial w_{ij}} = f'_{\log}(net_i) s_j \quad (4)$$

dimana  $w_{ij}$  adalah bobot penghubung dari neuron  $j$  ke neuron  $i$ ,  $s_i$  adalah output, dan  $net_i$  adalah jumlah hasil kali pada input dari neuron  $i$ .

## 2.4 Algoritma Quickprop

Pada algoritma Quickprop dilakukan pendekatan dengan asumsi bahwa kurva fungsi error terhadap masing-masing bobot penghubung berbentuk parabola yang terbuka ke atas, dan gradien dari kurva error untuk suatu bobot tidak terpengaruh oleh bobot-bobot yang lain [Fahlman, 1988]. Dengan demikian perhitungan perubahan bobot hanya menggunakan informasi lokal pada masing-masing bobot. Perubahan bobot pada algoritma Quickprop dirumuskan sebagai berikut:

$$\Delta w(t) = -\epsilon \frac{\partial E}{\partial w}(t) + \frac{\frac{\partial E}{\partial w}(t)}{\frac{\partial E}{\partial w}(t-1) - \frac{\partial E}{\partial w}(t)} * \Delta w(t-1) \quad (5)$$

di mana:

$\Delta w(t)$ : perubahan bobot

$\Delta w(t-1)$ : perubahan bobot pada epoch sebelumnya

$\epsilon$ : adalah learning rate

$\frac{\partial E}{\partial w}(t)$ : derivatif error

$\frac{\partial E}{\partial w}(t-1)$ : derivatif error pada epoch sebelumnya

Pada eksperimen dengan masalah XOR dan encoder/decoder [Fahlman, 1988], terbukti bahwa algoritma Quickprop dapat meningkatkan kecepatan training. Eksperimen dari [Schiffmann, 1993] juga menunjukkan peningkatan kecepatan training dan unjuk kerja yang signifikan.

## 3. Cara Penelitian

Dalam penelitian ini kami menggunakan data yang terdiri dari satu set citra untuk pelatihan (*training data set*) dan satu set citra untuk pengujian (*testing data set*). Untuk data pelatihan digunakan citra wajah berukuran 20x20 pixel sebanyak 3000 buah. Sedangkan untuk citra non-wajah diperoleh dari file-file citra yang tidak terdapat wajah manusia di dalamnya.

Sistem ini menggunakan jaringan syaraf tiruan (JST) dengan jenis multi-layer perceptron. Arsitektur yang digunakan diadaptasi dari hasil penelitian [Rowley, 1998], namun lebih disederhanakan.

Lapisan input terdiri dari 400 unit input, yang menerima masukan dari nilai grayscale pixel 20x20 dari subcitra yang akan dideteksi. Sebelum dijadikan input untuk JST, nilai grayscale yang berkisar dari 0 sampai 255 dinormalisasi menjadi antara -1 dan 1.

Lapisan output terdiri dari sebuah unit dengan nilai keluaran berkisar antara  $-1$  dan  $1$ . Pada training data set didefinisikan nilai  $1$  untuk data wajah dan  $-1$  untuk data non-wajah.

Lapisan tersembunyi (*hidden layer*) terdiri dari total  $25+16=41$  unit. Bagian pertama terhubung dengan lapisan input yang membentuk  $25$  area berukuran  $4 \times 4$  pixel. Bagian kedua terhubung dengan lapisan input yang membentuk  $16$  area berukuran  $5 \times 5$  pixel.

