

**SEGMENTASI *CONTINUOUS SPEECH* DENGAN MENGGUNAKAN
*SPECTRAL FEATURE EXTRACTION***



SKRIPSI

**Disusun Sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Komputer
pada Departemen Ilmu Komputer/ Informatika**

Disusun Oleh:

NURUL FADLILAH

24010315120008

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

2019

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini :

Nama : Nurul Fadlilah

NIM : 24010315120008

Judul : Segmentasi *Continuous Speech* dengan Menggunakan *Spectral Feature Extraction*

Dengan ini saya menyatakan bahwa dalam 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.

Semarang, 5 September 2019



Nurul Fadlilah

24010315120008

HALAMAN PENGESAHAN

Judul : Segmentasi *Continuous Speech* dengan Menggunakan *Spectral Feature Extraction*

Nama : Nurul Fadlilah

NIM : 24010315120008

Telah diujikan pada sidang tugas akhir pada tanggal 5 September 2019 dan dinyatakan lulus pada tanggal **5 September 2019**.

Semarang, 5 September 2019

Mengetahui,

a.n Ketua Departemen Ilmu Komputer/


Informatika FSM UNDIP

Sekretaris

Panitia Penguji Tugas Akhir

Ketua,




Dr. Eng. Adi Wibowo, S.Si, M.Kom

NIP. 198303092006041002

Dr. Retno Kusumaningrum, S.Si, M.Kom

NIP. 198104202005012001

HALAMAN PENGESAHAN

Judul : Segmentasi *Continuous Speech* dengan Menggunakan *Spectral Feature Extraction*

Nama : Nurul Fadlilah

NIM : 24010315120008

Telah diujikan pada sidang tugas akhir pada tanggal **5 September 2019**.

Semarang, 5 September 2019

Pembimbing



Sukmawati Nur Endah, S.Si, M.Kom

NIP. 19780502 200501 2 002

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan segala rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul “Segmentasi *Continuous Speech* dengan Menggunakan *Spectral Feature Extraction*”.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana komputer pada Departemen Ilmu Komputer/ Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.

Dalam pelaksanaan skripsi serta penyusunan dokumen skripsi ini, penulis menyadari banyak pihak yang membantu sehingga akhirnya dokumen ini dapat diselesaikan. Oleh karena itu, melalui kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Prof. Dr. Widowati, S.Si, M.Si, selaku Dekan Fakultas Sains dan Matematika.
2. Dr. Retno Kusumaningrum, S.Si, M.Kom, selaku Ketua Departemen Ilmu Komputer/Informatika.
3. Panji Wisnu Wirawan, S.T, M.T, selaku Koordinator Skripsi Departemen Ilmu Komputer/Informatika.
4. Sukmawati Nur Endah, S.Si, M.Kom, selaku dosen pembimbing skripsi.

Penulis menyadari bahwa dokumen skripsi ini masih jauh dari sempurna. Oleh karena itu, saran dan kritik yang membangun sangat penulis harapkan. Akhir kata, semoga skripsi ini dapat bermanfaat bagi semua pihak.

Semarang, 5 September 2019

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Diponegoro, saya yang bertanda tangan di bawah ini:

Nama : Nurul Fadlilah
NIM : 24010315120008
Program Studi : Informatika
Departemen : Ilmu Komputer/Informatika
Fakultas : Sains dan Matematika
Jenis karya : Skripsi

demikian pengembangan ilmu pengetahuan, menyetujui untuk memberikan **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*)** kepada Universitas Diponegoro atas karya ilmiah saya yang berjudul:

Segmentasi Continuous Speech dengan Menggunakan Spectral Feature Extraction

berserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-eksklusif ini Universitas Diponegoro berhak menyimpan, mengalihmedia/ formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis/ pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Semarang, 5 September 2019

