

SKRIPSI

**IMPLEMENTASI KOMUNIKASI DUA ARAH PADA
SERVICE ROBOT MENGGUNAKAN ALGORITMA
*TRANSFORMERS***



OLEH:

SYAUQI ZALFFA DAFFA

03041282025045

JURUSAN TEKNIK ELEKTRO

FAKULTAS TEKNIK

UNIVERSITAS SRIWIJAYA

2024

SKRIPSI

IMPLEMENTASI KOMUNIKASI DUA ARAH PADA *SERVICE ROBOT* MENGGUNAKAN ALGORITMA *TRANSFORMERS*



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

OLEH:

SYAUQI ZALFFA DAFFA

03041282025045

JURUSAN TEKNIK ELEKTRO

FAKULTAS TEKNIK

UNIVERSITAS SRIWIJAYA

2024

LEMBAR PENGESAHAN

IMPLEMENTASI KOMUNIKASI DUA ARAH PADA *SERVICE ROBOT* MENGUNAKAN ALGORITMA *TRANSFORMERS*



SKRIPSI

Disusun Untuk Memenuhi Syarat Memperoleh Gelar Sarjana
Teknik Pada Jurusan Teknik Elektro Fakultas Teknik
Universitas Sriwijaya

Oleh

SYAUQI ZALFFA DAFFA
NIM. 03041282025045

Palembang, 17 Juli 2024

Mengetahui

Ketua Jurusan Teknik Elektro



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

NIP.197108141999031005

Menyetujui

Dosen Pembimbing

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

IPM.

NIP. 198407302008122001

HALAMAN PERNYATAAN DOSEN

Saya sebagai pembimbing menyatakan bahwa telah membaca dan menyetujui skripsi ini dan dalam pandangan saya skop dan kuantitas skripsi ini mencakupi sebagai mahasiswa sarjana strata satu (S1).

Tanda Tangan

:  _____

Pembimbing Utama

: **Dr. Eng. Ir. Suci Dwijayanti. S.T., M.S., IPM.**


Tanggal

: **15/Juli/2024**

HALAMAN PERNYATAAN DOSEN

Saya sebagai pembimbing menyatakan bahwa telah membaca dan menyetujui skripsi ini dan dalam pandangan saya skop dan kuantitas skripsi ini mencakupi sebagai mahasiswa sarjana strata satu (S1).

Tanda Tangan

:  _____

Pembimbing Utama

: **Dr. Eng. Ir. Suci Dwijayanti. S.T., M.S., IPM.**

Tanggal

: **15/Juli/2024**

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

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

Nama : Syauqi Zalffa Daffa
NIM : 03041282025045
Jurusan : 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:

**IMPLEMENTASI KOMUNIKASI DUA ARAH PADA *SERVICE ROBOT*
MENGUNAKAN ALGORITMA *TRANSFORMERS***

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 sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Indralaya

Pada tanggal : 17 Juli 2024



Syauqi Zalffa Daffa

NIM. 03041282025045

KATA PENGANTAR

Puji syukur penulis ucapkan kepada Allah SWT. Atas Berkat dan dan Rahmat-Nya serta dukungan keluarga dan para sahabat penulis dapat menyelesaikan skripsi "Implementasi Komunikasi dua arah pada *service robot* menggunakan Algoritma *transformers*".

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

1. Orangtua serta keluarga yang selalu memberikan semangat dan memberi dukungan baik secara mental, fisik, maupun finansial.
2. Ibu Dr. Eng. Suci Dwijayanti S.T., M.S., IPM. selaku pembimbing utama tugas akhir ini yang telah memberikan bimbingan, arahan, dan masukan serta memberikan ilmu selama proses penulisan skripsi.
3. Bapak Dr. Bhakti Yudho Suprpto, S.T., M.T., IPM. dan Ibu Dr. Eng. Suci Dwijayanti S.T., M.S., IPM. selaku pencetus dan memberikan bimbingan pada tugas akhir ini serta pengembang ide.
4. Bapak Dr. H. Iwan Pahendra S, S.T., M.T. selaku Dosen pembimbing akademik
5. Bapak Muhammad Abu Bakar Sidik, S.T., M.Eng., Ph.D., IPU. Ketua Jurusan Teknik Elektro Universitas Sriwijaya dan Ibu Dr. Eng. Suci Dwijayanti S.T., M.S., IPM. selaku Sekretaris Jurusan Teknik Elektro Universitas Sriwijaya.
6. Bapak Dr. Bhakti Yudho Suprpto, S.T., M.T., IPM., Ibu Hera Hikmarika, S.T., M.Eng, Bapak Baginda Oloan Siregar, S.T., M.T, Bapak Irmawan, S.Si., M.T., dan Bapak Rendyansyah, S.Kom., M.T. yang telah memberikan kritik dan saran yang membangun dalam penelitian yang dilakukan agar menjadi lebih baik.

