

PENGARUH *PRINCIPLE COMPONENT ANALYSIS* TERHADAP TINGKAT IDENTIFIKASI *NEURAL NETWORK* PADA SISTEM SENSOR GAS

Muhammad Rivai

Jurusan Teknik Elektro, Institut Teknologi Sepuluh Nopember
Kampus ITS, Keputih, Sukolilo, Surabaya 60111
Telp. (031) 5947302, Fax. (031) 5931237
e-mail: muhammad_rivai@ee.its.ac.id

Abstract

In recently, it has been developed a gas identification system consists of a semiconductor sensor array and Neural Network pattern recognition. In this study, it has been implemented a method of Principle Component Analysis (PCA) as a preprocessing of the Neural Network algorithm. The sensory array is composed of eight commercial semiconductor sensors. Three layer-Neural Network was trained with the back propagation technique within 5000 epochs. PCA could reduce the eight-dimension into three-dimension without any information losses. The identification error rate was lower with the ratio of $\sim 10^{-4}$ and the training period was shorter with the ratio of ~ 0.6 . In generally, it can be concluded that the implementation of the PCA method into the Neural Network can enhance the performances of the neural include the identification rate and time consumed in the training phase.

Keywords: *Semiconductor sensor array, Principle Component Analysis, Neural Network*

Abstrak

Saat ini telah dikembangkan suatu sistim identifikasi jenis gas yang terdiri dari deret sensor semikonduktor dan pengenalan pola Neural Network. Pada penelitian ini telah dilakukan uji implementasi metoda Principle Component Analysis (PCA) sebagai pemrosesan awal pada algoritma Neural Network. Deret sensor terdiri dari delapan sensor gas semikonduktor yang tersedia secara komersial. Neural network tiga lapis dilatih dengan menggunakan teknik propagasi balik dengan perulangan iterasi 5000 epoch. PCA dapat mereduksi jumlah variabel masukan yang semula berdimensi delapan menjadi berdimensi tiga tanpa mengurangi tingkat informasinya. Taraf kesalahan identifikasi menjadi lebih kecil dengan rasio $\sim 10^{-4}$ dan periode pelatihan yang lebih pendek dengan rasio ~ 0.6 . Secara umum dapat disimpulkan bahwa implementasi PCA pada Neural Network dapat meningkatkan performanya yang meliputi taraf identifikasi dan waktu yang diperlukan pada fasa pelatihan.

Kata kunci: *Deret sensor semikonduktor, Principle Component Analysis, Neural Network*

1. PENDAHULUAN

Sensor kimia saat ini telah digunakan secara luas pada beberapa aplikasi seperti kontrol kualitas makanan dan minuman, monitoring tingkat polusi, diagnosa penyakit, dan keamanan. Pada umumnya sensor gas dibuat untuk mempunyai selektivitas yang tinggi terhadap jenis gas tertentu. Kelemahan sensor yang spesifik ini adalah jika digunakan untuk mendeteksi berbagai macam gas akan diperlukan sensor dalam jumlah yang besar. Para peneliti mencoba melakukan pendekatan alternatif dengan meniru prinsip kerja sistim penciuman mamalia. Pada pendekatan ini, sistim identifikasi jenis gas menggunakan sebuah deret sensor yang terdiri sejumlah elemen sensor dengan setiap elemen sensornya mempunyai tanggapan terhadap sejumlah gas tertentu. Tanggapan sebuah elemen sensor sebagian dapat tumpang tindih dengan tanggapan elemen sensor yang lain. Deret sensor tersebut dapat membentuk pola yang khas untuk setiap jenis gas. Dengan demikian sistim ini memerlukan algoritma pengenalan pola untuk mengetahui jenis gas yang terdeteksi.

Saat ini para peneliti menggunakan jenis sensor gas semikonduktor karena mempunyai keuntungan ukuran yang kecil, bobot yang ringan, sensitivitas dan kehandalan yang tinggi, serta penggunaan dalam jangka waktu yang lama [1]. Kelemahan dari sensor ini adalah tingkat selektivitas yang tidak terlalu tinggi. Sedangkan *Neural Network* sering digunakan untuk mengenal pola secara otomatis karena kemudahan algoritma dalam pengklasifikasian jenis gas [2]. Kelemahan dari algoritma ini adalah diperlukannya waktu yang lama dalam fasa pelatihannya.

