

# BAB I

## PENDAHULUAN

### I.1 Latar Belakang

Prinsip teknologi dikembangkan adalah untuk membuat alat atau sarana yang dapat membantu dan memberi kemudahan bagi manusia untuk melakukan kegiatan dalam hidup. Seiring dengan perkembangan teknologi, manusia selalu menginginkan peningkatan kualitas dan kepraktisan dari alat-alat tersebut. Oleh karena itu dibentuklah mesin-mesin yang dapat berinteraksi dengan manusia. Teknologi ini disebut teknologi *human machine*.

Teknologi *human machine* bertujuan menciptakan mesin yang memiliki kemampuan mengartikan informasi yang diucapkan manusia, bertindak sesuai dengan informasi tersebut dan berbicara untuk menyempurnakan pertukaran informasi. Maksudnya adalah menciptakan suatu mesin dengan kecerdasan buatan sehingga dapat berinteraksi dengan manusia melalui suara. Upaya atau penelitian kearah tersebut masih tetap dilakukan untuk mendapatkan hasil yang maksimal. Salah satu penelitian tentang suara (*Speech research*) yang berperan dalam teknologi *human machine* adalah *Automatic Speech Recognition* (ASR) [1].

Pengenalan ucapan atau lebih dikenal sebagai ASR adalah suatu proses dimana komputer (jenis mesin lainnya) dapat mengenal kata-kata yang diucapkan manusia. Proses ini disebut juga mengartikan ucapan manusia dalam komputer. Tujuan utama dari pengenalan ucapan adalah untuk mendapatkan cara yang efisien bagi manusia dalam berkomunikasi dengan komputer.

ASR merupakan suatu bidang yang sangat menarik dan memiliki peluang yang besar untuk dikembangkan. Dari segi kepentingan komersial, bidang ini berkembang dengan cepat . ASR memiliki banyak aplikasi dan keuntungan. Sebagai contoh *personal computer* (PC) dapat dikontrol dengan suara dan dapat digunakan untuk sistem diktasi. ASR dapat menjadi aplikasi yang penting bagi orang yang memiliki ketidakmampuan fisik, seperti tuna rungu . Aplikasi lain yang bisa dilakukan yaitu *environmental control* seperti menhidupkan lampu listrik, mengontrol TV, dan lain-lain.

Teknologi pengolahan suara atau *speech processing technology*, sinyal suara yang diproses harus diekstraksi terlebih dahulu. Hal ini bertujuan untuk mendapatkan parameter - parameter yang menjadikan setiap suara atau kata dari pembicara yang diproses memiliki perbedaan-perbedaan khusus dan bersifat unik. Ekstraksi parameter dapat dilakukan dengan menggunakan *MFCC (Mel Frequency Cepstrum Coefficient)*. Berdasarkan penelitian (2), MFCC memiliki kemampuan untuk menangkap karakteristik dan informasi-informasi penting yang terkandung dalam sinyal suara, kemudian menghasilkan data seminimal mungkin tanpa menghilangkan informasi – informasi penting yang ada. Oleh karena itu dipilih MFCC untuk melakukan ekstraksi parameter. [2]

Tahapan selanjutnya dari proses pengolahan suara khususnya pada aplikasi *speech recognition* yaitu proses pencocokan pola kata . Ada beberapa metoda yang sering digunakan dalam tahap ini. Diantaranya *Hidden Markov Model (HMM)*, *Fuzzy Logic*, *Neural Network*, dan *Genetic Algoritim (GA)*. Teknik *fuzzy logic* merupakan sistem yang paling sederhana, namun hasil yang didapat kurang akurat dibanding metode lainnya. Selain itu, *fuzzy logic* dapat memberikan suatu keputusan yang baik untuk kondisi yang tidak pasti (mutlak). Kondisi ini bisa kita golongan salahsatunya adalah sinyal suara yang memiliki pola berbeda-beda satu sama lain. Sedangkan teknik *neural network* memerlukan proses pembelajaran dan iterasi yang sangat banyak dan panjang, namun cukup akurat. Oleh karena itu, pada penelitian ini dilakukan pengembangan sistem neural network menjadi sistem *hybrid* dengan *fuzzy logic* . Sistem ini dinamakan sistem *neuro – fuzzy*. Teknik ini merupakan teknik yang cukup akurat , karena gabungan dari teknik *neural network* dan *fuzzy logic* yang saling melengkapi kekurangan masing – masing.[3]

