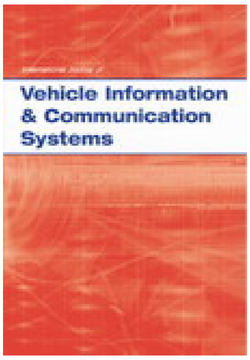


International Journal of Vehicle Information and Communication Systems

ISSN online: 1741-8208 - ISSN print: 1471-0242
https://www.inderscience.com/ijvics

Real-time voice-controlled human machine interface system for wheelchairs implementation using Raspberry Pi

Aymen Mnassri, Sihem Nasri, Mohammed Boussif, Adnane Cherif

DOI: 10.1504/IJVICS.2023.10060591

Article History:

Received:	21 October 2022
Last revised:	18 July 2023
Accepted:	02 August 2023
Published online:	25 January 2024

Real-time voice-controlled human machine interface system for wheelchairs implementation using Raspberry Pi

Aymen Mnassri*, Sihem Nasri,
Mohammed Boussif and Adnane Cherif

Department of Physics,
Faculty of Sciences of Tunis El Manar,
University of Tunis El Manar,
Belvedere, Tunis, Tunisia
Email: aymen.mnassri@gmail.com
Email: Nasri_Sihem@live.fr
Email: boussifmohamed1989@gmail.com
Email: adnen2fr@yahoo.fr
*Corresponding author

Abstract: The article describes the development of a wheelchair prototype designed to facilitate the mobility of people with disabilities. The proposed system is based on voice commands to ensure communication between humans and machines. The system consists of two modules. The first module involves the detection, processing and classification of actual voice signals acquired from a mobile phone. This module incorporates a robust and excellent speech recognition strategy. Indeed, the combination of Mel Frequency Cepstral Coefficients (MFCC) and the Discrete Wavelet Transform (DWT) in signal processing and feature extraction allows for better performance, achieving an effective recognition rate of 100% for an SNR of 5 db. The second module presents the mechanical design and development of the actual prototype, which enables real-time simulation of the first module. This module is based on a Raspberry Pi 3 board with a Linux operating kernel. Finally, tests carried out on the designed wheelchair prototype have demonstrated its efficiency, robustness and excellent response to critical situations, particularly obstacles.

Keywords: wheelchair; speech recognition; HMI; human-machine-interface; real-time; Raspberry Pi; mobile phone.

Reference to this paper should be made as follows: Mnassri, A., Nasri, S., Boussif, M. and Cherif, A. (2024) 'Real-time voice-controlled human machine interface system for wheelchairs implementation using Raspberry Pi', *Int. J. Vehicle Information and Communication Systems*, Vol. 9, No. 1, pp.81–102.

Biographical notes: Aymen Mnassri earned a PhD in Electrical Engineering from the Laboratory of Analysis and Processing of Electrical and Energy Systems Unit at the Faculty of Sciences, University of Tunis El Manar in 2020. With expertise in embedded systems, signal processing, voice recognition systems, and voice control, he currently serves as an associate teacher-trainer at Private Higher School of Engineering and Technology (ESPRIT).

Sihem Nasri earned a PhD in Electrical Engineering from the Laboratory of Analysis and Processing of Electrical and Energy Systems Unit at the Faculty of Sciences, University of Tunis El Manar in 2016. With expertise in renewable energy systems, power control, embedded system, power electronics, grid-connected system, modelling and control, she currently serves as an Assistant Professor at Faculty of Science of Tunis.

Mohammed Boussif earned a PhD in Electrical Engineering from the Laboratory of Analysis and Processing of Electrical and Energy Systems Unit at the Faculty of Sciences, University of Tunis El Manar in 2019. With expertise in embedded systems, signal processing, image watermarking, he currently serves as an assistant teacher at ISSAT-Sousse.

Adnane Cherif received the engineer, master and doctorate degrees from National Engineering School of Tunis (ENIT), in Tunisia. He is a university teacher in Electronics at Faculty of Sciences of Tunis (FST). He is responsible of the master specialty communication systems and networks. His field of interest concerns photovoltaic power system, digital signal processing, energy system designing and modelling.

1 Introduction

Nowadays, Automatic Speech Recognition (ASR) technology is undergoing tremendous development in various fields such as public, industry and society. Owing to these advanced technologies, on-board electronic devices used daily have become more complex and intelligent. In fact, the implementation of these technologies on embedded targets can allow humans to enter an era when Human-Machine Interaction (HMI) can be very flexible and enhanced via the ASR system. In this context, mobility aids are extensively used for people with limited mobility. Many research efforts are being made to design an HMI by exploiting physiological signals to improve the control of mechanical mobility aids, mostly wheelchairs.

The performance of the current voice control system is acceptable under well-controlled use conditions (no echo, no noise, etc.). However, these details are closely related to the difficulty and complexity of the planned activities. Currently, in a laboratory setting (assuming no noise constraints), a recognition rate of 99% for isolated words has been achieved in the single-speaker mode. However, when the voice control system is deployed in a real environment, several changes in learning conditions and system usage will occur, which can affect the recognition rate. The performance degradation depends on the level and type of noise present. Therefore, in general, these systems are not very robust to environmental fluctuations, even though they may not appear weak to the human ear.

There are several factors that can be regulated and optimised in an ASR system for real-time applications. One important factor is the Signal-to-Noise Ratio (SNR), which determines the accuracy of recognising the captured voice. To improve recognition, it is necessary to enhance the SNR by reducing background noise and improving the quality of the speech signal. Additionally, the size of the vocabulary plays a role in recognition accuracy. Reducing the vocabulary size, i.e., by shortening individual commands, can lead to improved recognition (Chevalier et al., 1995; Kamm et al., 1995). Another crucial aspect is the quality of the speech itself, as it directly influences recognition accuracy. Finally, real-time response is a necessary characteristic for an effective ASR system.

In this context, developing precise control commands such as 'backward', 'forward', 'right', 'left' and 'stop' within a suitable HMI presents a significant challenge. This challenge aims to improve the functionality and autonomy of individuals with severe disabilities, moving away from conventional wheelchair approaches.

Numerous advancements in the fields of rehabilitation and Artificial Intelligence (AI) have opened up new possibilities for the development of intelligent mobility aids (Champaty et al., 2014). These technologies specifically focus on the implementation of complex algorithms that are used to interpret and analyse HMIs and human cognition. Furthermore, conventional wheelchairs are unable to meet the significant needs of individuals who experience physical limitations in their lower limbs or suffer from conditions such as stroke or quadriplegia (Simpson, 2005). Therefore, a significant number of international research projects are aimed at studying the development of medical devices and rehabilitation aids for individuals with physical disabilities. The objective is to enhance their daily lives, enabling them to have minimal reliance on nurses and caregivers (Parikh et al., 2007). These projects seek to improve the functional capacity of patients and ameliorate their overall quality of life. This article specifically focuses on the development of a robust voice control system that can be easily integrated into a prototype wheelchair based on an electronic board. This wheelchair is designed to allow users to move freely by using voice commands in Arabic, specifying directions such as right, left, forward, backward and even adjusting the speed. To enhance safety, an HC-SR04 ultrasonic sensor is equipped to detect obstacles. When an obstacle is detected, the wheelchair immediately stops at a safe distance. Furthermore, the system can be promptly shut down using voice commands in critical situations. The structure of this paper is outlined as follows: Section 2 provides a literature review and highlights the contributions of this study. Section 3 presents the system model and design. Section 4 introduces the proposed approaches for monitoring and controlling the system, including an examination of important security scenarios. The results obtained are presented and discussed in Section 5 to demonstrate the reliability of the developed prototype. Finally, the paper concludes with a summary and outlines potential future work.