Metoda *Principle Component Analysis* (PCA) sering digunakan untuk visualisasi hasil klasifikasi pada sistem identifikasi. Metoda ini merupakan penurunan dari teknik *factor analysis* yang bertujuan untuk mengidentifikasi struktur dari banyak variable menjadi data yang lebih sederhana. Metoda ini juga dikenal sebagai transformasi Karhunen-Loève atau transformasi Hotelling. PCA merupakan transformasi linier ortogonal yang mentransformasi data ke dalam koordinat sistem yang baru yang mana variasi yang paling banyak diperoleh dari proyeksi data pada koordinat prioritas utama (*principle component*) yang pertama, variasi kedua yang terbanyak pada koordinat prioritas utama yang kedua dan seterusnya. PCA biasanya digunakan untuk mengurangi dimensi dari himpunan data tetapi juga mempertahankan karakteristik dari himpunan data tersebut dengan menjaga beberapa *principle component* yang tinggi prioritasnya dan membuang beberapa *principle component* yang rendah prioritasnya [3].

Ada beberapa metoda yang umum digunakan untuk mendapatkan *principle component* pada metoda PCA, yaitu metoda covariance, metoda korelasi dan *singular value decomposition*. Prosedur PCA dengan menggunakan metoda covariance [4] adalah sebagai berikut:

- a) Menghimpun data eksperimen $X(M,N)$ yang memiliki dimensi tertentu yang berkesesuaian dengan variabel atau jumlah sensor yang digunakan (M) dan jumlah data (N).
- b) Mengurangkan setiap data X dengan nilai *mean* \bar{X}_m untuk masing-masing variable atau sensor (m).

$$\bar{X}_m = \frac{1}{N} \sum_{n=1}^N X(M,n)$$

$$X = X - \bar{X}_m \quad (1)$$

- c) Mendapatkan matriks covariance (C)

$$\text{cov}(X_1, X_2) = \frac{\sum_{n=1}^N (X_1 - \bar{X}_1)(X_2 - \bar{X}_2)}{(N-1)}$$

$$C = \begin{pmatrix} \text{cov}(X_1, X_1) & \text{cov}(X_1, X_2) & \dots & \text{cov}(X_1, X_M) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_2) & \dots & \text{cov}(X_2, X_M) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_M, X_1) & \text{cov}(X_M, X_2) & \dots & \text{cov}(X_M, X_M) \end{pmatrix} \quad (2)$$

- d) Mendapatkan *eigenvalue* (λ) dan *eigenvector* (V) dari matrik covariance

$$(C - \lambda)V = 0 \quad (3)$$

Teknik untuk mendapatkan *eigenvalue* dan *eigenvector* ini dapat digunakan beberapa metoda yaitu *power method*, *orthogonal iteration*, *QL method*, *Lanczos method*, dan *Jacobi's method*.

e) Mendapatkan *principle component* (PC)

