INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue III, March 2025

Empowering Communication for Paralyzed Individuals and Spinal Cord Injury Patients: An Intelligent System With Eye Gaze Tracking, Voice Assistance, and Chat-Bot Integration

J. M. Chan Sri Manukalpa[1] , H. P. Dissanayake[2]

[1]Department of Industrial Management, University of Kelaniya, Sri Lanka

[2]Department of Physical Science, Rajarata University of Sri Lanka

DOI : https://doi.org/10.51583/IJLTEMAS.2025.140300011

Received: 17 March 2025; Accepted: 21 March 2025; Published: 02 April 2025

Abstract— Advancements in artificial intelligence (AI) and machine learning (ML) have paved the way for transformative solutions in healthcare, particularly for individuals with severe physical disabilities. This research presents an innovative assistive communication platform tailored to Sinhala-speaking patients in Sri Lanka, addressing the unique challenges faced by individuals with Amyotrophic Lateral Sclerosis (ALS) and spinal cord injuries (SCI). These conditions severely restrict mobility and communication, leading to social isolation and diminished quality of life. The proposed system integrates eye-gaze tracking, speech recognition, and natural language processing (NLP) technologies to provide inclusive and effective communication solutions. For ALS patients, real-time eye-gaze tracking powered by MediaPipe interprets gaze movements as inputs, processed by a dense neural network (DNN)-based chatbot that generates responses in both English and Sinhala text and voice outputs. For SCI patients, who often retain speech abilities but face physical limitations, the system leverages speech recognition and NLP to interpret spoken commands, translating them into Sinhala text and synthesized speech. The dual-system approach ensures inclusivity, catering to the larger SCI-affected population while remaining adaptable to the needs of ALS patients. Despite challenges such as precision in gaze tracking, complexities in Sinhala speech recognition, and cultural nuances in NLP, the system demonstrates significant potential to enhance independence, social inclusion, and overall quality of life for patients. Empirical validation highlights the system's effectiveness, with high accuracy in intent identification and response generation. This research underscores the importance of culturally adapted, scalable assistive technologies, offering a foundation for future innovations in underserved regions. By bridging communication barriers, this platform represents a transformative step toward empowering individuals with disabilities.

Keywords—Assistive Communication, Eye-Gaze Tracking, Speech Recognition, Natural Language Processing, Sinhala Localization

I. Introduction

Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) are changing healthcare. They are particularly helping those with severe physical disabilities. Conditions like Amyotrophic Lateral Sclerosis (ALS) and spinal cord injuries (SCI) limit mobility and communication. This often leads to social isolation and a lower quality of life. This research aims to create an assistive communication platform for Sinhala-speaking patients in Sri Lanka. It will use advanced technologies like eye-gaze tracking, voice recognition, and chatbot-based Natural Language Processing (NLP) [10, 11].

ALS is a progressive neurodegenerative disorder. It causes paralysis and eventually leads to the loss of communication abilities. In contrast, SCI is often caused by traumatic events like road accidents or falls. It results in severe physical limitations but usually preserves speech. In Sri Lanka, SCI is significantly more prevalent than ALS, reflecting global trends where SCI is a leading cause of disability. Worldwide, SCI affects millions annually, with approximately 20.6 million prevalent cases and 0.9 million new cases each year. Developing countries, including those in South Asia, report a high incidence of SCI, primarily due to traumatic causes. Sri Lanka reflects this trend, as road injuries are a major cause of SCI cases, especially among males and older adults. ALS, on the other hand, is rare and has a much lower prevalence both globally and in the region. Due to the significant impact of SCI in Sri Lanka, this research uses a dual-system approach. This approach addresses the needs of both SCI and ALS patients, ensuring inclusivity while focusing on the larger affected group [1].

The proposed system has two subsystems to meet the needs of different groups: eye-gaze tracking for ALS patients and speech recognition for SCI patients. This design is inclusive and adapts to each condition's specific requirements. For ALS patients, the system uses real-time eye-gaze tracking from MediaPipe to detect gaze movements. These movements act as user inputs. A chatbot, built using Dense Neural Networks (DNNs), processes these inputs. It generates responses in English, which are then converted into Sinhala text and voice outputs. This dual-output feature helps patients share their thoughts, needs, and emotions, even without speech. The system also converts Sinhala text into synthesized speech. This allows caregivers and family members to quickly grasp the patient’s intentions. Its intuitive design and focus on Sinhala-speaking users make it especially relevant in Sri Lanka [2].