7. Segenap Dosen Jurusan Teknik Elektro Fakultas Teknik Universitas Sriwijaya yang telah memberikan ilmu selama perkuliahan.
8. Teman-teman seperjuangan *Service robot* Karim, Arif, Iyan, Fahreza dan Akmal yang telah berkontribusi baik tenaga, waktu dan pikiran dalam menyelesaikan penelitian ini.
9. Seluruh teman-teman konsentrasi Teknik Kendali dan Robotika Teknik Elektro Universitas Siwijaya Angkatan 2020 selaku rekan kerja yang membantu dalam pengambilan data dan selalu bersemangat dalam pembuatan tugas akhir ini.
10. Teman-teman Klub Robotika UNSRI yang selalu membantu juga menyemangati selama proses penulisan skripsi ini berlangsung.
11. Orang Terdekat Putri Salma Ningayu yang telah memberi semangat serta motivasi dalam menyelesaikan penelitian ini
12. Seluruh pihak yang tidak mampu disebutkan satu-persatu yang telah membantu penulis dalam menyelesaikan tugas akhir hingga meraih gelar Sarjana Teknik.

Di dalam penyusunan skripsi ini, masih terdapat kekurangan karena keterbatasan penyusun, maka dari itu penyusun mengharapkan kritik dan saran yang membangun agar dapat menjadi evaluasi untuk penelitian yang lebih baik bagi penyusun dimasa yang akan datang. Akhir kata, penulis berharap agar skripsi ini menjadi lebih bermanfaat bagi pembaca terutama mahasiswa Jurusan Teknik Elektro Universitas Sriwijaya dan Masyarakat umum.

Palembang, 17 Juli 2024



Syauqi Zalfa Daffa

NIM. 03041282025045

ABSTRAK

Implementasi Komunikasi Dua Arah Pada *Service Robot* Menggunakan Algoritma *Transformers*

(Syauqi Zalffa Daffa, 03041282025045, 2024, 77 halaman)

Abstrak—*Humanoid robot* merupakan jenis robot yang dapat membantu berbagai tugas manusia sebagai *service robot*. Robot ini harus memiliki kemampuan komunikasi verbal untuk memungkinkan sistem komunikasi dua arah melalui *speech recognition*. Penelitian sebelumnya menunjukkan keterbatasan pada Bahasa Indonesia karena jumlah sampel data yang sedikit sehingga akurasi transkripsi bahasa Indonesia kurang maksimal. Penelitian ini bertujuan mengembangkan sistem komunikasi dua arah antara manusia dan robot secara *real-time*. *Dataset* yang digunakan pada penelitian ini terdiri dari 21 narasumber (15 laki-laki dan 6 perempuan). Sistem *speech-to-text* menggunakan algoritma *transformer* dengan model *Whisper*, sedangkan *text-to-speech* menggunakan Google Text to Speech (gTTS), library dari *Python* dan *command line interface (CLI) tool* sebagai antarmuka dengan Google Translates *text-to-speech API*. Model *transformer* hasil pelatihan terbaik diperoleh dengan menggunakan parameter *training step* 900. Hasil pengujian secara simulasi menunjukkan model tersebut menghasilkan *word error rate (WER)* dan *character error rate (CER)* untuk sampel suara laki-laki masing-masing sebesar 10% dan 4%, lalu untuk perempuan sebesar 11% dan 8%. Lalu, hasil pengujian secara *real-time* menunjukkan bahwa pada kondisi hening dengan kebisingan 47,1-59,0 dB, rata-rata WER untuk jarak 10 cm, 30 cm, dan 50 cm masing-masing adalah 6%, 12%, dan 15%, dengan rata-rata *character error rate (CER)* masing-masing 2%, 4%, dan 5%. Pada kondisi *noise* mesin dengan kebisingan 82,3-91,3 dB, rata-rata WER untuk jarak yang sama adalah 13%, 16%, dan 24%, dengan CER masing-masing 4%, 5%, dan 8%. Sementara pada kondisi *noise* keramaian dengan kebisingan 72,8-84,4 dB, rata-rata WER adalah 16%, 20%, dan 22%, serta rata-rata CER masing-masing 6%, 7%, dan 8% untuk jarak 10 cm, 20 cm, dan 50 cm. Penelitian ini menunjukkan bahwa sistem pengenalan ucapan dapat diimplementasikan untuk komunikasi antara robot dan manusia, memungkinkan robot memberikan respon yang sesuai terhadap input suara yang diterima.

Kata kunci: *Service robot, Speech recognition, Speech to text, Text to speech*

ABSTRACT

Implementation of Two-Way Communication on Service Robots Using Transformers Algorithm

(Syauqi Zalfha Daffa, 03041282025045, 2024, 77 pages)

