

SKRIPSI

**PROTOTYPE SISTEM KEAMANAN AKSES MASUK RUANGAN BERBASIS
VOICE RECOGNITION MENGGUNAKAN ALGORITMA DEEP LEARNING**



**Disusun untuk Memenuhi Syarat Mendapatkan Gelar Sarjana Teknik Pada
Jurusan Teknik Elektro Fakultas Teknik
Universitas Sriwijaya**

Oleh :

ADJI SULTHONI

03041181924008

JURUSAN TEKNIK ELEKTRO

FAKULTAS TEKNIK

UNIVERSITAS SRIWIJAYA

2023

LEMBAR PENGESAHAN

**PROTOTIPE SISTEM KEAMANAN AKSES MASUK RUANGAN BERBASIS
VOICE RECOGNITION MENGGUNAKAN ALGORITMA DEEP LEARNING**



**Disusun Untuk Memenuhi Syarat mendapatkan Gelar Sarjana Teknik
Pada Jurusan Teknik Elektro Fakultas Teknik
Univeristas Sriwijaya**

Oleh:

ADJI SULTHONI

03041181924008

Palembang, 24 Juli 2023

**Menyetujui,
Pembimbing Utama**

**Hera Hikmarika, S.T., M.Eng.
NIP : 197812072002122002**

Mengetahui

Ketua Jurusan Teknik Elektro



**Muhammad Abu Bakar Sidik, S.T., M.Eng., Ph.D., IPU.
NIP : 197108141999031005**

HALAMAN PERNYATAAN INTEGRITAS

Yang bertanda tangan di bawah ini:

Nama : Adji Sulthoni
NIM : 03041181924008
Fakultas : Teknik
Jurusan/Prodi : Teknik Elektro
Universitas : Universitas Sriwijaya

Hasil Pengecekan *Software iThenticate/Turnitin* : 12 %

Menyatakan bahwa tugas akhir saya yang berjudul “**Prototipe Sistem Keamanan Akses Masuk Ruang Berbasis *Voice Recognition* Menggunakan Algoritma *Deep Learning***”. 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, 24 Juli 2023



Adji Sulthoni
NIM. 03041181924008

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 : Hera Hikmarika, S.T., M.Eng.

Tanggal : 24 / Juli / 2023

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

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

Nama : Adji Sulthoni
NIM : 03041181924008
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:

**PROTOTIPE SISTEM KEAMANAN AKSES MASUK RUANGAN BERBASIS
VOICE RECOGNITION MENGGUNAKAN ALGORITMA DEEP LEARNING**

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 : 24 Juli 2023
Yang menyatakan,



Adji Sulthoni
NIM. 03041181924008

KATA PENGANTAR

Puji dan syukur penulis ucapkan kehadiran Allah SWT. serta shalawat dan salam penulis haturkan kepada Nabi Muhammad SAW beserta keluarga dan para sahabat. Berkat rahmat dan ridho Allah SWT. penulis dapat menyelesaikan skripsi yang berjudul “Prototipe Sistem Keamanan Akses Masuk Ruang Berbasis *Voice Recognition* Menggunakan Algoritma *Deep Learning*”.

Pembuatan skripsi ini disusun untuk memenuhi syarat mendapatkan gelar Sarjana Teknik pada Jurusan Teknik Elektro Fakultas Teknik Universitas Sriwijaya. Dalam kesempatan ini penulis mengucapkan terimakasih kepada:

1. Orang tua, saudara, dan keluarga yang selalu memberikan semangat dan memberi dukungan baik secara mental, fisik, maupun finansial.
2. Ibu Hera Hikmarika, S.T., M.Eng. selaku pembimbing utama tugas akhir ini yang telah memberikan bimbingan, memberikan ilmu dan masukan selama proses penulisan skripsi.
3. Ibu Dr. Eng. Ir. Suci Dwijayanti, S.T., M.S., IPM. yang selalu memberikan arahan secara teknis dalam proses menyelesaikan Prototipe Alat
4. Bapak Dr. Ir. Bhakti Yudho Suprpto, S.T., M.T. IPM dan Dr. Eng. Ir. Suci Dwijayanti, S.T., M.S., IPM selaku pencetus dan memberikan bimbingan pada tugas akhir ini serta pengembang ide.
5. Dosen pembimbing akademik, bapak Muhammad Abu Bakar Sidik, S.T., M.Eng., Ph.D., IPU. yang telah memberikan arahan serta bimbingan kepada penulis selama masa perkuliahan.
6. Bapak Muhammad Abu Bakar Sidik, S.T., M.Eng., Ph.D selaku Ketua Jurusan Teknik Elektro Universitas Sriwijaya dan Ibu Dr. Eng. Ir. Suci Dwijayanti, S.T., M.S., IPM selaku Sekretaris Jurusan Teknik Elektro Universitas Sriwijaya.
7. Segenap Dosen Jurusan Teknik Elektro Fakultas Teknik Universitas Sriwijaya yang telah memberikan ilmu selama perkuliahan.
8. Saudara Bima Pamungkas, Muhammad Yulwi Alwan, Immanuel Morries Pohan selaku rekan kerja yang selalu bersemangat dalam pembuatan *prototype* tugas akhir ini.

