

SKRIPSI

**PENGENALAN DIALEK DI SUMATERA SELATAN MENGGUNAKAN
ALGORITMA *DEEP NEURAL NETWORK***



Disusun untuk Memenuhi Syarat Mendapatkan Gelar Sarjana Teknik pada

Jurusan Teknik Elektro Fakultas Teknik

Universitas Sriwijaya

Oleh:

MUHAMMAD RIZKI PUTRA

03041181722022

JURUSAN TEKNIK ELEKTRO

FAKULTAS TEKNIK

UNIVERSITAS SRIWIJAYA

2021

LEMBAR PENGESAHAN
PENGENALAN DIALEK DI SUMATERA SELATAN MENGGUNAKAN
ALGORITMA DEEP NEURAL NETWORK



SKRIPSI

Disusun untuk Memenuhi Syarat Mendapatkan Gelar Sarjana Teknik pada

Jurusan Teknik Elektro Fakultas Teknik

Universitas Sriwijaya

Oleh:

MUHAMMAD RIZKI PUTRA

03041181722022

Mengetahui,

Ketua Jurusan Teknik Elektro



Muhammad Abu Bakar Sidik, S.T., M.Eng., Ph.D.

NIP. 197108141999031005

Indralaya, 23 November 2021

Menyetujui,

Pembimbing Utama

Dr. Eng. Suci Dwijavanti, S.T., M.Sc.

NIP. 198407302008122001

HALAMAN PERNYATAAN INTEGRITAS

Yang bertanda tangan di bawah ini :

Nama : Muhammad Rizki Putra
NIM : 03041181722022
Fakultas : Teknik
Jurusan/Prodi : Teknik Elektro
Universitas : Universitas Sriwijaya

Hasil Pengecekan *Software iThenticate/Turnitin*: 3 %

Menyatakan bahwa tugas akhir saya yang berjudul “Pengenalan Dialek di Sumatera Selatan Menggunakan Algoritma *Deep Neural Network*”. merupakan hasil karya sendiri dan benar keasliannya. Apabila ternyata dikemudian hari ditemukan unsur penjiplakan/plagiat dalam karya ilmiah ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tanpa paksaan.

Palembang, 17 November 2021



Muhammad Rizki Putra

NIM. 03041281722022

Saya sebagai Pembimbing dengan ini menyatakan bahwa saya telah membaca dan menyetujui skripsi ini dan dalam pandangan saya ruang lingkup dan kualitas skripsi ini mencukupi sebagai skripsi mahasiswa sarjana strata satu (S1).

Tanda Tangan :  . _____

Pembimbing Utama : SUEI DWIJAYANTI

Tanggal : 23 / 11 / 2021

**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK
KEPENTINGAN AKADEMIS**

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

Nama : Muhammad Rizki Putra
NIM : 03041181722022
Jurusan/Prodi : Teknik Elektro
Fakultas : Teknik
Jenis Karya : Skripsi

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

**Pengenalan Dialek di Sumatera Selatan Menggunakan
Algoritma Deep Neural Network**

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Sriwijaya berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan tulisan 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.

Dibuat di Palembang

Pada tanggal: 17 November 2021

Yang menyatakan,



Muhammad Rizki Putra

NIM. 03041181722022

KATA PENGANTAR

Puji dan syukur penulis ucapkan kepada Allah SWT serta shalawat dan salam kepada Nabi Muhammad SAW, keluarga dan para sahabat. Berkat rahmat dan ridho Allah, penulis dapat membuat skripsi ini dengan judul “Pengenalan Dialek di Sumatera Selatan Menggunakan Algoritma *Deep Neural Network*”. Pembuatan skripsi ini adalah salah satu syarat untuk memperoleh gelar Sarjana Teknik pada Jurusan Teknik Elektro Fakultas Teknik Universitas Sriwijaya. Dalam kesempatan ini penulis mengucapkan terima kasih kepada:

