



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
PROGRAM STUDI SI TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

FORM PENGAJUAN JUDUL



Nama : Gabryelle Ninna Deffanya Siahaan

NIM : 211402087

Judul diajukan oleh* : ☐ Dosen
☒ Mahasiswa

Bidang Ilmu (tuliskan dua bidang) :

1. Data Science and Intelligent System
2. Computer Graphics and Vision

Uji Kelayakan Judul** : ☒ Diterima ☐ Ditolak

Hasil Uji Kelayakan Judul :

Calon Dosen Pembimbing I: Annisa Fadhilah Pulungan, S.Kom., M.Kom
(Jika judul dari dosen maka dosen tersebut berhak menjadi pembimbing I)

Calon Dosen Pembimbing II: Rossy Nurhasanah, S.Kom., M.Kom

Paraf Calon Dosen Pembimbing I

Medan, 11 Januari 2025

Ka. Laboratorium Penelitian,

* Centang salah satu atau keduanya

** Pilih salah satu

(Fanindia Purnamasari, S.Tl., M.IT)

NIP. 198908172019032023



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

RINGKASAN JUDUL YANG DIAJUKAN

*Semua kolom di bawah ini diisi oleh mahasiswa yang sudah mendapat judul

Judul / Topik Skripsi	Implementasi Algoritma LSTM-RNN pada Speech Recognition dalam Aplikasi Mobile untuk Membantu Tunanetra Membaca Dokumen
Latar Belakang dan Penelitian Terdahulu	<p>Latar Belakang</p> <p>Tunanetra adalah individu yang tidak dapat melihat atau buta, sebagaimana dijelaskan dalam Kamus Besar Bahasa Indonesia (KBBI). Berdasarkan data dari <i>Australia-Indonesia Disability Research and Advocacy Network (AIDRAN)</i> pada tahun 2023, jumlah penyandang tunanetra di Indonesia mencapai 4 juta orang atau sekitar 1,5% dari total populasi. Namun, dari jumlah tersebut, hanya 1% yang bekerja di sektor formal. Hal ini menyatakan adanya kesenjangan aksesibilitas terhadap layanan publik dan literasi bagi tunanetra.</p> <p>Tunanetra dalam kesehariannya memiliki kesulitan untuk membaca dokumen fisik. Hal ini disebabkan oleh minimnya ketersediaan dokumen dalam format braille dan tidak semua orang menguasai bahasa braille. Selain itu, alat bantu baca yang ada cenderung memiliki harga yang tinggi. Akibatnya, tingkat kemandirian mereka dalam membaca dokumen menjadi rendah. (Surbakti, 2024).</p> <p>Pemanfaatan teknologi dapat membantu para tunanetra dalam meningkatkan kemandirian mereka untuk membaca dokumen. Beberapa teknologi yang dapat digunakan adalah <i>Optical Character Recognition (OCR)</i>, <i>Text-to-Speech (TTS)</i>, dan <i>speech recognition</i>. Penelitian yang dilakukan oleh Ab Wahab et al. (2021) mengembangkan aplikasi yang memungkinkan tunanetra mengambil gambar teks dengan kamera <i>smartphone</i> untuk kemudian diubah menjadi audio. Namun, aplikasi ini hanya mendukung pengenalan teks berbahasa Inggris dan tidak dilengkapi dengan fitur <i>speech recognition</i> untuk perintah (<i>voice command</i>). Proyek seperti iAssist (Sankar et al., 2023) telah mencoba mengintegrasikan OCR dan TTS, tetapi perangkat keras yang digunakan, masih Raspberry Pi, perangkat ini dapat dikembangkan dengan platform yang lebih umum seperti Android.</p> <p>Penelitian lain, seperti Sight Scribe (Pavithra et al., 2024), memanfaatkan kacamata pintar berbasis OCR untuk mengekstrak teks dan mengubahnya menjadi audio. Namun, perangkat ini memiliki harga yang tinggi dan belum familiar di kalangan tunanetra. Penelitian oleh Lee Jian (2024) mengembangkan aplikasi dengan OCR dan TTS menggunakan teknologi modern seperti Google ML Kit dan Flutter untuk pengembangan lintas platform. Penelitian ini juga mengusulkan penambahan fitur <i>speech recognition</i> untuk meningkatkan efisiensi aplikasi. Selain itu, penelitian oleh Fatimah et al. (2021) tentang <i>Digital Mathematics Learning Companion</i> menunjukkan bahwa penggunaan <i>speech recognition</i> yang digunakan untuk perintah sangat membantu kemandirian siswa tunanetra dalam aplikasi belajar matematika.</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Landasan untuk *voice command* adalah *speech recognition*. Penelitian yang dilakukan oleh Srivastava et al. (2021) mengenai *speech recognition* menggunakan *Fuzzy Hidden Markov Models* (HMM). Algoritma ini biasanya menggunakan teknik ekstraksi fitur seperti *Mel-Frequency Cepstral Coefficients* (MFCC), yang merupakan metode yang umum digunakan untuk merepresentasikan karakteristik suara dalam bentuk yang lebih diskriminatif. Penelitian terkait HMM juga dilakukan oleh Pooja (2021), penelitian ini bertujuan untuk menyediakan sistem yang memungkinkan interaksi dengan komputer melalui *voice command*. Dalam penelitian ini, HMM digunakan untuk mengenali pola dalam sinyal suara, sementara NLP digunakan untuk mengekstrak entitas dan memahami konteks dari perintah suara yang diberikan.

