

**SEGMENTASI *CONTINUOUS SPEECH*
DENGAN MENGGUNAKAN *DYNAMIC THRESHOLDING* DAN
*METODE BLOCKING BLOCK AREA***



SKRIPSI

**Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Komputer
Pada Departemen Ilmu Komputer/Informatika**

**Disusun Oleh:
BAYU ARASYI
24010311130042**

**DEPARTEMEN ILMU KOMPUTER/INFORMATIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO**

2016

HALAMAN PERNYATAAN KEASLLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini,

Nama : Bayu Arasyi

NIM : 24010311130042

Judul : Segmentasi *Continuous Speech* dengan Menggunakan *Dynamic Thresholding* dan Metode *Blocking Block Area*

Dengan ini saya menyatakan bahwa dalam tugas akhir atau skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.



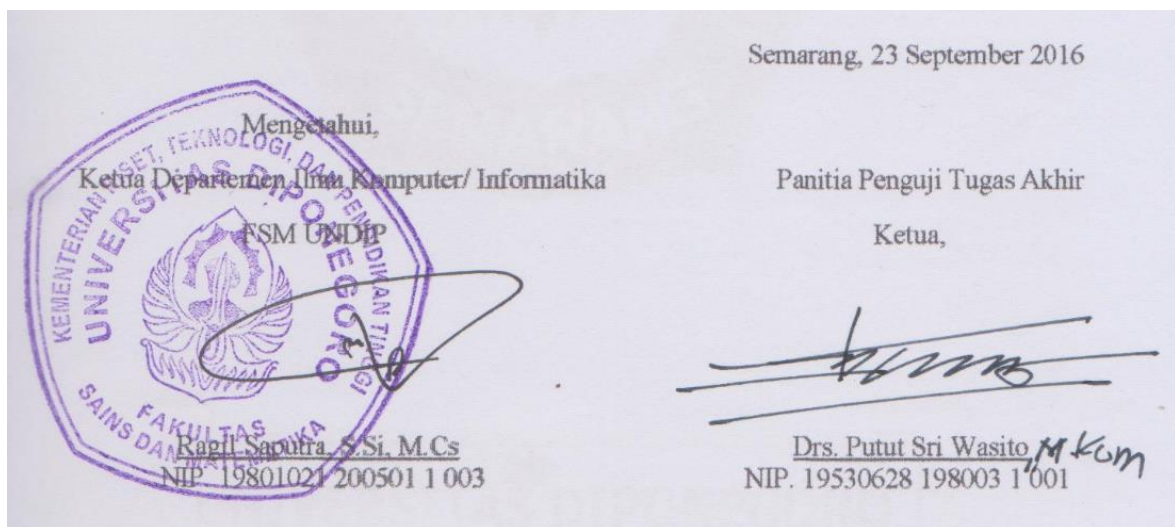
HALAMAN PENGESAHAN

Judul : Segmentasi *Continuous Speech* dengan Menggunakan Dynamic Thresholding dan Metode *Blocking Block Area*

Nama : Bayu Arasyi

NIM : 24010311130042

Telah diujikan pada sidang tugas akhir pada tanggal 23 September 2016 dan dinyatakan lulus pada tanggal 23 September 2016



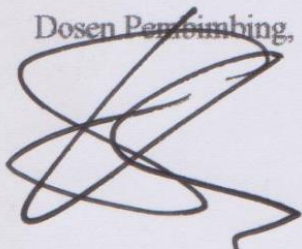
HALAMAN PENGESAHAN

Judul : Segmentasi *Continuous Speech* dengan Menggunakan Dynamic Thresholding dan Metode *Blocking Block Area*