Yang menyatakan



Nurul Fadlilah
Nurul Fadlilah

24010315120008

ABSTRAK

Segmentasi ucapan adalah suatu teknik untuk membagi sinyal ucapan ke dalam unit dasar dalam proses pengenalan ucapan. Tipe sinyal ucapan pada pengenalan ucapan dibagi menjadi *isolated word*, *connected word*, *continuous speech*, dan *spontaneous speech*. Penelitian ini berfokus pada tipe *continuous speech*. Segmentasi *continuous speech* digunakan untuk memecah kalimat menjadi kata, suku kata, atau fonem. Proses segmentasi *continuous speech* pada penelitian ini menggunakan pendekatan *spectral feature extraction* dan *dynamic thresholding* dalam melakukan deteksi segmen kata dalam domain bahasa Indonesia. Tujuan penelitian ini adalah melakukan perbandingan kinerja pada empat *spectral feature extraction*, yaitu *short time energy*, *zero crossing rate*, *spectral centroid* dan *spectral flux* dengan nilai parameter *threshold* dan *distance* untuk mendapatkan hasil optimum. Berdasarkan hasil penelitian, didapatkan akurasi kata terbaik 94.50% dan akurasi kalimat terbaik 83.75% menggunakan *zero crossing rate* dengan parameter *threshold* sebesar rata-rata dibagi empat dan *distance* kurang dari sama dengan 10 *frame*.

Kata Kunci: Segmentasi Ucapan, *Continuous Speech*, Segmentasi *Continuous Speech*, *Spectral Feature Extraction*, *Dynamic Thresholding*

ABSTRACT

Speech segmentation is a technique for dividing speech signals into basic units in the speech recognition process. The types of speech signals in speech recognition are divided into isolated words, connected words, continuous speech, and spontaneous speech. This study focuses on continuous speech types. Continuous speech segmentation is used to break sentences into words, syllables or phonemes. The continuous speech segmentation process in this study uses the spectral feature extraction approach and dynamic thresholding in detecting word segments on the Indonesian language domain. The purpose of this study is to compare the performance of the four spectral feature extractions, namely short time energy, zero crossing rate, spectral centroid and spectral flux with threshold and distance parameter values to obtain optimum results. Based on the results of the study, the best accuracy of word is 94.50% and accuracy of sentence is 83.75% using zero crossing rate with a threshold parameter of an average divided by four and a distance is less than 10 frames.

Keywords: Speech Segmentation, Continuous Speech, Continuous Speech Segmentation, Spectral Feature Extraction, Dynamic Thresholding

DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PENGESAHAN	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan dan Manfaat	3
1.4. Ruang Lingkup.....	3
1.5. Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA.....	6
2.1. Penelitian Terkait Segmentasi <i>Continuous Speech</i>	6
2.2. Sinyal Ucapan	7
2.3. <i>Continuous Speech</i>	9
2.4. <i>Pre-processing</i>	9
2.4.1. <i>Direct Current Removal</i>	10
2.4.2. <i>Pre-emphasizes</i>	10
2.4.3. <i>Framing</i>	11
2.4.4. <i>Windowing</i>	12
2.5. Segmentasi	13
2.5.1. Fast Fourier Transform (FFT)	14
2.5.2. <i>Spectral Feature Extraction</i>	15

2.6. Deteksi Segmen Ucapan.....	18
BAB III METODOLOGI PENELITIAN	20
3.1. Deskripsi Umum Penelitian	20
3.2. Tahapan Penelitian	20
3.2.1. Pengumpulan Data.....	20
3.2.2. <i>Pre-processing</i>	21
3.2.3. Segmentasi.....	27
3.3. Contoh Proses.....	33
3.3.1. Pengambilan Data.....	33
3.3.2. <i>Pre-Processing</i>	34
3.3.3. Segmentasi.....	39
3.3.4. Deteksi Segmen Ucapan.....	42
BAB IV HASIL DAN PEMBAHASAN.....	49
4.1. Data Penelitian	49
4.2. Skenario Penelitian.....	50
4.2.1. Skenario 1	50
4.2.2. Skenario 2	50
4.2.3. Skenario 3	51
4.2.4. Skenario 4.....	51
4.2.5. Skenario 5	51
4.3. Hasil Penelitian	51
4.4. Analisis Hasil Penelitian	52
4.4.1. Analisis Hasil Skenario 1	52
4.4.2. Analisis Hasil Skenario 2	54
4.4.3. Analisis Hasil Skenario 3	55
4.4.4. Analisis Hasil Skenario 4	56
4.4.5. Analisis Hasil Skenario 5	58
BAB V PENUTUP	60
4.1. Kesimpulan.....	60
4.2. Saran.....	60
DAFTAR PUSTAKA.....	62
LAMPIRAN-LAMPIRAN	65