For SCI patients, the system leverages speech recognition to interpret spoken commands in English. The chatbot handles commands and creates responses in English. These responses are then translated into Sinhala. The translated text appears on the

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue III, March 2025

interface and is also converted into Sinhala speech using text-to-speech technology. This dual-output method—showing Sinhala text and playing Sinhala voice—ensures clear communication. It helps caregivers and family members understand the patient's needs quickly. This solution makes communication easier and provides an accessible way for patients to express their needs, whether for medical help, social interaction, or emergencies.

Developing such a system presents unique difficulties. Eye-gaze tracking needs high precision and adaptability. This is vital for users with limited gaze control. Speech recognition in Sinhala adds more challenges because of the language's unique sounds and structure. The NLP chatbot must generate accurate responses. It also needs to adapt to Sinhala's cultural and linguistic details for a natural user experience. Additionally, text-to-speech conversion in Sinhala requires careful attention to pronunciation and clarity for effective communication.

This research marks a significant advance in assistive communication technologies for Sinhala-speaking patients in Sri Lanka. It empowers individuals with ALS and SCI to communicate better. This, in turn, boosts their independence, quality of life, and social inclusion. The next sections of this paper will cover the system's technical architecture, evaluate its performance through tests, and discuss its wider impact on healthcare and beyond.

II. Literature Review

This literature review explores Eye-Tracking Technologies and chat-bots in Education, Healthcare, and Mental Healthcare.

Eye tracking technologies

All Eye-tracking technologies digitize human interaction with their environment, including cognition, psychology, health, and computing. Three eye-tracking methods include direct-connected devices with infrared sensors, electrodes placed near eyes to measure electrical potentials, and computer vision algorithms that follow and identify human eyeballs in real-time. These methods are crucial for applications like eye-based communications and computer environment interactions. Tracking human eyes is essential for understanding cognition, psychology, computing, and health and for understanding human interactions with their environment [2].

Wassan et al. developed a home automation system based on an eye blink sensor for elderly and paralyzed patients. Using Arduino Nano and various sensors, the system recognizes the intent to blink, decodes the signal, and controls specific equipment accordingly. It includes Bluetooth, RF LINK Pair modules, a TCRT 5000 Eye Blink Sensor, and Arduino-compatible microcontrollers. The system also has a Bluetooth module for updates and alerts [3].

Heravian et al. Developed an IoT-based smart home system for Spinal Cord Injury (SCI) patients using a non-invasive technique that tracks eye movement. The system uses an algorithm that offers more accurate and faster connections compared to previous ones. Through eye tracking, individuals with SCI can control an IoT-based smart home. The command receipt, confirmation, and execution time is under 10 seconds. Face detection is based on features detected by MATLAB software and implemented with an Arduino Mega 2560 board for the IoT component [4].

Chatbots for Education

Chatbots are used in the educational sector, particularly in providing organizational support for exams, courses, and studies. The study focuses on developing a chatbot for professional guidance based on John Holland's theory and the RIASEC questionnaire and education. In the chatbot, students entering the job market with their undergraduate and graduate degrees are assessed for their dominant personality type [5]. Chatbot technology is one of the innovations that can bridge the education gap with technology in the e-learning context. A chatbot in learning can offer students the feeling of one-on-one interactions with the teachers [6]. A supportive system in friendly instancing that enables new students to make friends when joining a university. Chatbots are also helpful when it comes to social learning. Hence, students with various learning abilities can discuss various aspects of a given subject in front of the robot while returning to the previous version [7].

Chatbots in healthcare

The evolution of chatbots in artificial intelligence highlighted the shift to end-to-end neural networks and the encoder-decoder recurrent model around 2015. With examples such as "One Remission" for cancer patients and "Babylon" for common illnesses, it presents a survey of recent chatbot literature. It outlines a functional architecture for a healthcare assistance chatbot [8]. A chat messaging service known as 'Safedrugbot' can help by giving users and assisting health professionals,The doctor provided relevant information regarding drug use during breastfeeding [9]. Patients are reminded to take medications by the "Florence" chatbot, their pills [10].

Chatbots in Mental healthcare

There have been recent advancements in the use of chatbots in the mental care setting, mainly for improving social skills during a program involving the treatment of depression more than as a therapeutic tool. Much research has been conducted to determine the viability of employing chatbots in this domain. A review has been made to examine the available literature to see if chat agents or chatbots are currently used in psychiatry and their usefulness in diagnosing and treating different mental disorders. If chatbots are being utilized ethically and ethically, they could be a new type of tool in psychiatry [11]. A chatbot can also be utilized as a

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue III, March 2025

psychiatric counselor through emotional dialogue analysis and sentence generation [12]. A case study was performed on the efficacy of the chatbot in a sequence of mental health self-taught tutorials. This was mainly due to lack of sufficient number of psychological specialists despite the increasing need for mental counseling. The chatbot's effectiveness was highly beneficial, and it erased stressful feelings and kept motivating them, unlike a web course series [13]. Moreover, there is a concern that involves empathetic, health-oriented conversational chatbots for those with major depression [14].