Pengembangan model untuk pengenalan suara menggunakan *deep learning* juga telah dilakukan. Penelitian yang dilakukan oleh Oruh et al. (2022) menggunakan *Long Short-Term Memory Recurrent Neural Network* (LSTM-RNN) untuk *Automatic Speech Recognition* menunjukkan akurasi yang memuaskan untuk *speech recognition* yaitu mencapai 99,36% dibandingkan dengan model *deep learning* yang lain. Prasad et al. (2021) juga melakukan penelitian untuk penerjemahan bahasa dari bahasa Inggris ke Telugu, pada penelitiannya membandingkan *speech recognition* dengan model RNN-LSTM dan *Convolutional Neural Network* (CNN). Hasilnya RNN-LSTM memiliki hasil akurasi yang lebih memuaskan dengan 93% dan CNN 66,11%.

Dengan penerapan *deep learning*, konteks dalam *speech recognition* menjadi lebih luas dan fleksibel dibandingkan dengan model yang tidak menggunakannya. Dan model yang memanfaatkan kombinasi algoritma RNN dan LSTM mendapatkan hasil yang efektif. Kombinasi ini memiliki keunggulan dalam mengatasi tantangan yang dihadapi oleh RNN tradisional, seperti masalah *vanishing gradient*, yang dapat mengurangi kemampuan model untuk menangkap informasi dalam urutan data yang panjang. LSTM, dengan mekanisme memori khususnya, memungkinkan model untuk mempertahankan informasi penting dari data yang berurutan, seperti sinyal suara. Dengan algoritma ini, fitur *voice command* dapat diimplementasikan untuk mempermudah tunanetra dalam mengoperasikan aplikasi dan membaca dokumen secara mandiri.

Penelitian ini akan mengembangkan sebuah aplikasi mobile berbasis Android untuk membantu tunanetra membaca dokumen dengan mengintegrasikan teknologi OCR dan TTS serta memanfaatkan algoritma LSTM-RNN untuk fitur *speech recognition*. Aplikasi ini akan dirancang agar dapat menerima perintah suara (*voice command*), sehingga tunanetra dapat dengan mudah menggunakan aplikasi tanpa perlu bergantung pada bantuan orang lain. Implementasi ini diharapkan mampu meningkatkan kemandirian tunanetra dalam membaca dokumen.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Penelitian Terdahulu			
No.	Penulis	Judul	Tahun
1.	Mohd Nadhir Ab Wahab, Ahmad Sufiril Azlan Mohamed, Abdul Syafiq Abdull Sukor, dan Ong Chia Teng	Text Reader for Visually Impaired Person	2021
2.	Ariel Matius Surbakti	Implementasi Pure Data pada Inovasi Buku Berbasis Suara untuk Meningkatkan Literasi Penyandang Tunanetra di SLB - A Tanjung Morawa	2024
3.	Clara Fatimah, Dwi Parinata, Arieska Efendy, dan Yuli Santika	Digital Mathematics Learning Companion (DMLC): Aplikasi Android Guru Pendamping Khusus Matematika Bagi Penyandang Tunanetra Berbasis Suara	2021
4.	Dr. S. Perumal Sankar, Renjith P K, Ahammed Suhail P.I, Aswathy P S, Nithya Mary K J, dan Sharon K J	iAssist–An Intelligent Reading Assistant for Visually Impaired	2023
5.	Pavithra A, Janavarshini G, Nooril Afina T, Hemadharshini R, Keerthana S, dan Dr. R. Arthy	Sight Scribe–Accessing Text Through Sound	2024