2 Literature review and contributions

2.1 Literature review

In the context of voice-controlled systems for wheelchair applications, several studies have been conducted, focusing on recognition, control and voice command approaches. Researchers have explored various methodologies to enable wheelchairs to respond to voice commands and facilitate movement and directional changes. Notably, Hou et al. (2020), Kannan and Selvakumar (2015) and Mirza et al. (2015) employ the Arduino microcontroller to process voice commands and generate operational signals. Building upon these works, He et al. (2019) proposed an Arduino-based design incorporating an audio receiver to control a robot car using voice commands and IoT technologies, specifically integrating Google Voice API. Additionally, Anwer et al. (2020) presented a novel system design that combines Raspberry Pi and Arduino boards to create an eye and voice-controlled interface for wheelchair users with significant disabilities. Bluetooth module (HC-05 BT) is utilised for transmitting voice commands to the Arduino board. Moreover, Ullah et al. (2019) investigated a single-equipment, multiple-application system for controlling an Arduino-based robot car using hand gesture recognition and a mobile application with voice recognition. These works significantly contribute to the advancement of voice-controlled systems for wheelchair applications, demonstrating various approaches and technologies that enhance mobility and independence for individuals with disabilities. Furthermore, Bakouri et al. (2022) proposed a system that steers a robotic wheelchair using a voice recognition system based on convolutional neural networks. Their study focuses on developing an efficient and reliable control mechanism for wheelchair navigation (Bakouri et al., 2022). Voznenko et al. (2018) presented a control system for a robotic wheelchair that utilises an extended Brain-Computer Interface (BCI), enhancing the wheelchair's control and navigation capabilities. Joshi et al. (2019) described the design of a voice-controlled smart wheelchair, integrating voice recognition technology into the control system to enable individuals to navigate and control the wheelchair using voice commands. Xu et al. (2023) introduced an eye-gaze controlled wheelchair that utilises deep learning algorithms and eye-gaze tracking technology for controlling the wheelchair's movements and navigation. These articles collectively contribute to improving mobility and independence for individuals with disabilities through innovative voice-controlled wheelchair systems (Xu et al., 2023).

In order to extract the contribution of this work, a summary table has been drawn up to present the works recently published in the field of real-time voice recognition in the last decade (shown in Table 1)

Table 1 Speech recognition system summary in the last decade

Reference	*Speech recognition model*	*On-board card*	*Data sets*	*Recognition rate*
Martinek et al. (2020)	LMS-ICA	FPCGA	12 commands	83% to 91%
Ismail et al. (2020)	SVM	Smart phone/ Raspberry pi 4	10 commands	79%
	SVM–DTW	Smart phone/ Raspberry pi 4	–	97%
Sharifuddin et al. (2020)	MFCC/CNN	Raspberry pi	10 commands	95%
	MFCC/SVM	Raspberry pi	10 commands	72%
Sidiq et al. (2015)	MFCC/SVM	Android	10 commands	58%
Batista et al. (2020)	MFCC/PSO–SVM	FPGA	30 commands	73% to 98%
He and Dong (2020)	FTT/HMM	DSP TMS 320	30 commands	95%
Ding and Shi (2016)	–/GMM	KINECT	10 commands	85%

2.2 Motivations and contributions

The objective of this study is to develop and implement a wheelchair prototype that can be controlled through voice commands using a speech recognition system. This work introduces several advancements compared to the previously mentioned studies by enhancing the Support Vector Machine (SVM) classification technique through the integration of the Mel Frequency Cepstral Coefficients (MFCC) feature extraction method with the Discrete Wavelet Transform (DWT) wavelet transform. The proposed recognition method demonstrates improved performance compared to other recognition algorithms (Sharifuddin et al., 2020), particularly in terms of real-time response optimisation and recognition rates, especially in noisy environments. Consequently, this article thoroughly investigates the recognition system with the aim of enhancing the classification rate in real-time disability assistance applications.

The main contributions of this work are:

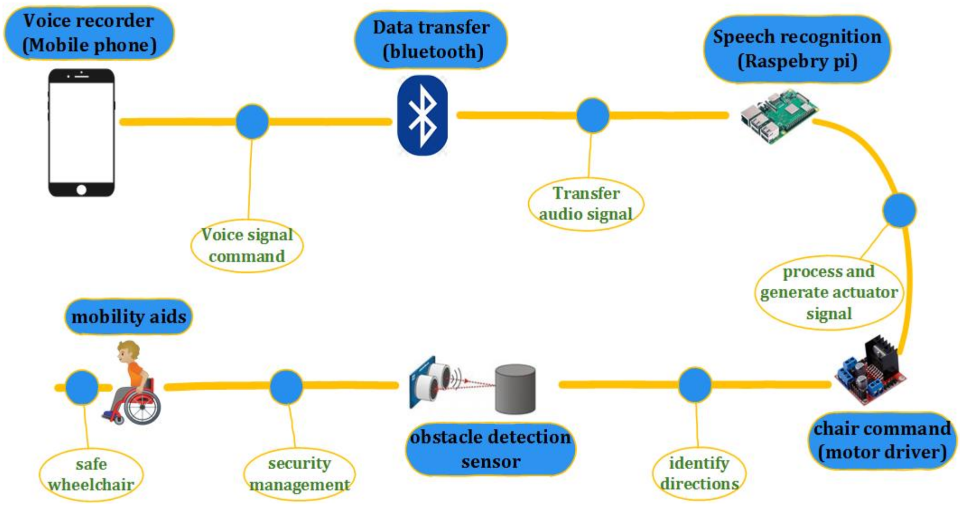
- *Utilisation of an Arabic speech database*: The work incorporates a database consisting of 10 voice commands recorded by 20 speakers, both male and female. This database enables the evaluation and testing of the proposed system.
- *Hybrid feature extraction approach*: The work introduces a novel approach that combines Mel Frequency Cepstral Coefficients (MFCC) and DWT to improve the recognition rate of the voice commands. This hybrid approach aims to enhance the accuracy and robustness of the speech recognition system.
- *Real-time implementation of the wheelchair prototype*: The wheelchair prototype is implemented using the Raspberry Pi3 board, providing a practical and efficient platform for the voice-controlled system.
- *Improvement of real-time response*: The proposed system utilises the Support Vector Machine (SVM) algorithm with the MFCC/DWT feature extraction approach to achieve improved real-time response in recognising voice commands. This enhancement contributes to the overall efficiency and effectiveness of the wheelchair control system.
- *Voice control via smartphone Bluetooth technology*: The wheelchair can be controlled using a smartphone application based on Bluetooth technology. This feature offers convenience and flexibility in operating the wheelchair remotely.
- *Real-time reliability measurement and obstacle detection*: The proposed prototype includes mechanisms for real-time reliability measurement and obstacle detection along a predefined trajectory. This ensures the safety and reliability of the wheelchair's navigation in various environments.

These contributions collectively demonstrate advancements in Arabic speech recognition, feature extraction techniques, real-time implementation, wheelchair control and obstacle detection, providing a comprehensive and practical solution for voice-controlled wheelchair systems.

3 System design and modelling

The proposed system consists of various equipment to better conceptualise and design the wheelchair prototype, on the one hand and to ensure communication between system elements to enhance its operation, on the other hand. The different system elements and their relationships and communication with each other are shown in Figure 1.