DAFTAR TABEL

Tabel	2.1	Penelitian Terkait Segmentasi <i>Continuous Speech</i>	6
Tabel	3.1	Data Hasil Sampling.....	34
Tabel	3.2	Hasil Proses <i>DC-Removal</i>	35
Tabel	3.3	Hasil Proses <i>Pre-Emphasizes</i>	36
Tabel	3.4	Hasil <i>Framing</i>	37
Tabel	3.5	Hasil <i>Windowing</i>	38
Tabel	3.6	Hasil FFT.....	39
Tabel	3.7	Nilai <i>Short Time Energy</i>	40
Tabel	3.8	Nilai <i>Zero Crossing Rate</i>	41
Tabel	3.9	Nilai <i>Spectral Centroid</i>	41
Tabel	3.10	Nilai <i>Spectral Flux</i>	42
Tabel	3.11	Hasil Perhitungan Deteksi Segmen STE.....	43
Tabel	3.12	Hasil Perhitungan Deteksi Segmen ZCR	44
Tabel	3.13	Hasil Perhitungan Deteksi Segmen <i>Spectral Centroid</i>	44
Tabel	3.14	Hasil Perhitungan Deteksi Segmen <i>Spectral Flux</i>	45
Tabel	4.1	Daftar Kalimat <i>Continuous Speech</i>	49
Tabel	4.2	Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 5.....	52
Tabel	4.3	Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 5	52

DAFTAR GAMBAR

Gambar 2.1	Contoh sinyal ucapan “It’s time” (Rabiner & Juang, 1993)	7
Gambar 2.2	Efek <i>direct current</i> pada sinyal ucapan	10
Gambar 2.3	Ilustrasi proses framing (Dewi, 2016)	11
Gambar 2.4	Ilustrasi <i>zero crossing</i> (Shete et al., 2014).....	16
Gambar 3.1	Blok proses segmentasi <i>continuous speech</i>	21
Gambar 3.2	Flowchart pengumpulan data.....	22
Gambar 3.3	<i>Flowchart</i> tahap <i>pre-processing</i>	22
Gambar 3.4	<i>Flowchart</i> proses <i>DC-Removal</i>	23
Gambar 3.5	<i>Flowchart</i> proses <i>Pre-emphasize</i>	24
Gambar 3.6	<i>Flowchart</i> proses <i>Framing</i>	25
Gambar 3.7	<i>Flowchart</i> proses <i>Windowing</i>	26
Gambar 3.8	<i>Flowchart</i> proses segmentasi.....	27
Gambar 3.9	<i>Flowchart</i> proses <i>time domain feature</i>	28
Gambar 3.10	<i>Flowchart</i> proses <i>short time energy</i>	29
Gambar 3.11	<i>Flowchart</i> proses <i>zero crossing rate</i>	29
Gambar 3.12	<i>Flowchart</i> proses <i>frequency domain feature</i>	30
Gambar 3.13	<i>Flowchart</i> proses FFT.....	31
Gambar 3.14	<i>Flowchart</i> proses <i>spectral centroid</i>	32
Gambar 3.15	<i>Flowchart</i> proses <i>spectral flux</i>	32
Gambar 3.16	<i>Flowchart</i> Deteksi Segmen.....	33
Gambar 3.17	Sinyal ucapan <i>original</i> “abang bercerita sesuatu yang bagus”	34
Gambar 3.18	Hasil <i>DC-Removal</i> pada sinyal ucapan "abang bercerita sesuatu yang bagus".....	35
Gambar 3.19	Hasil <i>pre-emphasize</i> pada sinyal ucapan "abang bercerita sesuatu yang bagus".....	36
Gambar 3.20	Hasil <i>short time energy</i> pada sinyal ucapan “abang bercerita sesuatu yang bagus”	46
Gambar 3.21	Hasil <i>zero crossing rate</i> pada sinyal ucapan “abang bercerita sesuatu yang bagus”	46

