

# Perancangan Sistem Pengenalan Suara Untuk Pengamanan Dan Pemantauan Fasilitas PLTA

Yufliha Dian Ayunisa, Ir. Achmad Affandi, DEA , Ir. Wirawan, DEA  
*Jurusan Teknik Elektro FTI – ITS*

**Abstrak** - Sistem pengawasan audio dapat menjadi aset yang besar bagi lingkungan perumahan atau bisnis. Dengan sistem ini dapat menambah penjagaan di lingkungan sekitar dan dapat membantu dalam menangkap pencuri dan perampok. Sistem keamanan pengawasan audio dapat juga diaplikasikan untuk melengkapi sistem pengawasan video yang sudah ada atau diaplikasikan sebagai perangkat terpisah yang berfungsi sebagai mikropon keamanan saja. Tugas akhir ini akan membahas mengenai pengenalan suara yang ada di suatu area pengawasan, dimana akan dibuat sebuah sistem yang dapat mendeteksi suara yang normal dan mencurigakan dari berbagai suara yang tertangkap mikropon pengawas di area pengawasan.

Dalam tugas akhir ini dilakukan pengenalan suara dengan suatu algoritma ekstraksi ciri yang disebut metode Linear Predictive Coding (LPC) untuk analisa sinyal. LPC merupakan salah satu metode analisis sinyal suara yang menyatakan ciri-ciri penting dari sinyal suara tersebut dalam bentuk koefisien-koefisien LPC. Sedangkan untuk pengenalan pola digunakan Jaringan Saraf Tiruan dengan algoritma pelatihan Learning Vector Quantization (LVQ). Jaringan saraf Tiruan dengan pelatihan LVQ memiliki kemampuan mengklasifikasikan vector masukan ke kelas target yang telah ditentukan sebelumnya. Eksperimen dilakukan dengan beberapa perubahan nilai parameter untuk memperoleh presentase pengenalan tertinggi.

**Kata kunci** : Pengenalan Suara, Linier Predictive coding (LPC), Jaringan saraf tiruan, Learning vector Quantization (LVQ)

## I. PENDAHULUAN

Dalam ilmu pengetahuan dan teknologi khususnya bidang komunikasi, pengolahan sinyal memegang peranan yang penting. Penelitian yang intensif dalam bidang pengolahan sinyal menyebabkan teknologi komunikasi berkembang dengan pesat. Salah satunya adalah pengenalan suara. Salah satu aplikasi teknologi pengenalan suara yang berkembang pesat saat ini adalah dalam bidang keamanan atau monitoring. Monitoring

berbasis pengenalan suara sangatlah berguna terutama pada situasi dimana pada saat sensor lain seperti video gagal dalam mendeteksi situasi tersebut. Sebagai contoh, ketika objek dalam kondisi yang gelap, sensor suara dapat lebih tepat dalam mendeteksi keberadaan objek tersebut dengan asumsi bahwa keberadaan objek tersebut menghasilkan beberapa suara.

Dalam tugas akhir ini dilakukan penelitian mengenai pengenalan suara yang berasal dari suara-suara yang tertangkap dari mikropon pengawas, dalam hal ini suara-suara yang diteliti adalah suara yang sering muncul di lingkungan penelitian yaitu PLTA kemudian diklasifikasikan sebagai suara yang normal dan mencurigakan. Tema ini diambil mengingat keterbatasan manusia dalam mengenali berbagai suara yang begitu sangat banyak ragamnya serta banyak bunyi yang kadang hampir serupa antara suara yang satu dengan yang lain yang ada di lingkungan penelitian. Dengan penelitian ini diharapkan masalah tersebut dapat terpecahkan. Pengenalan suara dilakukan dengan pengekstraksian koefisien LPC untuk analisa sinyal, sedangkan untuk pengenalan pola digunakan Jaringan Saraf Tiruan (JST) Learning Vector Quantization (LVQ). Terdapat 25 buah sampel suara dari 5 jenis suara digunakan sebagai input pada proses pelatihan JST. Setelah dilakukan proses pelatihan, sistem dicoba untuk mengenali data baru.