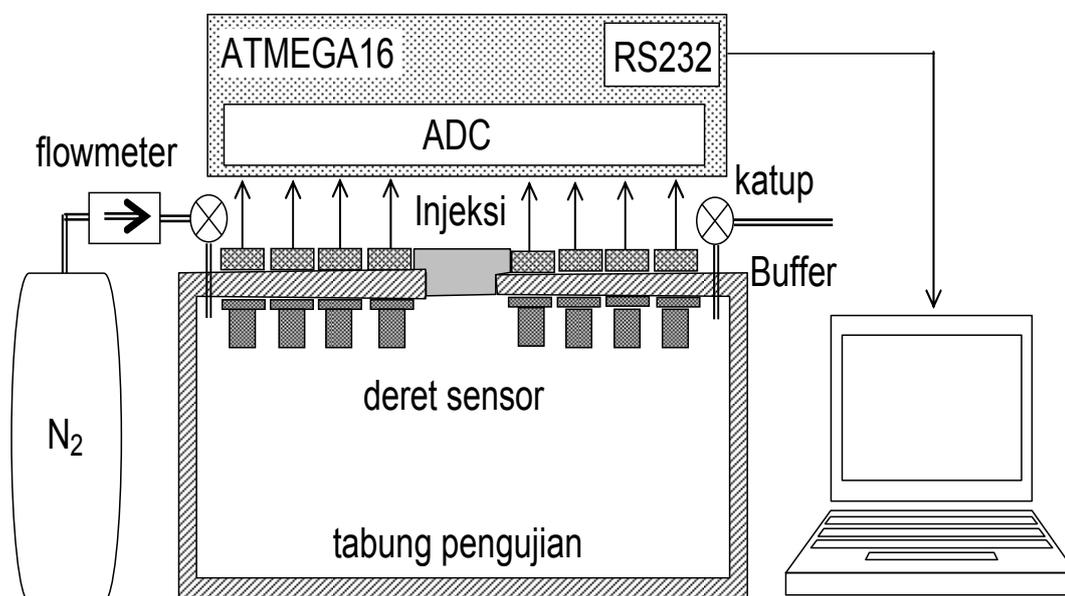
$$PC_i = X.V_i^T \quad (4)$$

dengan PC_1 adalah *principle component* pertama, V_1 *eigenvector* dengan nilai *eigenvalue* terbesar.

Pada penelitian ini dilakukan uji implementasi metoda PCA sebagai pemrosesan sinyal awal (*preprocessing*) pada algoritma *Neural Network*. Eksperimen meliputi penentuan jumlah *principle component* yang optimal, pengaruh PCA terhadap taraf kesalahan identifikasi dan lamanya fasa pelatihan *Neural Network*.

2. METODE PENELITIAN

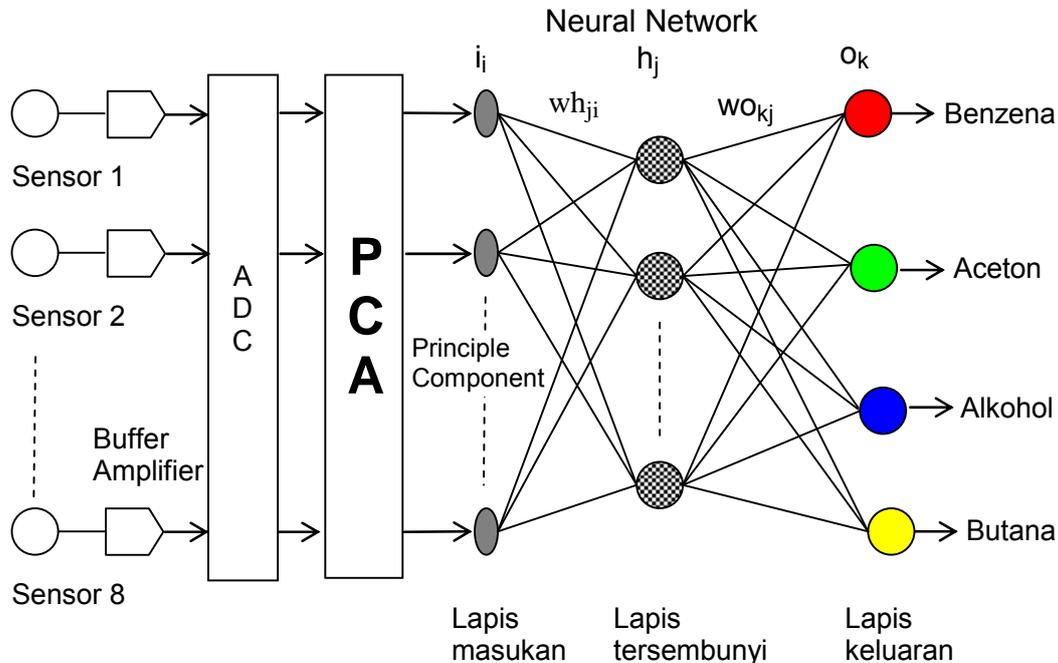
Sistem perangkat keras dalam penelitian ini ditunjukkan pada Gambar 1. Tabung pengujian mempunyai volume 0,6 L. Gas nitrogen (N_2) berkemurnian 99,999% dialirkan ke tabung pengujian dengan kecepatan aliran 1 L min^{-1} untuk melepaskan molekul uap atau gas pada tabung pengujian dan material sensor. Nilai yang stabil pada saat injeksi gas nitrogen ini digunakan sebagai nilai acuan pengukuran. Bahan sampel yang digunakan sebagai sumber gas adalah cairan yang mudah menguap seperti benzena, butana, alkohol dan aseton. Tiap sampel yang diinjeksikan kedalam tabung pengujian adalah sebesar $100 \mu\text{L}$. Semua pengukuran dilakukan pada suhu $32 \pm 1 \text{ }^\circ\text{C}$. Masing-masing sampel gas dilakukan pengukuran sebanyak lima kali pengukuran. Deret sensor yang digunakan terdiri dari delapan sensor gas semikonduktor TGS813, TGS2602, TGS2610, TGS2611, TGS2620, TGS2442, AF56 dan AF30. Tegangan sensor dan pemanas internal dicatu dengan tegangan 5 V. Masing-masing sinyal tegangan sensor dibaca oleh 10-bit *Analog to Digital Converter* (ADC) internal pada mikrokontroler ATMEGA16. Data masing-masing sensor dikirim ke komputer Pentium IV melalui komunikasi serial RS232. Visualisasi hasil pengukuran dibuat dengan menggunakan program *Visual Basic*.