Abstract—The humanoid robot is a type of robot that can assist in various human tasks as a service robot. This robot must have verbal communication abilities to enable two-way communication systems via speech recognition. Previous research indicates limitations in Indonesian language transcription accuracy due to the scarcity of data samples. This study aims to develop a real-time two-way communication system between humans and robots. The dataset used in this study consists of 21 speakers (15 males and 6 females). The speech-to-text system employs a transformer algorithm with the Whisper model, while text-to-speech utilizes Google Text to Speech (gTTS), a Python library, and a command-line interface (CLI) tool interfaced with Google's text-to-speech API. The best-trained transformer model was achieved using 900 training steps. Simulation results show that this model yields word error rates (WER) and character error rates (CER) of 10% and 4% for male speech samples, and 11% and 8% for female speech samples. Real-time testing under quiet conditions with noise levels of 47.1-59.0 dB shows average WERs of 6%, 12%, and 15% at distances of 10 cm, 30 cm, and 50 cm respectively, with corresponding average CERs of 2%, 4%, and 5%. In machine noise conditions with levels of 82.3-91.3 dB, the average WERs for the same distances are 13%, 16%, and 24%, with CERs of 4%, 5%, and 8%. Meanwhile, in noisy conditions with levels of 72.8-84.4 dB, average WERs are 16%, 20%, and 22%, with average CERs of 6%, 7%, and 8% for distances of 10 cm, 20 cm, and 50 cm. This research demonstrates that speech recognition systems can be implemented for communication between robots and humans, enabling robots to respond appropriately to received voice inputs.

Keywords: *Service robot, Speech recognition, Speech to text, Text to speech*

DAFTAR ISI

COVER DEPAN	i
HALAMAN JUDUL	ii
LEMBAR PENGESAHAN	iii
HALAMAN PERNYATAAN INTEGRITAS	iv
HALAMAN PERNYATAAN DOSEN	v
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	vi
KATA PENGANTAR	vii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xv
BAB I PENDAHULUAN	2
1.1 Latar Belakang	2
1.2 Perumusan Masalah	4
1.3 Tujuan Penulisan	4
1.4 Pembatasan Masalah.....	4
1.5 Keaslian Penelitian	5
BAB II TINJAUAN PUSTAKA	7
2.1 <i>State of the Art</i>	7
2.2 Teori Pendukung	11
2.2.1 <i>Humanoid Robot</i>	11
2.2.2 <i>Speech to Text</i>	11

2.2.3 <i>Text to speech</i>	12
2.2.4 <i>Natural Language Processing (NLP)</i>	12
2.2.5 <i>Transformers</i>	12
2.2.6 <i>Whisper</i>	15
2.2.7 <i>Robot Operating System (ROS)</i>	17
BAB III METODOLOGI PENELITIAN	19
3.1 Studi Literatur	20
3.2 Pengambilan Data.....	20
3.3 Perancangan Sistem Komunikasi Dua Arah pada <i>Humanoid Robot</i>	21
3.3.1 <i>Desain Service Robot</i>	21
3.3.2 Perangkat Keras dan Perangkat Lunak.....	22
3.3.3 Perancangan Algoritma Sistem Komunikasi Dua Arah pada <i>Service Robot</i>	27
3.4 Pengujian Sistem	32
BAB IV HASIL DAN PEMBAHASAN.....	34
4.1 <i>Preliminary Result</i>	34
4.2 Pengumpulan Data Suara.....	36
4.3 Pembuatan <i>Speech to Text</i>	37
4.3.1 <i>Training Model</i>	39
4.3.2 Pengujian Model.....	45
4.4 Pembuatan <i>Text to Speech</i> dengan gTTS	49
4.5 Hasil Desain Robot untuk Pengenalan Suara	50
4.6 Pengujian <i>Service Robot</i> untuk Komunikasi Dua Arah Secara <i>Real-Time</i>	53
4.6.1 Pengujian Pada Kondisi Ruangan Hening	53
4.6.2 Pengujian Pada Kondisi Ruangan <i>Noise</i> mesin.....	59

4.6.3	Pengujian Pada Kondisi Ruangan <i>Noise</i> Keramaian	65
4.7	Pengujian Algoritma saat <i>Service Robot</i> Menyapa	70
BAB V	72
KESIMPULAN DAN SARAN	72
DAFTAR PUSTAKA	74
LAMPIRAN	78

DAFTAR GAMBAR

Gambar 2. 1. Contoh Service Robot [26]	11
Gambar 2. 2. Arsitektur Transformer [32].	13
Gambar 2. 3. Arsitektur Wav2vec2 [34].	16
Gambar 3. 1 <i>Flowchart</i> penelitian	19
Gambar 3. 2 Perancangan peletakkan komponen pada Service Robot.....	22
Gambar 3. 3 <i>Microphone</i> RODE VIDEO MICRO	23
Gambar 3. 4 Kamera Nemesis A95 Albatros	23
Gambar 3. 5 <i>Speaker audio</i>	25
Gambar 3. 6 <i>Dot matrix 8x32</i>	25
Gambar 3. 7 Sensor suara.....	26
Gambar 3. 8 <i>Flowchart service robot</i> untuk <i>speech to text</i>	29
Gambar 3. 9 <i>Flowchart service robot</i> untuk <i>text to speech</i>	31
Gambar 3. 10 <i>Flowchart service robot</i> untuk menyapa.....	32
Gambar 4. 1 (a) Suara dalam suasana noise (b) Suara dalam suasana hening	37
Gambar 4. 2 Pembuatan file dataset dengan format .csv	37
Gambar 4. 3 Mel-filter bank.....	38
Gambar 4. 4 (a) Grafik <i>train/epoch</i> 900 steps, (b) Grafik <i>train/epoch</i> 500 steps .	42
Gambar 4. 5 (a) Grafik <i>learning rate</i> 900 steps, (b) Grafik <i>learning rate</i> 500 steps	Error! Bookmark not defined.
Gambar 4. 6 (a) Grafik <i>loss</i> 900 steps, (b) Grafik <i>loss</i> 500 steps	44
Gambar 4. 7 Hasil <i>testing</i> model untuk kalimat “Halo, selamat pagi”	45
Gambar 4. 8 Program dari <i>g tts</i> dengan menggunakan bahasa Indonesia untuk menjalankan komunikasi dua arah.....	49
Gambar 4. 9 (a) Algoritma saat suara diambil dari <i>database</i> (b) Respon dari API Gemini AI (c) algoritma saat suara diambil dari API Gemini AI	50
Gambar 4. 10 <i>Service Robot</i>	51
Gambar 4. 11 <i>Dot Matrix</i> menampilkan <i>spectrogram</i>	52
Gambar 4. 12 algoritma perintah untuk <i>speech recognition</i> dan deteksi wajah....	70