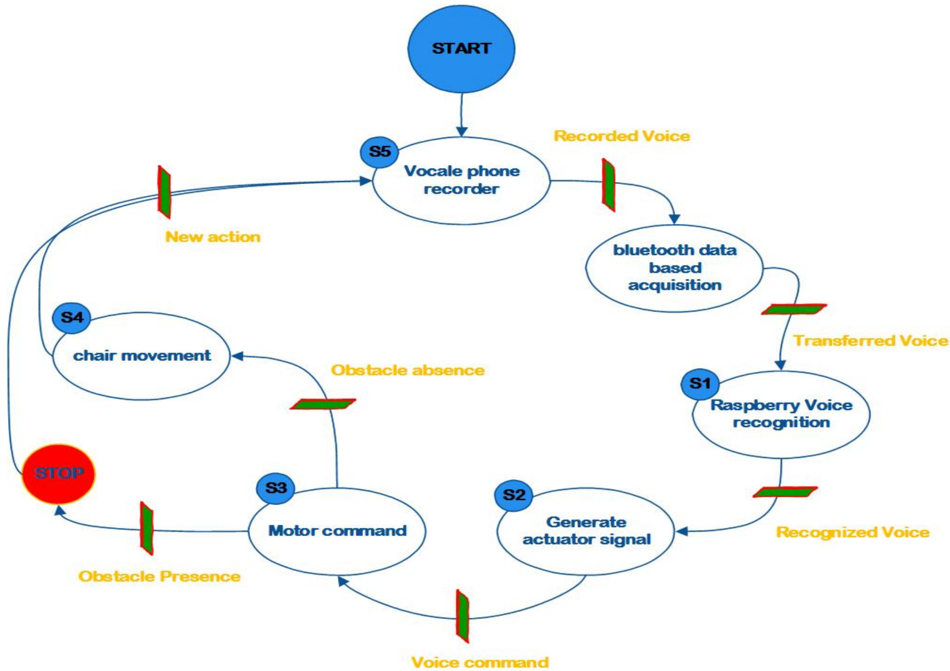
Figure 1 General system flow block diagram (see online version for colours)



To start the process, the wheelchair user initiates voice commands using an Android voice recording application on their smartphone. These commands are then wirelessly transmitted to the Raspberry Pi3 board via a Bluetooth module for further processing and classification. The Raspberry Pi3 board is specifically chosen for its affordability and suitability for real-time applications with simplified system complexity. It is equipped with a UNIX operating system and various packages that enable the analysis and processing of the received speech signal. This allows the generation of control signals to effectively manoeuvre the wheelchair based on the user's commands. The Raspberry Pi3 board is directly connected to the motor controller circuit, which is responsible for steering the wheelchair according to the provided Arabic commands. To ensure user safety, the system incorporates an HC-SR04 ultrasonic sensor for obstacle detection. Whenever an obstacle is detected, the wheelchair promptly comes to a halt at a predefined distance. Additionally, a security feature is implemented to enable immediate system shutdown through voice control in critical situations. In order to provide a more comprehensive understanding of the proposed system's operation, a step/transition diagram (see Figure 2) has been developed. This diagram visually represents the progression from one level to another once the corresponding transitions have been triggered. As illustrated in the diagram, the system's operation involves passing through five transient states, each with its own unique reaction:

- *S1:* Acquisition of voice commands from the smartphone to the Raspberry Pi board.
- *S2:* Voice recognition and decision-making process performed on the Raspberry Pi3.
- *S3:* Generation of actuator signals responsible for controlling the motor operation.
- *S4:* Motor response and execution of commands based on the voice input.
- *S5:* Wheelchair movement considering obstacle detection control for enhanced safety.

Figure 2 Functional flow chart of voice control system (see online version for colours)



4 System control and security

In this section, attention will be given to studying and analysing various operating system approaches that form the basis of the implemented prototype. Furthermore, an exploration of the methods adopted to guarantee the prototype's safety and provide a comfortable user experience will be conducted.

4.1 Speech recognition approach

The proposed speech recognition system incorporates a novel and robust algorithm that combines the extraction parameters of Mel-Frequency-Cepstral-Coefficients (MFCC) with the application of the DWT. The primary objective is to select and utilise significant features that effectively capture voice information. These extracted characteristics play a key role in constructing a reliable classification model, thereby enabling accurate recognition. Additionally, the DWT wavelet transform plays a crucial role in denoising the input signal, leading to improved classification performance (Aqil et al., 2017). Indeed, this method demonstrates its remarkable capability to effectively distinguish between signal and noise, surpassing the preservation of details compared to threshold denoising techniques. The applied noise elimination technique significantly improves the denoising effect, which is a key objective in real-time speech recognition systems.

The configuration of the adopted recognition system comprises various modules that fulfil necessary and crucial steps, such as: Processing the input voice signal from the smartphone and signals stored in the database for training and testing:

- Extraction of MFCC parameters based on DWT.
- SVM classification and decision-making.

The output of the ASR system is represented by a binary word, either 0 or 1. When the issued voice command corresponds to the analysed class, the system assigns a value of 1; otherwise, it assigns 0. The classification is achieved through the ‘One-vs.-All’ SVM approach (Duan et al., 2021), where each class is classified relative to the remaining data (other classes). This work focuses on ten distinct Arabic language voice commands, as depicted in Table. 2.

Table 2 Arabic voice command

Arabic commands	*Robot action*
تقدم	Go forward
أمام	Go forward with speed
أسرع	Speed up
يسار	Go left
يمين	Go right
استدر	Turn
توقف	Stop
تراجع	Go back with speed
خلف	Go back
واصل	Keep going

4.1.1 Feature extraction

The parameter extraction method involves computing the characteristics sequence of a given input speech signal over a short time frame. In this study, the feature extraction technique employed is a combination of MFCC and the DWT. This fusion of techniques aims to achieve robust and high-performance characteristics that are resilient to different types of noise, as illustrated in Figure 3. The developed parameter extraction method adheres to this algorithm to fulfil its objectives.

Algorithm:

Input: *Speech Signal s* (t)

Output: MFCC*-based DWT coefficients*

Function MFCC_DWT (input *parameters)*

Split into frames

Applying windowing (hamming)

Extract signal Approximation CA and Detail CD by applying DWT

Concatenate the obtained CA and CD.