Gambar 3.22 Hasil <i>spectral centroid</i> pada sinyal ucapan “abang bercerita sesuatu yang bagus”	46
Gambar 3.23 Hasil <i>spectral flux</i> pada sinyal ucapan “abang bercerita sesuatu yang bagus”	46
Gambar 3.24 Hasil deteksi segmen pada <i>short time energy</i>	47
Gambar 3.25 Hasil deteksi segmen pada <i>zero crossing rate</i>	47
Gambar 3.26 Hasil deteksi segmen pada <i>spectral centroid</i>	47
Gambar 3.27 Hasil deteksi segmen pada <i>spectral flux</i>	47
Gambar 3.28 Deteksi awal dan akhir setiap segmen pada kalimat ucapan "abang bercerita sesuatu yang bagus"	48
Gambar 4.1 Skenario Penelitian	50
Gambar 4.2 Grafik Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 1	53
Gambar 4.3 Grafik Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 1	53
Gambar 4.4 Grafik Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 2.....	54
Gambar 4.5 Grafik Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 2	55
Gambar 4.6 Grafik Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 3	55
Gambar 4.7 Grafik Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 3	56
Gambar 4.8 Grafik Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 4.....	57
Gambar 4.9 Grafik Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 4	57
Gambar 4.10 Grafik Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 5	58
Gambar 4.11 Grafik Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 5	58

DAFTAR LAMPIRAN

Lampiran 1. Tabel Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 1	66
Lampiran 2. Tabel Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 2	68
Lampiran 3. Tabel Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 3	69
Lampiran 4. Tabel Hasil Segmentasi <i>Continuous Speech</i> Berjeda Skenario 4	70
Lampiran 5. Tabel Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 1	71
Lampiran 6. Tabel Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 2	72
Lampiran 7. Tabel Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 3	73
Lampiran 8. Tabel Hasil Segmentasi <i>Continuous Speech</i> Natural Skenario 4	74
Lampiran 9. Tabel Hasil Segmentasi Pembandingan <i>Spectral Centroid</i> pada <i>Continuous Speech</i> Berjeda.....	75
Lampiran 10. Grafik Hasil Segmentasi Pembandingan <i>Spectral Centroid</i> pada <i>Continuous Speech</i> Berjeda.....	76
Lampiran 11. Tabel Detail Hasil Segmentasi pada <i>Continuous Speech</i> Berjeda Parameter 6 oleh <i>Speaker</i> Pertama.....	77
Lampiran 12. Tabel Detail Hasil Segmentasi pada <i>Continuous Speech</i> Natural Parameter 2 oleh <i>Speaker</i> Kedua	78
Lampiran 13. Tabel Hasil Segmentasi <i>Continuous Speech</i> Berjeda pada <i>Zero Crossing Rate</i> dengan <i>frame size</i> 30 ms.....	79
Lampiran 14. Grafik Hasil Segmentasi <i>Continuous Speech</i> Berjeda pada <i>Zero Crossing Rate</i> dengan <i>frame size</i> 30 ms.....	80
Lampiran 15. Tabel Hasil Segmentasi <i>Continuous Speech</i> Berjeda pada <i>Zero Crossing Rate</i> dengan <i>frame size</i> 40 ms.....	81
Lampiran 16. Grafik Hasil Segmentasi <i>Continuous Speech</i> Berjeda pada <i>Zero Crossing Rate</i> dengan <i>frame size</i> 40 ms.....	82
Lampiran 17. Tabel Hasil Segmentasi <i>Continuous Speech</i> Natural pada <i>Zero Crossing Rate</i> dengan <i>frame size</i> 30 ms	83
Lampiran 18. Grafik Hasil Segmentasi <i>Continuous Speech</i> Natural pada <i>Zero Crossing Rate</i> dengan <i>frame size</i> 30 ms.....	84
Lampiran 19. Tabel Hasil Segmentasi <i>Continuous Speech</i> Natural pada <i>Zero Crossing Rate</i> dengan <i>frame size</i> 40 ms	85

Lampiran 20. Grafik Hasil Segmentasi pada *Continuous Speech* Natural pada *Zero Crossing Rate* dengan *frame size* 40 ms 86

BAB I

PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, manfaat dan tujuan, ruang lingkup, serta sistematika penulisan skripsi yang berjudul Segmentasi *Continuous Speech* dengan Menggunakan *Spectral Feature Extraction*.