III. Methodology

This section explains how we developed an assistive communication system. It combines Natural Language Processing (NLP) and eye-tracking technologies. The system aims to improve interaction for people with severe physical disabilities, like paralysis and spinal cord injuries (SCI). Patients can use eye movements or voice commands to express their needs. An AI-driven chatbot processes these inputs to provide meaningful responses and alerts for caregivers.

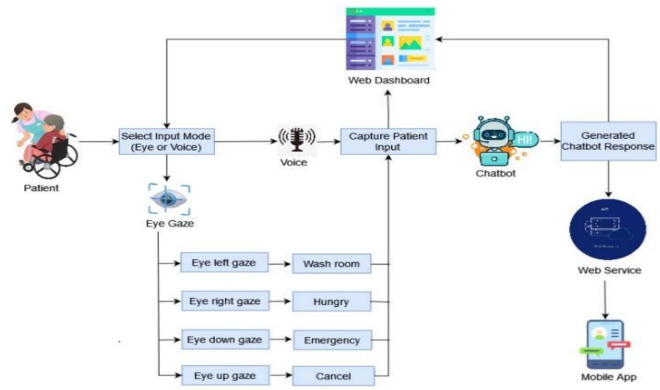


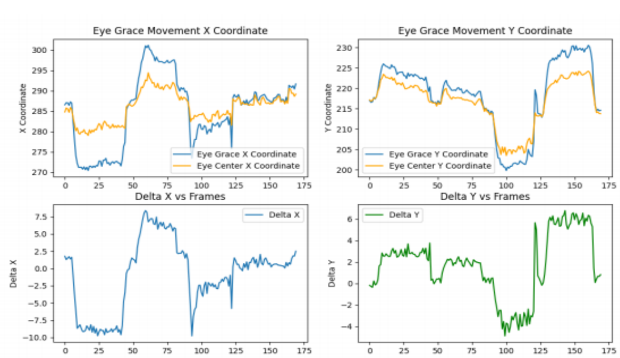
Fig. 1.System overview

Fig. 1. describes an assistive communication system that helps patients with mobility disabilities communicate with caregivers using their eyes or voice. Initially, the caregiver must set the mode according to the patient's preference. The patients communicate their wants, for example, help with food and the bathroom and call a caregiver through specific eye movements or vocalizations. It also controls and interprets eye or voice inputs to generate the required commands. The AI-integrated chatbot generates meaningful messages from the respective patients and transmits the messages to alert the caregiver through a mobile application. Which ensures a quick response is made. This system improves communication for mobility-restricted patients by validating information with AI and real-time notifications, thus making caregiving more accurate and efficient.

Eye gaze tracking

The system employs a webcam to capture video input, which serves as the primary source for gaze tracking. The captured frames are processed using the OpenCV library in conjunction with MediaPipe Face Mesh, enabling precise detection of facial landmarks. This setup is particularly focused on identifying key features in the eye region, which is essential for accurate gaze tracking. By leveraging the eye landmarks detected by MediaPipe, the system calculates the position of the iris relative to the eye socket. This calculation determines the relative gaze direction, such as up, down, left, or right, allowing the system to interpret sustained gaze positions as specific commands or requests.

To achieve this, a method was developed to determine the horizontal and vertical positions of the iris relative to the eye by comparing its x and y coordinates with those of the eye center. This method returns a string indicating the iris's position, which is used to infer the user's gaze direction. However, initial evaluations revealed that the method's accuracy and robustness were insufficient, prompting further refinement to enhance its performance. This step is critical for ensuring reliable interpretation of gaze-based inputs, which forms the foundation of the system's functionality.



ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue III, March 2025

Fig. 2. Action Triggers Based on Gaze Position

Fig. 2 illustrates the modified approach implemented to address the issues of accuracy and robustness observed in the initial system. The revised method calculates the horizontal and vertical deltas between the iris center and the eye center. The system then evaluates which delta—horizontal or vertical—has the greatest magnitude, ensuring that the primary direction of movement is prioritized. The iris position is subsequently categorized based on whether the absolute value of the largest delta exceeds a predefined threshold. This refinement enhances the precision of gaze direction detection, enabling a more reliable interpretation of user inputs.

Furthermore, the system tracks the gaze position for a duration exceeding two seconds to trigger corresponding actions. For instance, a prolonged gaze to the left may be interpreted as a request to use the washroom, while a gaze to the right could indicate hunger. This temporal threshold ensures that accidental or transient gaze movements are filtered out, allowing the system to respond only to deliberate and sustained user inputs. This approach significantly improves the system's functionality and usability for individuals relying on gaze-based communication.