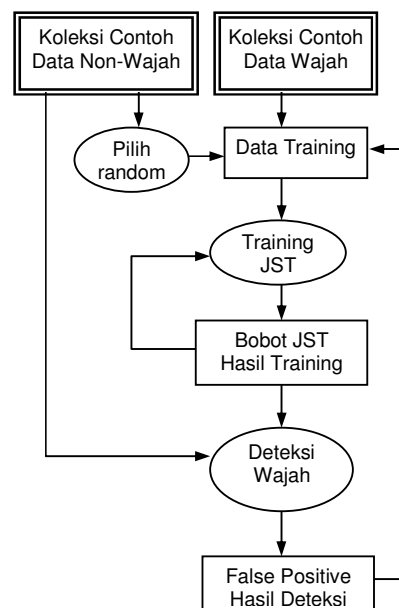
Secara keseluruhan jaringan ini memiliki  $883$  bobot penghubung, sudah termasuk bias. Jaringan ini lebih sederhana dibandingkan dengan sistem [Rowley, 1998] yang jumlah bobot penghubungnya mencapai  $4357$ .

### 3.1 Teknik Active Learning

Training dilakukan secara bertahap dengan menggunakan metode *active learning* [Sung, 1994]. Pada tahap pertama training dimulai dengan menggunakan sedikit data non-wajah. Pada tahap berikutnya, data training non-wajah ditambah sedikit demi sedikit. Namun data tambahan tersebut diseleksi hanya untuk data tertentu saja, yaitu data yang dideteksi sebagai wajah (*false positive*) pada hasil training tahap sebelumnya. Dengan demikian jumlah data training yang digunakan untuk jaringan syaraf tiruan akan lebih sedikit. Karena data training yang digunakan lebih sedikit, waktu yang diperlukan untuk proses training juga akan lebih singkat. Gambar 1 menunjukkan teknik *active learning* yang digunakan untuk sistem pendeteksi wajah.

### 3.2 Detektor Wajah

Bagian detektor wajah menggunakan arsitektur jaringan syaraf yang sama dengan yang digunakan untuk training. Bobot penghubung yang digunakan diambil dari bobot terakhir yang dihasilkan pada proses training. Hasil deteksi akan diputuskan sebagai wajah jika output dari JST lebih dari  $0$ , dan diputuskan sebagai non-wajah jika output JST kurang dari atau sama dengan  $0$ .



Gambar 1. Teknik Active Learning untuk Sistem Pendeteksi Wajah

### 3.3 Ekstraksi Subcitra

Posisi wajah bisa berada di mana saja pada citra yang akan dideteksi. Untuk itu digunakan window berukuran  $20 \times 20$  pixel yang akan digeser melalui seluruh daerah citra. Daerah citra yang dilewati oleh window tersebut akan diperiksa satu persatu apakah ada wajah atau tidak di area tersebut.

Untuk mengantisipasi ukuran wajah yang bervariasi di dalam citra yang dideteksi, citra diperkecil secara bertahap dengan skala perbandingan  $1:1,2$  sebagaimana dilakukan pada [Rowley, 1998]. Pada setiap ukuran citra yang diperkecil, window  $20 \times 20$  pixel akan digeser melalui seluruh area citra.

### 3.4 Preprocessing

Sebelum digunakan sebagai *training data set*, citra akan melalui tahap-tahap preprocessing berikut:

- Histogram Equalization, untuk memperbaiki kontras citra.
- Masking, yaitu menghilangkan bagian sudut-sudut citra untuk mengurangi variasi citra sehingga memperkecil dimensi data.
- Normalisasi, yaitu mengkonversi nilai intensitas grayscale citra sehingga memiliki range antara  $-1$  sampai dengan  $1$ .

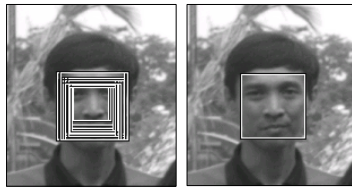
Tahap-tahap preprocessing ini juga digunakan pada saat proses pendeteksian wajah.



**Gambar 2.** Contoh data wajah yang telah mengalami preprocessing

### 3.5 Merging

Pada saat dilakukan deteksi wajah pada citra, biasanya sebuah wajah akan terdeteksi pada beberapa lokasi yang berdekatan. Lokasi-lokasi ini disebut dengan kandidat wajah. Untuk itu perlu dilakukan proses penggabungan (*merging*), yaitu menyatukan lokasi kandidat-kandidat wajah yang berdekatan.