1. Orang tua, saudara dan keluarga yang selalu memberikan doa dan dukungan selama masa studi.
2. Ibu Dr. Eng. Suci Dwijayanti, S.T., M.S. selaku pembimbing tugas akhir ini.
3. Bapak Dr. Bhakti Yudho Suprpto, S.T., M.T., Bapak Ir. Zainal Husein, M.S. dan Ibu Hera Hikmarika, S.T., M.Eng. selaku dosen penguji pada tugas akhir ini.
4. Dosen pembimbing akademik, Ir. Rudyanto Thayib, M.Sc. yang telah membimbing penulis selama masa perkuliahan dan memberi saran serta masukan dalam pengambilan mata kuliah.
5. Bapak Muhammad Abu Bakar Sidik, S.T., M.Eng., Ph.D. selaku Ketua Jurusan Teknik Elektro dan Ibu Dr. Eng. Suci Dwijayanti, S.T., M.S. selaku Sekretaris Jurusan Teknik Elektro.
6. Segenap Dosen Jurusan Teknik Elektro Fakultas Teknik Universitas Sriwijaya yang telah memberikan ilmu selama perkuliahan.
7. Fauzan, Zaid, Kak Vovo, Mutiyara, Sandy, Cepy, Andi, dan Nailah yang telah bersedia menjadi narasumber sampel audio penelitian.
8. Teman-teman TKK 2017, baik di Indralaya maupun di Palembang.
9. Koas beserta asisten-asisten Laboratorium LKR 2017, Kak Abid, Irvine, Markus, Yusuf, Aisyah, Delia, dan Sukan.
10. Serta pihak-pihak yang sangat membantu di dalam penyusunan skripsi ini, yang tidak dapat penyusun sebutkan satu persatu.

Penulis menyadari adanya kekurangan dalam penulisan usulan skripsi ini. Semoga skripsi ini dapat memberikan manfaat dan wawasan yang lebih luas kepada pembaca. Oleh karena itu, kritik, dan saran yang membangun sangat penulis harapkan agar dapat menjadi evaluasi yang baik dan berguna untuk perbaikan ke depannya.

Indralaya, 17 November 2021



Muhammad Rizki Putra

NIM. 03041181722022

ABSTRAK
PENGENALAN DIALEK DI SUMATERA SELATAN MENGGUNAKAN
ALGORITMA DEEP NEURAL NETWORK

(Muhammad Rizki Putra, 03041181722022, 2021, 81 halaman)

Suatu bahasa memiliki beragam dialek di setiap daerah. Hal ini dapat mempengaruhi perkembangan teknologi, khususnya dalam pengenalan suara. Namun, penelitian yang membahas tentang dialek Sumatera Selatan belum ada sehingga pada penelitian ini dikembangkan sistem pengenalan dialek daerah dari Sumatera Selatan dengan menggunakan model *deep neural network* (DNN). *Dataset* yang digunakan dalam penelitian ini adalah data primer dari 5 responden yang merupakan penutur asli dari dialek yang digunakan, yang terdiri dari dialek Beliti, dialek Sekayu, dialek Palembang, dialek Lahat, dialek Muara Enim, dan bahasa Indonesia baku. Ciri-ciri sinyal suara yang diperoleh dari *dataset* adalah *mel spectrogram*, *short time fourier transform* (STFT), dan *mel frequency cepstral coefficient* (MFCC). Hasil pengujian menunjukkan bahwa model DNN yang menggunakan *optimizer* Adam dan *loss crossentropy* memiliki nilai akurasi 0,575 dengan *input* berupa ekstraksi *mel spectrogram* dan 0,555 untuk input STFT. Akurasi tertinggi dicapai dalam mengenali dialek Beliti, yaitu 72,7% dan dialek Palembang 71,4% jika ekstraksi ciri yang digunakan adalah *mel spectrogram*. Sedangkan untuk Bahasa Indonesia, akurasi tertinggi adalah dengan menggunakan ekstraksi ciri STFT, yaitu 71,4%. Sedangkan model yang menggunakan ciri MFCC menunjukkan performansi yang paling rendah. Hasil ini menunjukkan bahwa *mel spectrogram* dan STFT dapat digunakan sebagai input DNN untuk pengenalan dialek. Hasil penelitian juga menunjukkan bahwa model DNN dapat memprediksi semua dialek, kecuali dialek Muara Enim. Hal ini dikarenakan dialek Muara Enim direkam pada ruang terbuka sehingga *background noise* mempengaruhi pengenalan dialek.

Kata kunci: *DNN, pengenalan dialek, mel spectrogram, STFT, MFCC*

Mengetahui,

Ketua Jurusan Teknik Elektro



Muhammad Abu Bakar Sidik, S.T., M.Eng., Ph.D.

NIP. 197108141999031005

Indralaya, 17 November 2021

Menyetujui,

Pembimbing Utama



Dr. Eng. Suci Dwijayanti, S.T., M.Sc.