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	6.	Lee Jian Shen	Empowering the Visually Impaired with a Mobile Application for Document Reader	2024
	7.	Dr. R.K. Srivastava dan Digesh Pandey	Speech recognition using HMM and Soft Computing	2021
	8.	Mrs. Pooja B S	Voice Assist and Control through Hidden Markov Model (HMM) and Natural Language Processing (NLP)	2021
	9.	Jane Oruh, Serestina Viriri, dan Adekanmi Adegun	Long Short-Term Memory Recurrent Neural Network for Automatic Speech Recognition	2022
	10.	Bathaloori Reddy Prasad1 dan N. Deepa	Classification of Analyzed Text in Speech Recognition Using RNN-LSTM in Comparison with Convolutional Neural Network to Improve Precision for Identification of Keywords	2021
Rumusan Masalah	Tunanetra memiliki hak yang sama dalam mengakses pekerjaan dan layanan secara setara, namun kenyataannya mereka masih menghadapi berbagai kendala dalam kehidupan sehari-hari, terutama dalam mengakses informasi dan dokumen fisik. Hal ini disebabkan oleh keterbatasan ketersediaan dokumen dalam format braille, mahalnya alat bantu baca, serta minimnya teknologi yang dirancang khusus untuk kebutuhan mereka. Teknologi seperti <i>Optical Character Recognition</i> (OCR) dan <i>Text-to-Speech</i> (TTS) memang telah berkontribusi dalam meningkatkan aksesibilitas, namun masih memiliki keterbatasan, terutama dalam mendukung pengoperasian berbasis perintah suara (<i>voice command</i>). Maka, diperlukan pengembangan aplikasi untuk membaca dokumen bagi tunanetra yang dilengkapi dengan fitur <i>voice command</i> untuk penggunaan yang lebih efisien.			



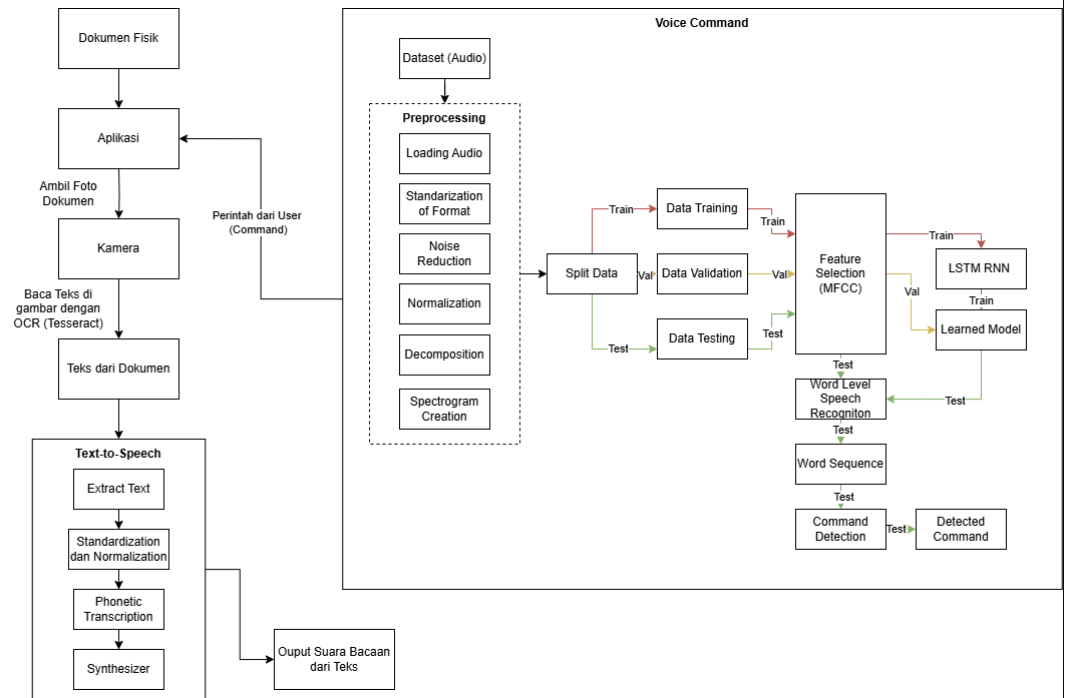
KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