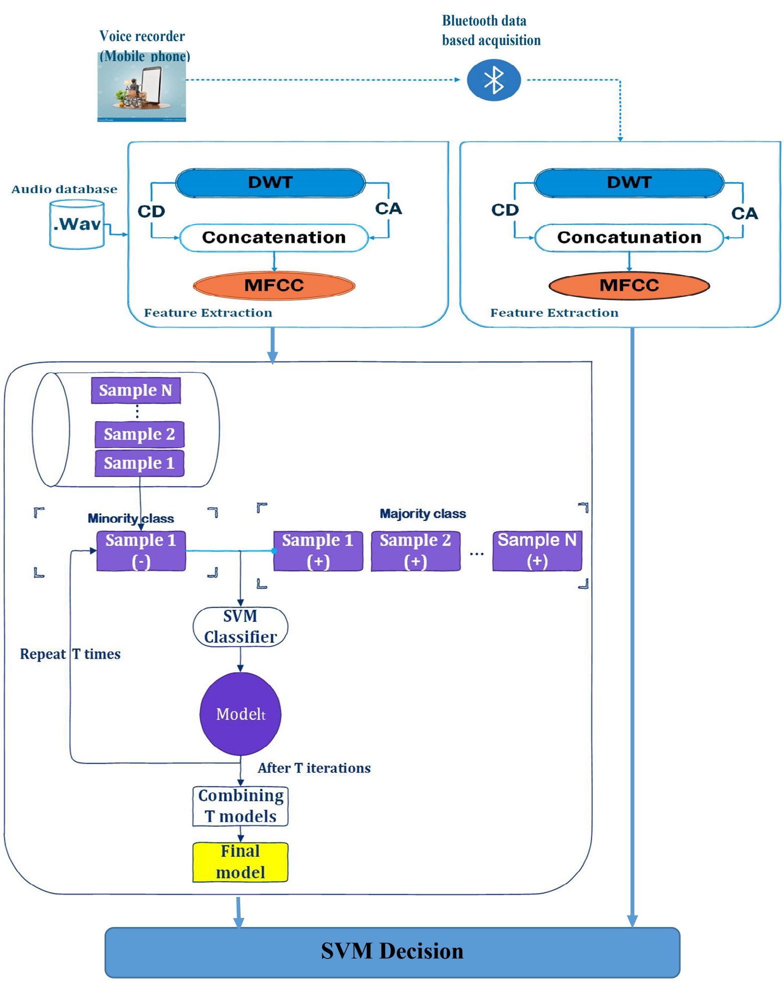
Calculation of the Mel scale according to the sampling frequency;

Output processing of log-energy

Discrete Cosine Transform (DCT) application

Organisation and establishment of frames to obtain the final MFCC coefficient

Figure 3 Proposed speech recognition approach (see online version for colours)



4.1.2 Classification approach

The primary objective of statistical learning is to identify the underlying relationship between the measurements of objects (referred to as ‘input variables’) and their corresponding properties (referred to as ‘output variables’). While the input variables are assumed to be observable for all objects of interest, the output variables are typically available only for a limited subset of objects. The goal is to estimate the dependency between the input and output variables accurately, enabling the prediction of output variable values for any object of interest.

In the context of classifying acquired speech command signals, the Support Vector Machine (SVM) technique is employed. The selection of SVM as the preferred technique is attributed to its robustness as a supervised machine learning method and its ability to efficiently classify diverse models. SVM has gained significant popularity due to its simplicity and its capability to achieve superior classification results. The core concept of SVM involves the separation of classes within a high-dimensional feature space, where a hyperplane is mapped to maximise the boundary between the classes, as defined (Yang et al., 2021):

$$\omega.x_i + b \geq +1 y_i = +1 \tag{1}$$

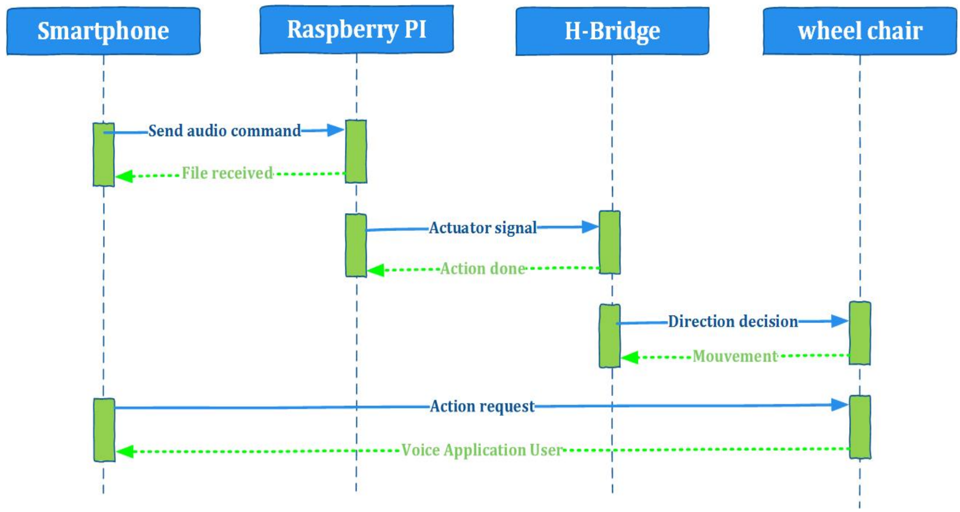
$$\omega.x_i + b \leq -1 y_i = -1 \ y_i = -1 \tag{2}$$

where ω is the normal to the hyperplane, $|b|/||\omega||$ is the perpendicular distance from the hyperplane to the origin and$||\omega||$ is the Euclidean norm of ω.

4.2 System devices communication

A sequence diagram is a descriptive diagram in the Unified Modelling Language (UML) that illustrates the flow of data and communication between different components within a system. It is used to specify the necessary operations and explain the interactions among the elements of the system under study. In the context of the given Figure 4, it depicts the information exchanges between the Smartphone, Raspberry Pi 3 card, H-bridge motor drive and the wheelchair prototype. The communication between the smartphone and the Raspberry Pi 3 is facilitated by a Bluetooth integrated module present in both devices. Once the input voice command is transmitted to the on-board card, it undergoes analysis and recognition using a real-time speech recognition algorithm that has been implemented. This process generates a command signal. The generated command signal is then sent to the H-Bridge motor drive, which is responsible for executing precise movements of the wheelchair. To ensure smooth data-flow exchange, each device involved in the communication sends a receipt confirmation to indicate the successful reception and processing of the information.

Figure 4 Sequence diagram of system control (see online version for colours)



4.3 System security

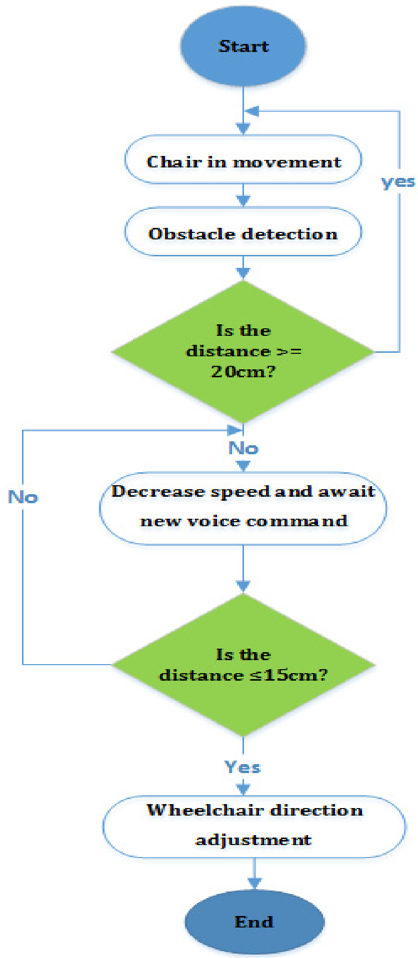
The primary objective of this study is to develop a highly effective control algorithm that prioritises the safety and comfort of individuals with disabilities during wheelchair movement. This algorithm relies on advanced obstacle detection technology using an ultrasonic sensor and operates in two distinct modes.

In the 'Automatic Mode', the algorithm takes immediate action when the ultrasonic sensor detects an obstacle at a predetermined safety distance of 20 cm. It activates a control audio signal to alert and protect the user while simultaneously reducing the wheelchair's speed. During this mode, the algorithm patiently awaits further voice commands from the user.

Upon approaching the obstacle and reaching a distance of 15 cm, the system seamlessly transitions to the second operating mode. In this mode, the algorithm doesn't necessarily stop the wheelchair immediately. Instead, it intelligently modifies the wheelchair's movement direction to navigate safely around the obstacle.