1.1. Latar Belakang

Ucapan adalah salah satu cara komunikasi antar manusia yang penting dan efektif untuk bertukar informasi. Seiring pesatnya perkembangan teknologi dan informasi, muncul teknologi ucapan (*speech technology*) seperti *speech synthesis*, *speaker recognition*, *speech compression*, *speaker verification*, *multimodal interaction*, dan *speech recognition* (Natarajan & Jothilakshmi, 2015). *Speech recognition* merupakan teknologi ucapan yang sering digunakan pada robot, mesin cerdas dan untuk komunikasi antar manusia dengan komputer. Adanya teknologi tersebut dapat digunakan secara efisien dalam berbagai keseharian untuk meningkatkan kinerja suatu pekerjaan dan mempermudah penderita cacat fisik dalam berkomunikasi. *Speech recognition* dapat diklasifikasi menjadi tiga tipe yaitu *speech recognition* berdasarkan sinyal ucapan, model pembicara dan perbendaharaan kata (Saksamudre et al., 2015). Adapun *speech recognition* berdasarkan sinyal ucapan mempunyai empat tipe, yaitu *isolated word*, *connected word*, *continuous speech*, dan *spontaneous speech* (Naziya & Deshmukh, 2016). Penelitian ini berfokus pada tipe *continuous speech* yang merupakan pengucapan kata secara natural yang terdiri dari dua bagian utama, yaitu membawa informasi ucapan dan membawa bagian sunyi atau bising yang berada di antara ucapan tanpa informasi verbal (Sakran et al., 2017).

Speech recognition yang dilakukan dalam *continuous speech* membutuhkan segmentasi untuk mempermudah dalam proses pengenalan. Segmentasi adalah proses memecah sinyal ucapan menjadi unit dasar (Rahman & Bhuiyan, 2012). Segmentasi ucapan pada *continuous speech* digunakan untuk mengidentifikasi kata, suku kata, atau fonem. Khususnya pada segmentasi kata yang mempunyai peran penting dalam *speech recognition*, yaitu mengurangi ukuran memori dan komputasi yang kompleks pada sistem dengan kosa kata yang besar (Sangeetha & Jothilakshmi, 2012). Proses segmentasi dilakukan dengan membedakan sinyal yang bersuara dan tidak bersuara (Abriyono & Harjoko, 2012). Selain

itu segmentasi juga digunakan untuk mengetahui awal dan akhir kata yang tepat (Kalamani et al., 2015).

Penelitian mengenai segmentasi ucapan dalam kata pada 30 tahun terakhir berkembang pesat dalam berbagai bahasa, diantaranya Rahman & Bhuiyan (2012), Kalamani et al. (2015), Londhe & Kshirsagar (2017) yang melakukan segmentasi pada bahasa Hindi dan Radman et al. (2015) yang melakukan segmentasi pada bahasa Arab. Adapun hasil segmentasi pada penelitian yang dilakukan Rahman & Bhuiyan (2012) mencapai 96.25%, Kalamani et al. (2015) mencapai 98.33% dan Londhe & Kshirsagar (2017) mencapai 96.00% dengan menggunakan *spectral feature extraction*. Segmentasi ucapan pada bahasa Indonesia masih minim dengan hasil yang belum maksimal, seperti pada penelitian yang dilakukan Arasyi (2016) menghasilkan akurasi segmentasi 90.00% dengan menggunakan metode *blocking block area*. Hal ini disebabkan oleh algoritma yang digunakan pada suatu proses menghasilkan sinyal suara yang bukan merupakan bagian dari kata menjadi ikut tersegmentasi dan mengakibatkan adanya *unvoice* pada suara hasil segmentasi. Selain itu, juga disebabkan oleh beberapa aspek, diantaranya pengucapan dan kesehatan yang melakukan rekaman.