Teknologi pengenalan ucapan saat ini telah mengalami perkembangan yang cukup pesat. Semakin berkembangnya aplikasi sistem pengenalan ucapan maka dibutuhkan sistem yang *robust*, yaitu sistem yang mampu bekerja dengan baik diberbagai kondisi lingkungan [4]. Sistem pengenalan ucapan yang *robust* masih merupakan permasalahan yang sangat sulit meskipun telah banyak dilakukan penelitian. Permasalahan tersebut muncul dengan adanya gangguan dari

lingkungan tempat sistem beroperasi, seperti derau. Sehingga tidak mudah bagi sistem pengenalan ucapan untuk memprediksi sinyal suara yang terganggu oleh derau. Derau pada lingkungan pengujian yang tidak terprediksi dan tidak teramati selama pelatihan dan pengujian menyebabkan ketidakcocokan antara kondisi pelatihan dan pengujian. Ketidakcocokan ini menyebabkan kinerja sistem menurun secara signifikan. Untuk mengatasi masalah tersebut, telah dikembangkan suatu metoda yang bertujuan untuk meningkatkan kualitas dan kejelasan pada sinyal yang mengandung derau yang disebut dengan metoda *speech enhancement*. [5]

Metoda *speech enhancement* yang digunakan pada tugas akhir ini adalah dengan menggunakan *filter* adaptif RLS (*Recursive Least Squares*). *Filter* ini memiliki keunggulan yaitu optimum di tiap kondisi dan mampu memberikan kinerja terbaik dalam proses pem-*filter*-an. Berbeda dengan *filter band* frekuensi tertentu seperti *low pass* atau *high pass*. Selain itu, *filter* digital ini juga mempunyai kemampuan untuk mengubah bobot koefisiennya secara otomatis, menyesuaikan dengan sinyal input untuk memperkecil *error*. Sistem adaptif juga mampu menyesuaikan dan beradaptasi langsung dengan setiap perubahan yang terjadi di lingkungan sekitar maupun dengan perubahan kondisi yang terjadi pada sistem itu sendiri. [6]

Beberapa penelitian yang berkaitan dengan *speech recognition* dan *speech enhancement*, diantaranya yaitu:

1. **Elsen Ronando dan M. Isa Irawan** dalam penelitiannya yang berjudul “Pengenalan Ucapan Kata Sebagai Pengendali Gerakan Robot Lengan Secara *Real Time* dengan Metode *Linear Predictive Coding – Neuro Fuzzy*” yang membahas tentang bagaimana mengendalikan robot lengan secara interaktif berdasarkan perintah suara ucapan, menggunakan LPC sebagai metode ekstraksi ciri dan *Neuro Fuzzy* untuk *recognize*-nya. [7]
2. **Hendra Suryana Putra** dalam penelitiannya yang berjudul “Perbandingan Metode Ekstraksi Ciri Segmentasi *Fast Fourier transform (FFT)* dan *Mel Frequency Cepstrum Coefficients (MFCC)* pada logika *fuzzy* untuk identifikasi penutur” . Penelitian ini membahas bagaimana

perbandingan ekstraksi ciri FFT dengan MFCC menggunakan *recognize* logika *fuzzy*. Hasilnya membuktikan bahwa ekstraksi ciri menggunakan MFCC lebih akurat dibandingkn FFT. [2]

3. **Yohanes TDS dkk** dalam penelitiannya yang berjudul “Aplikasi Sistem *Neuro Fuzzy* untuk Pengenalan Kata” yang membahas tentang bagaimana aplikasi sistem *hybrid* antara *neural network* dan *fuzzy logic* dalam mengenali kata. Sistem tersebut diekstraksi ciri dengan FFT dan diujikan dengan membandingkan struktur *neuro-fuzzy* dengan *neural network* . Hasilnya membuktikan bahwa *neuro-fuzzy* memberikan hasil yang lebih baik dibandingkan dengan *neural network*. Dengan nilai persentase kebenaran tertinggi dari sistem *neuro fuzzy* 96,36 % sedangkan sistem *neural network* hanya 62,86 %. [3]
4. **Ferdi Sukresno dkk** dalam penelitiannya yang berjudul “Reduksi Suara Jantung dari Rekaman Suara Paru - paru Menggunakan *Filter* Adaptif dengan Algoritma *Recursive Least Square*”. Yang membahas tentang bagaimana mereduksi suara jantung yang dianggap sebagai *noise* dari suara paru-paru sebagai suara yang diinginkan. Berdasarkan hasil penelitiannya, *filter RLS* efektif dan mampu untuk mereduksi suara jantung. [8]

Berdasarkan pemikiran diatas, penulis tertarik untuk membuat suatu sistem yang tidak hanya mengenali kata dengan baik, tetapi juga *robust* yang dapat digunakan untuk mengurangi atau menghilangkan derau yang telah bercampur dengan sinyal informasi menggunakan *filter adaptif*. Penulis akan menggunakan permasalahan ini sebagai tugas akhir dengan judul “**Perancangan sistem pengenalan ucapan menggunakan MFCC – Neuro fuzzy dengan filter RLS sebagai metoda perbaikan sinyal**”

## **I.2 Tujuan Penelitian**

Penelitian pada tugas akhir ini bertujuan untuk merancang sistem pengenalan ucapan yang tahan terhadap perubahan kondisi lingkungan dan untuk meningkatkan kinerja sistem pengenalan ucapan pada lingkungan berderau dengan menggunakan *filter adaptif RLS*.