Figure 5 illustrates the different controlled states and transition conditions of the proposed algorithm. The integration of obstacle detection and responsive control mechanisms in this algorithm aims to significantly enhance the overall safety and comfort of individuals using wheelchairs with disabilities.

Figure 5 System security: obstacle detection (see online version for colours)

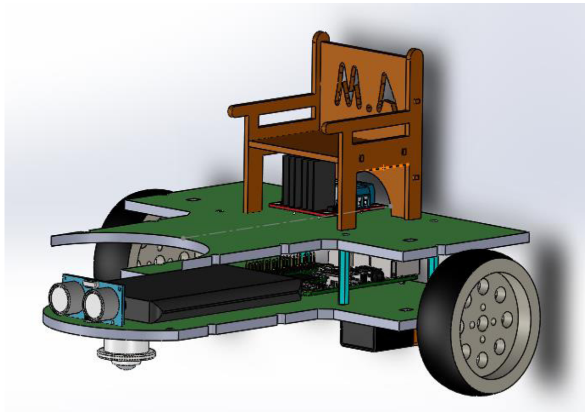


5 Findings and results

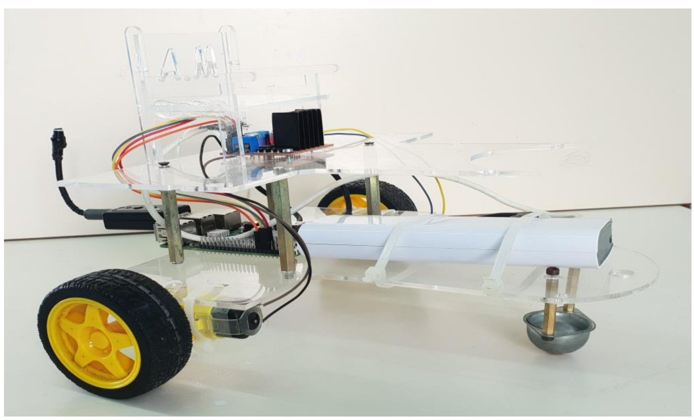
5.1 System implementation

The proper and successful distribution of mass among all essential components is crucial during the mechanical assembly of the prototype wheelchair. This is necessary to achieve a realistic appearance similar to that of a real wheelchair, considering the specified reduced metrics (Sprigle and Huang, 2015). During this phase, priorities include ensuring compatibility, aesthetics and patient comfort. In the design of the depicted wheelchair prototype shown in Figure 6, SolidWorks software is employed for modelling, planning, and distributing electronic components on the chassis. In Figure 6(a), all the necessary components, such as batteries, motors, controllers, and relays, are appropriately positioned on a designated platform as depicted in Figure 6(b).

Figure 6 (a) The solid works software model design (b) Complete wheelchair prototype mechanical assembly (see online version for colours)



(a)

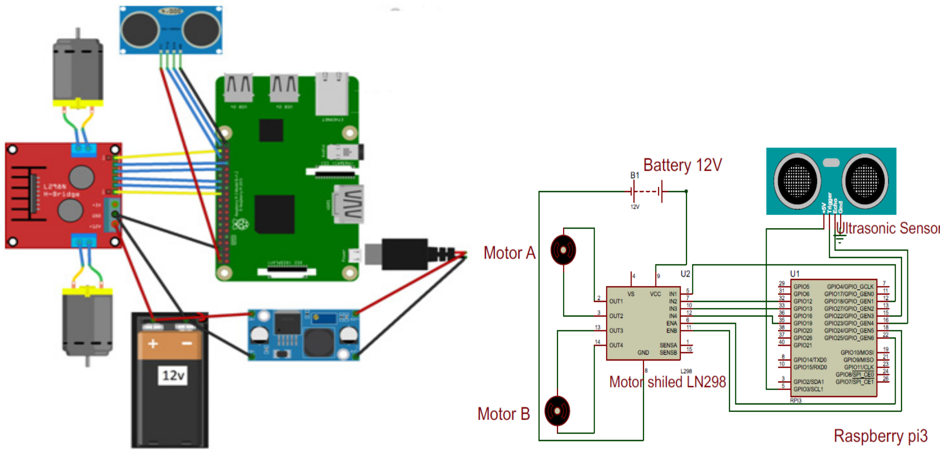


(b)

Figure 7 illustrates the circuit design that enables the control of wheelchair movement. In this design, the voice command is utilised to facilitate the shifting of the wheelchair. Figure 7(a) presents an external view of the circuit design, while Figure 7(b) showcases the circuit setup within the ISIS environment. The prototype integrates various hardware components, including a Raspberry Pi3 board, an LN298 motor module, an ultrasonic Obstacle Detector, a 12V battery and two DC motors. To establish the necessary connections, the Trigger and enable-pin of the ultrasonic Obstacle Detector are respectively linked to GPIO21 and GPIO22 pins of the Raspberry Pi. The VCC and ground connections are made to the 5V and ground Raspberry Pi pins. Similarly, the motor modules (IN1, IN2, IN3 and IN4) are connected to GPIO18, GPIO12, GPIO13 and GPIO19 pins of the Raspberry Pi board.

The motor module's power supply (12V) is connected to the positive terminal of the 12V battery, while the negative terminal is connected to the Raspberry Pi's ground port.

Figure 7 (a) Assembled prototype components (b) ISIS wheelchair circuit design (see online version for colours)



(a) (b)

Table 3 Specification of electronic components used in to design the proposed system

Components	*Specifications*
Raspberry Pi3	40-pins; Operating voltage: 4.75–5.25 volts
Android Mobile Application	Android compatible
L298N Motor Shield	Supply-Voltage range: 5V–35 V; Max Power: 25 W
HC-SR04 Ultrasonic Sensor	4 pins; Operating voltage: 3.3–5 V; Detection range: 2-400 cm
Power bank	Output Power 2 x 12 V/3.6 A
LM2596 buck converter	Input voltage: 4.5V-24V DC; Output current: 2A(MAX)

5.2 *Test results*

During this session, the performance of the recognition system is evaluated to demonstrate its effectiveness. The analysis is based on the results obtained from various tests. Firstly, different voice commands are tested under varying noise conditions (such as Babble, Circulation and Metro), with noise ratios (SNR) ranging from –10 to 5 dB. These tests aim to assess the system's performance in different noisy environments. Secondly, a comparative study is conducted to validate and establish the reliability of the proposed model innovations. The study compares the performance of the proposed model with existing approaches, using the provided Arabic data set. This analysis aims to highlight the superiority of the proposed model in terms of accuracy and effectiveness. Lastly, a system security test is performed, focusing on the obstacle detection and decision algorithm. The objective is to evaluate the system's ability to detect obstacles and make appropriate decisions to ensure user safety.

5.2.1 Voice command test

The figures provided below illustrate the test results conducted on 10 different voice commands across various noisy environments. The recognition system demonstrates high accuracy (RA) ranging from 98 to 100% when the Signal-to-Noise Ratio (SNR) is set to 5 dB, as shown in Figures 8, 9 and 10. These results highlight the system's effectiveness, which is influenced by SNR variations. Consequently, any decrease in SNR leads to a proportional reduction in system performance. To illustrate this further, let's consider the example of the Arabic voice command 'واصل' (meaning 'connect' in English). The system achieves a 100% RA at an SNR of 5 db. However, this accuracy decreases to 60% and 37%, respectively, at SNR values of 0 dB and –10 dB in the presence of babble noise. It's important to note that the accuracy values presented are impacted by the type of noise being simulated. For instance, with circulation and metro noise types, the RA remains at 99% for an SNR of 5 dB for the aforementioned voice command. However, at SNR values of 0 dB and –10 dB, the system demonstrates higher robustness against metro noise compared to other noise types. Figure 11 summarises the trend of RA evolution with varying SNR levels for different noise types. The results indicate that metro noise induces less performance degradation compared to other noise types, resulting in a lower error rate under such conditions.