DAFTAR TABEL

Tabel 4. 1 Hasil pengujian perbandingan LSTM dan Transformers	34
Tabel 4. 2 Parameter pada whisperfeatureextractor.....	38
Tabel 4. 3 Spesifikasi <i>google colab</i>	40
Tabel 4. 4 Parameter pada proses <i>training</i>	41
Tabel 4. 5 Akurasi sampel suara per kalimat dengan nama file “badar”	46
Tabel 4. 6 Akurasi sampel suara per kalimat dengan nama file “Shalput”	47
Tabel 4. 7 Parameter pengujian sistem komunikasi dua arah secara <i>real-time</i> pada kondisi hening	53
Tabel 4. 8 Hasil pengujian 5 data suara yang ada di dalam dataset dan 5 data suara dengan diluar dari dataset dengan jarak <i>microphone</i> 10 cm dan kondisi ruangan hening.....	54
Tabel 4. 9 Hasil pengujian 5 data suara yang ada di dalam dataset dan 5 data suara dengan diluar dari dataset dengan jarak <i>microphone</i> 30 cm dan kondisi ruangan hening.....	56
Tabel 4. 10 Hasil pengujian 5 data suara yang ada di dalam dataset dan 5 data suara dengan diluar dari dataset dengan jarak <i>microphone</i> 50 cm dan kondisi ruangan tenang	58
Tabel 4. 11 Parameter pengujian sistem komunikasi dua arah secara <i>real-time</i> pada kondisi <i>Noise</i> mesin	59
Tabel 4. 12 Hasil pengujian 5 data suara yang ada di dalam dataset dan 5 data suara dengan diluar dari dataset dengan jarak <i>microphone</i> 10 cm dan kondisi ruangan <i>Noise</i> mesin	60
Tabel 4. 13 Hasil pengujian 5 data suara yang ada di dalam dataset dan 5 data suara dengan diluar dari dataset dengan jarak <i>microphone</i> 30 cm dan kondisi ruangan <i>Noise</i> mesin	62
Tabel 4. 14 Hasil pengujian 5 data suara yang ada di dalam dataset dan 5 data suara dengan diluar dari dataset dengan jarak <i>microphone</i> 50 cm dan kondisi ruangan <i>Noise</i> mesin	63
Tabel 4. 15 Parameter pengujian sistem komunikasi dua arah secara <i>real-time</i> pada kondisi <i>Noise</i> Keramaian	65

Tabel 4. 16 Hasil pengujian 5 data suara yang ada di dalam dataset dan 5 data suara dengan diluar dari dataset dengan jarak microphone 10 cm dan kondisi ruangan <i>Noise Keramaian</i>	65
Tabel 4. 17 Hasil pengujian 5 data suara yang ada di dalam dataset dan 5 data suara dengan diluar dari dataset dengan jarak microphone 30 cm dan kondisi ruangan <i>Noise Keramaian</i>	67
Tabel 4. 18 Hasil pengujian 5 data suara yang ada di dalam dataset dan 5 data suara dengan diluar dari dataset dengan jarak microphone 50 cm dan kondisi ruangan <i>Noise Keramaian</i>	68

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada saat ini, teknologi dalam berbagai bidang telah mengalami pertumbuhan yang sangat cepat, termasuk pada bidang robotika[1][2]. Perkembangan ini dapat dilihat dari semakin banyaknya jenis robot yang diciptakan untuk membantu dan mendukung berbagai kegiatan yang dilakukan oleh manusia. Robot dapat didefinisikan sebagai alat yang beroperasi secara otomatis baik dengan bantuan manusia maupun tanpa bantuan manusia.

Dalam beberapa tahun terakhir ini, *humanoid robot* merupakan jenis robot yang menjadi fokus utama dalam dunia robotika baik untuk desain dan teknologi yang diterapkan[3], [4], [5]. *Humanoid robot* akan menjadi robot yang paling banyak digunakan pada masa yang akan datang karena memiliki fitur yang bisa melakukan kegiatan seperti manusia. *Humanoid robot* ini banyak membantu berbagai tugas manusia termasuk memasak, resepsionis hotel, dan menjadi pemandu di bandara. Sebagian kegiatan yang dilakukan robot ini merupakan bentuk layanan untuk membantu manusia atau dapat dikatakan sebagai *service robot*[6], [7], [8]. Dalam melakukan kegiatan tersebut, *service robot* perlu memiliki kemampuan berkomunikasi dengan manusia melalui bahasa verbal[9], [10], [11].