9. Teman – teman pejuang USEPT (Dani, Vidi, Ricky, Fanhar, Rian, Adam, Sandy, Fadil, Uyun, Zainal dan Juan) yang senantiasa membantu selama persiapan tes USEPT serta proses perkuliahan.
10. Teman - teman TKR 2019 yang selalu memberikan *support* dan dukungan dalam menyelesaikan skripsi.
11. Teman - teman Teknik Elektro 2019 yang selalu membantu, dan menyemangati dalam proses pembuatan skripsi.
12. Dan pihak-pihak yang sangat membantu di dalam penyusunan skripsi ini, yang tidak dapat penyusun sebutkan satu persatu.
13. *Last but not least, I want thank and appreciate myself for doing all this hard work with all suffer and pain, came back after days off, never stopping, and always being on point.*

Didalam penyusunan skripsi ini, masih terdapat kekurangan karena keterbatasan penulis, oleh karena itu penulis mengharapkan kritik dan saran yang membangun agar dapat menjadi evaluasi dan berguna untuk penulis dimasa yang akan datang.

Palembang, 24 Juli 2023



Adji Sulthoni

NIM. 03041181924008

ABSTRAK

PROTOTIPE SISTEM KEAMANAN AKSES MASUK RUANGAN BERBASIS VOICE RECOGNITION MENGGUNAKAN ALGORITMA DEEP LEARNING

(Adji Sulthoni, 03041181924008, 2023, 52 Halaman)

Sistem keamanan yang sering digunakan hingga sekarang adalah sistem keamanan konvensional. Sistem keamanan ini memiliki kelemahan yang memungkinkan untuk ditiru dan dibobol. Oleh karena itu, dibutuhkan sistem keamanan yang menggunakan biometrik. Suara manusia merupakan salah satu bentuk biometrik yang dapat digunakan untuk mengidentifikasi orang (*person identification*) yang dapat dimanfaatkan dalam bidang keamanan. Adapun tujuan dilakukannya penelitian ini adalah mengimplementasikan penggunaan suara sebagai sistem keamanan akses ruangan dengan menggunakan algoritma *deep learning* pada proses pembelajaran untuk mengenali suara secara *real-time*. Pada penelitian ini ada beberapa tahapan yang dilakukan yaitu pengambilan data suara, pengolahan data suara, pelatihan data suara menggunakan 3 arsitektur CNN, Pengujian secara *offline* menggunakan data testing dan pengujian *online* menggunakan microphone secara *real-time*. Pengolahan data suara yang digunakan yaitu *Short-Time Fourier Transform* (STFT) untuk proses ekstraksi ciri dalam bentuk spektogram. Kemudian, spektogram akan diolah oleh *Convolutional Neural Network* (CNN) sebagai metode pengidentifikasi. Penelitian ini menggunakan 3 Tipe Arsitektur CNN, yaitu arsitektur VGG16, AlexNet, dan Arsitektur Model Sendiri dan didapatkan hasil akurasi terbaik dari 3 arsitektur tersebut dengan epoch 100 sebesar 0.9993, 0.9915, dan 1. Pengujian dilakukan secara *offline* dan *online*, pada pengujian *offline* dilakukan dengan menggunakan data uji dan menggunakan 3 arsitektur tersebut, didapatkan hasil pengujian terbaik dengan menggunakan epoch 100 sebesar 61.76%, 94.11% dan 95,09%. Selanjutnya, pada pengujian *online* dilakukan secara *real time* pada prototipe di Laboratorium Dasar Sistem Kendali dan Robotika dengan menggunakan *microphone* terhadap 15 responden yang terdiri dari 8 responden yang berada dalam *dataset* dan 7 yang tidak ada dalam *dataset*. Hasil pengujian secara *real time* didapatkan nilai akurasi rata-rata pengenalan suara sebesar 80%. Hasil menunjukkan bahwa model dari arsitektur sendiri memiliki kinerja yang lebih baik untuk mengenali suara, dan dapat diimplementasikan pada sistem keamanan akses masuk ruangan.