Figure 8 RA of each command in babble noisy environment (see online version for colours)

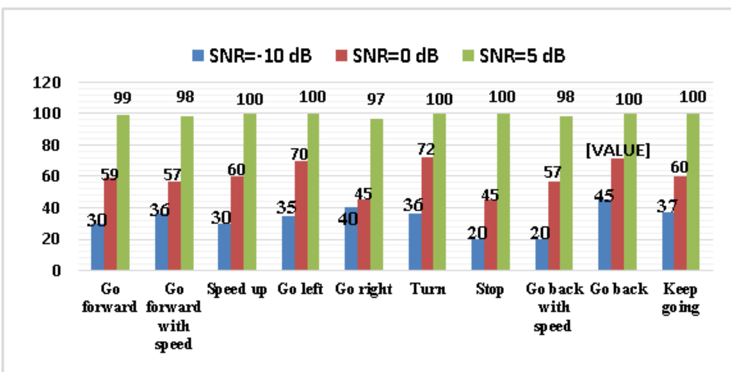


Figure 9 RA of each command in circulation noisy environment (see online version for colours)

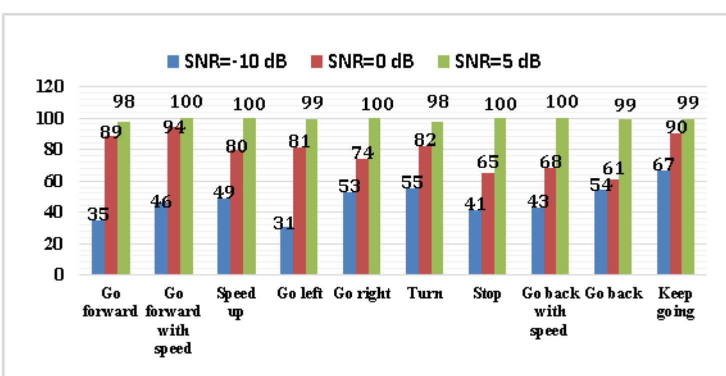


Figure 10 RA of each command in metro noisy environment (see online version for colours)

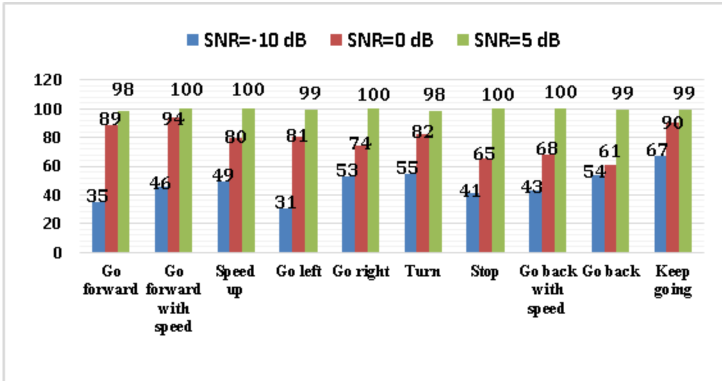
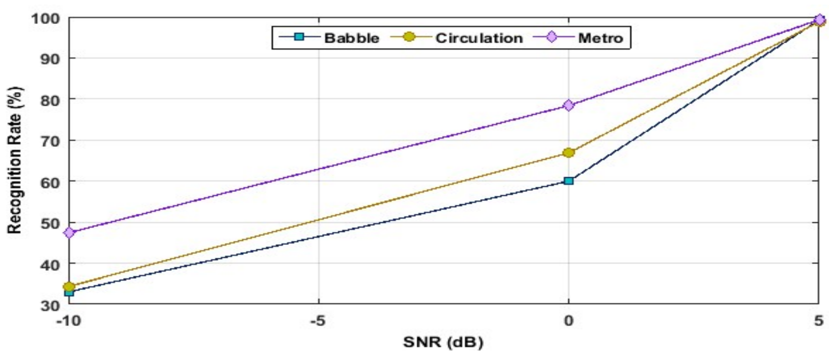


Figure 11 RA evolution versus SNR variation of different noise type (see online version for colours)



5.2.2 Comparative study

To conduct a thorough performance assessment of the proposed approach, a comparative analysis has been carried out against several existing state-of-the-art models for voice command recognition. This analysis aims to evaluate the effectiveness of the approach in relation to established methods. By comparing with these models, valuable insights can be gained into the strengths, weaknesses and overall performance of the approach. This rigorous evaluation framework provides a comprehensive assessment of the proposed approach's potential superiority in voice command recognition. The models considered for comparison include CNN (Sharifuddin et al., 2020), PSO-SVM (Batista et al., 2020), HMM (He and Dong, 2020) and GMM (Ding and Shi, 2016). To ensure a fair evaluation in an open-world scenario, the models have been re-evaluated using the data set. This re-evaluation allows for an assessment of the models' performance in a real-world context, considering the specific characteristics and challenges present in the data set. The aim is to provide an accurate and reliable assessment of the models' capabilities and their suitability for practical applications.

The evaluation of these systems is based on key metrics such as the True-Positive Rate (TPR) and False-Positive Rate (FPR). These rates are calculated using equations (3) and (4), respectively, enabling us to accurately measure the recognition accuracy and potential false detections.

$$TPR = \frac{TP}{TP + FN} \tag{3}$$

$$FPR = \frac{FP}{FP + FN} \tag{4}$$

The evaluation of the compared algorithms is based on various metrics, including TP (total number of correctly labelled monitored voice command test samples), TN (total number of correctly labelled unmonitored voice command test samples), FP (total number of unmonitored voice command test samples mislabelled as monitored voice commands) and FN (total number of monitored voice command test samples mislabelled as unmonitored voice commands) (Mao et al., 2022). To provide a comprehensive overview of the overall performance, the results of the compared algorithms are presented in Table 4. This table summarises the performance metrics and allows for a direct comparison between the algorithms in terms of their accuracy and effectiveness in correctly classifying voice commands.