Gambar 1. Desain perangkat keras eksperimental.

Sistem perangkat lunak dalam penelitian ini ditunjukkan pada Gambar 2. PCA mentransformasikan data sensor yang berkorelasi ke dalam bentuk yang bebas tidak berkorelasi. *Principle component* satu dengan yang lain tidak saling berkorelasi dan diurutkan

sedemikian rupa sehingga *principle component* yang pertama memuat paling banyak variasi dari data sensor. Hasil ekstraksi PCA tersebut diumpungkan ke *Neural Network* tiga lapis. Jumlah *node* masukan berkesesuaian dengan jumlah *principle component* yang digunakan, jumlah *neuron* tersembunyi adalah sepuluh dan jumlah *neuron* keluaran berkesesuaian dengan jumlah jenis gas yang dideteksi. *Neural Network* dilatih dengan menggunakan teknik propagasi balik: *batch gradient descent with momentum*, dengan parameter kecepatan dan momentum adalah 0,1 dan perulangan iterasi maksimum adalah 5000 *epoch*.



Gambar 2. Desain perangkat lunak eksperimental.

Perhitungan koreksi bobot antarkoneksi Lapis keluaran (Δw_o) dinyatakan:

$$\begin{aligned}\Delta w_{okj} &= \lambda(t_k - o_k) o_k (1 - o_k) h_j \\ \Delta w_{okj}(n+1) &= \varepsilon \Delta w_{okj} + \alpha \Delta w_{okj}(n)\end{aligned}\quad (5)$$

dengan t adalah vektor target pelatihan, ε adalah parameter yang menentukan kecepatan pelatihan, α adalah parameter momentum yang dapat mencegah osilasi sistem dan dapat membantu keluar dari minimum lokal fungsi kesalahan pada fasa pelatihan [5].

Kemudian proses dilanjutkan dengan melakukan perhitungan koreksi bobot antarkoneksi pada Lapis tersembunyi (Δw_h), dinyatakan:

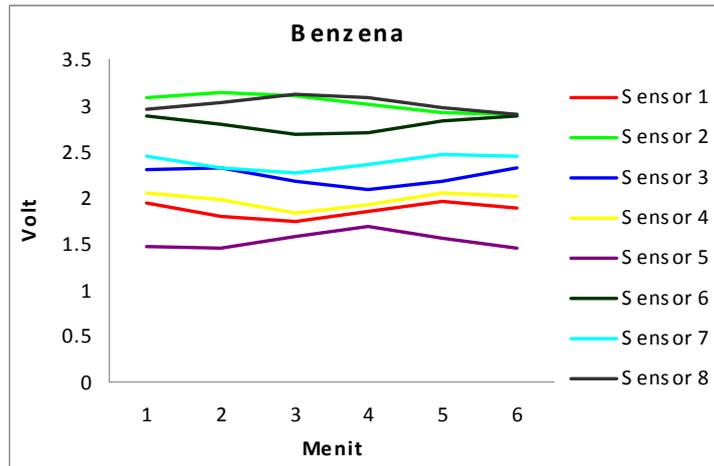
$$\begin{aligned}\Delta w_{hji} &= \lambda h_j (1 - h_j) \left\{ \sum_k \lambda (t_k - o_k) o_k (1 - o_k) w_{okj} \right\} i_i \\ \Delta w_{hji}(n+1) &= \varepsilon \Delta w_{hji} + \alpha \Delta w_{hji}(n)\end{aligned}\quad (6)$$