Kata Kunci: *Security System, Convolutional Neural Network (CNN), Deep Learning, Voice Recognition, Biometric, Spectogram.*

ABSTRACT

PROTOTYPE OF VOICE RECOGNITION-BASED ROOM ACCESS SECURITY SYSTEM USING DEEP LEARNING ALGORITHM

(Adji Sulthoni, 03041181924008, 2023, 52 Halaman)

The security system that is often used today is the conventional security system. This security system has weaknesses that allow it to be imitated and broken into. Therefore, a security system that uses biometrics is needed. The human voice is one form of biometrics that can be used to identify people (person identification) which can be utilized in the security field. The purpose of this research is to implement the use of voice as a room access security system by using deep learning algorithms in the learning process to recognize voices in real-time. In this research, there are several stages carried out, namely taking voice data, processing voice data, training voice data using 3 CNN architectures, offline testing using testing data and online testing using microphones in real time. The voice data processing used is Short-Time Fourier Transform (STFT) for the feature extraction process in the form of spectrograms. Then, the spectrogram will be processed by Convolutional Neural Network (CNN) as an identifier method. This research uses 3 types of CNN architecture, namely VGG16 architecture, AlexNet, and Own Model Architecture and obtained the best accuracy results from the 3 architectures with epoch 100 of 0.9993, 0.9915, and 1. Testing is done offline and online, offline testing is done using test data and using 3 architectures, obtained the best test results using epoch 100 of 61.76%, 94.11% and 95.09%. Furthermore, online testing is carried out in real time on a prototype in the Basic Laboratory of Control Systems and Robotics using a microphone against 15 respondents consisting of 8 respondents who are in the dataset and 7 who are not in the dataset. The real time test results obtained an average voice recognition accuracy value of 80%. The results show that the model of its own architecture has better performance for recognizing voices, and can be implemented in the room access security system.

Keywords: Security System, Convolutional Neural Network (CNN), Deep Learning, Voice Recognition, Biometric, Spectrogram.

DAFTAR ISI

LEMBAR PENGESAHAN	ii
HALAMAN PERNYATAAN INTEGRITAS.....	iii
HALAMAN PERNYATAAN DOSEN	iv
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	iv
KATA PENGANTAR.....	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	4
1.4 Batasan Masalah.....	4
1.5 Keaslian Penelitian	4
BAB II TINJAUAN PUSTAKA.....	7
2.1 State Of The Art	7
2.2 Security system.....	12
2.3 Pengenalan suara	12
2.4 Proses pre-processing	14
2.5 Short-Time Fourier Transform (STFT)	14
2.6 Convolutional Neural Network (CNN)	16
2.6.1 Convolution Layer	17
2.6.2 Pooling Layer	18
2.6.3 Rectified Linear Unit (ReLU)	18
2.6.4 Flattening Layer	19
2.6.5 Softmax	19
BAB III METODOLOGI PENELITIAN	20
3.1 Studi Literatur.....	20

3.2	Perancangan Sistem.....	21
3.3	Perancangan Sistem Pengenalan	22
3.4	Perancangan Prototype Security System Akses Ruangan.....	23
3.4.1	Microphone	24
3.4.2	Laptop Lenovo Ideapad Gaming 3.....	24
3.4.3	Board Arduino UNO R3 ATmega328 ATmega16U2	25
3.4.4	Solenoid.....	26
3.5	Pengambilan Data.....	26
3.6	Pelatihan Data dan Pengujian	26
3.7	Evaluasi	27
3.7.1	Pengujian Sistem Pengenalan Suara	27
3.7.2	Pengujian Sistem Keamanan	28
BAB IV	HASIL DAN PEMBAHASAN	29
4.1	Pengumpulan Data Latih dan Uji	29
4.2	Pengolahan Data Latih dan Uji.....	30
4.3	Pelatihan 3 Arsitektur CNN.....	31
4.4	Hasil Pelatihan 3 Arsitektur CNN.....	35
4.5	Pengujian Secara Offline	37
4.6	Prototipe Box Sistem Keamanan Biometrik	39
4.7	Pengujian Online Menggunakan microphone secara Real-Time	41
4.7.1	Pengujian dengan Machine noise.....	45
4.7.2	Pengujian dengan Bubble noise	46
4.8	Analisa Hasil Pengujian Sistem Pengenalan Suara	47
BAB V	KESIMPULAN DAN SARAN	49
5.1	Kesimpulan.....	49
5.2	Saran	49
	DAFTAR PUSTAKA	50
	LAMPIRAN.....	53
	LAMPIRAN KHUSUS	90