Komunikasi dua arah yang ada pada *service robot* dapat dilakukan karena adanya peran dari sistem *speech recognition* yang ditanamkan pada program robot. *Speech recognition* adalah suatu kemampuan mesin yang berfungsi untuk mendengar kata-kata yang diucapkan dan juga mampu untuk mengidentifikasinya serta mempunyai kemampuan untuk mengkonversi inputan suara ke dalam bentuk teks[12], [13], [14]. *Speech recognition* melibatkan penggabungan ilmu komputer dan linguistik untuk mengenali kata-kata yang diucapkan dan mampu mengubahnya ke dalam bentuk teks [15], [16], [17]. Oleh karena itu, metode untuk mengkonversi suara menjadi teks sangat diperlukan. Selanjutnya, teks tersebut diubah menjadi suara sebagai bentuk respons yang diberikan oleh robot.

Banyak penelitian terdahulu yang telah dilakukan untuk pengenalan suara pada *humanoid robot*. [18] melakukan penelitian tentang *speech recognition*

dengan menggunakan metode *hidden Markov model* (HMM). Pada penelitian tersebut, *speech recognition* berbasis android digunakan untuk mengendalikan *humanoid robot*. Namun, metode yang digunakan kurang komprehensif dalam menilai kinerja pengenalan suara yang mengakibatkan kemampuan untuk mengukur keakuratan dan ketahanan menjadi terbatas dan menjadi sulit untuk menentukan efektivitas metode pengenalan suara yang digunakan. Selanjutnya, Sushan Phodel dan Dr.R Anuradha merancang *speech command recognition* dengan menggunakan kombinasi algoritma *convolutional neural network* (CNN) dan *recurrent neural network* (RNN). Akan tetapi, kombinasi antara CNN dan RNN membutuhkan jumlah data pelatihan yang sangat banyak untuk mencapai hasil yang akurat [19]. Kemudian, [20] menggunakan metode algoritma *deep neural network* (DNN) untuk melaksanakan interaksi antara manusia dan robot. Namun, algoritma ini memiliki kompleksitas komputasi dan kebutuhan sumber daya yang tinggi yang menyebabkan adanya potensi *overfitting* atau penurunan kinerja. Penelitian selanjutnya dilakukan oleh Jianyou Wang dkk yang membuat *speech recognition* menggunakan metode kombinasi fitur *mel frequency cepstral coefficient* (MFCC) dan *mel-spectrogram* untuk melakukan pengenalan emosi dalam ucapan khusus menggunakan sinyal audio. Namun, metode ini tidak dapat menangani panjang variabel antara ucapan yang menyebabkan informasi penting hilang [21]. Lalu, Azhari Bastomi dan Umi Fadilah dalam penelitiannya menggunakan *pyttsx3* untuk melakukan proses *text to speech*, namun suara yang disediakan *pyttsx3* masih terbatas [22]. Selanjutnya, [23] menggunakan metode *shallow parsing* untuk *text to speech* tetapi metode ini tidak dapat menangani kata yang bermakna ganda dengan baik.

Pada penelitian-penelitian di atas sistem komunikasi antara manusia dan robot masih memiliki keterbatasan terutama pada metode yang dipakai, seperti pengenalan teks dengan akurasi yang kurang baik dan pelatihan dengan jumlah data yang sangat besar agar dapat menghasilkan akurasi yang baik [19]. Selain itu, bahasa yang digunakan masih terbatas pada bahasa Inggris [18]. Padahal implementasi *speech to text* perlu dibuat dalam Bahasa Indonesia jika *service robot* digunakan di Indonesia. Penelitian *speech to text* untuk Bahasa Indonesia masih sangat terbatas dan sebagian hanya menggunakan *dataset* dalam jumlah terbatas

[15]. Sehingga, pada penelitian ini dikembangkan suatu sistem komunikasi dua arah pada *service robot* secara *real-time* dengan melakukan konversi *speech to text* yang berfungsi untuk mengubah suara ke teks dengan menggunakan metode *transformer*. *Transformer* dapat mempercepat pembelajaran suatu data dan juga mempunyai sistem pengenalan ucapan yang memiliki tingkat kesalahan yang lebih rendah [24] . Kemudian, *text to speech* dilakukan dengan menggunakan modul *python* Google Text-to-Speech (gTTS) . *Text to speech* ini diperlukan agar *service robot* dapat berkomunikasi dan merespon apa yang diperintahkan ataupun diinformasikan.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, permasalahan yang dihadapi dalam *speech recognition* untuk implementasi di *service robot* adalah metode yang digunakan yang berdampak pada akurasi yang rendah. Selain itu, saat ini penerapan *speech recognition* yang dilakukan masih menggunakan *dataset* dalam bahasa Inggris yang belum diimplementasikan pada robot sehingga komunikasi dua arah belum dapat diaplikasikan.

