



# Engineering and Technology Journal

Journal homepage: [engtechjournal.org](http://engtechjournal.org)



## Wheelchair Movement Based on Convolution Neural Network

Hani S. Hassan <sup>a\*</sup>, Jammila H. Saud <sup>b</sup>, Maisa'a A. Kodher <sup>c</sup>

<sup>a</sup> Computer science Department, University of Technology, Baghdad, Iraq, [Hani.co81@gmail.com](mailto:Hani.co81@gmail.com)

<sup>b</sup> Computer science Department, Mustansiriyah University, Baghdad, Iraq, [dr.jameelahharbi@gmail.com](mailto:dr.jameelahharbi@gmail.com)

<sup>c</sup> Computer science Department, University of Technology, Baghdad, Iraq, [masiaa.Ali2015@gmail.com](mailto:masiaa.Ali2015@gmail.com)

\*Corresponding author.

Submitted: 21/02/2020

Accepted: 03/07/2020

Published: 25/06/2021

### KEY WORDS

Wheelchair, deep learning, amputee, disable people, Arduino UNO

### ABSTRACT

*This paper intends to develop a methodology for helping amputees and crippled people old, by ongoing voice direction and association between patient and personal computer (PC) where these blends offer a promising response for helping the debilitated people. The major objective of this work is accurately detected audio orders via a microphone of an English language (go, stop, right and left) in a noisy environment by the proposed system. Thus, a patient that utilizes the proposed system can be controlling a wheelchair movement. The venture depends on preparing an off-line dataset of audio files are included 10000 orders and background noise. The proposed system has two important steps of preprocessing to get accurate of specific audio orders, accordingly, the accurate direction of wheelchair movement. Firstly, a dataset was preprocessed to reduce ambient noise by using Butterworth (cutoff 500-5000 Hz) and Wiener filter. Secondly, in the input (a microphone) of the proposed discriminative model put a procedure of infinite impulse response filter (Butterworth), passband filter for cutoff input microphone from 150-7000 Hz for back-off the loud and environment noise and local polynomial approximation (Savitzky-Golay) smoothing filter that plays out a polynomial regression on the signal values. Thus, a better for filtering from ambient noise and keeping on a waveform from distortion that makes the discriminative model accurate when voice orders were recognized. The proposed system can work with various situations and speeds for steering; forward, stop, left and right. All datasets are trained by using deep learning with specific parameters of a convolutional neural network (CNN). These capacities are dependent on code written in MATLAB. The prototype uses Arduino UNO and a microphone (MIC).*

**How to cite this article:** H. S. Hassan, J. H. Saud and M. A. Kodher, "Wheelchair Movement Based on Convolution Neural Network", Engineering and Technology Journal, Vol. 39, No. 06, pp. 1019-1030, 2021.

DOI: <https://doi.org/10.30684/etj.v39i6.1615>

This is an open access article under the CC BY 4.0 license <http://creativecommons.org/licenses/by/4.0>

## 1. INTRODUCTION

Audio and voice communication, preparing frameworks have relentlessly ascended in insignificance in each day's life of a great many people in creating nations. From music frameworks, through radio to convenient music players, sound preparations immovably settled on giving stimulation to shoppers. Advanced sound procedures specifically had accomplished mastery in sound conveyance, with MP3 players, Internet radio, iPods and CD players being the frameworks of a decision by and large. Indeed, even inside TV and film studios, and in blending work areas for 'live' occasions, advanced preparing prevails [1]. Speech recognition is among the disciplinary sub-field of computational etymology that develops methods and advances that encourages the acknowledgment and interpretation of verbally expressed words into the printed configuration by PCs [2]. Developers with architects locate the sensible harmony between what can be modified inside the calendar and spending plan, and what might be perfect for people. The quality in the plan of human computer interaction (HCI) relies upon the usefulness and ease of use. The usefulness of a framework is a lot of administrations or activities given to the people [15].