NIP. 198407302008122001

ABSTRACT

**DIALECT RECOGNITION IN SOUTH SUMATRA USING DEEP
NEURAL NETWORK ALGORITHM**

(Muhammad Rizki Putra, 03041181722022, 2021, 81 pages)

A language has various dialects in each region. It can affect the development of technology, especially in speech recognition. However, the South Sumatran dialects have not been discussed yet. Thus, this study developed a method to recognize dialects using the deep neural network (DNN) model. The dataset used in this study is primary data from 5 respondents who are native speakers of the dialect used, which consists of Beliti dialect, Sekayu dialect, Palembang dialect, Lahat dialect, Muara Enim dialect, and standard Indonesian. The characteristics of the voice signal obtained from the dataset are mel spectrogram, short-time Fourier transform (STFT), and mel frequency cepstral coefficient (MFCC). The test results show that the DNN model that uses the Adam optimizer and loss cross-entropy has an accuracy value of 0.575 with input in the form of mel spectrogram and 0.555 for STFT input. The highest accuracy is achieved in recognizing the Beliti dialect at 72.7% and the Palembang dialect at 71.4% when using mel spectrogram features. As for Indonesian, the highest accuracy is by using STFT feature extraction, which is 71.4%. Meanwhile, MFCC shows the lowest performance. These results indicate that the mel spectrogram and STFT can be used as DNN input for dialect recognition. The results also show that the DNN model can predict all dialects, except the Muara Enim dialect. This is because the Muara Enim dialect was recorded in an open space so that background noise affects dialect recognition.

Keyword: *DNN, dialect recognition, mel spectrogram, STFT, MFCC*

Mengetahui,

Ketua Jurusan Teknik Elektro



Muhammad Abu Bakar Sidik, S.T., M.Eng., Ph.D.

NIP. 197108141999031005

Indralaya, 17 November 2021

Menyetujui,

Pembimbing Utama

Dr. Eng. Suci Dwijayanti, S.T., M.Sc.

NIP. 198407302008122001

DAFTAR ISI

HALAMAN SAMPUL	i
LEMBAR PENGESAHAN	ii
HALAMAN PERNYATAAN INTEGRITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
KATA PENGANTAR	vi
ABSTRAK	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penulisan	4
1.4 Pembatasan Masalah	4
1.5 Keaslian Penelitian	4
BAB II TINJAUAN PUSTAKA	7
2.1 <i>State of The Art</i>	7
2.2 Teori Pendukung	14
2.2.1. Sistem Pengenalan Dialek.....	14
2.2.2. Preprocessing	14
2.2.3. Ekstraksi Ciri dengan Short-Time Fourier Transform (STFT).....	19
2.2.4. Ekstraksi Ciri Mel Spectrogram.....	21
2.2.5. Ekstraksi Ciri MFCC (<i>Mel-Frequency Cepstrum Coefficient</i>).....	21

2.2.6. <i>Deep Neural Network</i> (DNN)	23
BAB III METODOLOGI PENELITIAN	29
3.1. Studi Literatur	29
3.2. Pengambilan Data	29
3.3. Perancangan Sistem.....	31
3.4. Pengujian.....	33
BAB IV HASIL DAN PEMBAHASAN	34
4.1. Pengambilan Data Rekaman Suara	34
4.2. Proses <i>Preprocessing</i> Sinyal Suara	34
4.3. Ekstraksi Ciri Sinyal Suara	36
4.4. Proses <i>Training Dataset</i>	38
4.4.1. Pelatihan Dataset Ciri Mel Spectrogram.....	41
4.4.2. Pelatihan Dataset Ciri STFT	42
4.4.3. Pelatihan Dataset Ciri MFCC	44
4.5. Pengujian Dataset.....	46
BAB V PENUTUP.....	56
5.1. Kesimpulan.....	56
5.2. Saran.....	56
DAFTAR PUSTAKA.....	57
LAMPIRAN.....	60

DAFTAR GAMBAR

Gambar 2.1 Hasil Perbandingan <i>Error</i> pada Uji Tertutup dan Uji Terbuka.....	7
Gambar 2.2 Sistem Pengenalan Dialek	14
Gambar 2.3 Ekstraksi Spektrogram dan Proses Komputasi Kepadatan Spektral Daya	20
Gambar 2.4 Contoh Arsitektur DNN	24
Gambar 3.1 Alur Proses Penelitian	32
Gambar 4.1 Contoh kalimat “Buku ini dijual dengan harga Rp150.000.” yang Diulang Sebanyak Dua Kali.....	34
Gambar 4.2 Perbandingan Spectrogram Sinyal Suara Sebelum <i>Filtering</i> dan Setelah <i>Filtering</i>	35
Gambar 4.3 Timeframe dari Salah Satu Sampel Suara	36
Gambar 4.4 Ciri Mel Spectrogram	37
Gambar 4.5 Ciri STFT (spektrogram)	37
Gambar 4.6 Ciri MFCC.....	38
Gambar 4.7 Grafik Akurasi <i>Training</i> dan Validasi (atas) dan <i>Loss Training</i> (bawah) pada Model DNN yang Menggunakan <i>Optimizer</i> Adam dan <i>Loss</i> <i>Function Categorical Crossentropy</i>	41
Gambar 4.8 Grafik Akurasi <i>Training</i> dan Validasi (atas) dan <i>Loss Training</i> (bawah) pada Model DNN yang Menggunakan <i>Optimizer</i> SGD dan <i>Loss Function</i> MSE	43
Gambar 4.9 Grafik Akurasi <i>Training</i> dan Validasi (atas) dan <i>Loss Training</i> (bawah) pada Model DNN yang Menggunakan <i>Optimizer</i> SGD dan <i>Loss Function</i> MSE	44