DAFTAR GAMBAR

Gambar 2.1 Perbandingan antara metode konvensional MFCC dan metode MFCC+PCA	7
Gambar 2.2 pre-emphasis sebelum dan sesudah[4].	8
Gambar 2.3 Spectrogram (diatas) dan Full-Mel Spectrogram (dibawah)[5].....	9
Gambar 2.4 Plot t-SNE berdasarkan vektor keluaran dari softmax lapisan L8 (kiri) dan lapisan padat pertama L5 (kanan)[9].	10
Gambar 2.5 perbandingan hasil pelatihan (a) arsitektur sederhana CNN, (b) arsitektur CNN model VGG-f[6].	11
Gambar 2.6 Sinyal Suara.....	13
Gambar 2.7 Image Spektrogram.....	16
Gambar 2.8 Contoh Jaringan CNN.	17
Gambar 2.9 Arsitektur Convolution Neural Network.....	17
Gambar 2.10 Max and Average Pooling.....	18
Gambar 2.11 Fungsi ReLU	18
Gambar 3.1 Flowchart Penelitian.....	20
Gambar 3.2 Diagram Alir Sistem	21
Gambar 3.3 Flowchart Sistem Akses Masuk Ruangan.....	22
Gambar 3.4 Perancangan Prototype Security System Akses Ruangan.....	23
Gambar 3.5 Perangkat untuk merekam suara.	24
Gambar 3.6 Mini PC.	24
Gambar 3.7 Arduino UNO RE Atmega328 Atmega16U2.	25
Gambar 3.8 Selenoid.....	26
Gambar 4.1 Sinyal Suara dengan pengucapan angka.	29
Gambar 4.2 Sinyal Suara sebelum preprocess signal.....	30
Gambar 4.3 Sinyal Suara sebelum preprocess signal.....	30
Gambar 4.4 Hasil Ekstraksi Ciri dengan metode Short-Time Fourier Transform (STFT)	31
Gambar 4.5 Training Accuracy 50 Epoch Pengenalan Suara	36
Gambar 4.6 Training Accuracy 100 Epoch Pengenalan suara.....	36
Gambar 4.7 Prototipe Sistem Keamanan Voice Recognition	40

Gambar 4.8 wiring Diagram Prototipe Sistem Keamanan voice Recognition.....	40
Gambar 4.9 Posisi Pada Saat Pengujian.....	41

DAFTAR TABEL

Tabel 2.1 Hasil percobaan pengelompokan.	10
Tabel 3.1 Arduino UNO R3 Atmega328 Atmega16U2.....	25
Tabel 3.2 Confussion Matrix	27
Tabel 4.1 Arsitektur VGG16 pada proses pelatihan	32
Tabel 4.2 Arsitektur AlexNet pada proses pelatihan.	33
Tabel 4.3 Arsitektur Modifikasi Vgg16 pada proses pelatihan.....	34
Tabel 4.4 Perbandingan Training Accuracy 3 Arsitektur Pada Pengenalan Suara	37
Tabel 4.5 Pengujian menggunakan data uji sample spectrogram	38
Tabel 4.6 Pengujian secara online.....	42
Tabel 4.7 Pengujian dengan machine noise	45
Tabel 4.8 Pengujian dengan Bubble Noise	46

BAB I PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi yang pesat memiliki efek yang menguntungkan dan merugikan bagi keberadaan manusia. Masalah sistem keamanan merupakan salah satu efek negatif yang menjadi perhatian. Salah satu sistem keamanan yang sering digunakan hingga sekarang adalah sistem keamanan konvensional, yang mengandalkan prinsip mekanik[1]. Sistem keamanan ini memiliki kelemahan, tingkat keamanan yang rendah memungkinkan pintu dapat dibobol dengan mudah untuk melakukan kejahatan seperti perusakan dan pencurian. Karena itulah berbagai macam pengembangan dilakukan untuk memberikan keamanan dengan memanfaatkan kemajuan teknologi yang dapat melindungi aset yang dimiliki, sehingga diharapkan dengan pengaplikasian sistem keamanan dapat memberikan rasa aman dan nyaman[2].