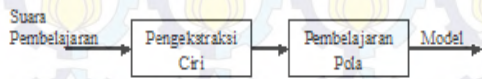
## II. TEORI PENUNJANG

### A. Pengenalan Suara Digital

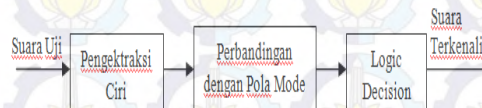
Pengenalan suara merupakan salah satu upaya agar suara dapat dikenali atau diidentifikasi sehingga dapat dimanfaatkan. Pengenalan suara dapat dibedakan ke dalam tiga bentuk pendekatan, yaitu pendekatan akustik-fonetik (*the acoustic-phonetic approach*), pendekatan kecerdasan buatan (*the artificial intelligence approach*), dan pendekatan pengenalan pola (*the pattern recognition approach*). Pendekatan pengenalan pola terdiri dari dua langkah yaitu pembelajaran pola suara dan pengenalan pola melalui



perbandingan pola. Tahap perbandingan pola adalah tahap bagi ucapan yang akan dikenali, dibandingkan polanya dengan setiap kemungkinan pola yang telah dipelajari dalam fase pembelajaran, untuk kemudian diklasifikasi dengan pola terbaik yang cocok. Blok diagram pembelajaran pola pada pengenalan suara ditunjukkan pada Gambar



(a) Blok diagram pembelajaran pola



(b) Blok diagram pengenalan suara

Gambar 1 Blok diagram pembelajaran pola dan pengenalan suara.

Berikut ini merupakan penjelasan dari masing-masing blok:

#### 1. Pengekstraksi ciri.

Merupakan proses mendapatkan sederetan besaran pada bagian sinyal masukan untuk menetapkan pola pembelajaran atau pola uji. Pada sinyal suara, ciri-ciri besaran biasanya merupakan keluaran dari beberapa bentuk teknik analisis spektrum seperti *filter bank analyzer*, LPC atau DFT (*Discrete Fourier Transform*).

#### 2. Pembelajaran Pola

Satu atau lebih pola pembelajaran yang berhubungan dengan bunyi suara dari kelas yang sama, digunakan untuk membuat pola representatif dari ciri-ciri kelas tersebut. Hasilnya yang biasa disebut dengan pola referensi, dapat menjadi sebuah model yang mempunyai karakteristik bentuk statistik dari ciri-ciri pola referensi.

#### 3. Perbandingan dengan Pola Model

Pola uji yang akan dikenali, dibandingkan dengan setiap kelas pola referensi. Kesamaan besaran antara pola uji dengan setiap pola referensi akan dihitung.

#### 4. Logic Decision

Menentukan kelas pola referensi mana yang paling cocok untuk pola uji berdasarkan klasifikasi pola.

Pengenalan suara secara umum dapat dibagi menjadi tiga tahap, yaitu : ekstraksi ciri, pemodelan, dan pengenalan. Ekstraksi ciri adalah upaya untuk memperoleh ciri dari sinyal suara.

Salah satu metode yang dapat digunakan untuk proses ekstraksi ciri adalah LPC. Setelah dilakukan ekstraksi ciri dan memperoleh ciri dari sinyal suara, kemudian dilakukan pengenalan pola, teknik pengenalannya melalui Jaringan Saraf Tiruan dengan algoritma pelatihan LVQ (*Learning Vector Quantization*).