Tarif kesalahan identifikasi *Neural Network* pada fasa pelatihan dinyatakan sebagai *Mean Square Error (MSE)* yang dinyatakan:

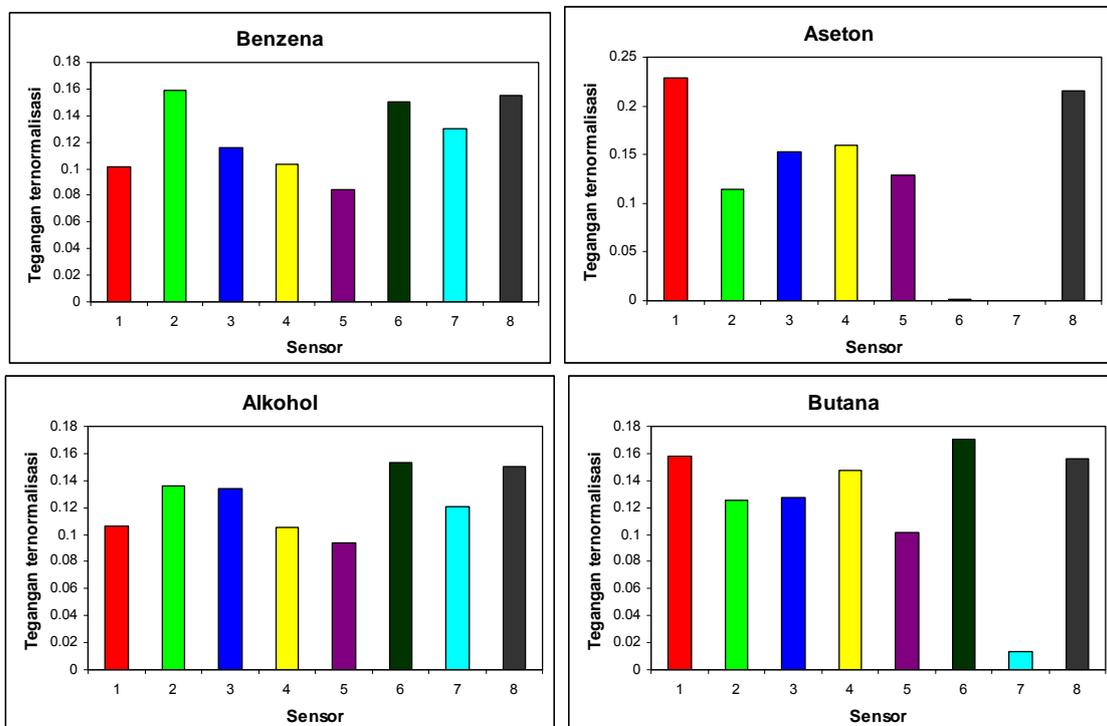
$$MSE = \frac{1}{N} \sum_{i=1}^N e_i^2 = \frac{1}{N} \sum_{i=1}^N (t_i - o_i)^2 \tag{7}$$

3. HASIL DAN PEMBAHASAN

Eksperimen diawali dengan penginjeksian N₂ ke dalam tabung pengujian sampai nilai tegangan keluaran sensor stabil. Kemudian sampel gas diinjeksikan ke tabung pengujian. Tanggapan waktu untuk tiap sensor ditunjukkan pada Gambar 3.



Gambar 3. Tanggapan sensor terhadap sebuah gas.

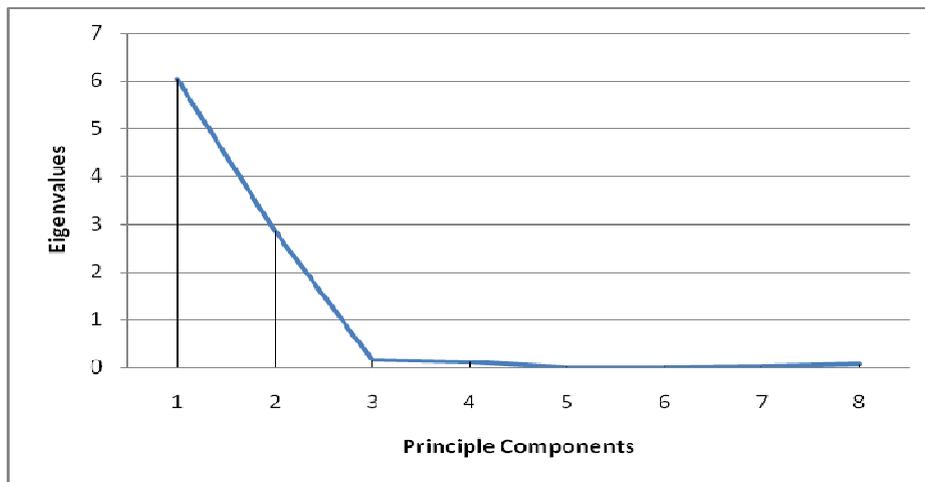


Gambar 4. Pola tanggapan sensor terhadap beberapa jenis gas.