Table 4 Performance evaluation of the tested algorithms

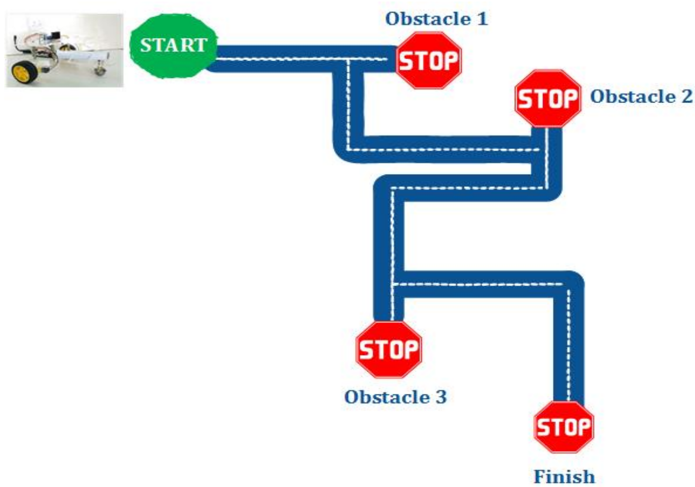
Metric	*Our Model*	*CNN*	*PSO–SVM*	*HMM*	*GMM*
Accuracy	99.4%	98.3%	76%	80.9%	67%
TPR	99.00%	97.4%	75.42	71.85%	65.09%
FPR	11.11%	11.11%	11.11%	11.11%	11.11%
Response time (MS)	201	216	209	225	223
Training Time (min)	10.1	12.4	20	8.2	11

Based on the results presented in Table 4, the model exhibits the highest accuracy of 99.4%, outperforming CNN with an accuracy of 89.73%. However, HMM, PSO-SVM and GMM achieve maximum accuracies of 80.9%, 76% and 67%, respectively. It is noteworthy that the performance of GMM and HMM in this study is lower compared to what has been reported in previous articles (He and Dong, 2020) and (Ding and Shi, 2016). This difference could possibly be attributed to variations in the implementation of hardware devices. Overall, the model consistently outperforms the other models by a margin of 1 to 32% in terms of classification accuracy. Additionally, when comparing the training times of the voice command using different models, PSO-SVM requires 20 minutes to achieve optimal accuracy, while CNN, HMM and GMM require 12.4, 8.2 and 11 minutes, respectively. In contrast, the model achieves an average time advantage of 2.8 minutes compared to the tested models, with a training time of approximately 10.1 minutes. Although HMM no longer exhibits the fastest training time, the performance achieved by the model remains significantly superior overall. It is important to note that while longer training times may be acceptable, as the learning phase is performed only once, the response time is a critical factor in evaluating real-time system quality. The proposed system demonstrates a faster response time for the pronounced commands, making it the most efficient among the tested models. This is primarily due to the less complex implementation of SVM, which results in shorter processing times.

5.2.3 System security test

This study presents a strategy for the safe and secure utilisation of the implemented wheelchair prototype. To assess and demonstrate the reliability of the investigated control strategy, a real-time operation test was conducted using the developed prototype, following the circuit path depicted in Figure 12. The path consists of three obstacles along the route from the departure point to the arrival point. The purpose of this test case is to analyse and discuss the system's reaction to obstacles in order to validate its performance. By evaluating the system's behaviour and responses during this test, we can effectively assess the effectiveness and reliability of the control strategy implemented in the wheelchair prototype.

Figure 12 Tested wheel chair trajectory (see online version for colours)



After detecting an obstacle, the system can respond in two different scenarios:

- The scenario 1 involves an immediate automatic stop command, which is activated when the distance between the wheelchair and the obstacle becomes approximately 20 cm.
- The scenario 2 is triggered when the system enters a critical shutdown state. This occurs when the distance decreases to 15 cm without any reaction. In such cases, the user is alerted and can intervene by providing an override voice command to assist in steering the wheelchair.

To further investigate the system's behaviour, several tests were conducted, considering two main factors: the speed of the wheelchair and the distance between the wheelchair and the obstacle. These tests aimed to assess how the system performs under different speed and distance conditions.

In Table 5, the performance of the system in scenario 1 is presented, focusing on two key parameters: time response and reactive distance. These parameters measure the system's effectiveness at different speeds (V1 and V2) when detecting three obstacles. The average time it takes for the system to detect the three different obstacles is

1.2 seconds when operating at speed V1. However, when the system operates at speed V2, the average detection time increases to 1.4 seconds. This implies that the system takes slightly longer to detect the obstacles at the higher speed. Reactive distance refers to the distance travelled by the system after receiving a stop order. In this scenario, the average stopping distance is approximately 2.4 cm when the system is operating at speed V1, while at speed V2, it increases to about 2.8 cm. This suggests that the system requires a slightly longer distance to come to a complete stop at the higher speed. These results provide valuable insights into the system's performance regarding detection time and stopping distance at different speeds. Considering these performance metrics is crucial when evaluating the efficiency and effectiveness of the system in scenario 1.

Table 5 Scenario 1 results

	Minimum starting speed V1		*Maximum starting speed V2*	
Obstacle	*Response time (s)*	*Reactive distance (cm)*	*Response time (s)*	*Reactive distance (cm)*
Obstacle 1	1	2	1.3	2.7
Obstacle 2	1.3	2.6	1.5	2.9
Obstacle 3	1.3	2.6	1.4	2.8
Average Value	1.2	2.4	1.4	2.8

In Table 6, the results obtained from scenario 2 are presented, highlighting the performance of the system. For each detected obstacle, three voice commands are considered. For example, when obstacle 1 is identified, the user has the option to choose from three commands: "تراجع" ",توقف," or "يمين". These commands are carefully selected based on the available paths the wheelchair can navigate. The table also includes data on the system's response time and reactive distance for speed values V1 and V2. In the case of obstacle 1, when the system is operating at speed V1, the average response time, measured from the launch of the voice command until the system reacts, is approximately 1.3 seconds. Additionally, the average reactive distance for obstacle 1 is about 2.6 centimetres. Similarly, for the speed value V2, the average response time decreases to 1.23 seconds and the mean reactive distance is approximately 2.5 centimetres for all three commands related to obstacle 1.

Table 6 Scenario 2 results

		V1		*V2*	
Obstacle	*Possible Commands*	*Response time (s)*	*Reactive distance (cm)*	*Response time (s)*	*Reactive distance (cm)*
	توقف	1	2	1.2	2.4
Obstacle 1	**تراجع**	1.6	3.2	1.4	2.8
	يمين	1.3	2.6	1.1	2.3
	توقف	1	2	1.2	2.4
Obstacle 2	**وراء**	1.3	2.7	1.5	3.1
	يمين	1.2	2.4	1.1	2.2
	توقف	1	2	1.3	2.6
Obstacle 3	**استدر**	1.5	3	1.	2.4
	وراء	1.3	2.5	1.6	3.2

These results provide insights into the system's performance in scenario 2, considering response time and reactive distance for different speeds. It is important to consider these metrics when evaluating the system's efficiency and effectiveness in real-world scenarios.

6 Conclusions

The development of a real-time voice control system devoted to help the disabled person to move independently in different environmental situations constitutes the main objective of this paper. In this context, intelligent algorithms based on speech processing tools for automatic speech recognition is investigated and deployed firstly. In second time, the proposed system algorithms are implemented in a real-time wheelchair prototype. Hence, this work deals with two important ideas. The first concerns the development of a new approach of voice recognition based on a combination between the MFCC feature extraction and the DWT wavelet transform to make the system robust against any noise environment. This approach has allowed to improve the performance of Support Vector Machine (SVM) model-based recognition technique which achieves a recognition rate of 100% in clean environment and about 98% in noisy one with different noise type at 5 dB of SNR. The second idea relies on the development of a smart control algorithm to ensure the safe service and user security. This algorithm has underlined the obstacle-detection-based approach to enhance the system performance and reaction according two different scenario cases study. Finally, after having designed and implemented the wheelchair prototype, several validation tests have been made to prove the efficiency and the robustness of the proposed system to better recognise voice commands on the one hand and to make sure and check the reliability of the adopted security approach. According to the results obtained, it is demonstrated that the developed prototype has presented a high efficiency and an acceptable response time, which makes it a good and suitable product for use.

In future developments, the aspiration is to improve the prototype by integrating an intelligent strategy based on vocal emotion detection. This integration aims to prioritise user security and enable individual and personal use through voice print technology. By analysing vocal emotions, the system will provide a more personalised experience for users. Additionally, the incorporation of voice print technology will enhance security by authenticating individuals based on their unique vocal characteristics.