Speech Recognition and Response Generation

The system integrates advanced technologies to enable seamless and efficient interaction, beginning with speech recognition. User inputs are captured via a microphone and processed using the Speech Recognition library, which converts spoken commands into text through Google's Speech Recognition API. This ensures high accuracy in transcribing user inputs, forming the foundation for subsequent processing.

The transcribed text is then passed to the Natural Language Processing (NLP) component, which employs a trained neural network to perform tokenization, lemmatization, and text classification. These steps enable the system to accurately identify the user's intent by breaking down the input into meaningful components and analyzing its context. Once the intent is determined, the system proceeds to response generation and action execution. It selects an appropriate response from a predefined set of options and triggers corresponding actions, such as sending alerts to caregivers or executing specific commands.

This streamlined workflow ensures that communication is both efficient and accurate, tailored to the unique needs of the user. By combining speech recognition with advanced NLP techniques, the system provides a robust and user-friendly solution for enhancing communication capabilities in individuals with mobility impairments.

Chat bot creation

This section outlines the process of designing and training a neural network-based chatbot model capable of interpreting natural language inputs to accurately identify user intent. The development of the chatbot involves several critical steps, including dataset preparation, data processing, and the construction of a neural network model

Dataset preparation: The foundation of the chatbot lies in a structured dataset comprising predefined user intents formatted in JSON. Each intent is associated with multiple typical user expressions (patterns) and corresponding system-generated responses. The initial dataset, sourced from myintentnew.json, includes a variety of intent categories such as greetings, emergency requests, and general inquiries. This dataset enables the chatbot to recognize and respond to a diverse range of user inputs effectively.

Data processing: To prepare the data for training, a series of preprocessing steps are applied. Tokenization is performed using the NLTK library, which segments the text into individual words. Lemmatization is then carried out using NLTK's WordNetLemmatizer to reduce words to their base or dictionary form, ensuring consistency in the dataset. Filtering is applied to remove non-essential characters and punctuation, thereby cleaning the data. Finally, feature extraction involves creating a vocabulary list (words.pkl) from the lemmatized words, ensuring the uniqueness of each word. Each pattern is converted into a "bag of words" vector, where the presence or absence of a word is represented by 1 or 0, respectively. This transformation converts textual data into a numerical format suitable for machine learning models.

Neural Network Model: The chatbot model is constructed using a sequential architecture developed with Keras. This architecture consists of multiple layers designed to process the feature vectors derived from the text input. The model is structured to handle the complexities of natural language understanding, ensuring accurate intent classification and response generation. The following sections provide further details on the model's architecture and training process.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 128)               38528