1.3 Tujuan Penulisan

Penelitian ini bertujuan untuk mengembangkan sistem komunikasi dua arah antar manusia dan *service robot* dengan menggunakan Bahasa Indonesia dengan menerapkan metode *transformer* untuk *speech to text* dan gTTS untuk *text to speech*. Selain itu, performansi dari *transformer* dalam melakukan *speech to text* secara *real-time* dengan menggunakan Bahasa Indonesia dievaluasi pada penelitian ini.

1.4 Pembatasan Masalah

Pembatasan masalah pada penelitian ini meliputi :

1. Penerapan algoritma *transformer* digunakan untuk mengubah suara menjadi teks.
2. *Service robot* mampu memberi tanggapan terhadap pertanyaan lawan bicara melalui modul dari *python*, yaitu gTTS
3. Bahasa Indonesia yang digunakan adalah bahasa baku.

1.5 Keaslian Penelitian

Penelitian ini merupakan lanjutan dari studi-studi sebelumnya dengan fokus pada komunikasi dua arah antara robot dan manusia. Ada beberapa penelitian sebelumnya yang melakukan penelitian dengan tema yang sama. Hendy Rudiansyah menggunakan metode *hidden Markov model* (HMM) untuk *speech recognition* yang digunakan untuk mengendalikan *humanoid robot*[18]. Hasil pengujian menunjukkan nilai akurasi sebesar 98,14% dan *error* sebesar 1,86%. Pada pengujian pergerakan robot, *error* pada perintah suara sebesar 4,26% untuk kaki kiri dan 5,52% untuk kaki kanan sedangkan untuk perintah mundur menghasilkan nilai *error* sebesar 2,3% untuk kaki kiri dan 3,01% untuk kaki kanan. Hal ini disebabkan oleh proses penangkapan suara berasal berbagai jenis suara yang berbeda dan ketidakjelasan suara, seperti aksen yang berbeda dan tinggi rendahnya suara.

Kemudian, [19] menggunakan kombinasi metode *convolutional neural network* (CNN) dan *recurrent neural network* (RNN). Pengujian pengenalan suara dilakukan dengan menggunakan *dataset* yang disediakan oleh TensorFlow yang terdiri dari 65.000 suara *wav* dengan durasi satu detik di setiap perintah suara. Akurasi validasi yang didapatkan pada penelitian sebesar 94,79% dan akurasi pengujian sebesar 96.66% dengan *loss* sebesar 0.117. Penelitian ini menunjukkan akurasi yang sangat baik akan tetapi untuk mendapatkan hasil yang terbaik harus menyediakan *dataset* yang sangat banyak.

Penelitian selanjutnya dilakukan oleh Mustafa Can Bingol dan Omur Aydogmus yang membahas tentang algoritma *deep neural network* (DNN)[20] untuk pengenalan suara pada robot manipulator. Pengujian klasifikasi dengan menggunakan DNN menunjukkan hasil yang kurang maksimal dikarenakan terbatasnya jumlah kata pada saat pengujian.

Pada penelitian lain digunakan pendekatan kombinasi fitur *mel frequency cepstral coefficient* (MFCC) dan *mel-spectrogram* yang dihasilkan dari sinyal audio, serta penggunaan arsitektur DNN yang dirancang untuk belajar dari data yang diambil dan menghasilkan model prediksi penyakit berbasis suara yang akurat. Berdasarkan hasil pengujian didapatkan akurasi tertinggi yang dicapai adalah

77,49%. Selain itu, model dari DNN mencapai akurasi 88,01% dengan mengintegrasikan informasi gender pembicara[25].

Penelitian selanjutnya adalah *text to speech* dengan menggunakan *pyttsx3* [22]. Pada penelitian ini didapatkan hasil pengujian efektivitas perangkat penerjemah, dimana pengucapan yang baik terdapat pada huruf A dan E dengan akurasi 90% dan pada huruf B, C, F, H, I, J, K, L, dan P memiliki akurasi yang rendah dengan akurasi pengucapan sebesar 60%.

Dari hasil penelitian-penelitian sebelumnya dapat dilihat bahwa penelitian-penelitian tersebut terbatas pada jumlah *sample data* yang sedikit sehingga akurasi yang didapatkan masih rendah saat mencoba sintesis bahasa Indonesia. Selain itu, kinerja metode yang digunakan masih kurang maksimal. Sehingga, pada penelitian ini memanfaatkan algoritma *transformer* sebagai metode *speech to text*. Algoritma *transformer* mampu mempercepat proses pembelajaran data dan juga dapat diandalkan dengan *dataset* yang sedikit dapat memiliki akurasi yang tinggi [24]. Sedangkan untuk *text to speech* menggunakan gTTS yang merupakan *library* Python dan *command line interface (CLI) tool* sebagai antar muka dengan Google Translate's text-to-speech API. Sehingga, *service robot* dapat memberikan respon dan bisa berkomunikasi secara dua arah dengan manusia.