Penelitian ini membahas proses segmentasi kata pada *continuous speech* dalam domain bahasa Indonesia dengan menggunakan *spectral feature extraction* seperti dalam penelitian Rahman & Bhuiyan (2012) yang dilakukan pada domain bahasa Hindi. Domain bahasa Hindi dan bahasa Indonesia merupakan bahasa yang mempunyai induk bahasa Sansekerta, sehingga pengucapan nama Sansekerta dalam bahasa Indonesia mempunyai kemiripan dengan pelafalan bahasa Hindi.

Berdasarkan uraian tersebut, penelitian ini membahas tentang segmentasi ucapan pada *continuous speech* domain bahasa Indonesia dengan menggunakan *spectral feature extraction* untuk menghasilkan potongan kata yang sesuai dengan ucapan dari masukan *continuous speech*.

1.2. Rumusan Masalah

Berdasarkan uraian latar belakang diatas, maka dapat dirumuskan masalah yaitu bagaimana melakukan segmentasi kata pada *continuous speech* bahasa Indonesia dengan menggunakan *spectral feature extraction*.

1.3. Tujuan dan Manfaat

Tujuan umum dari penelitian skripsi ini adalah melakukan segmentasi kata pada *continuous speech* bahasa Indonesia menggunakan *spectral feature extraction*. Adapun tujuan khusus dari penelitian skripsi ini adalah mengetahui kinerja *spectral feature extraction* dalam segmentasi *continuous speech* bahasa Indonesia.

Manfaat yang diharapkan dari penelitian skripsi ini adalah sebagai dasar pengembangan pengenalan ucapan pada *continuous speech* bahasa Indonesia menggunakan *spectral feature extraction* dalam proses segmentasi kata.

1.4. Ruang Lingkup

Ruang lingkup untuk menerapkan metode *spectral feature extraction* dalam proses segmentasi *continuous speech* bahasa Indonesia adalah :

1. Dataset berupa rekaman suara yang dilakukan dengan perekaman secara langsung menggunakan headset Logitech Stereo H150.
2. Saat perekaman suara, jarak headset dengan sumber suara (mulut seseorang) adalah ± 3 cm.
3. Objek perekaman adalah 8 orang laki-laki dengan umur 19 – 25 tahun dalam kondisi suara normal.
4. Inputan suara berupa kalimat *continuous speech* berjeda dan *continuous speech* natural dengan kata dalam bahasa Indonesia.
5. Implementasi dilakukan pada bahasa pemrograman Python dan berbasis web.
6. Input suara berekstensi file *.wav hasil perekaman dan output berupa file suara tersegmentasi berekstensi *.wav.

1.5. Sistematika Penulisan

Sistematika penulisan dalam laporan penelitian ini dapat dibagi dalam beberapa pokok bahasan, yaitu :

BAB I PENDAHULUAN

Bab ini membahas latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup serta sistematika penulisan penelitian ini.

BAB II TINJAUAN PUSTAKA

Bab ini menjabarkan studi pustaka terkait teori-teori yang mendasari topik yang dibahas pada penelitian ini yang meliputi Sinyal Ucapan, *Continuous Speech*, *Pre-*

Processing, Segmentasi, Spectral Feature Extraction dan Deteksi Segmen Ucapan.

BAB III METODOLOGI PENELITIAN

Bab ini membahas urutan langkah dalam penelitian ini diantaranya adalah pengambilan dataset, *pre-processing*, segmentasi, dan deteksi segmen ucapan, serta contoh perhitungannya.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil penelitian dan analisis terhadap proses segmentasi dan deteksi segmen ucapan.

BAB V PENUTUP

Bab ini berisi kesimpulan dari pembahasan pada bab-bab sebelumnya dan saran untuk pengembangan berikutnya.