DAFTAR TABEL

Tabel 2.1 Jumlah Penutur dalam Set Training dan Uji bagi Tiga Dialek Makro dan Bahasa Slovak Standar	8
Tabel 2.2 Jumlah Ujaran dalam Set Training dan Uji bagi Tiga Dialek Makro dan Bahasa Slovak Standar	8
Tabel 2.3 <i>Confusion Matrix</i> – Klasifikasi Kelompok Dialek Makro dan Bahasa Slovak Standar.....	9
Tabel 2.4 Hasil Perbandingan Data Latih dengan Data Uji pada Penelitian Bahasa Sunda	10
Tabel 2.5 Hasil Pengujian Per Kalimat Hasil Pengujian Per Kalimat pada Penelitian Bahasa Sunda	11
Tabel 2.6 Perbandingan Presisi dan <i>Recall</i> pada Sistem A2DID-DNN dan A2DID-SVM	13
Tabel 3.1 Kalimat-kalimat untuk Data Rekaman.....	30
Tabel 4.1 Parameter Pelatihan DNN	39
Tabel 4.2 Arsitektur Model DNN untuk Pelatihan Ciri Mel Spectrogram	39
Tabel 4.3 Arsitektur Model DNN untuk Pelatihan Ciri STFT	39
Tabel 4.4 Arsitektur Model DNN untuk Pelatihan Ciri MFCC	40
Tabel 4.5 Perbandingan Kinerja Pelatihan Model DNN Menggunakan Fitur Mel Spectrogram.....	42
Tabel 4.6 Perbandingan Kinerja Pelatihan Model DNN Menggunakan Fitur STFT	44
Tabel 4.7 Perbandingan Kinerja Pelatihan Model DNN Menggunakan Fitur MFCC.....	45

Tabel 4.8 <i>Confusion Matrix</i> Hasil Prediksi Model DNN Menggunakan Optimizer SGD dan Fungsi Loss MSE.....	46
Tabel 4.9 Nilai F1-Score pada Prediksi Model DNN Menggunakan Optimizer SGD dan Fungsi Loss MSE.....	48
Tabel 4.10 <i>Confusion Matrix</i> Hasil Prediksi Model DNN Menggunakan Optimizer Adam dan Fungsi Loss MSE.....	49
Tabel 4.11 Nilai F1-Score pada Prediksi Model DNN Menggunakan Optimizer Adam dan Fungsi Loss MSE.....	50
Tabel 4.12 <i>Confusion Matrix</i> Hasil Prediksi Model DNN Menggunakan Optimizer SGD dan Fungsi Loss Categorical Crossentropy.....	51
Tabel 4.13 Nilai F1-Score pada Prediksi Model DNN Menggunakan Optimizer SGD dan Fungsi Loss Categorical Crossentropy	52
Tabel 4.14 <i>Confusion Matrix</i> Hasil prediksi Model DNN dengan menggunakan Optimizer Adam dan Fungsi Loss Categorical Crossentropy	53
Tabel 4.15 Nilai F1-Score pada Prediksi Model DNN Menggunakan Optimizer Adam dan Fungsi Loss Categorical Crossentropy	54

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bahasa merupakan sistem komunikasi yang terdiri atas kumpulan suara dan simbol-simbol tertulis yang digunakan manusia pada negara atau daerah tertentu untuk berbicara atau menulis. Beberapa bahasa memiliki sistem tulisan yang terdiri dari simbol-simbol yang melambangkan suara-suara, suku-suku kata, gestur, arti dari bahasa tersebut ataupun gabungannya [1].