Selain hal tersebut tentunya dibutuhkan Sistem keamanan yang hanya memungkinkan seseorang diidentifikasi berdasarkan karakteristik yang dimilikinya merupakan salah satu solusi untuk masalah ini. Ada beberapa contoh sistem keamanan berdasarkan identifikasi seseorang, antara lain: *Pin Lock*, *Password Lock* dan *RFID*. Namun, dari beberapa metode tersebut diketahui bahwa seseorang yang tidak berkepentingan dapat mudah meniru atau mencuri dan itu merupakan kelemahan dari metode tersebut. Memanfaatkan karakteristik alami manusia sebagai input ke dalam sistem keamanan adalah salah satu pendekatan untuk mengatasi masalah ini. Pengenalan identitas seseorang menggunakan ciri khas dari karakteristik alami manusia sebagai dasar disebut sebagai biometrik. Biometrik pada dasarnya merupakan pengenalan seseorang berdasarkan kepribadian unik yang dimiliki setiap orang seperti pengenalan wajah, sidik jari, pengenalan suara, pengenalan retina, dan tanda tangan[3]. Identifikasi dan verifikasi adalah dua tujuan penting yang dipenuhi oleh pengenalan biometrik. Identitas seseorang harus dipastikan melalui sistem identifikasi. Sementara itu, sistem verifikasi bertujuan untuk menerima atau menolak klaim identitas seseorang. Teknologi biometrik beroperasi menggunakan Teknik *pattern recognition* atau Teknik pengenalan pola.

Banyak pola, seperti sidik jari, garis telapak tangan, wajah, iris mata, dan pengenalan suara, yang dapat diidentifikasi oleh teknologi ini.

Salah satu biometrik yang dapat digunakan untuk mengidentifikasi seseorang adalah suara manusia. *Speech Recognition* (SR) memiliki sejumlah keunggulan dibandingkan metode pengenalan biometrik lainnya, termasuk fakta bahwa metode ini tidak memerlukan perangkat keras yang mahal dan memiliki ukuran file yang kecil. Banyak metode yang telah dilakukan untuk mendapatkan pengenalan suara yang akurat. Penelitian yang dilakukan Hariz Zakka Muhammad, membahas pengenalan suara dengan menggunakan metode HMM[4]. Pada penelitiannya penulis menggunakan *Mel-Frequency Cepstrum Coefficients* (MFCC) untuk proses ekstraksi fitur. Pada penelitian ini menjelaskan sebuah sistem yang dapat menerjemahkan dari bahasa inggris ke bahasa indonesia berdasarkan pengenalan ucapan menggunakan *mel frequency cepstral coefficients* (MFCC) untuk ekstraksi fitur dan metode hidden markov model untuk klasifikasi. Untuk pengucapan kalimat yang terdiri dari 2 kata atau 3 kata masih kurang baik dalam memproses kalimat. Hal ini dikarenakan dalam klasifikasi HMM, semakin banyak kata yang diproses maka model perhitungan probabilitas akan terbagi. cara ini tidak efisien karena, dalam kinerja komputasi, semakin banyak kata yang diucapkan juga mempengaruhi lamanya waktu komputasi untuk memproses sistem.

Pada penelitian yang berkaitan dengan Pemilihan fitur ekstraksi yang efektif adalah salah satu aspek kunci dalam pengenalan suara karena fitur tersebut berdampak pada akurasi pengenalan. Beberapa peneliti telah mempelajari ekstraksi fitur secara mendalam. Salah satu dari penelitian ini dilakukan oleh Ruchao Fan, yang membahas perbandingan fitur input. Dalam penelitiannya, dia mengusulkan penggunaan fitur *Full-Mel Spectrogram* (FBANK) sebagai metode ekstraksi fitur. Alasan pemilihan *FBANK* adalah karena fitur ini memiliki struktur spektrum yang lebih jelas pada frekuensi rendah dan kurang terlihat pada frekuensi tinggi, yang sesuai dengan cara persepsi nonlinier telinga manusia[5]. Meskipun demikian, penelitian ini memiliki kelemahan dalam hal kompleksitas proses komputasi. Untuk mengatasi kelemahan fitur FBANK, penulis mencoba mengimplementasikan algoritma *Short-Time Fourier Transform* (STFT). STFT memiliki keunggulan dibandingkan *Fast Fourier Transform* (FFT) karena dapat memberikan informasi

frekuensi yang akurat dan juga informasi waktu dari suatu sinyal. Dengan menggabungkan kedua konsep ini, STFT menawarkan data waktu dan frekuensi yang tepat. Oleh karena itu, penulis memutuskan untuk menggunakan fitur mentah berupa spektrogram sebagai metode ekstraksi fitur pada pengenalan ucapan (Speech Recognition).