Nama : Bayu Arasyi

NIM : 24010311130042

Telah diujikan pada sidang tugas akhir pada tanggal 23 September 2016

Semarang, 23 September 2016
Dosen Pembimbing,

Sukmawati Nur Endah, S.Si, M.Kom
NIP. 19780502 200501 2 002

ABSTRAK

Pengenalan ucapan adalah suatu teknik yang memungkinkan komputer untuk dapat mengerti kata yang diucapkan oleh seseorang. Aplikasi pengenalan ucapan dibagi menjadi dua jenis berdasarkan input yang digunakan yaitu *isolated word* dan *continuous speech*. Pada pengenalan ucapan *continuous speech* dibutuhkan tahap segmentasi yang berfungsi untuk memecah kalimat yang diucapkan menjadi kata-kata yang dapat dikenali oleh komputer. Kualitas hasil segmentasi, dapat mempengaruhi hasil pengenalan yang dilakukan. Penelitian ini, meneliti mengenai *threshold* dinamis yang digunakan pada proses segmentasi *continuous speech* dan juga perbaikan metode *Blocking Block Area* pada domain Bahasa Indonesia. Pada penelitian ini, dilakukan perbandingan tiga algoritma (*K-Means*, *Fuzzy C-Means*, dan Otsu) untuk mencari *threshold* dinamis terbaik dan dilakukan penambahan proses morfologi serta kolom *overlapping* pada metode *blocking block area* sehingga diperoleh akurasi segmentasi terbaik. Berdasarkan hasil penelitian, algoritma *Fuzzy C-Means* memberikan hasil *threshold* terbaik dibandingkan dengan dua algoritma lainnya. Secara keseluruhan dengan menggunakan algoritma *Fuzzy C-Means*, dan penambahan morfologi serta kolom *overlapping*, penelitian ini dapat meningkatkan akurasi segmentasi *continuous speech* Bahasa Indonesia dari 24% menjadi 90%.

Kata Kunci: *Continuous Speech*, Pengenalan Ucapan, *Mel-Frequency Cepstral Coefficient*, *Hidden Markov Model*, Segmentasi Ucapan

ABSTRACT

Speech Recognition is a technique that allows computer to understand words that had been spoken by a person. Based on the type of speech, speech recognition application was divided into two, they are isolated word and continuous speech. Segmentation are needed in continuous speech recognition, its function is to segment continuous speech into words that can be recognized by computer. The quality of segmentation results can influence the recognition results. This research, study about dynamic thresholding that used on segmentation process and also fixing *Blocking Block Area* method in Indonesia Domain speech recognition. In this research, researchers compare three algorithms (*K-Means*, *Fuzzy C-Means*, dan Otsu) to find best dynamic threshold and also adding morphological process and implementation of column overlapping concept on *Blocking Block Area* method to achieve the best segmentation accuracy. Based the result of this research, *Fuzzy C-Means* algorithm, gives best threshold result than the other two algorithms. Overall by using *Fuzzy C-Means* algorithm, adding mophological process, and column overlapping implementation, this research increases segmentation accuracy of Bahasa Indonesia continuous speech from 24% to 90%.

Key Word: Continuous Speech, Speech recognition, Mel-Frequency Cepstral Coefficient, Hidden Markov Model, Speech Segmentation.

KATA PENGANTAR

Puji syukur penulis haturkan kehadirat Allah SWT atas rahmat, anugerah, dan kesempatan yang diberikan-Nya penulis dapat menyelesaikan laporan tugas akhir.

Tugas akhir yang berjudul “**Segmentasi *Continuous Speech* dengan Menggunakan *Dynamic Thresholding* dan Metode *Blocking Block Area***” disusun sebagai salah satu syarat untuk memperoleh gelar sarjana strata satu pada Jurusan Ilmu Komputer/ Informatika Fakultas Sains dan Matematika Universitas Diponegoro Semarang.

Dalam penyusunan tugas akhir ini penulis banyak mendapat bimbingan, bantuan, dan dukungan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis menyampaikan terima kasih kepada:

1. Ragil Saputra, S.Si, M.Cs, selaku Ketua Departemen Ilmu Komputer / Informatika FSM UNDIP
2. Helmie Arif Wibawa, S.Si, M.Cs, selaku Koordinator Tugas Akhir Departemen Ilmu Komputer / Informatika
3. Sukmawati Nur Endah, S.Si, M.Kom, selaku dosen pembimbing
4. Semua pihak yang telah membantu hingga selesainya tugas akhir ini, yang tidak dapat penulis sebutkan satu persatu. Semoga Allah membalas segala kebaikan yang telah diberikan kepada penulis

Penulis menyadari bahwa dalam laporan ini masih banyak terdapat kekurangan baik dari penyampaian materi maupun isi dari materi itu sendiri. Hal ini dikarenakan keterbatasan kemampuan dan pengetahuan dari penulis. Oleh karena itu, kritik dan saran yang bersifat membangun sangat penulis harapkan.