Sementara itu, dialek merupakan variasi bahasa yang merupakan karakteristik kelompok tertentu pada penutur suatu bahasa [2]. Jenis dialek terbagi menjadi empat, yaitu: dialek regional, dialek etnis, *sociolect*, dan aksen. Dialek regional merupakan variasi bahasa yang dipengaruhi oleh kondisi geografis suatu wilayah. Dialek etnis merupakan variasi bahasa yang berhubungan dengan kelompok etnis tertentu. *Sociolect* merupakan jenis dialek yang dipengaruhi oleh kelompok-kelompok sosial. Sedangkan, aksen adalah hasil dari perbedaan fonetik atau pengucapan antara kelompok satu dengan yang lain [3].

Salah satu cara untuk mengetahui asal usul seorang penutur adalah dengan mendengar cara/gaya bicarannya. Setiap penutur memiliki dialek tersendiri yang didapatkan dari keluarga, kerabat dan lingkungan sekitarnya. Terkadang, dialek kedua orang dalam sebuah dialog dapat terdengar sama meskipun memiliki latar belakang daerah yang berbeda ketika berbicara. Selain itu, ada kemungkinan seseorang yang berasal dari daerah lain berbicara dengan dialek yang berbeda dengan daerah asalnya. Dialek akan mempengaruhi pengenalan suara dan berbagai aplikasi suara lainnya.

Dalam survei pengenalan dialek dan bahasa yang dilakukan oleh A. Etman dan A. A. Beex, 12 percobaan dilakukan dengan berbagai metodologi dalam banyak bahasa dan dialek. Hasil penelitian tersebut menunjukkan beragam nilai akurasi dan observasi yang diperoleh dari 12 percobaan tersebut [4]. Terdapat pula penelitian mengenai pengenalan dialek antara lain bahasa Slovak dengan menggunakan

gaussian mixture model (GMM) [5] dan bahasa Arab menggunakan jaringan syaraf tiruan [6].

Selain itu, Qian Zhang dan John H. L. Jansen melakukan percobaan pengenalan dialek pada *unsupervised deep learning* dengan menggunakan tiga korpus: korpus dialek bahasa China, korpus dialek bahasa Arab dan korpus *multigenre challenge* (MGB-3) bahasa Arab. Pada dekade sebelumnya, ciri-ciri *bottleneck* telah digunakan sebagai pengembangan terkini dari pengenalan dialek/bahasa. Bagaimanapun, ekstraksi ciri *bottleneck* tradisional memerlukan informasi ucapan yang telah dicatat. Sebagai pilihan, dua jenis metode *unsupervised deep learning* diperkenalkan dalam penelitian ini. Untuk mengatasi batasan tersebut, maka diusulkan pendekatan ekstraksi fitur *bottleneck unsupervised* yang berasal dari struktur *bottleneck* tradisional, namun dilatih dengan perkiraan label fonetik [7]. Ada juga penelitian yang dilakukan oleh Soumia Bougrine yang menggunakan DNN dan *support vector machine* (SVM) untuk mengenali dialek bahasa Arab Aljazair diantaranya adalah dialek Hilali, Sulaymiyyah, Ma'liqi, dan Algiers-blanks [8]. Pada penelitian ini diusulkan sebuah sistem yang berdasarkan pada informasi suara prosodik seperti intonasi dan ritme untuk identifikasi dialek dalam negeri. Fitur suara ini diekstrak setelah segmentasi konsonan/vokal butiran-kasar (*coarse-grained*) dan model dialeknnya dibangun menggunakan DNN dan SVM.

Penelitian dialek untuk bahasa Indonesia juga sangat penting karena Indonesia sendiri memiliki lebih dari 700 bahasa yang dituturkan [9]. Keberagaman etnis dan bahasa daerah di Indonesia merupakan salah satu faktor terbentuknya keberagaman dialek pada bahasa Indonesia. Kondisi geografis Indonesia juga memberi keunikan dialek dimana terdapat banyak bahasa daerah yang berbeda satu sama lain, baik pada pulau yang sama atau antar pulau.

Salah satu penelitian yang membahas pengenalan dialek di Indonesia yaitu dialek Jawa dan Sunda dilakukan oleh R. Rahmawati dan D. P. Lestari. Pada penelitian tersebut digunakan metode GMM dan *I-Vector* [10]. Namun, hasil yang diperoleh belum akurat karena GMM sangat bergantung pada ekstraksi ciri yang digunakan.

Lalu, penelitian yang telah dilakukan oleh Warrohmah *et al.* mengidentifikasi dialek penutur pada bahasa Indonesia dimana penutur memiliki dialek Jawa, Batak, dan Minang. Metode yang digunakan adalah *back propagation neural network* (BPNN) dengan ekstraksi ciri berupa *mel frequency cepstral coefficient* (MFCC) [11]. Ada pula penelitian yang dilakukan oleh Hakim *et al* mengenai pengenalan dialek Garut pada bahasa Sunda Selatan dengan memakai metode *recurrent neural network* (RNN) dan ekstraksi ciri yang sama [12].