 dropout (Dropout)           (None, 128)               0

 dense_1 (Dense)             (None, 64)                8256

 dropout_1 (Dropout)         (None, 64)                0

 dense_2 (Dense)             (None, 10)                650

=================================================================
Total params: 47434 (185.29 KB)
Trainable params: 47434 (185.29 KB)
Non-trainable params: 0 (0.00 Byte)
```

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue III, March 2025

Fig. 3. DNN for model.

The neural network model is designed with a structured architecture to effectively process input data and classify user intents. The input layer accepts input vectors whose dimensions correspond to the length of the vocabulary, ensuring that the model can handle the full range of textual inputs. To mitigate overfitting, the architecture incorporates two hidden dense layers with 128 and 64 neurons, respectively. These layers are followed by a dropout layer, which randomly deactivates a portion of neurons during training to enhance the model's generalization capabilities. The output layer consists of a dense layer with a neuron for each intent class. This layer employs the softmax activation function to generate probability distributions over the classes, enabling the model to predict the most likely intent based on the input data.

For model optimization, the Stochastic Gradient Descent (SGD) optimizer is utilized, configured with a learning rate of 0.01, momentum of 0.9, and Nesterov acceleration. This configuration ensures efficient convergence during training while maintaining stability. The model is compiled using the categorical cross-entropy loss function, which is well-suited for multiclass classification tasks. This loss function measures the discrepancy between the predicted probability distribution and the true distribution, guiding the model toward accurate intent classification. Together, these components form a robust framework for training the chatbot model, ensuring high performance and reliability in interpreting user inputs.

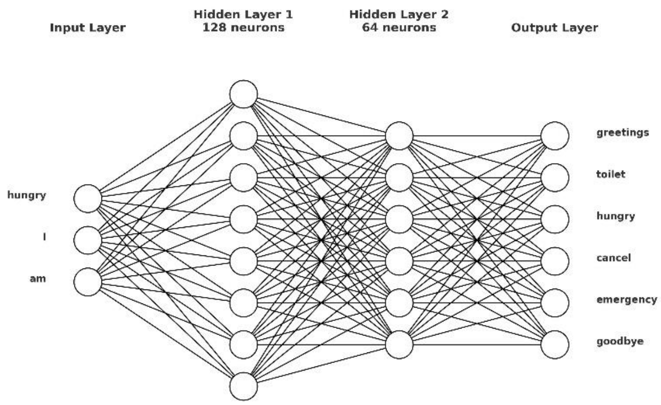


Fig. 4. DNN for model.

The architecture of the model is illustrated in Fig. 3. It is a feed-forward neural network comprising two hidden layers, where each neuron in the current layer is fully connected to every neuron in the previous layer, ensuring comprehensive information flow across the network. The input layer receives the bag-of-words representation of the patterns, with the number of neurons in this layer corresponding to the size of the vocabulary. The output layer consists of neurons representing specific intent tags, such as greetings, hungry, toilet, cancel, emergency, and goodbye. Each neuron in the output layer computes probabilities for the respective intents using the softmax function as the activation function. The softmax function, which serves as the final output of the network, is defined by the following formula:

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

where:

- σ = softmax function
- $\mathbf{z}$ = input vector
- e^{z_i} = standard exponential function for the input vector element z_i
- K = number of classes in the multi-class classifier
- e^{z_j} = standard exponential function for the output vector element z_j

Fig. 5. Softmax Function.

User input is first preprocessed and transformed into a bag-of-words representation before being fed into the neural network. The network processes this input and assigns probabilities to each neuron in the output layer, corresponding to different intent classes. These probabilities vary, reflecting the model's confidence in classifying the user's intent. The neuron with the highest probability is selected, and a random response from its associated intent category is generated and returned to the user. In cases where the user's input does not align with any predefined intents, a default response from the "greetings" category is provided. To ensure meaningful and confident responses, the system displays the final output only if the maximum probability assigned to the selected

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue III, March 2025

neuron exceeds a threshold of 0.75. This approach ensures that the chatbot provides accurate and contextually appropriate responses while minimizing irrelevant or low-confidence outputs.

Training: The training process begins by shuffling the dataset to prevent the model from learning any potential sequential biases inherent in the data order. The model is trained over 300 epochs with a batch size of five, striking an optimal balance between computational efficiency and convergence speed. This configuration ensures that the model learns effectively without excessive computational overhead.

The neural network training was conducted using Google Colaboratory (Colab), a cloud-based platform that provides free access to high-performance GPUs and TPUs. The training process leveraged the NVIDIA Tesla K80 GPU available within Colab's managed environment, which significantly accelerated computations due to its high memory bandwidth and parallel processing capabilities. The use of Colab's Jupyter notebook interface further enhanced flexibility, enabling seamless development, testing, and iteration of the machine learning model without the need for physical hardware infrastructure.

Throughout the training process, progress was monitored using loss and accuracy metrics across epochs. This monitoring ensures that the model learns effectively from the data, minimizing loss and maximizing accuracy. This approach optimized the training process for both performance and efficiency, ensuring robust model development within a scalable and accessible environment.

IV. Results & Evaluation

Preprocessing and training progress result of chatbot model

The preprocessing phase is critical in preparing the data for model training. It involves tokenization, lemmatization, and the creation of a bag-of-words representation. These steps are essential for transforming raw text into a structured format suitable for machine learning algorithms, enabling the extraction of meaningful insights from the data.

Tokenization was performed using the NLTK tokenize module, which splits the training data into individual tokens based on white spaces. Each token represents a word or punctuation mark. Fig. 6. illustrates the tokenization results, displaying the tokenized words alongside their corresponding classes.

Lemmatization is a linguistic normalization process that reduces words to their base or root form, known as the lemma. This step was applied to each word in the dataset, followed by the removal of punctuation marks. The resulting lemmas were then sorted alphabetically, and duplicates were eliminated to ensure uniqueness. This process enhances the efficiency of natural language processing tasks by reducing redundancy and maintaining consistency. The results of the lemmatization process are depicted in Fig. 7.

Neural networks require numerical inputs, as raw text cannot be directly processed. To address this, the bag-of-words approach was employed, where each sentence is represented as a vector of binary values corresponding to the presence or absence of words in the model's vocabulary. Specifically, a word present in the vocabulary is encoded as "1," while absent representation for the input string: "['wishing,' 'you,' 'a', 'fantastic,' 'day,' '!'].''words are encoded as "0." Fig. 8. demonstrates the bag-of-words.