Semoga laporan tugas akhir ini dapat bermanfaat bagi penulis dan juga pembaca pada umumnya.

Semarang, September 2016

Penulis

DAFTAR ISI

SEGMENTASI <i>CONTINUOUS SPEECH</i>	i
HALAMAN PERNYATAAN KEASLLIAN SKRIPSI	ii
HALAMAN PENGESAHAN	iii
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI	viii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan dan Manfaat	3
1.4 Ruang Lingkup	3
1.5 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Penelitian Terdahulu Tentang Segmentasi <i>Continuous Speech</i>	6
2.2 Tinjauan Pustaka.....	6
2.2.1. Bahasa Indonesia	6
2.2.2. Sinyal Ucapan.....	7
2.2.3. Pengenalan Ucapan.....	8
2.2.4. <i>Pre-Processing</i>	9
2.2.5. <i>Mel-Frequency Cepstral Coefficient (MFCC)</i>	17
2.2.6. <i>Hidden Markov Model (HMM)</i>	23

2.2.7.	Proses Pengembangan <i>Perangkat Lunak</i>	25
2.2.8.	Unified Modeling Language (UML)	26
BAB III METODOLOGI PENELITIAN		29
3.1.	Gambaran Umum Peneitian	29
3.1.1.	Deskripsi Umum Penelitian	29
3.2.	Tahapan Penelitian	30
3.2.1.	Pengumpulan Data	30
3.2.2.	Pelatihan	32
3.2.3.	Pengenalan	55
3.3.	Contoh Proses	56
BAB IV PENGEMBANGAN PERANGAT LUNAK		90
4.1.	Fase Inception	90
4.1.1.	Kebutuhan Fungsional	90
4.1.2.	Kebutuhan Non-Fungsional	90
4.1.3.	Model <i>Use Case</i>	91
4.2.	Fase Elaboration	95
4.2.1.	Design Model	96
4.2.2.	Data Model	99
4.2.3.	Menyusun Prototype Antarmuka	100
4.2.4.	Menyusun Rencana Pengujian	104
4.3.	Fase Construction	104
4.3.1.	Spesifikasi Perangkat	104
4.3.2.	Implementasi Kelas	105
4.3.3.	Implementasi Antarmuka	105
BAB V HASIL DAN ANALISA		108
5.1.	Skenario Pengujian	108
5.1.1.	Pengujian Segmentasi	108

5.1.2. Pengujian Pengenalan Suara.....	108
5.2. Data Penelitian	109
5.3. Hasil.....	110
5.3.1. Hasil Pengujian Segmentasi	110
5.3.2. Hasil Pengujian Pengenalan Suara	111
5.4. Analisa Hasil.....	112
5.4.1. Analisa Hasil Segmentasi	112
5.4.2. Analisa Hasil Pengenalan Suara	113
BAB VI PENUTUP.....	114
6.1. Kesimpulan	114
6.2. Saran	114
DAFTAR PUSTAKA.....	115
LAMPIRAN-LAMPIRAN	117
Lampiran 1. Deskripsi Tabel Hasil Uji.....	118