## 2. RELATED WORK

There are numerous researchers that considered the movement chair. Rakhi, et al. (2013), planned design a significant number of the necessities of people with incapacities, a human who can't utilize their arms to quality a manual wheelchair this work utilizes head developments and infrared sensor incorporated with a wheelchair [3]. In all actuality, this structure unseemly for an amputee. Patil, et al. (2014), planned a structure for assisting crippled patients who can move the fingers. It relies upon a picture dealing with and Vegard RISC processor (AVR) microcontroller for wheelchair movement. Wheelchair movement by finger movement based on picture getting ready by picture area and affirmation subject to change of finger picture from RGB to HSV. Without a doubt, the system exhibits acquiring pixels of a nose area [4]. This arrangement suits some incapacitated individuals are not an accident with their fingers and doesn't oblige individuals who have encountered an accident in the preferences and fingers and lower limbs as a result of this arrangement rely upon picture preparing and this proposition suite it for amputee on account of it relies upon discourse acknowledgment by utilizing the discriminative model. Buvanswari et, al. (2015), this work presents steering arranges a system for the comprehensive network that is influenced by Tetraplegia. This work subjected to pictures taking care of in two phases. In the first stage, iris affirmation by using Gaussian dimness work. In the second stage, determine the point of convergence of understudy and send information to find a course. The Arduino board gets orders from PC for controlling wheelchair courses [5]. Hani Saeed Hassan et, al. (2017), the association among humans and PC by recognizing a human face, tailing it and as necessities be setting propelled stick of the Arduino UNO stop, low or high. By using MATLAB, face acknowledgment estimation and CAM Shift figuring have been done with the strategy for thresholding [6]. Likewise, the proposed framework work with stop, forward, left and right. Amiel Hartman et, al. (2019), his work displays a wheelchair. It depicts the compromise of hardware and programming with sensor development and PC taking care of to develop the forefront astute wheelchair. This plan is a PC bunch setup to test tip-top preparing for shrewd wheelchair action and human affiliation. The LabVIEW pack is delivered for progressing self-administering way orchestrating and sensor data taking care of. Four little structure factor PCs are related over a Gigabit Ethernet neighborhood to shape the PC gathering [7]. Finally, this work required four PCs that are associated with Gigabit Ethernet in this way, as opposed to the proposed framework that is a single PC and vigorous its condition clamor.

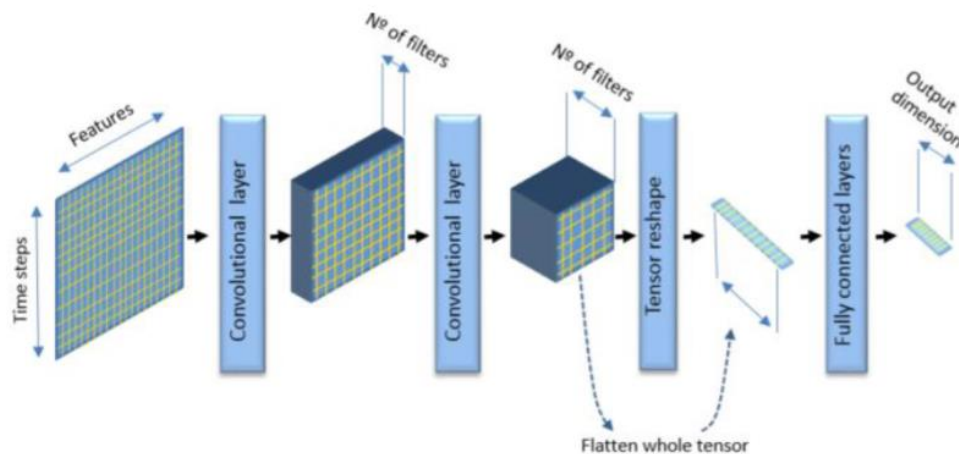
## 3. DEEP LEARNING TECHNIQUE

There are two essential sorts of logical model discriminative and generative models the capability between them Depends on likelihood dissemination. Discriminative models, directly register the probability of a yield given data. On the other hand, the generative model gives joint probability dispersal of the yield and the info [8]. Discourse acknowledgment involves a wide field and is normally orchestrated by two or three key clear articulations: Automatic discourse acknowledgment (ASR), Continuous discourse acknowledgment and Natural language preparation (NLP). A constant discourse acknowledgment. Portrays a discourse acknowledgment system that can see relentless sentences of talk. On a fundamental level, this would not require a customer to postpone when talking and would fuse correspondence and translation systems. The alternative is a discrete word affirmation structure,

used on a very basic level for dealing with vocal bearings, that sees single words delimited by stops [9].