Namun, penelitian yang membahas tentang dialek Sumatera Selatan belum ditemukan. Padahal, Sumatera Selatan memiliki keanekaragaman dialek yang dituturkan, diantaranya adalah penutur bahasa Melayu Palembang, bahasa Ogan, bahasa Komerling, bahasa Semendo, bahasa Lahat, dan bahasa-bahasa lain di Sumatera Selatan. Terlebih lagi, Indonesia memiliki 718 bahasa daerah berdasarkan data dari Dapobas [13]. Sehingga, pada penelitian ini akan dikembangkan sistem identifikasi dialek yang ada di Sumatera Selatan. Berbeda dengan penelitian-penelitian sebelumnya yang sangat bergantung pada ekstraksi ciri [10]-[12], penelitian ini menggunakan metode *deep neural network* (DNN).

Penelitian berkaitan dengan pengenalan dialek menggunakan DNN telah dilakukan oleh Ruan *et al* untuk pengenalan dialek Lhasa pada bahasa Tibet. Fitur yang diekstrak dengan metode *deep learning* memiliki kemampuan pemodelan yang lebih baik daripada sinyal-sinyal model *hidden markov model* (HMM) fonem tradisional, sehingga performa model pengenalan suara telah meningkat. Penelitian tersebut menunjukkan metode yang lebih baik dalam menggantikan GMM dengan DNN untuk pengenalan dialek bahasa Tibet Lhasa [14]. Sehingga, pada penelitian ini akan digunakan DNN untuk mengenali dialek yang ada di Sumatera Selatan. Selain itu, data suara yang digunakan berasal langsung dari subjek yang berasal dari daerah asal penutur bahasa asli.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang dijelaskan di atas, pengenalan dialek di Indonesia terutama di Sumatera Selatan masih sangat terbatas. Sedangkan pengenalan dialek berpengaruh pada sistem pengenalan suara. Selain itu, pengenalan dialek yang ada masih sangat dipengaruhi pemilihan ekstraksi ciri yang

tepat sebagai *input* pada sistem pengklasifikasi dialek. Sehingga, penelitian ini akan membahas mengenai pengenalan dialek bahasa Indonesia pada penutur asli Sumatera Selatan yang belum pernah diteliti sebelumnya dengan menggunakan metode *deep learning* yang dapat melakukan pembelajaran dengan menggunakan *raw feature* dari suara itu sendiri.

1.3 Tujuan Penulisan

Tujuan penelitian ini adalah untuk mengembangkan sistem pengenalan dialek Sumatera Selatan sebagai salah satu dialek yang ada di Indonesia dengan menggunakan algoritma DNN sehingga bahasa ibu dari penutur dapat diketahui. Selain itu, penelitian ini dilaksanakan untuk mengetahui kinerja dari DNN dalam melakukan pengenalan dialek.

1.4 Pembatasan Masalah

Pada penelitian ini ada beberapa hal yang menjadi batasan masalah diantaranya:

1. Penutur merupakan orang asli/tumbuh besar di Sumatera Selatan.
2. Bahasa/dialek yang akan diuji pada penelitian ini berjumlah enam (dialek Palembang, dialek Muara Enim, dialek Lahat, dialek Sekayu, dialek Musi Rawas dan bahasa Indonesia baku).
3. Bahasa pemrograman yang digunakan adalah Python.

1.5 Keaslian Penelitian

Ada beberapa penelitian yang telah dilakukan berkaitan dengan pengenalan dialek. Rahmawati *et al* dalam penelitiannya membahas tentang pengenalan dialek Jawa dan Sunda menggunakan GMM. Pada penelitian tersebut terdapat dua pengujian, yaitu pengujian pada data uji yang diturunkan dari data yang dipakai untuk membangun model (*closed test*) dan pengujian pada data uji yang tidak dimasukkan ke dalam data yang dipakai untuk membuat model (*open test*). Hasil yang diperoleh dari kedua tes tersebut menyimpulkan bahwa fitur yang terbaik untuk mengenali dialek Jawa dan Sunda pada Bahasa Indonesia adalah kombinasi MFCC dan fitur nada dengan teknik pemodelan I-Vector [10].

Selanjutnya, Ayu Mawaddah Warrohmah meneliti tentang pengenalan dialek Batak, Jawa, dan Minang dengan menggunakan BPNN. Hasil penelitian

menunjukkan bahwa persentase yang diperoleh pada pengenalan dialek Batak terbilang lebih kecil daripada dialek lainnya [11]. Kelemahan pada penelitian ini terletak pada dialek Batak yang dipakai tidak seragam sehingga nilai akurasi pengenalan dialek Batak lebih rendah dari dialek yang lain.