DAFTAR GAMBAR

Gambar 2.1	Contoh Sinyal Ucapan “it’s time”.....	7
Gambar 2.2	Sinyal Suara Bahasa Indonesia “bapak ibu pergi bersama adik”	10
Gambar 2.3	Spectrogram Sinyal Suara “bapak ibu pergi bersama adik”	11
Gambar 2.4	(a) Struktur Elemen Citra Sebelum proses Erosi, (b) Citra Setelah Proses Erosi	14
Gambar 2.5	(a) Struktur Elemen Citra Sebelum proses Dilatasi, (b) Citra Setelah Proses Dilatasi	15
Gambar 2.6	Hasil Spectrogram yang Telah Diubah Menjadi Citra Biner dengan Menggunakan Hasil <i>Dynamic Thresholding</i>	16
Gambar 2.7	Hasil Morfologi Terhadap Citra Biner.....	16
Gambar 2.8	Hasil Blocking Block Area pada Citra Hasil Morfologi.....	17
Gambar 2.9	Block Diagram Untuk MFCC.....	18
Gambar 2.10	Frame Blocking.....	19
Gambar 2.11	Contoh Markov Chain.....	24
Gambar 2.12	Arsitektur Umum HMM	24
Gambar 2.13	Contoh Class Diagram Pemesanan Barang.....	27
Gambar 2.14	Simbol Use Case	27
Gambar 2.15	Simbol Aktor.....	27
Gambar 2.16	Contoh Sequence Diagram untuk Proses Pemesanan Barang	28
Gambar 3.1	Arsitektur Sistem Secara Umum.....	30
Gambar 3.2	Flowchart Pengambilan Data Pengenalan	31
Gambar 3.3	Flowchart Pengambilan Data Pelatihan	31
Gambar 3.4	Flowchart Proses Pelatihan	32
Gambar 3.5	Flowchart Proses Segmentasi	33
Gambar 3.6	Sinyal Suara “dia punya dua mobil hitam”.....	33
Gambar 3.7	Spectrogram Sinyal Suara “dia punya dua mobil hitam”	34
Gambar 3.8	<i>Grayscale</i> Citra Spectrogram Sinyal Suara “dia punya dua mobil hitam” ..	34
Gambar 3.9	Flowchart Algoritma K-Means Clustering – Menghitung Threshold.....	35
Gambar 3.10	Flowchart Algoritma FCM Clustering – Menghitung Threshold	36
Gambar 3.11	Flowchart Algoritma Otsu – Menghitung Threshold.....	37

Gambar 3.12	Flowchart Proses Erosi	38
Gambar 3.13	Flowchart proses Dilatasi.....	39
Gambar 3.14	Flowchart <i>Blocking Block Area</i>	41
Gambar 3.15	Contoh Batas Awal dan Akhir Citra <i>Block</i>	40
Gambar 3.16	Contoh Pemotongan Sinyal Suara	40
Gambar 3.17	Flowchart <i>DC-Removal/</i> Normalisasi.....	42
Gambar 3.18	Flowchart MFCC	43
Gambar 3.19	Flowchart Proses <i>Pre-Emphasize</i>	44
Gambar 3.20	Flowchart <i>Frame Blocking</i>	45
Gambar 3.21	Flowchart Proses <i>Windowing</i>	46
Gambar 3.22	Flowchar proses FFT	47
Gambar 3.23	Proses Mel-Frequency <i>Wrapping</i>	47
Gambar 3.24	Flowchart Proses DCT.....	48
Gambar 3.25	Flowchart Proses Cepstral Liftering	49
Gambar 3.26	Flowchart Pelatihan HMM	50
Gambar 3.27	Flowchart Proses Pelatihan Parameter HMM.....	51
Gambar 3.28	Flowchart Inisialisasi Parameter B	53
Gambar 3.29	Flowchart Proses Forward	52
Gambar 3.30	Flowchart Proses Backward.....	53
Gambar 3.31	Flowchart Proses <i>Normalize</i>	54
Gambar 3.32	Flowchart Proses <i>Symetrize</i>	54
Gambar 3.33	Flowchart Proses Stochastic	55
Gambar 3.34	Flowchart Proses Pengenalan Ucapan	56
Gambar 3.35	Citra Spectrogram Sinyal Suara “kapan kita main bola pantai”	57
Gambar 3.36	Citra <i>Grayscale Spectrogram</i> Sinyal Suara “kapan kita main bola pantai”	58
Gambar 3.37	Contoh Citra untuk Metode Otsu	64
Gambar 3.38	Histogram Cita 3.36	65
Gambar 3.39	Contoh Proses Erosi pada Citra Biner dengan Struktur Elemen 3x3.....	66
Gambar 3.40	Contoh Proses Dilatasi pada Citra Biner dengan Struktur Elemen 3x3.....	67
Gambar 3.41	Contoh Citra Biner Hasil Morfologi Berukuran 20x16	68
Gambar 3.42	Citra Biner Hasil Morfologi Dibagi Menjadi Beberapa Frame	68
Gambar 3.43	Contoh Penandaan Warna pada Frame	69
Gambar 3.44	Hasil Akhir Metode <i>Blocking Block Area</i>	69