```
[(['Hello', '!'], 'greetings'), (['Hi'], 'greetings'), (['Hey'], 'greetings'),
(['Good','morning'], 'greetings'), (['Good', 'afternoon'], 'greetings'),
(['How','are','you', '?'], 'greetings'), (['Good', 'Evening'], 'greetings'),
(['Hi','there'], 'greetings'), (['Howdy', '!'], 'greetings'), (['Nice', 'to', 'see',
'you'],'greetings'), (['How', "'s", 'it', 'going', '?'], 'greetings'), (['It', "'s", 'nice', 'to'
'meet', 'you'], 'greetings'), (['It', "'s", 'a', 'pleasure', 'to', 'meet', 'you'],
'greetings'), (['What', '", 's', 'new', '?'], 'greetings'), (['It', "'s", 'good', 'to', 'see',
'you'], 'greetings'), (['What', "'s", 'up', '?'], 'greetings'), (['I', 'need', 'to','use','the'
'washroom'], 'toilet'), (['I', 'have', 'to', 'go', 'to', 'the', 'toilet'], 'toilet'),
(['I','need', 'to', 'use', 'the', 'restroom'], 'toilet'), (['Can', 'you', 'help',
'me','get','to', 'the', 'bathroom', '?'], 'toilet'), (['I', 'need', 'to', 'take', 'a', 'quick',
'restroom', 'break'], 'toilet'), (['I', 'am', 'hungry'], 'hungry'),
(['I','have',"n't",'eaten', 'in', 'a', 'while'], 'hungry'), (['Can', 'you', 'help', 'me',
'with', 'some','food', '?'], 'hungry'), (['I', 'need', 'something', 'to', 'eat'], 'hungry')
(['I',"'m",'feeling', 'peckish'], 'hungry'), (['I', "'m", 'in', 'need', 'of', 'a', 'snack'],
'hungry'), (['I', 'could', 'really', 'use', 'something', 'to', 'eat'], 'hungry'), (['I', "'m"
'feeling', 'famished'], 'hungry'), (['I', "'m", 'craving', 'some', 'food'], 'hungry'),
(['My', 'stomach', 'is', 'growling'], 'hungry'), (['I', "'m", 'feeling', 'hungry', 'as',
'a', 'wolf'], 'hungry'), (['I', "'m", 'starving'], 'hungry'), (['Cancel', 'the', 'help',
```

Fig. 6. Tokenizing Result.

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue III, March 2025

```
["'m", "'s", "'ve", ',', 'Abort', 'Actually', 'Adios', 'Call', 'Can', 'Cancel',
'Emergency', 'Evening', 'Farewell', 'Good', 'Goodbye', 'Have', 'Hello', 'Help',
'Hey', 'Hi', 'How', 'Howdy', 'I', 'It', 'My', 'Nevermind', 'Nice', 'Quick', 'Reverse',
'See', 'Stay', 'Take', 'Thank', 'This', 'Undo', 'Until', 'Urgent', 'What', 'Wishing',
'a', 'action', 'afternoon', 'ahead', 'am', 'an', 'and', 'anymore', 'anyway', 'are',
'assistance', 'bathroom', 'be', 'break', 'care', 'changed', 'command', 'could',
'craving', 'critical', 'day', 'decided', 'do', 'eat', 'eaten', 'emergency', 'famished',
'fantastic', 'feeling', 'filled', 'food', 'for', 'get', 'go', 'going', 'good', 'great',
'growling', 'happiness', 'have', 'help', 'here', 'hungry', 'if', 'immediate', 'in', 'is',
'it', 'last', 'later', 'matter', 'may', 'me', 'meet', 'mind', 'morning', 'my', "n't",
'need', 'needed', 'new', 'next', 'nice', 'no', 'of', 'off', 'peckish', 'please', 'pleasure',
'pressing', 'previous', 'quick', 'really', 'reconsidered', 'remember', 'request',
'require', 'required', 'restroom', 's', 'safe', 'see', 'situation', 'snack', 'some',
'something', 'soon', 'starving', 'stay', 'stomach', 'take', 'thanks', 'the', 'there',
'time', 'to', 'toilet', 'up', 'urgent', 'use', 'washroom', 'well', 'while', 'with', 'wolf',
'wonderful', 'you', 'your', '']
```

Fig. 7.Lemmatization Result.