Pada penelitian ini, metode pengenalan yang digunakan masih bersifat konvensional seperti HMM (Hidden Markov Model). Namun, metode ini memiliki kelemahan terutama ketika menghadapi suara yang terdistorsi oleh *noise*. Jika ada banyak noise pada latar belakang dalam pengaturan sumber suara, kualitas suara dapat terganggu. Prosedur identifikasi terhambat oleh rekaman suara yang terdistorsi oleh noise. Di samping itu, penelitian sebelumnya terbatas pada eksplorasi suara biometrik untuk sistem keamanan. Oleh karena itu, penelitian ini akan menerapkan pendekatan menggunakan Convolutional Neural Network (CNN). Meskipun CNN sering digunakan untuk pengenalan objek atau gambar, namun untuk pengenalan suara, metode CNN belum banyak digunakan. Alvio Yunita Putri melakukan penelitian yang membahas sistem identifikasi suara, dimana pada penelitian ini digunakan *Short-Time Fourier Transform* (STFT) untuk proses ekstraksi ciri dalam bentuk spektrogram. Selanjutnya, spektrogram akan diolah oleh *Convolutional Neural Network* (CNN) sebagai metode pengidentifikasi[6].

Berdasarkan penelitian-penelitian sebelumnya, masih belum banyak yang menggunakan suara sebagai sistem keamanan untuk mengakses ruangan. Sehingga, pada penelitian ini akan dikembangkan suatu *prototype security system* untuk mengakses pintu masuk menggunakan suara secara *real-time*. *Prototype* ini sendiri nantinya akan digunakan sebagai alat sistem keamanan di Laboratorium Sistem Kendali sebagai sistem keamanan. Pengenalan suara pada *prototype* ini menggunakan algoritma CNN untuk meningkatkan kualitas dalam pengenalan suara dan terbukanya pintu masuk secara *real-time*. Algoritma ini telah menunjukkan performa yang baik dalam mengenali suara.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, penelitian sebelumnya telah memanfaatkan suara sebagai identifikasi untuk sistem keamanan. Namun

penelitian tersebut masih menggunakan algoritma biasa (*machine learning*) dan belum di implementasikan secara real. Disamping itu, performansi dari sistem identifikasi suara dengan menggunakan ekstraksi ciri membutuhkan waktu komputasi yang lebih besar sehingga pada penelitian ini, pengidentifikasian suara seseorang akan diimplementasikan pada sistem keamanan ruangan dan dilakukan secara real – time dengan menggunakan algoritma *deep learning*.

1.3 Tujuan Penelitian

Adapun tujuan dilakukannya penelitian ini adalah mengimplementasikan penggunaan suara sebagai sistem keamanan akses ruangan dengan menggunakan algoritma *deep learning* pada proses pembelajaran untuk mengenali suara secara *real-time*.

1.4 Batasan Masalah

Agar penelitian ini lebih terfokus dan tidak meluas dari pembahasan, maka unsur penelitian ini dibatasi pada:

- a. Algoritma pemodelan *deep learning* yang digunakan adalah CNN (*Convolutional Neural Network*)
- b. Diimplementasikan pada Laboratorium Dasar Sistem Kendali dan Robotika, Teknik Elektro, Fakultas Teknik, Universitas Sriwijaya.
- c. Dataset yang digunakan berupa *image spektogram* mahasiswa pranata laboratorium dan dosen di Teknik Elektro Universitas Sriwijaya.

1.5 Keaslian Penelitian

Penelitian ini merupakan penelitian yang berdasarkan pada studi-studi sebelumnya, yaitu Anggun Winursito, Risanuri Hidayat dan Agus Bejo. Membahas membahas untuk sistem pengenalan suara menggunakan *Mel Frequency Cepstral Coefficients* (MFCC) dengan Principal Component Analysis (PCA) untuk meningkatkan akurasi dalam sistem pengenalan suara bahasa Indonesia[7]. Namun, penelitian ini tetap memiliki kelemahan pada fitur MFCC yang digunakan karena penggunaan MFCC dapat menyebabkan kehilangan informasi suara asli akibat proses pengurangan atau reduksi dimensi. Akan tetapi, Dengan menggabungkan MFCC dan PCA, diharapkan dapat meningkatkan meningkatkan akurasi sistem dan mengurangi dimensi data fitur.