### B. Proses Analisis Sinyal

Langkah-langkah analisis untuk pengenalan suara adalah:

- **Preemphasis.** Pada tahap ini digunakan untuk mendatarkan spektral sinyal dan meningkatkan keaslian sinyal pada pemrosesan sinyal yang selanjutnya.
- **Frame Blocking.** Pada tahap ini sinyal suara yang telah teremphasi dibagi menjadi frame-frame dengan masing-masing frame memuat N cuplikan dan frameframe yang berdekatan dipisahkan sejauh M cuplikan.
- **Windowing.** Pada langkah ini dilakukan fungsi weighting pada setiap frame yang telah dibentuk pada langkah sebelumnya.

$$w(n) = 0,54 - 0,46 \cos\left(\frac{2\pi n}{N-1}\right), 0 \leq n \leq N-1$$

- **Autokorelasi.** Pada tahap ini masing-masing frame yang telah di windowing diautokorelasikan dengan nilai autokorelasi yang tertinggi adalah orde dari analisa LPC
- **Analisis LPC.** Langkah selanjutnya adalah analisa LPC, dimana pada tahap ini nilai autokorelasi pada setiap frame diubah menjadi satu set LPC parameter dalam hal ini yang dipakai adalah koefisien LPC.

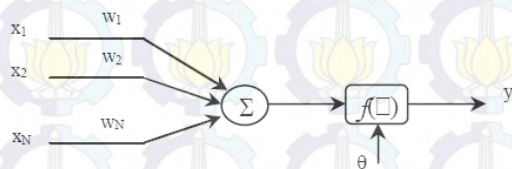
### C. Jaringan Saraf Tiruan

JST bisa didefinisikan kurang lebih sebagai system komputasi yang didasarkan pada pemodelan saraf biologis (*neuron*) melalui pendekatan dari sifat-sifat komputasi biologis (*biological computation*). JST bisa dibayangkan berupa jaringan dengan elemen pemroses sederhana yang saling terhubung. Elemen pemroses berinteraksi melalui sambungan yang variabel, disebut bobot, dan bila diatur secara tepat dapat menghasilkan sifat yang diinginkan. Pada tahun 1943 Warren McCulloch dan Walter Pitts memperkenalkan model matematis sederhana dari neuron, yang disebut model McCulloch-Pitts, terdiri dari banyak masukan dan satu keluaran. Model dari neuron pada Gambar 2 mempunyai persamaan



$$y = f \left( \sum_{i=1}^N w_i x_i - \theta \right)$$

Dengan  $x_i$  = Sinyal masukan,  $i = 1, 2, 3, \dots, N$  ( $N$  = jumlah simpul masukan)  
 $w_i$  = Bobot hubungan atau sinapsis  
 $\theta$  = Treshold atau bias  
 $f(\bullet)$  = Fungsi aktivasi  
 $y$  = Sinyal keluaran dari neuron



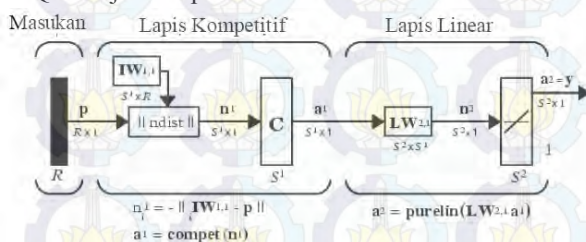
Gambar 2 Model Mc-Culloch-Pitts untuk neuron

Ide dasar JST adalah konsep belajar. Jaringan-jaringan belajar melakukan generalisasi karakteristik tingkah laku obyek. JST menyimpan ilmu pengetahuannya dalam nilai bobot sambungan (seperti *synapsis* dalam otak manusia) dan elemen-elemen (*neuron*) yang menghasilkan keluaran.

Terminologi belajar artinya membentuk pemetaan (fungsi) yang menggambarkan hubungan antara vector masukan dan vektor keluaran. Untuk menyelesaikan permasalahan, JST memerlukan algoritma untuk belajar, yaitu bagaimana konfigurasi JST dapat dilatih untuk mempelajari data historis yang ada. Dengan pelatihan ini, pengetahuan yang terdapat pada data bisa diketahui dan direpresentasikan dalam bobot sambungannya.

