MAKERERE UNIVERSITY

COLLEGE OF ENGINEERING, DESIGN, ART AND TECHNOLOGY

SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

BACHELOR OF SCIENCE IN COMPUTER ENGINEERING

Project

LUGANDA VOICE-CONTROLLED WHEELCHAIR

Students

OKIROR SAMUEL VINALD	**20/U/6640/PS**
ANKUNDA A ADDI	**20/U/6702/PSA**

Supervisor

DR. ANDREW KATUMBA

A FINAL YEAR ABBREVIATED PROJECT PROPOSAL SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF BACHELOR OF SCIENCE IN COMPUTER ENGINEERING

Background

Voice control technology traces its origins back to Automatic Speech Recognition (ASR) and has since undergone evolution, with emphasis on the field of Conversational Artificial Intelligence (AI). This evolution centers on the concept of speech intent classification, also referred to as speech command classification [1]. In recent years, voice control technology has emerged as a transformative force in human-machine interaction which enables users to interact with devices and services through natural language leading to voice-based interactions. The evolution of voice control technology led to applications in many areas of consumer electronics, home automation, and virtual assistive technology like Google Home, Siri, and Alexa [2]. The ability of these systems to understand and execute spoken commands relies heavily on speech intent recognition which is a fundamental component of conversational AI. Speech intent recognition aims to understand the user intent from spoken words which enables devices to perform actions based on the user's voice commands. Intent refers to the specific action a user wishes to communicate to a device through spoken words [3, 4]. Traditionally, speech intent recognition involves the use of Automatic Speech Recognition to transcribe spoken language into text, followed by intent classification based on the transcribed text [5]. However, another method involves a direct speech-to-intent approach that can directly map spoken language to intent, bypassing the intermediary step of transcribing the speech into text [6]. This technique involves the extraction of distinctive features from audio signals, emphasizing the spectral characteristics of the audio, such features include Mel Frequency Cepstral Coefficient (MFCC), Spectrogram, Chromagram, Spectral Centroid, and Spectral Roll-off. [7, 8, 9]. Significant advancements in assistive technology have led to the development of voice-controlled wheelchairs, which have proven highly effective in enabling users to operate them through spoken commands. These voice-controlled wheelchairs consist of various components, with the process commencing in speech recognition technology, where spoken words are transformed into digital data. This is followed by speech intent recognition, in which commands are subsequently relayed to a microcontroller integration like Arduino or Raspberry Pi [9] .The embedded systems act as the bridge between the trained model and the wheelchair's motors, responsible for translating the commands into actions, thus facilitating effective navigation [10]. However, the integration of speech intent classification into wheelchairs has primarily focused on major languages such as English, French, and German, with limited efforts directed toward low-resource languages [11]. Low resource languages are those that have relatively less resources for training Conversational AI systems. In these Luganda-speaking communities, the existing English voice-controlled wheelchairs create linguistic and cultural gaps and hinder the accessibility and community involvement for Luganda speaking individuals with mobility impairments.

The project aims to address this linguistic and cultural gap through building a foundation of Luganda voice-controlled wheelchairs. By recognizing and processing Luganda voice commands, Luganda is brought into the equation empowering Luganda-speaking individuals with mobility impairments to independently control their wheelchairs and navigate their surroundings with confidence and ease without anyone's assistance.

Problem statement

Physically disabled individuals in Luganda-speaking communities, primarily located in central Uganda, face barriers to independent mobility and community participation, due to the lack of accessible infrastructure and assistive technologies designed for their linguistic and cultural context [12]. The absence of Luganda language support in current assistive technologies, particularly voice-controlled wheelchairs, restricts the independence and quality of life of this population. The project expands the work on speech intent recognition in the form of wake word and keyword commands in low-resource languages by direct conversion of Luganda speech signals into motor maneuvering commands hence addressing the linguistic and cultural gap through building a foundation for Luganda voice-controlled wheelchair.