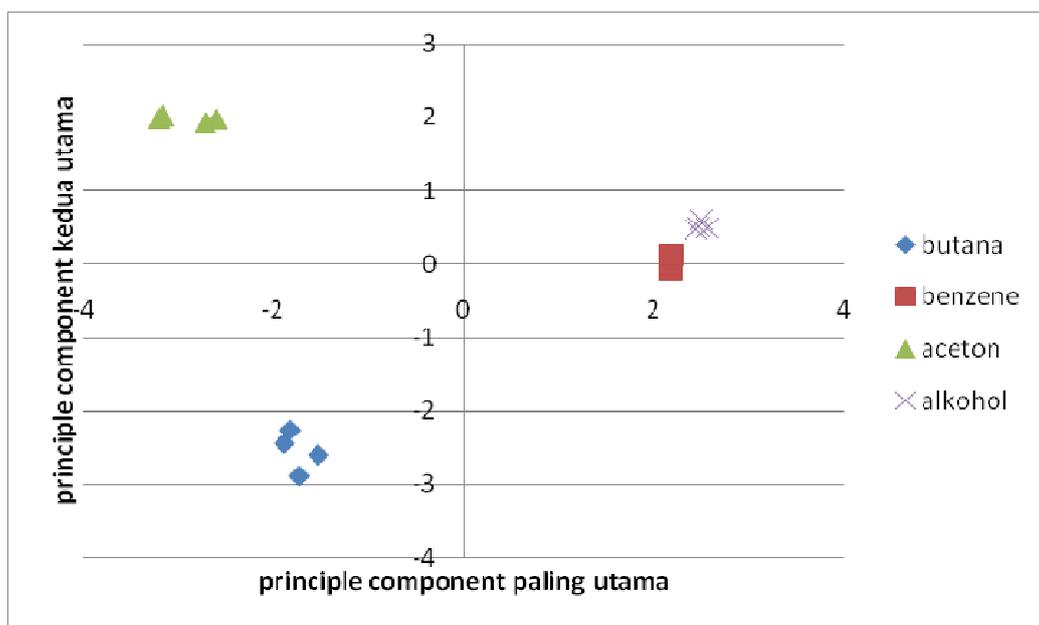
Untuk sampel yang sama terlihat bahwa masing-masing sensor mempunyai tanggapan yang berbeda. Pola keseluruhan sensor ini dapat dinyatakan dalam nilai tegangan ternormalisasi:

$$V_i^n = \frac{V_i}{\sum_{j=1}^8 V_j} \quad i = 1, 2, \dots, 8 \quad (8)$$

dengan V adalah tegangan keluaran sensor. Gambar 4 merupakan pola keluaran dari deret sensor untuk pemaparan beberapa sampel gas. Masing-masing gas dapat diwakili oleh pola yang khas tanpa terpengaruh besarnya konsentrasi [6].



Gambar 5. Grafik *eigenvalue* pada PCA.



Gambar 6. Pemetaan jenis gas menggunakan PCA.

Nilai pola ternormalisasi digunakan sebagai nilai masukan PCA. *Scree test* digunakan untuk melihat grafik *eigenvalue* hasil PCA, ditunjukkan pada Gambar 5. Penurunan secara drastis terjadi antara *principle component* kesatu sampai ketiga. Sedangkan setelah *principle component* keempat tidak ada perubahan nilai yang berarti. Sehingga tiga buah *principle component* yang utama bisa dianggap dapat mewakili variasi dari keseluruhan data. Untuk mengetahui pemetaan dua dimensi tiap-tiap gas yang diujikan maka digunakan dua *principle component* yang utama, ditunjukkan pada Gambar 6. Terlihat bahwa masing-masing jenis gas menempati lokasi yang berbeda sehingga dapat dikatakan bahwa deret sensor ini dapat membedakan jenis gas yang dideteksinya.