DAFTAR PUSTAKA

- [1] Ramadhani Prianka Kevin, Amrozi Yusuf, and Adi Irzaldi, "INOVASI SISTEM ROBOTIKA PADA PERPUSTAKAAN," *JEECOM*, vol. 2, p. 2, 2020.
- [2] A. F. Nurqolbi, "Penerapan Robotika Dalam Perpustakaan Masa Depan," *Journal UII*, vol. 14, no. 1, pp. 57–64, 2023, doi: 10.20885/unilib.Vol14.iss1.art6.
- [3] P. H. Kuo and K. L. Chen, "Two-stage fuzzy object grasping controller for a humanoid robot with proximal policy optimization," *Eng Appl Artif Intell*, vol. 125, Oct. 2023, doi: 10.1016/j.engappai.2023.106694.
- [4] P. H. Kuo and K. L. Chen, "Two-stage fuzzy object grasping controller for a humanoid robot with proximal policy optimization," *Eng Appl Artif Intell*, vol. 125, Oct. 2023, doi: 10.1016/j.engappai.2023.106694.
- [5] A. A. S. G. Heri Ngarianto, "Pengembangan Sistem Kendali Robot Humanoid untuk Edukasi berdasarkan Robosapien dan Raspberry Pi Zero," *EMACS (Engineering, MAtematics and Computer Science)*, vol. Vol.2 No.2, pp. 85–90, 2020, Accessed: Oct. 12, 2023. [Online]. Available: file:///C:/Users/User/Downloads/Documents/6473-Article%20Text-31362-6-10-20210105.pdf
- [6] J. Liao and J. Huang, "Think like a robot: How interactions with humanoid service robots affect consumers' decision strategies," *Journal of Retailing and Consumer Services*, vol. 76, Jan. 2024, doi: 10.1016/j.jretconser.2023.103575.
- [7] X. (Stella) Liu, L. C. Wan, and X. (Shannon) Yi, "Humanoid versus non-humanoid robots: How mortality salience shapes preference for robot services under the COVID-19 pandemic?," *Ann Tour Res*, vol. 94, May 2022, doi: 10.1016/j.annals.2022.103383.
- [8] J. A. Gonzalez-Aguirre *et al.*, "Service robots: Trends and technology," *Applied Sciences (Switzerland)*, vol. 11, no. 22. MDPI, Nov. 01, 2021. doi: 10.3390/app112210702.
- [9] C. Li, D. Chrysostomou, and H. Yang, "A speech-enabled virtual assistant for efficient human–robot interaction in industrial environments," *Journal of Systems and Software*, vol. 205, Nov. 2023, doi: 10.1016/j.jss.2023.111818.
- [10] A. Marin Vargas, L. Cominelli, F. Dell'Orletta, and E. P. Scilingo, "Verbal Communication in Robotics: A Study on Salient Terms, Research Fields and Trends in the Last Decades Based on a Computational Linguistic Analysis," *Frontiers in Computer Science*, vol. 2. Frontiers Media S.A., Feb. 15, 2021. doi: 10.3389/fcomp.2020.591164.

- [11] U. Zabala, I. Rodriguez, J. M. Martínez-Otzeta, and E. Lazkano, "Expressing robot personality through talking body language," *Applied Sciences (Switzerland)*, vol. 11, no. 10, May 2021, doi: 10.3390/app11104639.
- [12] S. Jiang and Z. Chen, "Application of dynamic time warping optimization algorithm in speech recognition," *Heliyon*, p. e21625, Oct. 2023, doi: 10.1016/j.heliyon.2023.e21625.
- [13] N. Nedjah, A. D. Bonilla, and L. de Macedo Mourelle, "Automatic speech recognition of Portuguese phonemes using neural networks ensemble," *Expert Syst Appl*, vol. 229, Nov. 2023, doi: 10.1016/j.eswa.2023.120378.
- [14] R. Arief, N. A. Iriawan, D. A. Lawi, and E. Coprespondent, "KLASIFIKASI AUDIO UCAPAN EMOSIONAL MENGGUNAKAN MODEL LSTM," *Konferensi Nasional Ilmu Komputer (KONIK)*, 2021.
- [15] F. Adnan, I. Amelia, D. Sayyid ', and U. Shiddiq, "Implementasi Voice Recognition Berbasis Machine Learning," *Edu Elekrika Journal*, vol. 11, no. 1, 2022.
- [16] R. Aditya, A. Muid, and U. Ristian, "Tempat Sampah Otomatis Speech Recognition Menggunakan Pocketsphinx," *Jurnal Ilmiah Ilmu Komputer*, vol. 15, no. 1, 2020, doi: 10.30872/jim.v15i1.2107.
- [17] K. Amin, L. Elvitaria, and L. Trisnawati, "Jurnal Politeknik Caltex Riau Artificial Intelligence Automatic Speech Recognition (ASR) untuk pencarian potongan ayat Al-Qu'ran," 2022. [Online]. Available: <https://jurnal.pcr.ac.id/index.php/jkt/>
- [18] H. Rudiansyah, A. P. Th, and P. M. Bandung, "Pengendalian Robot Humanoid menggunakan Metode Speech Recognition Berbasis Android," *Jurnal Teknologi dan Rekayasa Manufaktur JTRM /*, vol. 2, no. 1, 2020.
- [19] S. Poudel and R. Anuradha, "Speech Command Recognition using Artificial Neural Networks," *INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION*, vol. 4, p. 2, 2020.
- [20] M. C. Bingol and O. Aydogmus, "Performing predefined tasks using the human–robot interaction on speech recognition for an industrial robot," *Eng Appl Artif Intell*, vol. 95, Oct. 2020, doi: 10.1016/j.engappai.2020.103903.
- [21] Wang Jianyou, Xue Michael, Culhane Ryan, Diao Enmao, Ding Jie, and Tarokh Vahid, "SPEECH EMOTION RECOGNITION WITH DUAL-SEQUENCE LSTM ARCHITECTURE," *IEEE*, 2020.
- [22] R. M. Azhari Bastomi, U. Fadlilah, and M. Eng, "Alat Penerjemah Bahasa Isyarat Penyandang Tunarungu Berbasis Raspberry-Pi Kamera," *Simposium Nasional RAPI XIX*, 2020.