#### D. Learning Vector Quantization (LVQ)

Jaringan LVQ mempunyai target yang akan dicapai. Lapisan kompetitif belajar mengenali dan mengklasifikasikan vektor-vektor masukan. Jika ada 2 vektor yang hampir sama, maka lapisan kompetitif akan menempatkan keduanya pada kelas yang sama. Dengan demikian LVQ belajar mengklasifikasikan vektor masukan ke kelas target yang ditentukan oleh pengguna. Arsitektur jaringan LVQ ditunjukkan pada Gambar 3.



Gambar 3 Arsitektur jaringan LVQ

Jaringan LVQ terdiri dari 2 lapis yaitu lapis kompetitif dan lapis linear. Lapis kompetitif disebut juga *Self Organizing Map* (SOM). Disebut lapis kompetitif karena neuron neuron berkompetisi dengan algoritma kompetisi yang akan menghasilkan neuron pemenang (*winning neuron*).

### III. PERANCANGAN DAN IMPLEMENTASI SISTEM

Program pengenalan suara ini disusun melalui beberapa tahapan, yaitu perancangan, pembuatan, dan pengujian. Diagram alir program pengenalan suara dapat dilihat pada Gambar 4.



Gambar 4 Diagram alir program pengenalan suara

Masukan suara yang akan diolah dalam program simulasi terdiri dari 2 jenis, yaitu data suara terekam dan suara masukan baru.

Data suara terekam diperoleh dari perekaman lima buah suara yang sering muncul di lokasi pemantauan, yaitu suara ledakan, teriakan, langkah kaki, percakapan dan suara – suara alam yang akan dimasukkan dalam database. Hasil perekaman kemudian dibaca untuk mendapatkan bentuk diskret dari sinyal suara untuk mempermudah pengolahan selanjutnya. Sinyal suara kemudian diekstraksi sampai didapatkan orde LPC. Hasil ekstraksi merupakan masukan untuk proses pelatihan pada jaringan saraf tiruan dengan algoritma LVQ. Pengujian data rekaman dan data baru dilakukan setelah didapat perubahan bobot pada jaringan.

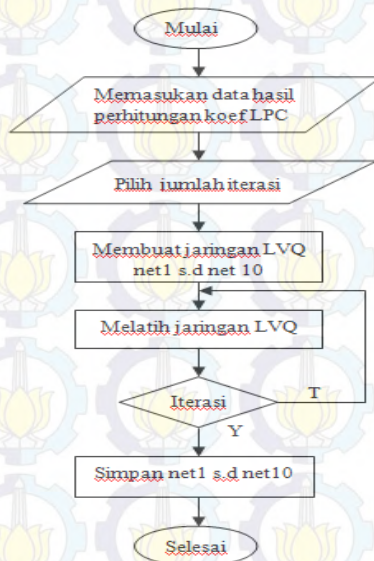
#### A. Pelatihan Jaringan LVQ

Pelatihan Jaringan Saraf Tiruan yang dibuat dalam program simulasi ini dilatih dengan menggunakan metode LVQ. Data hasil



penghitungan koefisien LPC menjadi masukan untuk pelatihan jaringan agar diperoleh perubahan bobot pada neuron.

Pengklasifikasian suara oleh JST menggunakan skema Pembagian Pasangan Biner (*Binary Pair Partitioned scheme*) sehingga 1 buah jaringan mengklasifikasikan 2 buah tipe suara. Dengan demikian digunakan 10 buah net atau jaringan yang masing masing merupakan variasi dari 2 jenis tipe suara, yaitu /bang/-/scream/, /bang/-/nature/, /bang/-/speech/, /bang/-/steps/, /scream/-/nature/, /scream/-/speech/, /scream/-/steps/, /nature/-/speech/, /nature/-/steps/, dan /speech/-/steps/. Proses pelatihan jaringan ditunjukkan diagram alir pada Gambar 5.