```
[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0]
```

Fig. 8.Bag of words

Table I. Training Progress by Epoch.

Epoch	Loss	Accuracy
50	0.2135	0.8873
100	0.1733	0.9155
300	0.1106	0.9577

As shown in TABLE I. the model exhibited significant improvement over the training period. At Epoch 50, the loss was 0.2135, with an accuracy of 0.8873. By Epoch 100, the loss decreased to 0.1733, and the accuracy increased to 0.9155. Further training up to Epoch 300 resulted in the lowest loss of 0.1106 and the highest accuracy of 0.9577, indicating successful model convergence and optimal performance.

Mode Selection Interface

The system offers two interaction modes: Eye Mode and Talking Mode, which can be selected via voice commands. For instance, if the user says "eye mode," the system activates the eye-tracking interface.

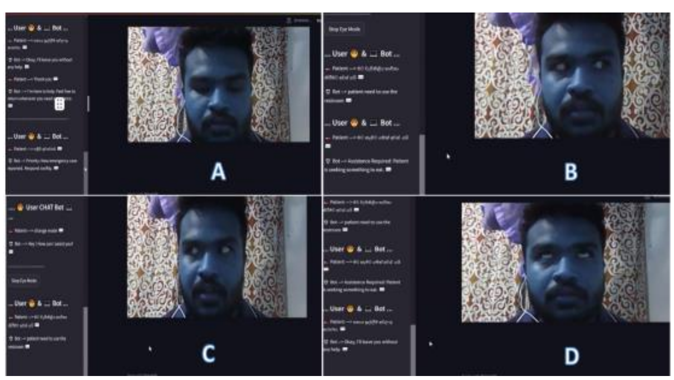


Fig. 9.Examples of eye gaze tracking emergency (A), hungry (B), restroom (C), cancel (D).

Fig. 9 demonstrates the eye gaze tracking functionality for various scenarios: (A) emergency, (B) hunger, (C) restroom needs, and (D) cancellation. The system interprets specific eye movements to generate corresponding messages when Eye Mode is enabled. For example, moving the eye to the left triggers a message indicating the need to use the washroom, as shown in (C). Moving the eye to the right generates a message indicating hunger, as depicted in (B). Moving the eye downward signals an emergency, as illustrated in (A). Moving the eye upward cancels the previous request, as shown in (D). This intuitive interface enables paralyzed individuals to communicate effectively using eye movements, enhancing their ability to express needs and interact with caregivers.

Validation by Medical Professionals

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue III, March 2025

The proposed system, titled "Empowering Communication for Paralyzed Individuals and Spinal Cord Injury Patients: An Intelligent System with Eye Gaze Tracking, Voice Assistance, and Chatbot Integration," has been rigorously evaluated by medical professionals with extensive experience in treating patients with spinal cord injuries, paralysis, and Amyotrophic Lateral Sclerosis (ALS).

Dr. Gayan Abeygunawardana, a medical officer specializing in anesthesia at Teaching Hospital, Karapitiya, Galle, has endorsed the system, emphasizing its alignment with the urgent needs of paralyzed and ALS patients in Sri Lanka. Dr. Abeygunawardana highlighted the system's potential to address both the physical and emotional challenges faced by these individuals, thereby improving their quality of life and access to essential resources.

Similarly, Dr. W.K. Jayani, also a medical officer at Teaching Hospital, Karapitiya, commended the system's innovative integration of IoT and machine learning technologies. She noted that the system effectively bridges communication gaps in resource-constrained settings, particularly through its use of eye-tracking technology and Natural Language Processing (NLP). Both experts praised the system's ability to empower patients to communicate more efficiently and independently, underscoring its transformative potential for individuals with disabilities in Sri Lanka.

User Feedback and System Evaluation

To ensure the system's usability and effectiveness, a structured evaluation process was conducted with caregivers, who assessed various aspects such as overall design, user interface intuitiveness, functional understanding, and perceived effectiveness in facilitating communication for paralyzed individuals. An open-ended section in the evaluation form captured detailed suggestions for improvement, enabling iterative refinements to align the system with the specific needs of its target users. This user-centered design approach highlights the critical role of stakeholder input in developing assistive technologies, enhancing their practicality and impact. The proposed system represents a significant advancement in assistive communication technologies, offering a robust and user-friendly solution for individuals with severe physical disabilities. Its validation by medical professionals and positive caregiver feedback underscore its potential to transform patient care and improve the quality of life for paralyzed individuals.

V. Conclusion

This research represents a significant advancement in assistive communication technologies, offering a transformative solution for individuals with Amyotrophic Lateral Sclerosis (ALS) and spinal cord injuries (SCI). By integrating eye-gaze tracking, speech recognition, and natural language processing (NLP), the system addresses the unique communication challenges faced by these patients. The eye-gaze subsystem provides ALS patients with a precise, non-invasive method to communicate, while the speech recognition and response generation system ensures effective interaction for SCI patients who retain speech abilities. The dual-system approach underscores the system's inclusivity, catering to the diverse needs of different patient groups. The NLP-based chatbot enhances usability by generating accurate and contextually appropriate responses, fostering greater social inclusion, independence, and an improved quality of life for users.