### **I.3 Manfaat Penelitian**

Tugas akhir ini diharapkan dapat :

1. Memberikan gambaran tentang konsep system untuk mengatasi pengaruh derau terhadap kejelasan informasi ucapan dengan menggunakan *filter RLS*.
2. Penelitian ini diharapkan dapat digunakan untuk meminimalisir derau sehingga didapatkan sinyal audio yang lebih jelas dan dapat dikenali.
3. Penelitian ini diharapkan dapat menunjang proses pembelajaran untuk mata kuliah Pengolahan Sinyal Digital (PSD) dan Pengolahan Suara pada jurusan teknik elektro.
4. Hasil tugas akhir ini nantinya dapat dikembangkan dan dimanfaatkan pada sistem *command & control* untuk pengontrolan alat - alat di industri ataupun pengontrolan robotik yang mampu bekerja pada perubahan kondisi lingkungan terutama pada lingkungan yang berderau.
5. Penelitian ini diharapkan selanjutnya dapat diimplementasikan (secara *hardware*) pada system pengenalan ucapan untuk mendapatkan akurasi yang lebih baik dan *robust*.
6. Memberikan gambaran tentang konsep *neuro - fuzzy* sehingga dapat dikembangkan dalam jenis aplikasi bidang lain seperti kedokteran, bioteknologi dan lain-lain.

### **I.4 Perumusan Masalah**

Berdasarkan latar belakang diatas, maka yang menjadi rumusan masalah untuk tugas akhir ini yaitu bagaimana pengaruh *filter RLS* dalam upaya meminimalisir pengaruh derau terhadap kualitas informasi yang diberikan dalam meningkatkan kinerja sistem pengenalan ucapan pada lingkungan berderau dan bagaimana membangun sistem pengenalan ucapan yang *robust* ketika digunakan pada lingkungan berderau agar akurasi pengenalan tidak menurun secara signifikan .

## 1.5 Batasan Masalah

Untuk mengarahkan pokok permasalahan dalam penelitian ini, penulis mengambil batasan pembahasan sebagai berikut :

1. *Input* kata yang dikenali hanya satu kata dalam satu perintah.
2. Data yang digunakan terdiri dari empat ucapan, yaitu: biru, hijau, kuning, dan merah.
3. Menggunakan derau lingkungan yang berasal dari AWGN, derau suara pengering rambut, dan derau suara mobil dengan SNR 0 dB, 5dB, 10dB, dan 15dB.
4. Ekstraksi parameter suara menggunakan MFCC
5. Pembentukan dan pengujian sistem pengenalan ucapan menggunakan *Neuro-fuzzy*.
6. Sistem perbaikan sinyal menggunakan *filter* RLS .
7. Pembuatan GUI dan simulasi menggunakan *software* Matlab R2013a

## 1.6 Sistematika Penulisan

Sistematika penulisan laporan Tugas Akhir ini adalah sebagai berikut :

- a. BAB I : Pendahuluan  
Bab ini berisi tentang tentang latar belakang, tujuan penelitian, manfaat penelitian, perumusan masalah, batasan masalah dan sistematika penulisan.
- b. BAB II : Tinjauan Pustaka  
Bab ini berisi teori dasar yang mendukung penelitian tugas akhir ini
- c. BAB III: Metodologi Penelitian  
Bab ini berisi metodologi yang digunakan dalam penelitian tugas akhir.
- d. Bab IV : Perancangan Sistem  
Bab ini berisi penjelasan mengenai sistem yang akan dirancang dan tahapan pada rancangan sistem tersebut.
- d. Bab V : Analisis Hasil dan Pembahasan  
Bab ini akan berisi analisis terhadap hasil kerja sistem dan output yang diperoleh dari pengujian sistem itu sendiri.
- e. Bab VI : Penutup  
Bab terakhir ini berisi simpulan dari hasil penelitian dan saran yang disampaikan berdasarkan hasil analisis dan pembahasan dari penelitian ini.