Selain itu, Lutfi Abdul Hakim membahas tentang pengenalan dialek Garut pada bahasa Sunda dengan metode RNN. Penelitian tersebut menyatakan bahwa RNN dapat digunakan untuk pengenalan Bahasa Sunda dialek Garut. Selain itu, persentase kecocokan yang diperoleh dari perbandingan database pada setiap data uji terbilang tinggi, yaitu 100% pada orang Sunda berdialek Garut, 77,8% pada orang Sunda selain dialek Garut, dan 66% pada orang non-Sunda [12].

Kemudian, Darjaa *et al* membahas tentang pengenalan dialek pada bahasa Slovak menggunakan GMM. Hasil penelitian ini menunjukkan bahwa nilai presisi mengindikasikan bahwa nilai keputusan *false positive* dan *false negative* lebih rendah dari pengklasifikasi keputusan acak serta volume informasi dialek yang cukup tinggi untuk pengenalan otomatis [5]. Kelemahan dari penelitian ini diantaranya jumlah data yang kurang untuk semua dialek yang diuji, tidak ada pemisahan jenis kelamin dan usia pada pemodelan data, dan pengambilan data yang menggunakan perangkat yang berbeda-beda.

Namun, ketiga penelitian di atas sangat bergantung pada ekstraksi ciri yang digunakan seperti MFCC [5,10-12]. Sehingga, metode yang lebih sederhana dan dapat mempelajari fitur yang dibawa oleh dialek secara langsung tanpa proses ekstraksi ciri yang membutuhkan komputasi yang lebih kompleks sangat dibutuhkan. Metode yang dapat digunakan untuk mengatasi hal tersebut adalah *deep learning*. Salah satu penelitian pengenalan dialek yang menggunakan *deep learning* dilakukan oleh Qian Zhang dan John H.L dengan *unsupervised deep learning*. Data yang digunakan pada penelitian tersebut berasal dari tiga korpora: korpus dialek bahasa China, korpus dialek bahasa Arab dan korpus *multigenre challenge* (MGB-3) bahasa Arab. Fitur yang diusulkan terbukti mengungguli MFCC fitur akustik tradisional secara konsisten pada tiga korpora [7].

Lalu, pada penelitian W. Ruan mengenai pengenalan suara bahasa Tibet Lhasa yang menggunakan DNN menunjukkan hasil bahwa fitur-fitur yang diekstrak oleh metode *deep learning* memiliki kemampuan yang lebih kuat dalam

pemodelan daripada sinyal model HMM fonem tradisional, sehingga performa model pengenalan suara telah dikembangkan [13]. Pada penelitian pengenalan dialek bahasa Arab Aljazair yang dilakukan oleh Soumia Bougrine, Hadda Cherroun dan Ahmed Abdelali menunjukkan bahwa pemodelan SVM dan DNN sangat baik untuk mengidentifikasi dialek bahasa Arab. Sistem yang digunakan pada penelitian menunjukkan keunggulan kinerjanya ketika ukuran ucapan pada pengujian dibuat lebih pendek. Bila dibandingkan dengan pemodelan SVM, sistem A2DID-DNN memperoleh presisi yang hampir sama. [8].

Selain itu, S. Shon *et al* pada penelitian pengenalan dialek *end-to-end* menunjukkan bahwa sistem end-to-end berbasis fitur akustik mengungguli i-Vector dan *embedding* bahasa dari *siamese neural network*. Penggunaan *convolutional neural network* (CNN) memungkinkan penggunaan spektogram untuk dipakai sebagai fitur akustik apabila dataset pelatihan cukup besar. Penelitian ini juga menunjukkan bahwa sistem identifikasi dialek *end-to-end* dapat ditingkatkan secara signifikan dengan menggunakan segmentasi acak dan perturbasi volume/kecepatan untuk meningkatkan keberagaman dan jumlah data latihan [6].

