

BAB I

PENDAHULUAN

1.1 Latar Belakang

Speech recognition atau pengenalan pola suara juga dikenal sebagai *Automatic Speech Recognition (ASR)* merupakan sistem yang digunakan untuk mengenali perintah kata dari suara manusia dan kemudian diterjemahkan menjadi suatu data yang dimengerti oleh komputer. Sistem *speech recognition* dapat diartikan juga sebagai proses mengubah inputan suara ke media lain misalnya teks oleh karena itu *speech recognition* terkadang disebut sebagai *speech-to-text (STT)*. Sistem pengenalan pola suara dibangun oleh beberapa komponen utama di antaranya *microphone* untuk input suara, perangkat lunak pengenalan suara, komputer untuk mengambil data suara, dan *soundcard* untuk *input* atau *output*.

Speech recognition adalah salah satu bidang kecerdasan mesin yang sedang berkembang pesat, hal itu ditandai oleh hampir semua *device* teknologi dilengkapi oleh *voice command*. Hal ini telah menarik bagi para peneliti untuk menjadikan *speech recognition* sebagai disiplin ilmu yang penting untuk menciptakan dampak teknologi pada masyarakat dan diharapkan akan berkembang lebih jauh di bidang interaksi mesin dengan manusia.

Tujuan penelitian pengenalan suara adalah untuk menciptakan sistem cerdas, sebuah sistem pengenalan pola suara yang dapat mendengarkan apa yang pengguna ucapkan kemudian melakukan instruksi yang diperintahkan. Salah satu contoh sederhananya adalah penggunaan *voice command* untuk menggantikan fungsi *keyboard* pada komputer ataupun *smartphone*. Proses mengubah pola suara ke dalam teks terlihat sederhana, namun kenyataannya melalui proses transkripsi yang rumit.

Salah satu masalah utama dalam pengenalan dan pemahaman ucapan komputer adalah mengatasi banyak ruang pencarian. Keakuratan sebuah pengenalan pola suara didasari pada banyaknya jumlah kata yang terdapat dalam sebuah sistem, jika jumlah kosa kata meningkat maka ukuran ruang pencarian juga meningkat. Seiring bertambahnya ukuran ruang pencarian berakibat pada kinerja pengenalan pola suara menurun. Masalah lain dalam pengenalan suara adalah kompleksitas yang sangat besar yang terlibat dalam menganalisis input ucapan. Variasi dalam pengucapan, aksen, fisiologi pembicara, penekanan dan karakteristik lingkungan akustik biasanya menghasilkan ratusan klasifikasi fonem yang berbeda untuk setiap suara. Hal tersebut menyebabkan banyak klasifikasi fonem yang menghasilkan banyak probabilitas kata pada setiap titik yang dihasilkan. Semua pilihan kata ini kemudian dapat digabungkan untuk menghasilkan ratusan kemungkinan kalimat untuk setiap ucapan. Ruang pencarian yang dihasilkan menjadi sangat besar. Oleh karena itu, dibutuhkan sistem pengenalan pola suara yang dapat memproses sinyal suara secara cepat untuk menyaring semua probabilitas yang salah dan dapat mengenali secara input suara secara *real time* dengan hasil yang akurat.

Salah satu metode yang dapat digunakan untuk ekstraksi suara adalah *Mel Frequency Cepstral Coefficients* (MFCC) karena metode ini dapat dengan baik dalam mempresentasikan sinyal. Sedangkan untuk pengenalan pola suara, metode yang dapat digunakan *Vector Quantitation* karena metode ini mampu mengklasifikasi sinyal suara dari hasil ekstraksi suara. *Vector Quantitation* adalah sebuah proses dalam memetakan vektor dari ruang vektor yang besar menjadi bentuk terbatas dalam ruang tersebut [1]. *Vector Quantitation* digunakan dibanyak aplikasi seperti pengenalan gambar dan pengenalan pola suara.