Justification

Luganda language, spoken by over 17% of the Ugandan population, is often overlooked in the development of voice control technology [13]. The existing voice-controlled applications and technology are designed for major languages like English, French, and German [11]. This creates a linguistic and cultural gap, hindering accessibility and community involvement, further limiting the independence and the quality of the lives for Luganda-speaking individuals with mobility impairments. Uganda's population includes an estimated 12% of physically disabled individuals, who face challenges in accessing education, healthcare, employment, and social activities [14]. The "Luganda Voice-Controlled Wheelchair" project aims to enhance the daily lives of these marginalized individuals by providing them with a voice-controlled solution that aligns with their linguistic and cultural context. By focusing on Luganda, a low-resource language, this project contributes to the artificial intelligence (AI) landscape since the field of speech intent recognition has majorly focused on major languages, leaving speakers of low-resource languages at a disadvantage. It also makes technology more accessible to a wider range of linguistic communities leading to a sense of independence, confidence, and community participation. The impact of the project exceeds the development of a voice-controlled wheelchair. It addresses linguistic and cultural gaps, and promotes inclusivity in technology. To the disabled Luganda speakers in Uganda, this project offers a means of independent mobility leading to improved quality of life and greater participation in their communities.

Objectives

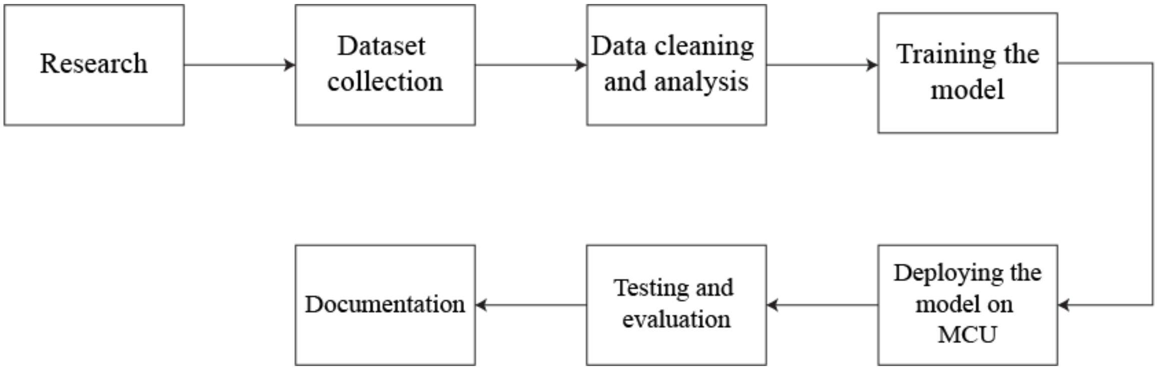
The main objective of this project is to develop a Luganda voice-controlled wheelchair system, enhancing mobility and independence for Luganda speakers with physical disabilities.

Specific objectives

- To collect and curate a Luganda voice dataset for common wheelchair movement commands.
- To develop, train and evaluate a speech intent recognition model to classify spoken Luganda commands.
- To integrate the trained model with a wheelchair to form a voice-controlled wheelchair system.
- To evaluate the system.

Methodology

Literature review will be conducted to study existing voice-controlled wheelchair technologies and their implementations. A further study will be conducted to identify common wheelchair control commands and an infrastructure for data collection will be to collect and store the audio dataset. The collected dataset will be split into training, testing and validation sets and then preprocessing will be done on these subsets. A speech intent recognition model based on neural networks will be developed and trained on the train set to recognize intents from the Luganda voice commands and the model performance will be evaluated using the test set to ensure its effectiveness in real-world scenarios. The significant attention will be given to the development and implementation of wake word and keyword commands to make the voice-controlled system more responsive to user needs. The trained model will be converted to the TensorFlow Lite and will be integrated with the wheelchair through a microcontroller board like Arduino or Raspberry Pi. enabling voice control of the wheelchair using Luganda commands. With all these steps, the main and specific objectives will be achieved and all the research activities, design documentation, code, data and results will be thoroughly documented in a report.



