

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era teknologi informasi yang semakin maju saat ini, proses pembelajaran menjadi lebih fleksibel dan praktis. Para siswa tidak perlu hadir ke kelas untuk mendapatkan materi yang mereka butuhkan, dan hanya cukup mencari materi tersebut di internet. Melalui internet, mereka dapat menemukan media pembelajaran yang variatif seperti teks, gambar, suara, bahkan video.

Dari berbagai macam media pembelajaran yang ditawarkan, video menjadi media yang populer dan efektif untuk mendukung proses pembelajaran. Hal ini dikarenakan siswa mampu melihat sekaligus mendengar materi yang disajikan. Untuk mempermudah proses pemahaman, maka setiap video perlu diberi *caption* atau teks keterangan dari setiap kata yang diucapkan sang pengajar. Oleh karena itu pada penelitian kali ini akan dibangun sistem pembuat *caption* otomatis (*automatic video captioning*) menggunakan teknologi *speech recognition*.

Sinyal suara dari video masukan diekstrak dan diproses dengan *speech recognition* untuk menghasilkan teks yang sesuai. Sistem *speech recognition* sendiri terdiri dari dua modul utama, yaitu *feature extraction* untuk mengkonversi sinyal suara menjadi beberapa jenis representasi parametrik dan *feature matching* untuk membandingkan atau mencocokkan data dari sinyal masukan yang telah melalui proses *feature extraction* dengan data latih yang ada dalam database. Pada penelitian ini digunakan *Linear Predictive Coding* untuk *feature extraction* dan *Hidden Markov Model* untuk *feature matching*. Teks yang dihasilkan dari *speech recognition* tersebut kemudian disimpan dalam file dengan format (*.srt) dan dijalankan pada aplikasi pemutar video.

Parameter yang akan diukur dalam penelitian ini adalah tingkat akurasi (*Syllable Error Rate*) dalam mengkonversi suara menjadi teks.

1.2 Rumusan Masalah

Permasalahan yang diangkat dalam pembuatan tugas akhir ini antara lain:

1. Bagaimana merancang sistem *Speech Recognition* yang handal untuk mengkonversi sinyal suara menjadi teks?
2. Bagaimana mendapatkan vektor ciri dari sinyal suara menggunakan metode *Linear Predictive Coding*?
3. Bagaimana membandingkan data uji hasil ekstraksi ciri dengan data latih menggunakan *Hidden Markov Model*?
4. Bagaimana menulis kata hasil konversi kedalam file (*.srt) dengan *timestamp* yang tepat?
5. Bagaimanakah performansi sistem *automatic video captioning* yang telah dibuat?

1.3 Tujuan

Tujuan dari pembuatan tugas akhir ini adalah untuk menghasilkan sistem *Automatic Video Captioning* yang mampu membuat *caption* dari video secara otomatis menggunakan *Speech Recognition* dan menganalisis faktor – faktor yang dapat mempengaruhi performansi sistem.

1.4 Batasan Masalah

Adapun batasan ruang lingkup permasalahan dari pengerjaan tugas akhir ini adalah sebagai berikut:

1. Video masukan dibuat dalam ruangan yang hening untuk meminimalisir *noise*.
2. *Caption* yang dihasilkan disimpan dalam file dengan format (*.srt).
3. Sistem dibangun dengan menggunakan bahasa pemrograman C/C++.
4. Intonasi suara yang direkam sesuai dengan nada bicara standar.
5. Kata yang dapat dikonversi adalah kata-kata dalam bahasa Indonesia.
6. Sinyal suara dikenali per-suku kata.
7. Sinyal suara untuk data latih dan data uji adalah suara dengan 1 kanal (*mono*).
8. Sistem bersifat *speaker dependent* (hanya mengenali satu pembicara).
9. Sinyal audio untuk data latih dan data uji berformat 16000 Hz 16 bit mono.

1.5 Metodologi Penelitian

Pembuatan tugas akhir ini dilakukan dengan menggunakan metodologi sebagai berikut:

1. Studi literatur

Pada tahap ini dilakukan pencarian pustaka yang dapat menunjang pengerjaan Tugas Akhir. Pustaka yang dikumpulkan berupa *paper*, *e- book*, *textbook*, dan tutorial-tutorial mengenai Speech Recognition, *Linear Predictive Coding (LPC)*, serta teori-teori tentang *Hidden Markov Model (HMM)*.

2. Perancangan perangkat lunak

Pada tahap ini diawali dengan melakukan analisis awal terhadap permasalahan utama yang muncul pada topik tugas akhir ini kemudian dilakukan perancangan aplikasi.

3. Implementasi dan pembuatan sistem

Pada tahap ini dilakukan implementasi metode yang diusulkan dari rancangan yang telah dibuat pada tahap sebelumnya dengan menggunakan bahasa pemrograman C++.

4. Uji coba dan evaluasi

Pada tahap ini dilakukan uji coba dengan menggunakan bermacam masukan yang bervariasi untuk mencoba jalannya aplikasi telah sesuai dengan rancangan dan desain implementasi yang dibuat dan juga untuk mencari kesalahan program yang mungkin terjadi untuk selanjutnya dapat dilakukan perbaikan.

5. Penyusunan laporan tugas akhir

Pada tahap ini dilakukan penyusunan laporan yang berisi dasar teori, dokumentasi dari perangkat lunak, dan hasil-hasil yang diperoleh selama pengerjaan tugas akhir.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Berisi latar belakang, rumusan masalah, tujuan, batasan masalah, metodologi penelitian, sistematika penulisan, dan rencana kerja.

BAB II DASAR TEORI

Berisi teori-teori dan materi yang digunakan dalam pelaksanaan tugas akhir.

BAB III PERANCANGAN SISTEM

Berisikan bahasan perancangan sistem *Automatic Video Captioning*.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Berisi dokumentasi dan implementasi penelitian yang dilakukan sesuai dengan rancangan yang sudah dibuat serta hasil uji coba dari aplikasi yang dibuat termasuk analisa yang telah dicapai dalam Tugas Akhir ini.

BAB V KESIMPULAN DAN SARAN

Berisi kesimpulan yang dapat diambil dari tugas akhir ini beserta saran untuk pengembangan selanjutnya