Gambar 5. Diagram alir pelatihan jaringan LVQ

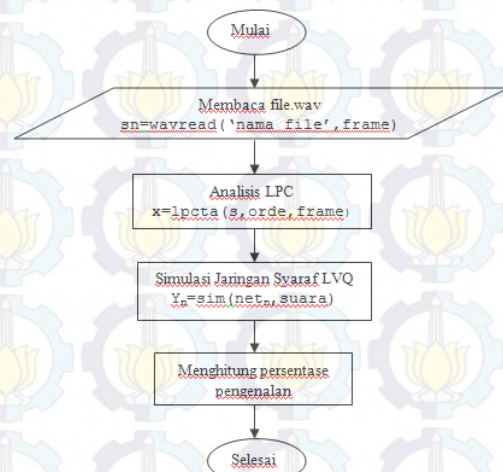
Pembuatan dan pelatihan jaringan memerlukan nilai variabel-variabel elemen masukan, jumlah *hidden neuron*, laju pembelajaran, algoritma pembelajaran, dan *epoch* /iterasi.

Penentuan jumlah neuron dilakukan dengan cara eksperimen. Ada 2 cara yang dapat dipilih yaitu metode konstruktif / *evolutile* dan destruktif / *subtractive*. Pada metode konstruktif pertama-tama dipilih jumlah neuron yang kecil kemudian diuji hasilnya, jika hasilnya tidak sesuai dengan yang diharapkan maka jumlah neuron ditambah lalu diuji lagi. Penambahan jumlah neuron dihentikan bila tidak lagi dihasilkan peningkatan hasil. Metode destruktif merupakan kebalikan dari metode konstruktif. Laju pembelajaran juga ditentukan dengan cara eksperimen, dimulai dari 0,05 sampai 0,9.

## B. Pengujian Data Latihan dan Data Baru

Dalam program simulasi, umumnya setiap pembacaan suara, diawali dengan proses membunyikan *file* suara tersebut dan menampilkan grafik.

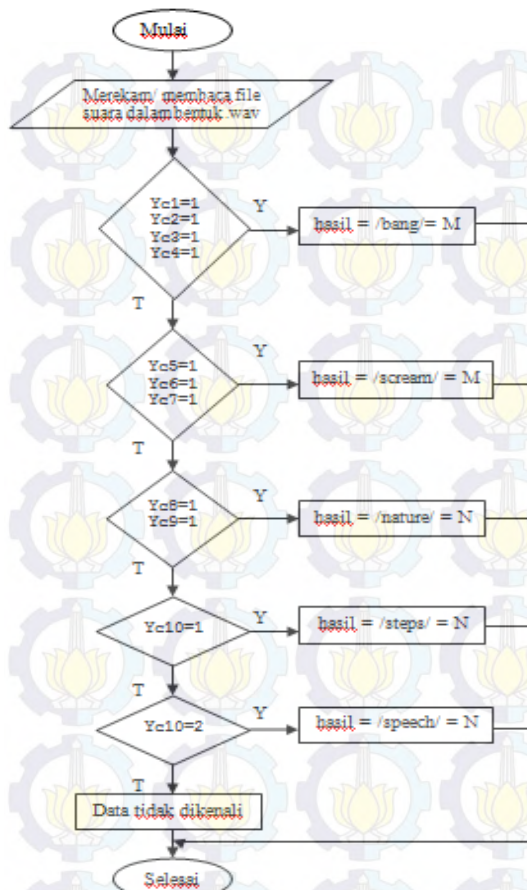
Data latihan yang akan diujikan di sini berupa satu paket masukan koefisien LPC untuk masing-masing suara, kemudian disimulasikan ke seluruh jaringan dari net1 sampai net10 untuk mendapatkan hasil pengenalan masing-masing suara. Diagram alir untuk proses pengujian data latihan dan data baru ditunjukkan oleh Gambar 6.