References

Anwer, S., Waris, A., Sultan, H., Butt, S.I., Zafar, M.H., Sarwar, M. and Pujari, A.N. (2020) 'Eye and voice-controlled human machine interface system for wheelchairs using image gradient approach', *Sensors*, Vol. 20, No. 19. Doi: 10.3390/s20195510.

Aqil, M., Jbari, A. and Bourouhou, A. (2017) 'ECG signal denoising by discrete wavelet transform', *International Journal of Online Engineering*, Vol. 13, No. 9, pp.51–68. Doi: 10.3991/ijoe.v13i09.7159.

Bakouri, M., Alsehaimi, M., Ismail, H.F., Alshareef, K., Ganoun, A., Alqahtani, A. and Alharbi, Y. (2022) 'Steering a robotic wheelchair based on voice recognition system using convolutional neural networks', *Electronics*, Vol. 11, No. 1. Doi: 10.3390/electronics11010168.

Batista, G.C., Oliveira, D.L., Saotome, O. and Silva, W.L.S. (2020) 'A low-power asynchronous hardware implementation of a novel SVM classifier, with an application in a speech recognition system', *Microelectronics Journal*, Vol. 105. Doi: 10.1016/j.mejo.2020.104907.

Champaty, B., Jose, J., Pal, K. and Thirugnanam, A. (2014) 'Interface control system for motorized wheelchair', *Proceedings of the Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/ICMMD)*, pp.1–7.

Chevalier, H., Ingold, C., Kunz, C., Moore, C., Roven, C., Yamron, J., Baker, B., Bamberg, P., Bridle, S., Bruce, T. and Weader, A. (1995) 'Large-vocabulary speech recognition in specialized domains', *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 1, pp.217–220. Doi: 10.1109/icassp.1995.479403.

Ding, I. and Shi, J. (2016) 'Kinect microphone array-based speech and speaker', *Computers and Electrical Engineering*, pp.1–11. Doi: 10.1016/j.compeleceng.2015.12.010.

Duan, Y., Zou, B., Xu, J., Chen, F., Wei, J. and Tang, Y.Y. (2021) 'OAA-SVM-MS: a fast and efficient multi-class classification algorithm', *Neurocomputing*, Vol. 454, pp.448–460. Doi: 10.1016/j.neucom.2021.04.115.

He, S., Zhang, A. and Yan, M. (2019) 'Voice and motion-based control system: proof-of-concept implementation on robotics via internet-of-things technologies', *Proceedings of the ACM Southeast Conference*, pp.102–108. Doi: 10.1145/3299815.3314440.

He, Y. and Dong, X. (2020) 'Microprocessors and microsystems real time speech recognition algorithm on embedded system based on continuous Markov model', *Microprocessors and Microsystems*, Vol. 75. Doi: 10.1016/j.micpro.2020.103058.

Hou, T.K., Yagasena, Y. and Chelladurai, C. (2020) 'Arduino based voice controlled wheelchair', *Journal of Physics: Conference Series*. Doi: 10.1088/1742-6596/1432/1/012064.

Ismail, A., Abdlerazek, S. and El-Henawy, I. M. (2020) 'Development of smart healthcare system based on speech recognition using support vector machine and dynamic time warping', *Sustainability (Switzerland)*, Vol. 12, No. 6. Doi: 10.3390/su12062403.

Joshi, K., Ranjan, R., Sravya, E. and Baig, M.N.A. (2019) 'Design of voice-controlled smart wheelchair for physically challenged persons', *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS*, Springer, Singapore, Vol. 3, pp.79–91. Doi: 10.1007/978-981-13-1501-5_7.

Kamm, C.A., Shamieh, C.R. and Singhal, S. (1995) 'Speech recognition issues for directory assistance applications', *Speech Communication*, Vol. 17, Nos. 3/4, pp.303–311. Doi: 10.1016/0167-6393(95)00023-H.

Kannan, K. and Selvakumar, J. (2015) 'Arduino based voice controlled robot', *International Research Journal of Engineering and Technology*, Vol. 2, pp.49–55.

Mao, J., Wang, C., Guo, Y., Xu, G., Cao, S., Zhang, X. and Bi, Z. (2022) 'A novel model for voice command fingerprinting using deep learning', *Journal of Information Security and Applications*, Vol. 65. Doi: 10.1016/j.jisa.2021.103085.

Martinek, R., Vanus, J., Nedoma, J., Fridrich, M., Frnda, J. and Kawala-Sterniuk, A. (2020) 'Voice communication in noisy environments in a smart house using hybrid LMS+ ICA algorithm', *Sensors*, Vol. 20, No. 21. Doi: 10.3390/s20216022.

Mirza, I.A., Tripathy, A., Chopra, S., D'sa, M., Rajagopalan, K., D'souza, A. and Sharma, N. (2015) 'Mind-controlled wheelchair using an EEG headset and Arduino microcontroller', *Proceedings of the International Conference on Technologies for Sustainable Development*, pp.2–6. Doi: 10.1109/ICTSD.2015.7095887.

Parikh, S.P., Grassi, V., Kumar, V. and Okamoto, J. (2007) 'Integrating human inputs with autonomous behaviors on an intelligent wheelchair platform', *IEEE Intelligent Systems*, Vol. 22, No. 2, pp.33–41. Doi: 10.1109/MIS.2007.36.

Sharifuddin, M.S.I., Nordin, S. and Ali, A.M. (2020) 'Comparison of CNNs and SVM for voice control wheelchair', *IAES International Journal of Artificial Intelligence*, Vol. 9, No. 3, pp.387–393. Doi: 10.11591/ijai.v9.i3.pp387-393.

Sidiq, M., Agung Budi, W.T. and Sa'Adah, S. (2015) 'Design and implementation of voice command using MFCC and HMMs method', *Proceedings of the 3rd International Conference on Information and Communication Technology*, pp.375–380. Doi: 10.1109/ICoICT.2015.7231454.

Simpson, R.C. (2005) 'Smart wheelchairs: a literaturereview', *The Journal of Rehabilitation Research and Development*, Vol. 42, pp.423–436.

Sprigle, S. and Huang, M. (2015) 'Impact of mass and weight distribution on manual wheelchair propulsion torque', *Assistive Technology*, Vol. 27, No. 4, pp.226–235. Doi: 10.1080/10400435.2015.1039149.

Ullah, S., Mumtaz, Z., Liu, S., Abubaqr, M., Mahboob, A. and Madni, H.A. (2019) 'Single-equipment with multiple-application for an automated robot-car control system', *Sensors*, Vol. 19, No. 3. Doi: 10.3390/s19030662.

Voznenko, T.I., Chepin, E.V. and Urvanov, G.A. (2018) 'The control system based on extended BCI for a robotic wheelchair', *Procedia Computer Science*, Vol. 123, pp.522–527. Doi: 10.1016/j.procs.2018.01.079.

Xu, J., Huang, Z., Liu, L., Li, X. and Wei, K. (2023) 'Eye-gaze controlled wheelchair based on deep learning', *Sensors*, Vol. 23, No. 13. Doi: 10.3390/s23136239.

Yang, B.Q., Guan, X.P., Zhu, J.W., Gu, C.C., Wu, K.J. and Xu, J.J. (2021) 'SVMs multi-class loss feedback based discriminative dictionary learning for image classification', *Pattern Recognition*, Vol. 112. Doi: 10.1016/j.patcog.2020.107690.