- [23] L. S. Fitriawati, A. Bijaksana, P. Negara, and R. D. Nyoto, "Implementasi Text To Speech Pada Website Menggunakan Metode Shallow Parsing," *InfoTekJar : Jurnal Nasional Informatika dan Teknologi Jaringan*, vol. 5, no. 1, 2020, doi: 10.30743/infotekjar.v5i1.2141.
- [24] M. Orken, O. Dina, A. Keylan, T. Tolganay, and O. Mohamed, "A study of transformer-based end-to-end speech recognition system for Kazakh language," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-12260-y.
- [25] C. and M. Methods in Medicine, "Retracted: An Analytical Study of Speech Pathology Detection Based on MFCC and Deep Neural Networks," *Comput Math Methods Med*, vol. 2023, pp. 1–1, Dec. 2023, doi: 10.1155/2023/9829813.
- [26] J. Song, Y. Gao, Y. Huang, and L. Chen, "Being friendly and competent: Service robots' proactive behavior facilitates customer value co-creation," *Technol Forecast Soc Change*, vol. 196, p. 122861, Nov. 2023, doi: 10.1016/j.techfore.2023.122861.
- [27] H. Kreuzen, D. Dull, V. de Rover, and R. Span, "Can AI Powered Speech-to-Text and Text-to-Speech techniques limit the interviewer bias in sensory and consumer research?," *Food Qual Prefer*, vol. 107, Apr. 2023, doi: 10.1016/j.foodqual.2023.104828.
- [28] X. Li, D. Ma, and B. Yin, "Advance research in agricultural text-to-speech: the word segmentation of analytic language and the deep learning-based end-to-end system," *Computers and Electronics in Agriculture*, vol. 180. Elsevier B.V., Jan. 01, 2021. doi: 10.1016/j.compag.2020.105908.
- [29] A. M. Mutawa, "An end-to-end Tacotron model versus pre trained Tacotron model for Arabic text-to-speech synthesis," *Journal of Engineering Research*, Aug. 2023, doi: 10.1016/j.jer.2023.08.016.
- [30] H. A. Younis, N. I. R. Ruhaiyem, W. Ghaban, N. A. Gazem, and M. Nasser, "A Systematic Literature Review on the Applications of Robots and Natural Language Processing in Education," *Electronics (Switzerland)*, vol. 12, no. 13. Multidisciplinary Digital Publishing Institute (MDPI), Jul. 01, 2023. doi: 10.3390/electronics12132864.
- [31] R. Fu *et al.*, "Transformer-BLS: An efficient learning algorithm based on multi-head attention mechanism and incremental learning algorithms," *Expert Syst Appl*, vol. 238, Mar. 2024, doi: 10.1016/j.eswa.2023.121734.
- [32] A. Vaswani *et al.*, "Attention Is All You Need," Jun. 2017, [Online]. Available: <http://arxiv.org/abs/1706.03762>
- [33] R. S. A. Pratama and A. Amrullah, "ANALYSIS OF WHISPER AUTOMATIC SPEECH RECOGNITION PERFORMANCE ON LOW

RESOURCE LANGUAGE,” *Jurnal Pilar Nusa Mandiri*, vol. 20, no. 1, pp. 1–8, Mar. 2024, doi: 10.33480/pilar.v20i1.4633.

- [34] J. Peng, H. Ye, Q. He, Y. Qin, Z. Wan, and J. Lu, “Design of Smart Home Service Robot Based on ROS,” *Mobile Information Systems*, vol. 2021, 2021, doi: 10.1155/2021/5511546.
- [35] Ahmad Reinaldi Akbar, “PERANCANGAN SISTEM KOMUNIKASI DUA ARAH PADA HUMANOID ROBOT BERBASIS ALGORITMA DEEP LEARNING,” Palembang, 2022.
- [36] R. S. Alkhaldeh, “DGR: Gender Recognition of Human Speech Using One-Dimensional Conventional Neural Network,” *Sci Program*, vol. 2019, 2019, doi: 10.1155/2019/7213717.