Gambar 4.1	Use Case Diagram Aplikasi Continuous Speech Recognition	92
Gambar 4.2	Class Diagram Aplikasi Continuous Speech Recognition Bahasa Indonesia	96
Gambar 4.3	Sequence Diagram Melakukan Pelatihan	97
Gambar 4.4	Sequence Diagram Menambah Data Latih	97
Gambar 4.5	Sequence Diagram Melakukan Pengujian	98
Gambar 4.6	Sequence Diagram Menambah Data Uji.....	98
Gambar 4.7	Sequence Diagram Melakukan Segmentasi.....	99
Gambar 4.8	Desain Antarmuka Halaman Beranda.....	100
Gambar 4.9	Desain Antarmuka Halaman Pelatihan	101
Gambar 4.10	Desain Antarmuka Halaman Pengujian	101
Gambar 4.11	Desain Antarmuka Halaman Segmentasi.....	102
Gambar 4.12	Desain Antarmuka Detail Segmentasi I.....	102
Gambar 4.13	Desain Antarmuka Detail Segmentasi II.....	103
Gambar 4.14	Desain Antarmuka Detail Segmentasi III	103
Gambar 4.15	Tampilan Halaman Beranda	106
Gambar 4.16	Tampilan Halaman Pelatihan.....	106
Gambar 4.17	Tampilan Halaman Pengujian.....	107
Gambar 4.18	Tampilan Halaman Segmentasi	107
Gambar 5.1	(a) Sinyal Suara Asli, (b) Hasil Blocking Block Area	112
Gambar 5.2	(a) Sinyal Suara “dia punya dua mobil hitam” tanpa noise. (b) Sinyal Suara “dia punya dua mobil hitam” dengan noise	113

DAFTAR TABEL

Tabel 2.1	Perkembangan Penelitian Tentang Segmentasi Suara.....	6
Tabel 2.2	Tabel Parameter pengenalan Suara.....	8
Tabel 2.3	Tabel Jenis <i>Relationship</i> pada <i>Use Case</i>	27
Tabel 3.1	Contoh Data K-Means	58
Tabel 3.2	Contoh Cluster K-Means	59
Tabel 3.3	Jarak Data Terhadap Pusat Cluster K-Means	59
Tabel 3.4	Pengelompokan Data K-Means	60
Tabel 3.5	Pusat Cluster Baru K-Means	60
Tabel 3.6	Hasil Akhir Cluster	61
Tabel 3.7	Contoh Data FCM	61
Tabel 3.8	Matrik Uik Stokastik.....	62
Tabel 3.9	Cluster Center FCM.....	62
Tabel 3.10	<i>Euclidean Distance</i> FCM	63
Tabel 3.11	Keanggotaan <i>Fuzzy</i> Baru	63
Tabel 3.12	Pengelompokan Data FCM.....	63
Tabel 3.13	Hasil Akhir Cluster	64
Tabel 3.14	Contoh Data Otsu	64
Tabel 3.15	Persebaran Skala Keabuan.....	65
Tabel 3.16	Hasil Metode Otsu	66
Tabel 3.17	Hasil Ekstraksi Ciri Data Latih Sinyal Ucapan “yang”	76
Tabel 3.18	Hasil Ekstraksi Ciri Data Latih Sinyal Ucapan “ibu”.....	77
Tabel 3.19	Nilai Parameter Σ untuk State 1	77
Tabel 3.20	Nilai Parameter B	78
Tabel 3.21	Nilai α	79
Tabel 3.22	Nilai β	80
Tabel 3.23	Nilai γ	80
Tabel 3.24	Nilai $\gamma_{\text{observasi}}$ pada State 1	81
Tabel 3.25	Nilai $\gamma_{\text{observasi}}$ pada State 2.....	81
Tabel 3.26	Nilai $\gamma_{\text{observasi}}$ pada State 3	82
Tabel 3.27	Nilai μ baru.....	82