Hariz Zakka Muhammad, Muhammad Nasrun, Casi Setianingsih, dan Muhammad Ary Murt. Melakukan penelitian yang membahas sebuah sistem yang dapat menerjemahkan dari bahasa Inggris ke bahasa Indonesia berdasarkan pengenalan ucapan menggunakan Mel Frequency Cepstral Coefficients (MFCC) untuk ekstraksi fitur dan metode Hidden Markov Model untuk klasifikasi[4]. Sistem ini berfokus pada implementasi pengenalan suara, sedangkan untuk menerjemahkan bahasa menggunakan modul terjemahan yang ada seperti, Google Translation atau Microsoft Translation. Pengenalan ucapan menggunakan Hidden Markov Model didapatkan akurasi sistem untuk mengenali kata sebesar 70,10% sedangkan untuk mengenali kalimat sebesar 68,42%. Berdasarkan hasil pengujian, untuk pengucapan kalimat masih kurang baik dalam pemrosesan ketika terdapat lebih dari 3 kata dalam satu kalimat. Hal ini dikarenakan dalam klasifikasi HMM klasifikasi HMM, semakin banyak kata yang diproses maka model perhitungan probabilitas model akan terbagi dan juga semakin banyak kata yang diucapkan maka akan mempengaruhi lamanya waktu komputasi untuk memproses sistem.

Penelitian yang dilakukan oleh Ruchao Fan dan Gang Liu membahas tentang penggunaan Convolutional Neural Network (CNN) dalam pemilihan fitur input dan desain parameter CNN[5]. Dalam membandingkan fitur input, penelitian ini mengusulkan penggunaan FBANK (Full-Mel Spectrogram feature) karena sesuai dengan karakteristik persepsi nonlinier telinga manusia, struktur spektrum pada frekuensi rendah lebih jelas dan lebih jarang terlihat pada frekuensi tinggi. Namun, penelitian ini juga mencatat beberapa kelemahan terkait penggunaan CNN. Misalnya, apabila data yang tersedia terbatas, sejumlah besar parameter dalam CNN dapat menyebabkan overfitting. Selain itu, penggantian model bahasa, seperti dari bahasa n-gram ke model bahasa Recurrent Neural Network (RNN), dan pelatihan model bahasa bersamaan dengan model akustik juga menjadi tantangan yang sulit dalam penggunaan CNN.

Pankaj H. Chandankhede, Abhijit S. Titarmare dan Sarang Chauhan membahas pengenalan suara dengan menggunakan metode CNN. Pada penelitian penulis menggunakan *Mel-Frequency Cepstrum Coefficients* (MFCC) untuk proses ekstraksi ciri[8]. Pada penelitian ini mengambil analisis statistik suara dengan menggunakan scaled up dan scaled down spektrogram, bersamaan dengan

eksploitasi Google Speech-to-text API mengubah ucapan menjadi kode sandi, itu akan diverifikasi silang untuk tujuan keamanan yang lebih luas. Peneliti telah menemukan bahwa kombinasi frekuensi *Mel* dan *Convolutional Neural Network* memberikan akurasi terbaik dan kinerja yang paling efektif. Ini menunjukkan bahwa untuk mendapatkan hasil yang memuaskan bahwa maksimum epochs perlu meningkat karena informasi dalam sistem akan meningkat dengan bertambahnya jumlah pembicara.

Alvio Yunita Putri membahas sistem identifikasi suara, dimana pada penelitian ini digunakan Short-Time Fourier Transform (STFT) untuk proses ekstraksi ciri dalam bentuk spektrogram[6]. Kemudian, spektrogram akan diolah oleh *Convolutional Neural Network* (CNN) sebagai metode pengidentifikasi. Data suara yang digunakan dalam penelitian ini diperoleh dari rekaman suara menggunakan microphone. Rekaman suara berjumlah 780 data primer yang diambil dari 78 orang mahasiswa. Pada penelitian ini menggunakan 2 tipe arsitektur CNN, yaitu arsitektur sederhana CNN dan arsitektur CNN-VGG model VGG-f. Arsitektur CNN yang dirancang pada penelitian ini, digunakan dua tipe arsitektur Convolutional Neural Network (CNN), yaitu arsitektur sederhana CNN dan arsitektur CNN-VGG model VGG-f. Arsitektur CNN yang dirancang dalam penelitian ini mengadopsi model VGG-f yang terdiri dari lapisan konvolusi (convolutional layer), lapisan pengecilan (pooling layer), dan lapisan softmax loss untuk klasifikasi. Pengaturan parameter yang digunakan adalah ukuran 224 x 224 piksel, learning rate 0.001, dan ukuran batch 256. Hasil pengujian nilai akurasi rata-rata identifikasi suara adalah 98.7%. Hal ini menunjukkan bahwa kombinasi spektrogram dan CNN dapat menghasilkan akurasi identifikasi yang baik.

Berdasarkan beberapa penelitian yang telah dibahas diatas, maka pada penelitian ini akan dikembangkan dan diimplementasikan sebuah sistem keamanan pintu berbasis pengenalan suara dengan *Convolutional Neural Network* (CNN) sebagai metode pengenalan suara dikarenakan tingkat akurasi yang cukup tinggi. Penerapan secara real-time pada sistem keamanan pintu masih belum dilakukan pada penelitian-penelitian sebelumnya. Sistem ini diharapkan dapat meningkatkan keamanan dalam suatu ruangan.