Tabel 1. Pengaruh PCA pada fasa pelatihan *Neural Network*.

Parameter	Jumlah <i>Principle Component</i>				
	0	1	2	3	4
Nilai MSE	0,014	0,121	0,05	$1,07 \cdot 10^{-6}$	$0,99 \cdot 10^{-6}$
Waktu training (detik)	39,2	39,6	23,5	22,0	23,4

Catatan: Jumlah iterasi adalah 5000 epoch.

Pengaruh implementasi PCA pada pengenalan pola *Neural Network* dilakukan pada saat fasa pelatihan dengan jumlah iterasi 5000 *epoch*, ditunjukkan pada Tabel 1. Jika tanpa memakai PCA yang mana *Neural Network* langsung mendapatkan sinyal ternormalisasi dari ke delapan sensor secara langsung, maka terlihat bahwa taraf kesalahan identifikasi sebesar 0,014 dan memerlukan waktu iterasi sebesar 39,2 detik. Jika *Neural Network* mendapat masukan dari PCA, maka terlihat bahwa semakin besar jumlah *principle component* yang digunakan akan semakin kecil taraf kesalahan identifikasi dan waktu pelatihan yang diperlukan semakin menurun. Pada eksperimen ini terlihat bahwa jumlah optimal *principle component* yang digunakan adalah tiga komponen utama yang memperkuat hasil dari *scree test* sebelumnya. Hal ini membuktikan bahwa metoda ini dapat digunakan sebagai pereduksi dimensi dari delapan dimensi menjadi tiga dimensi. Selain itu taraf kesalahan identifikasi menjadi lebih kecil yaitu $1,07 \cdot 10^{-6}$ atau dengan ratio $\sim 10^{-4}$ dan waktu iterasi yang lebih pendek yaitu 22,0 detik atau dengan rasio ~ 0.6 . Jumlah *principle component* diatas tiga akan menunjukkan performa yang tidak berbeda secara signifikan dan akan lebih menambah kekompleksan sistem identifikasi yang digunakan.

4. SIMPULAN

Telah dilakukan pengujian pengaruh implementasi PCA sebagai pemrosesan awal pada *Neural Network* dalam pengidentifikasian jenis gas yang menggunakan deret sensor semikonduktor. Hasil eksperimen menunjukkan bahwa PCA dapat mereduksi jumlah variabel masukan yang semula berdimensi delapan menjadi berdimensi tiga tanpa mengurangi tingkat informasinya. Taraf kesalahan identifikasi menjadi lebih kecil dengan rasio $\sim 10^{-4}$ dan periode pelatihan yang lebih pendek dengan rasio ~ 0.6 . Secara umum dapat disimpulkan bahwa implementasi PCA pada *Neural Network* dapat meningkatkan performanya yang meliputi taraf identifikasi dan waktu yang diperlukan pada fasa pelatihan.

DAFTAR PUSTAKA

- [1] Fort A, Machetti N, Rocchi S, Serrano B, Tondi L, Ulivieri N, Vignoli V, Sberveglieri G, "Tin Oxide Gas Sensing: Comparison Among Different Measurement Techniques for Gas Mixture Classification". IEEE Instrumentation and Measurement Technology Conference, Budapest, 2001
- [2] Bermak A, Belhouari SB, Shi M, Martinez D, "Pattern Recognition Techniques for Odor Discrimination in Gas Sensor Array", Encyclopedia of Sensors Vol X, pp.1-17, American Scientific Publishers, 2006.
- [3] Mirmohseni A, Rostamizadeh K, "Quartz Crystal Nanobalance in Conjunction with Principle Component Analysis for Identification of Volatile Organic Compounds", Sensors, 6, pp.324-334, 2006

- [4] Smith LI, "**A Tutorial on Principle Component Analysis**", Cornell University, 2005.
- [5] Sinha RK, "**Back propagation Artificial Neural Network to Detect Hyperthermic Seizures in Rat**", Online Journal of Health, Allied Scs 4:1, 2002.
- [6] M Rivai, "**Pengenalan Uap Campuran Menggunakan Sensor Kristal SiO₂**", Industrial Electronics Seminar, EEPIS-ITS, pp.D51-D55, 2007.