Tabel 3.28	Nilai Parameter Σ baru pada State 1	84
Tabel 3.29	Nilai Parameter Σ baru pada State 2	84
Tabel 3.30	Nilai Parameter Σ baru pada State 3	84
Tabel 3.31	Nilai Parameter A setelah Pelatihan	86
Tabel 3.32	Nilai Parameter Π setelah Pelatihan	86
Tabel 3.33	Nilai Parameter μ setelah Pelatihan	86
Tabel 3.34	Nilai Parameter Sigma pada State 1 setelah Pelatihan	86
Tabel 3.35	Nilai Parameter Sigma pada State 2 setelah Pelatihan	87
Tabel 3.36	Nilai Parameter Sigma pada State 3 setelah Pelatihan	87
Tabel 3.37	Contoh sinyal ucapan yang akan diuji	87
Tabel 3.38	Nilai Parameter B berdasarkan Parameter μ dan Σ	87
Tabel 3.39	Nilai α sinyal ucapan terhadap Model	88
Tabel 3.40	Nilai likelihood sinyal ucapan terhadap Model	89
Tabel 4.1	Kebutuhan Fungsional Aplikasi	90
Tabel 4.2	Kebutuhan Non-Fungsional Aplikasi	91
Tabel 4.3	Tabel Karakteristik Pengguna.....	91
Tabel 4.4	Daftar Use Case	92
Tabel 4.5	Use Case Detail untuk Melakukan Pelatihan.....	93
Tabel 4.6	Use Case Detail untuk Menambah Data Latih	93
Tabel 4.7	Use Case Detail untuk Melakukan Pengujian	94
Tabel 4.8	Use Case Detail untuk Menambah Data Uji.....	94
Tabel 4.9	Use Case Detail untuk Melakukan Segmentasi	95
Tabel 4.10	Struktur Folder data_pelatihan	99
Tabel 4.11	Struktur Folder database_pelatihan.	99
Tabel 4.12	Rencana Pengujian Aplikasi Continuous Speech Recognition Bahasa Indonesia.....	104
Tabel 4.13	Implementasi Class	105
Tabel 5.1	Daftar Kalimat yang Diucapkan	109
Tabel 5.2	Hasil Segmentasi Metode Rahman & Bhuiyan	110
Tabel 5.3	Hasil Segmentasi Proposed Method	110
Tabel 5.4	Prosentase Hasil Segmentasi	111
Tabel 5.5	Hasil Pengujian Pengenalan Suara	111

BAB I

PENDAHULUAN

Bab ini membahas latar belakang, rumusan masalah, tujuan dan manfaat, dan ruang lingkup penelitian tugas akhir mengenai Segmentasi *Continuous Speech* dengan Menggunakan *Dynamic Thresholding* Metode *Blocking Block Area*.

1.1 Latar Belakang

Teknologi yang berkembang pesat menyebabkan kemudahan untuk berkomunikasi dengan orang lain. Komunikasi yang paling efektif untuk menyampaikan maksud seseorang adalah secara lisan. Penerapan teknologi untuk menjalin hubungan manusia dan komputer sudah banyak berkembang, sehingga muncul aplikasi – aplikasi yang menerapkan suara sebagai sarananya, diantaranya *voice command*, identifikasi suara untuk keamanan, *voice input*, dan lain-lain. Suara dianggap sebagai cara paling efisien untuk menyampaikan maksud/ perintah.

