



**PENERAPAN ALGORITMA CONVOLUTIONAL NEURAL  
NETWORK DENGAN PENGGUNAAN FEATURES  
EXTRACTOR MFCC UNTUK KLASIFIKASI BAHASA  
DAERAH INDONESIA (BALI, BATAK, JAWA, MADURA,  
MELAYU, MINANG, dan SUNDA)**

**LAPORAN SKRIPSI**

**UNIVERSITAS  
BRAMA TRI HADI SYAPOETRA  
MERCU BUANA  
41519010001**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA JAKARTA  
2023**



**PENERAPAN ALGORITMA CONVOLUTIONAL NEURAL  
NETWORK DENGAN PENGGUNAAN FEATURES  
EXTRACTOR MFCC UNTUK KLASIFIKASI BAHASA  
DAERAH INDONESIA (BALI, BATAK, JAWA, MADURA,  
MELAYU, MINANG, dan SUNDA)**

**LAPORAN SKRIPSI**

**BRAMA TRI HADI SYAPOETRA  
41519010001**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA JAKARTA  
2023**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Brama Tri Hadi Syapoetra  
NIM : 41519010001  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma Convolutional Neural Network Dalam Klasifikasi Bahasa Daerah Indonesia (Bali, Batak, Jawa, Madura, Melayu, Minang, dan Sunda)

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

UNIVERSITAS  
MERCU BUANA

Jakarta, 14 Juni 2023



Brama Tri Hadi Syapoetra

## HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Brama Tri Hadi Syapoetra  
NIM : 41519010001  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma Convolutional Neural Network Dengan Penggunaan Features Extractor MFCC Untuk Klasifikasi Bahasa Daerah Indonesia (Bali, Batak, Jawa, Madura, Melayu, Minang, dan Sunda).

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Drs. Achmad Kodar, M.T.   
NIDN : 0323085801  
Ketua Penguji : Dr. Hadi Santoso, S.Kom., M.Kom (  )  
NIDN : 0225067701  
Penguji 1 : Mohamad Yusuf, S.Kom., M.C.S (  )  
NIDN : 0307097606

Jakarta, 7 Agustus 2023

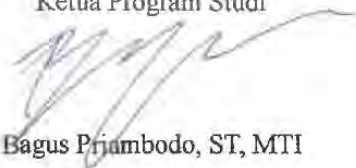
Mengetahui,

Dekan



Dr. Bambang Jokonowo, S.Si., M.T.I

Ketua Program Studi



Bagus Priambodo, ST, MTI

## KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan Tugas Akhir yang berjudul Penerapan Algoritma Convolutional Neural Network Dengan Penggunaan Features Extractor MFCC Untuk Klasifikasi Bahasa Daerah Indonesia (Bali, Batak, Jawa, Madura, Melayu, Minang, dan Sunda). Penulisan Tugas Akhir ini dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana.

Selama penulisan Tugas Akhir ini, penulis menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir ini, sangatlah sulit bagi penulis untuk menyelesaikan Tugas Akhir ini. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M. Eng selaku Rektor Universitas Mercu Buana
2. Bapak Dr. Bambang Jokonowo, S.Si,MTI selaku Dekan Fakultas Ilmu Komputer
3. Bapak Bagus Priambodo, ST, MTI selaku Ketua Program Studi Teknik Informatika
4. Bapak Drs. Achmad Kodar, M.T. selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan penulis dalam penyusunan tugas akhir ini.
5. Orang tua, keluarga dan rekan kuliah atas dukungan, doa, dan semangat yang diberikan sepanjang perjalanan penulisan Skripsi ini.
6. Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.

Akhir kata, penulis berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tugas Akhir ini dapat membawa manfaat bagi pengembangan ilmu.