Metodologi



Tahapan Metodologi:

1. Proses OCR

- Dokumen Fisik: Dokumen fisik adalah sumber *input* utama yang berisi teks yang perlu diproses.
- Aplikasi: Aplikasi berfungsi untuk mengambil gambar dokumen menggunakan kamera dan memprosesnya lebih lanjut.
- Kamera: Kamera digunakan untuk menangkap gambar dokumen fisik yang akan diproses.
- OCR (Tesseract): Proses OCR dilakukan menggunakan Tesseract, sebuah mesin OCR *open-source*. Fungsi OCR adalah mengekstrak teks dari gambar yang diambil. Hasil dari proses OCR berupa teks digital yang akan digunakan lebih lanjut dalam proses TTS

2. Proses TTS

- Extract Text*: Teks yang diekstrak dari gambar akan diproses lebih lanjut untuk dibaca secara suara.
- Standardization dan Normalization*: Teks yang diekstrak distandarisasi dan dinormalisasi untuk memastikan keterbacaan dan keakuratan suara yang dihasilkan.
- Phonetic Transcription*: Proses ini mengubah teks menjadi transkripsi fonetik untuk digunakan dalam proses sintetisasi suara.
- Synthesizer*: proses mengubah transkripsi fonetik menjadi output suara yang dapat didengar.
- Output*: Hasil akhir dari proses TTS berupa suara bacaan teks dari dokumen fisik.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

3. Pengolahan Perintah Suara (*Voice Command*)

- a. *Dataset (Audio)*: *Dataset* yang digunakan audio yang diambil dari 7 suara wanita dan 7 suara laki-laki. Setiap orang mengatakan 7 kata yaitu: “Halo”, “Foto”, “Info”, “Baca”, “Ulang”, “Berhenti”, dan “Kembali”. Satu kata diucapkan sebanyak 10 kali, sehingga total audio yang akan diolah sebanyak 980 audio.
- b. *Preprocessing*: *Preprocessing* yang dilakukan adalah *loading audio* (membaca audio), *standardization of format* (menyamakan format file), *noise reduction*, *normalization* (menyamakan sinyal audio), *decomposition* (sinyal audio diuraikan menjadi komponen yang lebih sederhana untuk diolah), *spectrogram creation* (mengubah sinyal audio dari domain waktu ke domain frekuensi dalam bentuk representasi visual).
- c. *Split Data*: *Dataset* dibagi menjadi *Data Training* (70%), *Data Validation* (10%), dan *Data Testing* (20%).
- d. *Feature Extraction (MFCC)*: *Mel Frequency Cepstral Coefficients (MFCC)* digunakan untuk mengekstrak fitur dari sinyal audio yang dapat digunakan untuk analisis audio.
- e. *LSTM RNN*: Model pembelajaran mesin berbasis RNN (*Recurrent Neural Network*) dan LSTM (*Long Short-Term Memory*) digunakan untuk melatih model.
- f. *Learned Model*: Model yang telah selesai dilatih menggunakan data training dan divalidasi dengan validation set.
- g. *Word Level Speech Recognition*: Proses mengenali urutan kata dari data audio yang diuji menggunakan model yang telah dilatih.
- h. *Word Sequence*: Urutan kata yang dihasilkan oleh model diolah lebih lanjut untuk membentuk makna atau perintah.
- i. *Command Detection*: Proses mendeteksi perintah berdasarkan urutan kata yang telah dikenali.
- j. *Detected Command*: *Output* berupa perintah dari *user*.