### Convolutional Neural Network (CNN)

CNN has had profound results over the early decade in a collection of fields related to configuration design acknowledgment; from voice, taking care of picture find. The most favorable piece of CNNs is decreasing the number of parameters. This achievement has incited the two researchers and architects to move toward greater models in order to light up complex endeavors. To procure great acknowledgment of CNN must be started with components Convolution, Stride, Padding, filters, and Feature of CNNs as shown in Figure 1. [16,10]. In the preparation alternative there are numerous calculations to play out this procedure, for example, got from versatile minute estimation (Adam), root mean square engendering (RMSprop) [11] and the stochastic slope plummet with momentum (SGDM), the last calculation utilized in the proposed framework. The SGDM calculation apple to influence a route of steepest plunge towards the optimum [12]. Time weighed by its clearness, clarity, and request. Voice improvement is a primer structure in the discussion managing a territory, including talk mix, talk coding, talk acknowledgment, and talk examination. Voice sign recorded in a predictable situation may contain undesirable sound, for example, playing uproarious articulation by individuals, the sound of fan, atmosphere control framework, and so forth. These are considered under the class of racket. To the gathering of the onlooker's people, these obstructions are altogether unpalatable and ought to be reduced in the sales to improve the quality and soundness of talk signal. Additionally, the talking sign preparing tallies to rely upon the uncertainty that the voice sign is liberated from foundation commotion. The closeness of foundation turmoil in voice sign will chop down the presentation of the discussion managing structure fundamentally [13]. The PC part contains a PC or microcomputer, for instance, Arduino Atmel ATmega328 for remarkable explanation and propelled contraption speak with a human. The interface part is where the two articles meet. Correspondence occurs among individuals and PCs. The human association joins both programming and equipment. Individuals ought to use PCs or installed gadgets that are differing for various purposes. Researchers have created various methods and interfaces for those product specialists and fashioners to find a reasonable amicability between what can be changed in the plan and spending plan and would be perfect for individuals [14].



**Figure 1: Convolutional Neural Network**

### 4. PROPOSED SYSTEM

The proposed system of training voice orders was based on the cooperation of computer and sick people, where it's a prototype of a movement chair. A proposal was depending on real-time audio. A proposal consists of recognizing direction and navigation of movement chair that show in Figure 2. that depicts a diagram proposal of gathering dataset, preprocessing of the dataset, training all datasets and recognizing voice orders, respectively. The proposed system acquires a dataset from gathering the dataset. The second stage is preprocessing of the human voice dataset. The third stage is (recognizing the voice command) based on input microphone and trained dataset by using deep learning technique.