Jakarta, 14 Juni 2023



Brama Tri Hadi Syapoetra

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA  
ILMIAH**

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama	: Brama Tri Hadi Syapoetra
NIM	: 41519010001
Program Studi	: Teknik Informatika
Judul Laporan Skripsi	: Penerapan Algoritma Convolutional Neural Network Dalam Klasifikasi Bahasa Daerah Indonesia (Bali, Batak, Jawa, Madura, Melayu, Minang, dan Sunda)

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalty-Free Right)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 14 Juni 2023

Yang menyatakan,



Brama Tri Hadi Syapoetra

## ABSTRAK

Nama	:	Brama Tri Hadi Syapoetra
NIM	:	41519010001
Program Studi	:	Teknik Informatika
Judul Laporan Skripsi	:	Penerapan Algoritma Convolutional Neural Network Dengan Penggunaan Features Extractor MFCC Untuk Klasifikasi Bahasa Daerah Indonesia (Bali, Batak, Jawa, Madura, Melayu, Minang, dan Sunda).
Pembimbing	:	Drs. Achmad Kodar, M.T.

Asisten suara telah menjadi fitur yang penting dalam produk unggulan perusahaan teknologi besar seperti Google, Apple, Amazon, dan Microsoft. Namun, salah satu tantangan utama yang dihadapi dalam pengembangan asisten suara adalah pengenalan bahasa yang digunakan oleh pengguna. Untuk meningkatkan pengalaman pengguna dan mendukung tugas seperti deteksi ucapan otomatis atau transkripsi ucapan ke teks, klasifikasi bahasa otomatis menjadi langkah penting yang diperlukan. Di Indonesia, dengan lebih dari 700 bahasa daerah yang beragam, sulit bagi orang luar dan orang asing untuk membedakan bahasa yang digunakan. Oleh karena itu, dalam penelitian ini, dilakukan pengembangan metode menggunakan algoritma *Convolutional Neural Network* (CNN) untuk mengklasifikasikan bahasa daerah Indonesia berdasarkan rekaman suara dalam format .wav. Bahasa daerah yang dipelajari dalam penelitian ini meliputi Bahasa Bali, Bahasa Batak, Bahasa Jawa, Bahasa Madura, Bahasa Melayu, Bahasa Minang, dan Bahasa Sunda. Penelitian ini bertujuan untuk menghasilkan informasi yang dapat digunakan untuk mencapai akurasi klasifikasi yang baik dalam pengenalan bahasa daerah Indonesia dengan menggunakan metode *Deep Learning* dan algoritma *Convolutional Neural Network* (CNN).

**Kata Kunci** : Algoritma *Convolutional Neural Network*, Bahasa Daerah, *Deep Learning*, Klasifikasi Bahasa Daerah, Audio atau Suara

## ABSTRACT

Name	:	Brama Tri Hadi Syapoetra
NIM	:	41519010001
Study Program	:	Teknik Informatika
Title Thesis	:	Penerapan Algoritma Convolutional Neural Network Dengan Penggunaan Features Extractor MFCC Untuk Klasifikasi Bahasa Daerah Indonesia (Bali, Batak, Jawa, Madura, Melayu, Minang, dan Sunda).
Counsellor	:	Drs. Achmad Kodar, M.T.

Voice assistants have become crucial features in flagship products of major technology companies such as Google, Apple, Amazon, and Microsoft. However, one of the main challenges in voice assistant development is the recognition of the user's language. To enhance user experience and support tasks like automated speech detection or speech-to-text transcription, automatic language classification is an important necessary step. In Indonesia, with over 700 diverse regional languages, it is difficult for outsiders and foreigners to differentiate the language being used. Therefore, in this research, a method is developed using the Convolutional Neural Network (CNN) algorithm to classify Indonesian regional languages based on audio recordings in .wav format. The regional languages studied in this research include Balinese, Batak, Javanese, Madurese, Malay, Minang, and Sundanese. This study aims to provide information that can contribute to achieving accurate language classification in recognizing Indonesian regional languages using Deep Learning methods and the Convolutional Neural Network (CNN) algorithm.

**Keywords:** *Convolutional Neural Network* (CNN) algorithm, regional languages, Deep Learning, regional language classification, audio or sound.