Pengenalan ucapan atau lebih dikenal dengan *speech recognition*, merupakan salah satu contoh perkembangan penerapan teknologi pada media suara. Banyak faktor yang mempengaruhi keberhasilan suatu aplikasi pengenalan ucapan, diantaranya adalah kesehatan, umur dan jenis kelamin penguji serta algoritma yang digunakan pada aplikasi tersebut. Perkembangan *speech recognition (speech to text)* sudah berjalan cukup pesat, sudah banyak penelitian bermunculan yang membahas *speech recognition* (Fawziah, 2013), (Andriana, 2013), (Park, 2009). Akan tetapi, ketiga penelitian tersebut hanya membahas mengenai *speech recognition* yang diterapkan pada *isolated word*.

Pengenalan ucapan jenis *isolated word* adalah pengenalan ucapan yang dapat mengenali kata secara individu bukan merupakan bagian dari satu kesatuan kalimat, sedangkan pengenalan ucapan jenis *continuous speech* merupakan pengenalan ucapan yang dapat mengenali kata sebagai bagian dari suatu kalimat (Fawziah, 2013).

Pada aplikasi *continuous speech recognition* terdapat beberapa tahapan yaitu *pre-processing*, ekstraksi ciri, dan pengenalan. *Pre-Processing* merupakan suatu tahapan dimana suara akan diproses sedemikian rupa sehingga siap untuk diekstraksi ciri. *Pre-Processing* sendiri memiliki beberapa tahapan dan diantaranya adalah

segmentasi. Proses Segmentasi berfungsi sebagai pemecah kalimat menjadi kata. Sudah ada penelitian yang dilakukan untuk melakukan segmentasi ucapan, diantaranya adalah (Rahman & Bhuiyan, 2013) yang melakukan segmentasi suara pada bahasa Hindi, menggunakan *dynamic thresholding* dan metode *blocking block area*. Pada penelitian ini, peneliti menerapkan metode yang digunakan oleh Rahman & Bhuiyan pada domain bahasa Indonesia. Akan tetapi, metode yang digunakan kurang sesuai jika diterapkan pada domain bahasa Indonesia, hal ini dikarenakan perbedaan fonem pada kata bahasa Hindi dan bahasa Indonesia. *Dynamic Thresholding* menjadi pendekatan secara umum yang sering dilakukan untuk menentukan zona identifikasi kata (Rahman & Bhuiyan, 2013). Oleh karena itu pada penelitian kali ini, peneliti membandingkan hasil dari ketiga algoritma yaitu K-Means, Fuzzy C-Means, dan Otsu untuk dijadikan sebagai nilai *threshold*. Hasil dari proses *dynamic thresholding* adalah citra biner *spectrogram* yang akan diproses lebih lanjut menggunakan metode *blocking block area*. Tujuan penggunaan metode *Blocking Block Area* adalah untuk membuat *block* kata dari citra biner *spectrogram*. Pada kenyataannya hasil segmentasi yang dilakukan kurang akurat, oleh karena itu penelitian ini mencoba memperbaiki kelemahan yang ada yaitu dengan menerapkan proses morfologi pada citra biner *spectrogram*, serta dengan menerapkan konsep kolom *overlapping* untuk menentukan jenis *block*. Pengujian terkait penelitian ini dilakukan menggunakan dua tahap, yaitu pengujian akurasi segmentasi kata dan pengujian hasil segmentasi terhadap aplikasi pengenalan suara Bahasa Indonesia.

Proses ekstraksi ciri dilakukan setelah melakukan proses segmentasi, metode yang sering digunakan dalam tahapan ini adalah metode *Mel Frequency Cepstral Coefficients* (MFCC). Metode MFCC sering digunakan dalam proses ekstraksi ciri karena proses – prosesnya yang menyerupai pendengaran manusia (Mustofa, 2007). Hasil ekstraksi ciri selanjutnya akan dikenali dengan menggunakan metode HMM. HMM memiliki 5 proses dasar dalam melakukan pengenalan suara proses-proses inilah yang menyebabkan HMM mempunyai tingkat akurasi yang lebih tinggi dibanding metode lain (Rabiner, 1991).