Gambar 6. Diagram alir pengujian data latihan dan data baru

Masukan untuk pengujian data tunggal berupa matriks koefisien LPC dari suara baru yang baru direkam. Matriks koefisien kemudian disimulasikan pada semua net di jaringan yang memberikan persentase pengenalan terbaik. Gambar 7 menunjukkan diagram alir pengujian data tunggal yang akan memberikan hasil akhir berupa tampilan dalam bentuk teks suara hasil pengenalan terhadap suara masukan.





Gambar 7 Diagram alir pengujian data tunggal

#### IV. ANALISA DAN PENGUJIAN

Pada program ini terdapat beberapa parameter yang dapat divariasi, seperti yang ditunjukkan pada Tabel 1.

Tabel 1  
Variasi nilai parameter

Parameter	Variasi
Orde LPC	8
	10
	14
	16
Iterasi	500
	1000
	2000

Dengan variasi dari ketiga jenis parameter di atas, maka akan diperoleh 12 kombinasi jaringan.

Variasi 12 kombinasi nilai parameter diuji satu persatu hingga menghasilkan 12 set jaringan yang memiliki persentase keberhasilan tertentu, baik untuk data latihan maupun data baru.

Tabel 2

Hasil pengenalan data latihan untuk orde 8, dan iterasi 500

Klasifikasi suara	Presentase pengenalan	
m dan n	m_m=8	m_n=2
	n_m=0	n_n=15

Pada Tabel 2, variabel m\_m berarti suara masukan /mencurigakan/ yang dikenali sebagai suara /mencurigakan/, m\_n berarti suara masukan /mencurigakan/ yang dikenali sebagai suara /normal/, dan seterusnya.

Tabel 3

Statistik pengenalan untuk orde 8, dan iterasi 500

Suara	Normal	Mencurigakan
Normal	100	0
Mencurigakan	20	80

Kemudian untuk total presentase pengenalan dihitung dengan :

$$\% \text{pengenalan} = \frac{n_n + m_m}{2} = \frac{100 + 80}{2} = 90\%$$

Selain menggunakan data latihan (data yang digunakan untuk melatih jaringan), perlu juga dilakukan pengujian dengan data uji untuk menilai kinerja jaringan terhadap data dari luar. Proses pengujian ini adalah sama dengan pengujian data latihan, tetapi tanpa melalui proses pelatihan jaringan.

Pengujian untuk setiap jaringan dilakukan 3 kali. Rata-rata hasil pengenalan dari ketiga pengujian ditunjukkan pada Tabel 4

Tabel 4

Rata-rata persentase pengenalan untuk data latihan dan data baru

Orde	Iterasi					
	Data Latihan			Data Baru		
	500	1000	2000	500	1000	2000
8	91.67	95	100	88.3	86.67	93.3
10	93.3	96.67	100	88.3	83.3	85
14	93.3	91.67	91.67	88.3	85	85
16	96.67	91.67	95	86.67	86.67	85

Menurut hasil yang ditunjukkan Tabel 4, maka pengenalan terbaik untuk data latihan diperoleh untuk orde LPC 8 dan 10, dan iterasi 2000. Sedangkan untuk data baru pengenalan terbaik diperoleh pada orde LPC 8 dan iterasi 2000.



Perbedaan antara hasil pengenalan data latihan dan data baru disebabkan data baru memiliki karakteristik sinyal yang berbeda dengan data latihan sehingga akan memberikan koefisien LPC yang berbeda pula. Akibatnya jaringan tidak dapat mengklasifikasikan suara masukan dengan tepat. Masalah ini dapat diatasi dengan menambah data yang dilatihkan pada jaringan sehingga referensi ciri suara yang dimiliki jaringan bertambah banyak.