Dalam pengembangan teknologi *speech recognition* terdahulu, fokus tujuannya hanya pada cara mengubah suara kedalam teks. Dewasa ini diperlukan fokus penelitian untuk menjadikan teknologi *speech recognition* menjadi sistem pengenalan pola suara yang terpadu, yakni hasil dari output pengenalan pola suara

dapat digunakan untuk menyelesaikan sebuah tugas atau sebuah kontrol. Salah satu implementasinya adalah pada *voice command* pada robot kontrol.

Dari latar belakang yang telah diuraikan mendasari dilakukannya penelitian untuk mengimplementasikan sebuah konsep pengenalan suara menggunakan *Mel Frequency Cepstral Coefficients* (MFCC) dan *Vector Quantitation* dengan output dari pengenalan suara tersebut dijadikan perintah untuk navigasi robot kontrol.

1.2. Rumusan Masalah

Dari latar belakang yang telah diuraikan sebelumnya didapat rumusan masalah penelitian ini adalah bagaimana cara mengimplentasikan *speech recognition* pada robot kontrol dengan menggunakan *Mel Frequency Cepstral Coefficients* (MFCC) dan *Vector Quantitation*.

1.3. Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

- a. Sistem navigasi pada robot kontrol berupa perintah untuk menggerakkan atau mengontrol motor DC berupa kata “MAJU”, “MUNDUR”, “KANAN”, “KIRI”, dan “STOP”
- b. Data latih suara diperoleh dari proses perekaman suara menggunakan Matlab 2016a dari hasil Laptop Dell Core i5
- c. Metode yang digunakan untuk pengenalan pola suara adalah MFCC dan *Vector Quantitation*
- d. *Tools* yang digunakan untuk menganalisis pola pengenalan suara adalah Matlab 2016a
- e. *Mikrokontroler* yang digunakan pada penelitian ini adalah arduino nano
- f. Pengujian pola suara tidak secara *real time*.

1.4. Tujuan penelitian

Tujuan dari penelitian ini adalah untuk menerapkan konsep pengenalan suara yang diimplementasikan dalam sebuah sistem kontrol. Secara lebih rinci tujuan penelitian ini adalah untuk:

- a. Untuk menerapkan *speech recognition* dengan metode MFCC dan *Vector Quantitation* dengan difasilitasi oleh robot kontrol.
- b. Untuk mengetahui tingkat akurasi pengenalan pola suara dengan metode MFCC dan *Vector Quantitation*.

1.5. Manfaat Penelitian

Manfaat pada penelitian ini adalah dapat mengimplementasikan teknologi *speech recognition* untuk *voice command* yang pada penelitian ini difasilitasi robot kontrol.

1.6. Kontribusi Tesis

Penelitian ini berkontribusi dalam membuat sebuah pola pengenalan suara untuk mengontrol sebuah robot dengan menggunakan metode *Mel Frequency Cepstral Coefficients* (MFCC) untuk ekstraksi suaranya dan *vector quantitation* untuk pengenalan pola dari hasil ekstraksi suara.

1.7. Sistematika Penulisan

Sistematika penulisan pada penelitian ini adalah sebagai berikut:

Bab I Pendahuluan

Merupakan pendahuluan yang berisi latar belakang, perumusan masalah, tujuan penelitian, batasan-batasan masalah, metodologi penelitian dan sistematika penulisan dari penelitian ini.

Bab II Tinjauan Pustaka

Membahas teori-teori yang mendukung dalam perancangan dan pembuatan alat dalam penyusunan penelitian ini.

Bab III Metodologi Penelitian

Membahas tentang metode pengenalan pola suara yang akan digunakan pada penelitian ini yakni MFCC dan *Vector Quantitation*. Pada bab ini juga dijelaskan bagaimana cara merancang robot kontrol serta cara menganalisis data yang akan diujikan pada Matlab

Bab IV Hasil dan Pembahasan

Berisi tentang pembahasan mengenai metode dan hasil pengujian yang dilakukan pada alat/sistem yang telah dibuat.

Bab V Kesimpulan dan Saran

Berisi kesimpulan dari pembuatan alat yang telah dibuat serta saran-saran untuk perbaikan dan pengembangan alat/sistem pada penelitian selanjutnya.