**Gambar 3.** Hasil deteksi sebelum dan sesudah proses penggabungan (*merging*)

## 4. Hasil dan Pembahasan

Unjuk kerja dari detektor wajah pada umumnya diukur dengan menggunakan dua parameter, yaitu *detection rate* dan *false positive rate* [Yang, 2002]. *Detection rate* adalah perbandingan antara jumlah wajah yang berhasil dideteksi dengan jumlah seluruh wajah yang ada. Sedangkan *false positive rate* adalah banyaknya subcitra non-wajah yang dideteksi sebagai wajah.

Contoh hasil deteksi yang dilakukan pada beberapa citra pengujian ditunjukkan pada gambar 4. Pengujian dilakukan dengan data uji citra yang berasal dari Massachusetts Institute of Technology (MIT) yang terdiri dari 23 file citra yang secara keseluruhan berisi 149 wajah (data uji MIT-23). Kumpulan citra ini pertama kali dipublikasikan pada [Sung, 1994]. Pada data uji ini diperoleh hasil *detection rate* sebesar 71,14% dan *false positives* sebanyak 62. Hasil ini diperoleh dari training yang menggunakan 3000 data wajah dan 5200 data non-wajah yang diperoleh melalui metode *active learning*.

Tabel 1 menunjukkan hasil deteksi yang pernah dilakukan oleh para peneliti lain dengan menggunakan data uji MIT-23. Perbandingan ini tidak bisa dijadikan patokan mutlak untuk menyimpulkan bahwa satu metode lebih baik dari metode yang lain, karena faktor-faktor berikut [Hjelmas, 2001]:

- data set yang digunakan untuk training tidak sama
- jumlah data yang digunakan untuk training tidak sama



**Gambar 4.** Contoh hasil deteksi wajah pada beberapa citra uji yang juga digunakan pada [Sung, 1994]

**Tabel 1.** Beberapa hasil deteksi wajah pada data uji MIT-23

Metode	Detection Rate	False Positive
Support vector machines (SVM) (Osuna, 1997)	74,2%	20
Distribution-based dan clustering (Sung, Poggio, 1994)	79,9%	5
Neural Networks (Rowley, 1998)	84,5%	8
Kullback relative information (Lew, Huijismans, 1996)	94,1%	64

### 4.1 Pengaruh Jumlah Data Training

Tabel 2 menunjukkan pengaruh jumlah data training yang digunakan terhadap hasil deteksi. Tabel ini berdasarkan hasil deteksi pada suatu citra berisi 15 wajah dan memiliki total 790.797 window. Terlihat bahwa semakin banyak data training non-wajah yang digunakan, semakin kecil angka *false*

positive yang dihasilkan. Hal ini menunjukkan bahwa dengan data yang semakin lengkap, hasil belajar sistem akan semakin baik.

**Tabel 2.** Pengaruh Jumlah Data Training pada Unjuk Kerja Deteksi Wajah

Jumlah data training			Detection Rate	False Positive
wajah	non-wajah	total		
1000	1000	2000	15/15	42
1000	1200	2200	15/15	32
1000	1400	2400	15/15	31
1000	1600	2600	15/15	24
1000	1800	2800	15/15	23
1000	2000	3000	15/15	19
1000	2200	3200	15/15	10
1000	2400	3400	15/15	5
1000	2600	3600	15/15	4
1000	2800	3800	15/15	2
1000	3000	4000	15/15	2
1000	3200	4200	15/15	2
1000	3400	4400	15/15	2
1000	3600	4600	15/15	1
1000	3800	4800	15/15	1
1000	4000	5000	15/15	1

#### 4.2 Pengaruh Algoritma Quickprop pada Kecepatan Training

Tabel 3 menunjukkan perbandingan waktu training yang diperlukan antara training yang menggunakan algoritma backpropagation standar dengan training yang menggunakan algoritma Quickprop. Eksperimen dilakukan dengan menggunakan 2000, 3000, dan 4000 data. Untuk setiap proses training, iterasi dihentikan pada saat nilai error mencapai 0,05. Terlihat bahwa untuk jumlah data training yang semakin besar, algoritma Quickprop memberikan peningkatan kecepatan yang signifikan.