## V. PENUTUP

### 1. KESIMPULAN

Kesimpulan yang dapat diambil dari pembuatan dan pengujian program simulasi pengenalan suara ini adalah sebagai berikut :

1. Orde LPC yang digunakan dalam analisis LPC menentukan hasil pengenalan suara, dimana orde yang memberikan hasil terbaik adalah LPC-10.
2. Tingkat pengenalan data latihan lebih tinggi dibandingkan data baru.
4. Keadaan lingkungan dan kondisi pada saat perekaman suara untuk database dan perekaman untuk pengujian akan memberikan pengaruh terhadap hasil pengenalan.

### 2. SARAN

1. Penambahan data latihan (untuk database) dari jenis suara yang lain akan semakin melengkapi referensi pengenalan.
2. Penelitian ini dapat dikembangkan dengan metode ekstraksi ciri lain seperti Mel-Frequency Cepstrum Coefficients (MFCC).
3. Penelitian ini dapat dilakukan dengan menggunakan metode pencocokan pola yang lain seperti HMM (Hidden Markov model), JST dengan algoritma lain, dan sebagainya.
4. Penerapan system keamanan dan monitoring berdasarkan pengenalan suara pada PLta dapat dilakukan pada penelitian selanjutnya dengan menggunakan beberapa microphone surveillance yang diletakkan di beberapa lokasi pengawasan.

## DAFTAR PUSTAKA

- [1] Atrey, P.K, Maddage, Namunnu C, Kankanali, Mohan S., 2006, "Audio Based Event Detection for Multimedia Surveillance", *IEEE Accoustics, Speech and Signal Processing*, pp. V-V.
- [2] Bradbury, Jeremy, Linear Predictive Coding, Paper, 2002
- [3] Bullimaria, John A., Introduction to Neural Network  
<http://www.cs.bham.ac.uk/~jbx/inn.htm>
- [4] C . De Stefano, C. Samsone and M.Vento, Comparing Generalization and Recognition Capability of Learning Vector Quantization and Multi-Layer Perceptron Architectures, University of Naples, Napoli
- [5] CLAVEL, C, EHRETTE, T, RICHARD. G, THALES RES, TECHNOL. FRANCE, ORSAY, "EVENTS DETECTION FOR AN AUDIO-BASED SURVEILLANCE SYSTEM", IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, 2005
- [6] Harma, Aki, Mckinney, Martin F., "Automatic Surveillance of The Acoustic Activity in Our Living Environment ", IEEE Multimedia and Expo, pp 4.
- [7] Lasse L Mølgaard, Kasper W Jørgensen, Speaker Recognition: Special Course; IMM-DTU; 2005
- [8] RADHAKRISHNAN. R, DIVAKARAN. A, SMARAGDIS. A, "AUDIO ANALYSIS FOR SURVEILLANCE APPLICATIONS", IEEE WORKSHOP ON APPLICATIONS OF SIGNAL PROCESSING TO AUDIO AND ACOUSTICS, 2005
- [9] Vergara, Luis, "Audio signal processing for surveillance applications", International Conference on Emerging Security Information, Systems and Technologies SECURWARE, Valencia, 2007

## BIODATA PENULIS



Yufliha Dian Ayunisa, dilahirkan di Brebes 17 Desember 1988. Memulai pendidikan Sekolah Dasar di SDN Pesantunan 2 Brebes, kemudian meneruskan pendidikan di SLTPN 2 Brebes dan SMA Negeri 2 Brebes. Kemudian meneruskan pendidikan D3 Teknik Elektro bidang studi teknik telekomunikasi di Politeknik Negeri Semarang dan melanjutkan ke jenjang S1 di Teknik Elektro bidang studi Telekomunikasi Multimedia ITS. Sekarang sedang mengerjakan tugas akhir di Bidang Studi Telekomunikasi Multimedia, Jurusan Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember (ITS) Surabaya.