DAFTAR PUSTAKA

- [1] Collins Dictionary, "Language." [Online]. Available: <https://www.collinsdictionary.com/dictionary/english/language/>. [Accessed: Jul. 14, 2020].
- [2] Lexico, "Dialect." [Online]. Available: <https://www.lexico.com/definition/dialect/>. [Accessed: 14-Jul-2020].
- [3] JTA Technology Consulting. "English Structures: Sociolinguistics" (pg. 2) [Online]. Available: <http://web.mnstate.edu/houtsli/tesl551/Socio/page2.htm>. [Accessed: Jan 6, 2021]
- [4] A. Etman and A. A. L. Beex, "Language and Dialect Identification: A survey," *2015 SAI Intelligent Systems Conference (IntelliSys)*, London, UK, 2015, pp. 220-231.
- [5] S. Darjaa, R. Sabo, M. Trnka, M. Rusko, and G. Múcskova, "Automatic recognition of slovak regional dialects," *DISA 2018 - IEEE World Symp. Digit. Intell. Syst. Mach. Proc.*, pp. 305–308, 2018.
- [6] S. Shon et al, *Convolution Neural Network and Language Embedding for End-to-End Dialect Recognition*. Massachusetts Institute of Technology. 2018.
- [7] Q. Zhang and J. H. L. Hansen, *Language/Dialect Recognition Based on Unsupervised Deep Learning*. Northwest Normal University, Lanzhou, China. 2017.
- [8] S. Bougrine et al, *Spoken Arabic Algerian Dialect Identification*. Universite Amar Telidji Laghoua. 2018.
- [9] M. P. Lewis, *Ethnologue: Languages of the World* (sixteenth ed.), SIL International. 2009.

- [10] R. Rahmawati and D. P. Lestari, "Java and Sunda dialect recognition from Indonesian speech using GMM and I-Vector," *Proceeding 2017 11th Int. Conf. Telecommun. Syst. Serv. Appl. TSSA 2017*, vol. 2018-Janua, pp. 1–5, 2018.
- [11] A. Mawadda Warohma, P. Kurniasari, S. Dwijayanti, Irmawan and B. Yudho Suprpto, "Identification of Regional Dialects Using Mel Frequency Cepstral Coefficients (MFCCs) and Neural Network," *2018 International Seminar on Application for Technology of Information and Communication*, Semarang, Indonesia, 2018, pp. 522-527.
- [12] L. A. Hakim, A. B. Osmond, and R. E. Saputra, "Recurrent Neural Network Untuk Pengenalan Ucapan Pada Recurrent Neural Network for Speech Recognition", 2018.
- [13] Kementerian Pendidikan dan Kebudayaan, "Bahasa Daerah di Indonesia", *Kementerian Pendidikan dan Kebudayaan*. [Online]. Available: <https://dapobas.kemdikbud.go.id/homecat.php?show=url/petabahasa&cat=6/>. [Accessed: Jul. 14, 2020]
- [14] W. Ruan, Z. Gan, B. Liu, and Y. Guo, "An Improved Tibetan Lhasa Speech Recognition Method Based on Deep Neural Network," *Proc. - 10th Int. Conf. Intell. Comput. Technol. Autom. ICICTA 2017*, vol. 2017-Octob, pp. 303–306, 2017.
- [15] P. Khunarsal, C. Lursinsap, and T. Raicharoen, "Very short time environmental sound classification based on spectrogram pattern matching" *Inf. Sci. (Ny)*., vol. 243, pp. 57–74, 2013.
- [16] Q. Zhou et al., "Cough Recognition Based on Mel-Spectrogram and Convolutional Neural Network," *Front. Robot. AI*, vol. 8, no. May, pp. 1–7, 2021, doi: 10.3389/frobt.2021.580080.
- [17] M. A. Hossan, S. Memon, and M. A. Gregory, "A novel approach for MFCC feature extraction," *4th Int. Conf. Signal Process. Commun. Syst. ICSPCS'2010 - Proc.*, 2010.

- [18] D. O'Shaughnessy: *Speech communication: human and machine*, pp. 150, Addison-Wesley, 1987.
- [19] Dong Yu, Li Deng, *Automatic Speech Recognition: A Deep Learning Approach*. 2020.
- [20] LeCun, Y., Bottou, L., Orr, G.B., Müller, K.R.: *Efficient backprop. Neural Networks: Tricks of The Trade*, pp. 9–50. Berlin: Springer, 1998.
- [21] Y. A. Ibrahim, J. C. Odiketa, and T. S. Ibiyemi, “Preprocessing technique in automatic speech recognition for human computer interaction: an overview,” *Ann. Comput. Sci. Ser.*, vol. 15, no. 1, pp. 186–191, 2017.
- [22] L. Lamel, L. Rabiner, A. Rosenberg, J. Wilpon, *An Improved Endpoint Detector for Isolated Word Recognition*, *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 29, pp. 777– 785, 1981.
- [23] timsainb, T. Sainburg, “noisereducer”. [Online]. Available: <https://github.com/timsainb/noisereducer>. [Accessed Sept. 26, 2021].