DAFTAR PUSTAKA

- [1] C. Hung, Y. Fanjiang, K. Chung, and C. Kao, "A Door Lock System with Augmented Reality Technology," no. Gcce, 2017.
- [2] Y. Hasan, Y. Wijanarko, S. Muslimin, and R. Maulidda, "The Automatic Door Lock to Enhance Security in RFID System The Automatic Door Lock to Enhance Security in RFID System," 2020, doi: 10.1088/1742-6596/1500/1/012132.
- [3] S. Ariyanti, S. S. Adi, and S. Purbawanto, "Sistem Buka Tutup Pintu Otomatis Berbasis Suara Manusia," vol. 3, no. May, pp. 83–91, 2018, doi: 10.21831/elinvo.v3i1.19076.
- [4] H. Z. Muhammad, M. Nasrun, C. Setianingsih, and M. A. Murti, "Speech Recognition for English to Indonesian Translator Using Hidden Markov Model," pp. 255–260, 2018.
- [5] R. Fan and G. Liu, "CNN-Based Audio Front End Processing on Speech Recognition," *2018 International Conference on Audio, Language and Image Processing (ICALIP)*, pp. 349–354, 2018.
- [6] S. Dwijayanti, A. Y. Putri, and B. Y. Suprpto, "Speaker Identification Using a Convolutional Neural Network," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 140–145, 2022, doi: 10.29207/resti.v6i1.3795.
- [7] A. Winursito, "Improvement of MFCC Feature Extraction Accuracy Using PCA in Indonesian Speech Recognition," pp. 379–383, 2018.
- [8] P. H. Chandankhede, A. S. Titarmare, and S. Chauhvan, "Voice recognition based security system using convolutional neural network," *Proceedings - IEEE 2021 International Conference on Computing, Communication, and Intelligent Systems, ICCIS 2021*, pp. 738–743, 2021, doi: 10.1109/ICCIS51004.2021.9397151.
- [9] Y. Lukic, C. Vogt, O. Durr, and T. Stadelmann, "Speaker identification and clustering using convolutional neural networks," *IEEE International*

Workshop on Machine Learning for Signal Processing, MLSP, vol. 2016-Novem, 2016, doi: 10.1109/MLSP.2016.7738816.

- [10] W. Dadang, "Security System Using A Robot Based On Speech Recognition," pp. 25–29, 2020.
- [11] T. Juwariyah, A. C. Dewi, P. Studi, T. Industri, and J. Selatan, "RANCANG BANGUN SISTEM KEAMANAN SEPEDA MOTOR," vol. 13, pp. 223–227, 2017.
- [12] H. Isyanto, A. Solikhin, and W. Ibrahim, "Perancangan dan Implementasi Security System pada Sepeda Motor Menggunakan RFID Sensor Berbasis Raspberry Pi," vol. 2, no. 1, pp. 29–38.
- [13] D. Aji, F. Astuti, M. K. E. Ronando, and M. Si, "Pengenalan Suara Manusia Menggunakan Metode Support Vector Machine (Svm)," no. 2010, pp. 1–5, 2012.
- [14] D. Aji, "Sistem Pengenalan Suara Manusia Menggunakan Metode Support Vector Machine (SVM)," 2018.
- [15] A. rio Adriansyah, K. D. Prasetyo, and H. A. Atmam Al Faruqi, "Pengenalan Pola Fonem Vokal menggunakan Short Time Fourier Transform (STFT) dan Fitur Mel Frequency Cepstral Coefficient (MFCC)," *Jurnal Teknologi Terpadu*, vol. 7, no. 1, pp. 1–6, 2021, doi: 10.54914/jtt.v7i1.298.
- [16] A. Rajagopalan, "Real-Time Deep Learning-Based Face Recognition System," 2022, [Online]. Available: https://repository.stcloudstate.edu/ece_etds/8
- [17] S. Tammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 9, no. 10, p. p9420, Oct. 2019, doi: 10.29322/ijsrp.9.10.2019.p9420.
- [18] M. N. Stolar, M. Lech, R. S. Bolia, and M. Skinner, "Real time speech emotion recognition using RGB image classification and transfer learning," in *2017, 11th International Conference on Signal Processing and*

Communication Systems, ICSPCS 2017 - Proceedings, Institute of Electrical and Electronics Engineers Inc., Jan. 2018, pp. 1–8. doi: 10.1109/ICSPCS.2017.8270472.