Referensi

- Fatimah, C., Parinata, D., Efendy, A., & Santika, Y. (2021). DIGITAL MATHEMATICS LEARNING COMPANION (DMLC): APLIKASI ANDROID GURU PENDAMPING KHUSUS MATEMATIKA BAGI PENYANDANG TUNANETRA BERBASIS SUARA. *Jurnal Ilmiah Matematika Realistik*, 2(1), 40–46.
- Kalkhoran, L. S., Tabibian, S., & Homayounvala, E. (2022). Detecting Persian speaker-independent voice commands based on LSTM and ontology in communicating with the smart home appliances. *Artificial Intelligence Review*, 56(7), 6039–6060. <https://doi.org/10.1007/s10462-022-10326-x>
- Lee, J. S. (2024). *Empowering the visually impaired with a mobile application for document reader* (Doctoral dissertation, UTAR). <http://eprints.utar.edu.my/id/eprint/6649>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

Nadhir Ab Wahab, M., Sufiril Azlan Mohamed, A., Syafiq Abdull Sukor, A., & Chia Teng, O. (2021). Text Reader for Visually Impaired Person. *Journal of Physics: Conference Series*, 1755(1), 012055. <https://doi.org/10.1088/1742-6596/1755/1/012055>

Oruh, J., Viriri, S., & Adegun, A. (2022). Long Short-Term Memory Recurrent Neural Network for Automatic Speech Recognition. *IEEE Access*, 10, 30069–30079. <https://doi.org/10.1109/access.2022.3159339>

Pavithra A, Janavarshini G, Nooril Afina T, Hemadharshini R, Keerthana S, & Dr. R. Arthy. (2024). Sight Scribe – Accessing Text Through Sound. *International Research Journal on Advanced Engineering Hub (IRJAEH)*, 2(05), 1157–1165. <https://doi.org/10.47392/irjaeh.2024.0159>

Pooja, B. S. (2021). Voice assist and control through hidden Markov model [HMM] and natural language processing [NLP]. *Int J Eng Res Technol (IJERT)*, 10(08).

Prasad, B. R. (2021). Classification of Analyzed Text in Speech Recognition Using RNN-LSTM in Comparison with Convolutional Neural Network to Improve Precision for Identification of Keywords. *Revista Gestão Inovação e Tecnologias*, 11(2), 1097–1108. <https://doi.org/10.47059/revistageintec.v11i2.1739>

Redwan, S. M., Md Rashed-Al-Mahfuz, & Hamid, M. E. (2023). Recognizing Command Words using Deep Recurrent Neural Network for Both Acoustic and Throat Speech. *European Journal of Information Technologies and Computer Science*, 3(2), 7–13. <https://doi.org/10.24018/compute.2023.3.2.88>

Sankar, S. P., PI, A. S., KJ, N. M., Renjith, P. K., Aswathy, P. S., & Sharon, K. J. iAssist–An Intelligent Reading Assistant for Visually Impaired.

Srivastava, Dr. R. K., & Pandey, D. (2022). Speech recognition using HMM and Soft Computing. *Materials Today: Proceedings*, 51, 1878–1883. <https://doi.org/10.1016/j.matpr.2021.10.097>

Surbakti, A. M. (2024). *Implementasi Pure Data pada Inovasi Buku Berbasis Suara untuk Meningkatkan Literasi Penyandang Tunanetra di SLB-A Tanjung Morawa* (Doctoral dissertation, Universitas Sumatera Utara).

Zaidi, B. F., Selouani, S. A., Boudraa, M., & Sidi Yakoub, M. (2021). Deep neural network architectures for dysarthric speech analysis and recognition. *Neural Computing and Applications*, 33(15), 9089–9108. <https://doi.org/10.1007/s00521-020-05672-2>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Zargar, S. (2021). Introduction to sequence learning models: RNN, LSTM, GRU. <i>Department of Mechanical and Aerospace Engineering, North Carolina State University.</i>

Medan, Januari 2025
Mahasiswa yang mengajukan,

(Gabryelle Ninna Deffanya Siahaan)

NIM. 211402087