Expected results

A voice-controlled wheelchair system that can be navigated through spoken Luganda voice commands demonstrating that the integrated system correctly translates Luganda voice inputs to corresponding wheelchair movements and a final report documenting proof of concept, lessons learned and roadmap for applied research.

Proposed Time Schedule

Time	**2023**				**2024**			
Activity	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Literature review								
Data collection								
Analysis and prototyping								
Implementation								
Presentation and report								

References

[1] A. H. F. I. H. T. M. A. R. F. R. B. N. K. .. &. Z. Y. B. Pervaiz, "Incorporating noise robustness in speech command recognition by noise augmentation of training data.," 2020.

[2] S. a. O. A. a. T. M. Venkatraman, "Smart home automation—use cases of a secure and integrated voice-control syste," *Systems,* vol. 9, no. venkatraman2021smart, p. 77, 2021.

[3] D. C. a. L. S. a. V. M. L. D. S. a. B. C. De Andrade, "A neural attention model for speech command recognition," *arXiv preprint arXiv:1808.08929,* no. de2018neural, 2018.

[4] M. a. K. H. a. K. K. a. K. Y. Ayache, "Speech command recognition using deep learning," in *021 sixth international conference on advances in biomedical engineering (ICABME)*, IEEE, 2021, pp. 24--29.

[5] D. X. W. G. W. Y. Z. D. Guiming, "Speech recognition based on convolutional neural networks.," *ICSIP,* p. 708–711, 2019.

[6] A. L. A. O. J. S. a. M. T. N. Zeghidour, ""Soundstrearn: An end-to-end neural audio codec, IEEE/ACM "," *Transactions on Audio, Speech, and Language Processing,* vol. 30, pp. 495-507, 2021.

[7] D. a. S. W. a. N. M. a. M. M. Anggraeni, "The implementation of speech recognition using mel-frequency cepstrum coefficients (MFCC) and support vector machine (SVM) method based on python to control robot arm," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, 2018, pp. 012-042.

[8] M. R. a. H. M. M. a. H. M. Z. Hasan, "How many Mel-frequency cepstral coefficients to be utilized in speech recognition? A study with the Bengali language," *The Journal of Engineering,* vol. 2021, no. Wiley Online Library, pp. 817-827, 2021.

[9] Y. a. P. A. a. J. E. Singh, "Features of speech audio for deep learning accent recognition," 2019.

[10] T. K. Hou, "Arduino based voice controlled wheelchair.," *Journal of Physics,* vol. Vol. 1432, 2020.

[11] I. a. P. A. a. G. R. Galkin, "Data evaluation of novel modular wheelchair motor," in *2020 27th International Workshop on Electric Drives: MPEI Department of Electric Drives 90th Anniversary (IWED),* IEEE, 2020, pp. 1-4.

[12] S. N. W. M. R. A. G. Ray, "Listen with intent: Improving speech recognition with audio-to-intent front-end," 2021.

[13] J. A.-O. a. Hasheem, "Progress and challenges," [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5443041/. [Accessed 22 10 2023].

[14] "The-official-and-unofficial-languages-of-uganda," [Online]. Available: https://bunnystudio.com/blog/the-official-and-unofficial-languages-of-uganda/#:~:text=Luganda.,up%2016.7%25%20of%20Uganda's%20population.. [Accessed 23 October 2023].

[15] "Living with disability," [Online]. Available: https://disabilityin.org/country/uganda/#:~:text=From%20the%20Uganda%20Bureau%20of,is%20living%20with%20a%20disability.. [Accessed 23 Oct 2023].

[16] C. e. a. Yi, "Applying wav2vec2. 0 to speech recognition in various low-resource languages," *arXiv preprint,* 2020.