Berdasarkan uraian di atas penelitian ini melakukan perbandingan hasil dari algoritma K-Means, Fuzzy C-Means, dan Otsu sebagai *dynamic thresholding* dan metode *blocking block area* sebagai penentu batas kata, serta dilakukan perbaikan

sehingga diperoleh hasil segmentasi yang akurat, sebelum akhirnya suara akan dikenali menggunakan metode MFCC dan HMM.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, dapat dirumuskan permasalahan yang dihadapi yaitu membandingkan algoritma-algoritma untuk mencari nilai *threshold* terbaik serta memperbaiki metode *blocking block area* pada proses segmentasi jenis *continuous speech* bahasa Indonesia, sehingga didapatkan metode terbaik untuk segmentasi kata pada aplikasi pengenalan *continuous speech* bahasa Indonesia.

1.3 Tujuan dan Manfaat

Tujuan yang ingin dicapai dalam penelitian tugas akhir ini adalah mengetahui algoritma terbaik pada *dynamic thresholding* dan metode *blocking block area* untuk proses segmentasi *continuous speech* bahasa Indonesia, serta mencoba untuk memperbaiki algoritma yang ada untuk meningkatkan akurasi.

Manfaat dari penelitian tugas akhir ini adalah:

1. Sebagai dasar dalam penelitian pengembangan aplikasi *continuous speech*
2. Membantu konversi suara ke teks untuk pemberian *subtitle* pada video berbahasa indonesia.

1.4 Ruang Lingkup

Adapun ruang lingkup dari aplikasi pengenalan ucapan jenis *continuous speech* dengan menggunakan metode MFCC dan HMM adalah:

1. Suara yang di-*input*-kan berupa hasil rekaman langsung dan *output* berupa *file* suara hasil segmentasi (*.wav) dan teks hasil pengenalan suara.
2. Objek perekam berusia antara 19 tahun - 25 tahun
3. Objek perekaman untuk data pelatihan menggunakan 5 orang *speaker* (5 laki-laki)
4. Perekaman inputan suara menggunakan headset Logitech Stereo H150 dengan jarak ± 3 cm dari sekitaran mulut seseorang.
5. Inputan suara dilakukan pada speaker dengan kondisi normal.
6. Sistem ini akan diimplementasikan berbasis desktop, dengan bahasa pemrograman MatLab

7. Ucapan berupa kalimat dalam bahasa Indonesia dengan basis data berupa kata dalam bahasa Indonesia
8. Pengucapan kalimat diberi jeda pada tiap kata.

1.5 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam tugas akhir ini terbagi dalam beberapa pokok bahasan, yaitu:

BAB I PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup dan sistematika penulisan tugas akhir Segmentasi *Continuous Speech* dengan Menggunakan *Dynamic Thresholding* dan Metode *Blocking Block Area*

BAB II TINJAUAN PUSTAKA

Bab ini memaparkan studi pustaka yang berhubungan dengan topik tugas akhir ini meliputi Bahasa Indonesia, Sinyal Ucapan, Pengenalan Ucapan, *Pre-Processing*, *Mel-Frequency Cepstral Coefficient* (MFCC), Hidden Markov Model (HMM), dan Unified Modelling Language (UML).

BAB III METODOLOGI PENELITIAN

Bab ini membahas langkah-langkah yang dilakukan pada penelitian Tugas Akhir. Penyelesaian masalah tersebut diawali dengan pengumpulan data, Pelatihan yang meliputi pre-processing, ekstraksi ciri menggunakan MFCC dan pelatihan menggunakan HMM, langkah selanjutnya yaitu melakukan pengenalan menggunakan HMM.

BAB IV PENGEMBANGAN PERANGKAT LUNAK

Bab ini memaparkan tahapan pembangunan perangkat lunak menggunakan model pengembangan Unified Modelling Language (UML). Model pengembangan UML melalui beberapa fase yaitu fase inception, fase elaboration, fase construction, dan fase transition.

BAB V HASIL DAN ANALISA

Bab ini memaparkan tahapan pengujian aplikasi meliputi pengujian tahap segmentasi dan pengujian pengenalan.

BAB VI PENUTUP

Penutup berisi tentang kesimpulan dari penulisan tugas akhir dan saran – saran untuk pengembangan selanjutnya.