A key strength of this system lies in its localization for Sinhala and English-speaking users, making it particularly relevant to the Sri Lankan healthcare context. This culturally adapted approach not only addresses immediate communication barriers but also serves as a scalable model for other underserved regions facing similar challenges. The integration of real-time gaze tracking with an NLP-driven chatbot represents a novel contribution to the field, demonstrating the potential of combining advanced technologies to create impactful assistive solutions.

Looking ahead, further development of a fully Sinhala voice recognition and response system could significantly enhance the platform's functionality and accessibility. Continued research and investment are essential to expand the system's capabilities, improve its adaptability to different linguistic and cultural contexts, and promote its widespread adoption. By bridging communication gaps and empowering individuals with severe physical disabilities, this research lays the foundation for future innovations in assistive technologies, ultimately contributing to more inclusive and equitable healthcare solutions globally.

Acknowledgments

The Author would like to thank the Department of Industrial Management, University of Kelaniya, Sri Lanka, for providing support as required in completing the master of Information Technology thesis.

References

1. H, H, N, Kalyani., S, Dassanayake., Upul, Senarath. "Effects of paraplegia on quality of life and family economy among patients with spinal cord injuries in selected hospitals of Sri Lanka." Spinal Cord, 53 (2015).:446-450. doi: 10.1038/SC.2014.183
2. M.Ezzat, M.Maged, Y. Gamal, M. Adel, M. Alrahmawy and S. El-Metwally, "Blink-ToLive eye-based communication system for users with speech impairments," scientific reports, 17 May 2023..
3. Z. A. Wassan, "IoT Based Smart Home for Paralyzed Patients through Eye Blink," International Journal of Advanced Trends in Computer Science and Engineering, vol. 10, no. 2, 2021.
4. S.Heravian, N.Nouri, M.B.Taghadosi and S.M.H.Seyedkash, "Implementation of Eye Tracking in an IoT-Based Smart Home for Spinal Cord Injury Patients," Modern Care Journal ·, October 2019.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue III, March 2025

5. O. Zahour, E. H. Benlahmar, A. Eddaoui, H. Ouchra, and O. Hourrane, "A system for educational and vocational guidance in Morocco: Chatbot EOrientation," Procedia Computer Science, vol. 175, pp. 554–559, 2020.
6. F. Colace, M. D. Santo, M. Lombardi, F. Pascale, A. Pietrosanto, and S. Lemma, "Chatbot for E-Learning: A Case of Study," IJMERR, pp. 528– 533, 2018.
7. F. Clarizia, F. Colace, M. Lombardi, F. Pascale, and D. Santaniello, "Chatbot: An Education Support System for Student," Cyberspace Safety and Security, vol. 11161, A. Castiglione, F. Pop, M. Ficco, and F. Palmieri, Eds. Cham: Springer International Publishing, 2018, pp. 291– 302.
8. S. Ayanouz, B. A. Abdelhakim, and M. Benhmed, "A Smart Chatbot Architecture based NLP and Machine Learning for Health Care Assistance," Proceedings of the 3rd International Conference on Networking, Information Systems & Security, Marrakech Morocco, pp. 1– 6, March 2020.
9. I. Poola, "Making Artificial Intelligence (AI) and Disrupted Business Intelligence (BI) truly Conversational with Humanity Touch, Automated Descriptions and Talking Bots," International Journal of Advance Research, Ideas and Innovations in Technology, 2017.
10. B. Sharma, H. Puri, and D. Rawat, "Digital Psychiatry - Curbing Depression using Therapy Chatbot and Depression Analysis," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, pp. 627–631, April 2018.
11. A. N. Vaidyam, H. Wisniewski, J. D. Halamka, M. S. Kashavan, and J. B. Torous, "Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape," Can J Psychiatry, March 2019.
12. K. J. Oh, D. Lee, B. Ko, and H. J. Choi, "A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation," 2017 18th IEEE International Conference on Mobile Data Management (MDM), Daejeon, South Korea, pp. 371–375, May 2017.
13. T. Kamita, T. Ito, A. Matsumoto, T. Munakata, and T. Inoue, "A Chatbot System for Mental Healthcare Based on SAT Counseling Method," Mobile Information Systems, vol. 2019, pp. 1–11, March 2019.
14. N. Harilal, R. Shah, S. Sharma, and V. Bhutani, "CARO: An Empathetic Health Conversational Chatbot for People with Major Depression," Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, pp. 349–350, January 2020.