## DAFTAR ISI

<b>HALAMAN COVER</b> .....	i
<b>HALAMAN PERNYATAAN KARYA SENDIRI</b> .....	iii
<b>HALAMAN PENGESAHAN</b> .....	iv
<b>KATA PENGANTAR</b> .....	v
<b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH</b> .....	vi
<b>ABSTRAK</b> .....	vii
<b>ABSTRACT</b> .....	viii
<b>DAFTAR ISI</b> .....	ix
<b>DAFTAR TABEL</b> .....	xi
<b>DAFTAR GAMBAR</b> .....	xii
<b>DAFTAR LAMPIRAN</b> .....	xiii
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	2
1.3 Tujuan Penelitian.....	2
1.4 Manfaat Penelitian.....	2
1.5 Batasan Penelitian.....	2
<b>BAB II TINJAUAN PUSTAKA</b> .....	4
2.1 Penelitian Terdahulu.....	4
2.2 Teori Pendukung.....	17
<b>BAB III METODE PENELITIAN</b> .....	22
3.1 Jenis Penelitian.....	22
3.2 Tahapan Penelitian.....	22
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	25
4.1 <i>Dataset</i> .....	25
4.2 <i>Pre-processing</i> .....	26
4.3 Pembuatan Model.....	30
4.4 Visualisasi Data.....	31
4.5 Pengujian.....	34
4.6 Analisis Hasil.....	37

<b>BAB V KESIMPULAN DAN SARAN</b> .....	46
5.1 Kesimpulan.....	46
5.2 Saran.....	46
<b>DAFTAR PUSTAKA</b> .....	48
<b>LAMPIRAN</b> .....	50



## DAFTAR TABEL

Table 1. Penelitian Terkait.....	4
Table 2. Hasil Akurasi Tanpa Menggunakan Split Validation .....	40
Table 3. Hasil Akurasi, Precision, Recall dan F1-score tanpa menggunakan split validation.....	44
Table 4. Hasil, Akurasi, Precision, Recall dan F1-score menggunakan split validation.....	45



## DAFTAR GAMBAR

Gambar 1. Contoh Proses Convolutional Neural Network.....	18
Gambar 2. Tahapan Penelitian .....	22
Gambar 3. Subfolder dari folder Dataset .....	25
Gambar 4. Data Audio Dari Subfolder .....	26
Gambar 5. Visualisasi Gelombang Suara Dari Bahasa Bali .....	33
Gambar 6. Visualisasi Gelombang Suara Dari Bahasa Batak.....	33
Gambar 7. Visualisasi Gelombang Suara Dari Bahasa Jawa .....	33
Gambar 8. Visualisasi Gelombang Suara Dari Bahasa Madura.....	33
Gambar 9. Visualisasi Gelombang Suara Dari Bahasa Melayu.....	34
Gambar 10. Visualisasi Gelombang Suara Dari Bahasa Minang.....	34
Gambar 11. Visualisasi Gelombang Suara Dari Bahasa Sunda .....	34
Gambar 12. Nilai Loss Model.....	38
Gambar 13. Confusion Matrix Tanpa Split Validation.....	40
Gambar 14. Nilai Loss Model Menggunakan Split Validation.....	41
Gambar 15. Confusion Matrix Menggunakan Split Validation .....	44



## DAFTAR LAMPIRAN

Lampiran 1. Lampiran Bimbingan *didapat dari SIA .....	50
Lampiran 2. Halaman Persetujuan .....	51
Lampiran 3. Luaran Tugas Akhir .....	52
Lampiran 4. Bukti Submit / Published Artikel Ilmiah / HKI.....	53
Lampiran 5. Naskah Artikel Jurnal .....	54
Lampiran 6. Curriculum Vitae.....	72
Lampiran 7. Plagiarisme Check .....	74
Lampiran 8. Surat Pernyataan HKI.....	81
Lampiran 9. Sertifikat BNSP/Bukti Daftar .....	82