**Tabel 3.** Pengaruh Algoritma Quickprop pada Kecepatan Training

Jumlah data training	Error	Waktu training (detik)	
		Backprop standar	Quickprop
2000	0,05	49	12
3000	0,05	649	96
4000	0,05	4664	867

#### 4.3 Pengaruh Metode Active Learning

Tabel 4 menunjukkan perbandingan antara hasil training yang menggunakan metode *active learning* untuk memilih contoh data non-wajah, dengan hasil training yang menggunakan data non-wajah yang dipilih secara random. Pada eksperimen pertama digunakan 6000 data yang terdiri dari 3000 data wajah dan 3000 data non-wajah. Sedangkan

pada eksperimen kedua digunakan 8200 data yang terdiri dari 3000 data wajah dan 5200 data non-wajah. Terlihat bahwa teknik *active learning* memberikan hasil yang lebih baik. Ini berarti bahwa teknik *active learning* dapat memilih data yang benar-benar perlu, sehingga dapat meminimalkan jumlah data training yang digunakan.

**Tabel 4.** Hasil Perbandingan Kinerja antara Metode Active Learning dengan Random Data

Jumlah Data Training	Random Data		Active Learning	
	Detection Rate	False Positive	Detection Rate	False Positive
6000	62,42%	1160	71,14%	201
8200	63,76%	732	71,14%	62

### 5. Kesimpulan

Dari penelitian ini dapat diambil kesimpulan bahwa jaringan syaraf tiruan dengan jenis multi layer perceptron dapat digunakan untuk melakukan deteksi wajah pada citra digital. Pada sistem pendeteksi wajah yang berbasis contoh, kinerja hasil deteksi yang didapatkan sangat tergantung dari kualitas dan kuantitas dari data contoh yang diberikan. Untuk training dengan jumlah data yang besar, algoritma Quickprop memberikan peningkatan kecepatan training yang signifikan. Metode *active learning* dapat digunakan untuk memilih data contoh yang lebih tepat, sehingga meminimalkan jumlah data training yang digunakan.

### Daftar Pustaka

- [1] S.E. Fahlman, "An Empirical Study of Learning Speed in Back-Propagation Networks", *Technical Report CMU-CS-88-162*, Carnegie Mellon University, USA, 1988.
- [2] L. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*, Prentice-Hall Inc., USA, 1994.
- [3] E. Hjelmas, B.K. Low, "Face Detection: A Survey", *Computer Vision and Image Understanding*. 83, pp. 236-274, 2001.
- [4] H. Rowley, S. Baluja, T. Kanade, "Neural Network-Based Face Detection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, 1998.
- [5] D.E. Rumelhart, G.E. Hinton, R.J. Williams, "Learning Internal Representations by Error Propagation", *Parallel Distributed Processing*, chapter 8, MIT Press, Cambridge, MA, 1986.
- [6] W. Schiffmann, M. Joost, R. Werner, "Comparison of Optimized Backpropagation Algorithms", *Proc. of the European Symposium on Artificial Neural Networks (ESANN) '93*, Brussels, 1993.

- [7] K.K. Sung, "Learning and Example Selection for Object and Pattern Detection", *AITR 1572*, Massachusetts Institute of Technology AI Lab, 1996.
- [8] K.K. Sung, T. Poggio, "Example-Based Learning for View-Based Human Face Detection", *Technical Report AI Memo 1521*, Massachusetts Institute of Technology AI Lab, 1994.
- [9] M.H. Yang, D. Kriegman, N. Ahuja, "Detecting Faces in Images: A Survey", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, 2002.