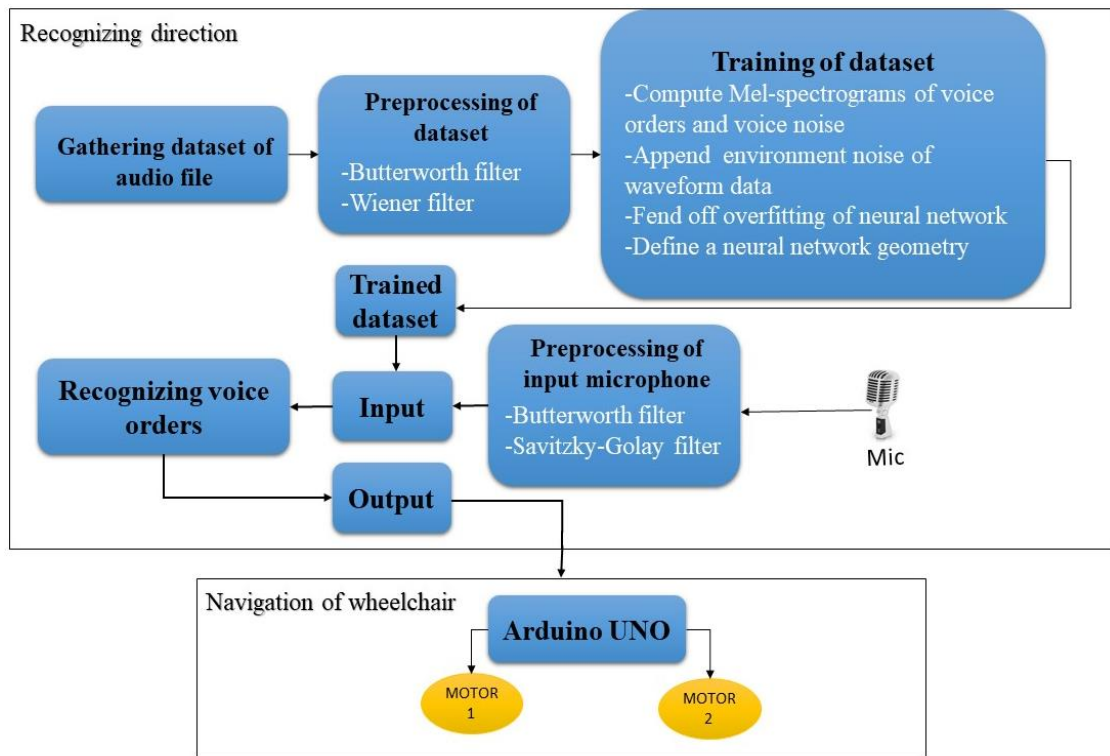


Figure 2: Diagram of proposed system

### I. Gathering of audio dataset

The proposed system uses a dataset which is an offline dataset where these datasets are divided into an isolated voice of the word command (voice orders), environment noise and different sounds that split from environmental noise (voice noise). The isolated words are English languages with their properties such as length: 1second, .wav format, bitrate: 256 Kbps, sampling rate: 16000, number of bit 16 and the number of channels is 1. The offline dataset is collected from Google's TensorFlow. All datasets are equivalent to 10000 audio files.

### II. Preprocessing of the dataset

Preprocessing of the datasets was carried out by using Butterworth (bandpass used in a range of frequency between 500-5000 Hz) and wiener filters for acquiring a pure waveform of isolated English voice orders ("go", "left", "right" and "stop").

### III. Training of the dataset

Characterize a neural system geometry by utilizing convolutional 2D layers, clump standardization layers and max-pooling layers. Our neural system has system geometry comprised of 10 layers with 40 filters and utilizing fascinating nonlinear enactment capacities, the parametric corrected straight unit at preparing time it used to duplicate info esteem that under zero by scaler learned. The proposed preparing system utilized 50 iterations for preparing and select SGDM calculations with force 0.95 as shown in Equation 1 that shows how the parametric rectified linear unit works, any input value less than zero is multiplied by scaler learned. The proposed framework utilized the mix (approval and preparing information) at every age before preparing and learning rate each number of age by projection learning rate during preparing that dependent on duplicated with factor (0.1) at every (20) number of age. Perceiving voice order that characterization a voice direction depends on preparing datasets at the previous stage. Right now, the examining rate to (16e3) and utilizing a receiver to enter voice from the present patient that utilizing a wheelchair. Likewise, show parameters for the spilling spectrogram figuring's and present an introduce a cradle for an input voice that is mean snatch the order mark and arrange the cushion for input the human voice.

$$\sigma(z) = \begin{cases} z, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (1)$$

#### IV. Preprocessing of input microphone

The hybrid filter was composed of two filters which were Butterworth and Savitzky-Golay finite impulse response filters. This suggested hybrid filter was applied by two steps which were first, Butterworth filter type passband filter for cut-off input microphone from 150-7000 Hz in back-off the loud and environment noise. Secondly, the Savitzky-Golay smoothing filter used its low pass filter for filtering, smoothing and keeping the original form of the wave.

#### V. Recognizes voice orders

Recognizing the voice orders was the third stage in this proposed system where algorithm 1 shows the recognize voice command stage procedure based on input microphone and trained dataset by using deep learning technique. Accordingly, it was detecting the voice orders in the English language (go, stop, left and right).

##### Algorithm 1: Recognize voice orders

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Input</b></p> <p>Fs:16*10<sup>3</sup></p> <p>Classification rate =13</p> <p>Input MIC\\cleaned by hybrid filters (Butterworth (cutoff 150-7000 Hz) and Savitzky-Golay)</p> <p>Label \\trained network</p> <p>B= Buffer of voice noise;</p> <p>Count = 8\\number of orders</p> <p>Threshold1= Classification rate x 0.2 \\round to largest integer</p> <p>Threshold2 = Buffer of maximum of predicate probabilities of label</p> <p>Frame of duration = 0.025;</p> <p>Hop of duration = 0.010;</p> <p><b>Output</b></p> <p>True command</p>                                                                                                                                                                                                                                                                    |
| <p><b>Begin</b></p> <p>Step1: Specify the parameter</p> <p>begin</p> <p>1.1: Hop of length = floor (hop of duration x Fs);</p> <p>1.2: Frame of length = floor (frame of duration x Fs);</p> <p>1.3: Buffer of the input wave = zeros([Fs,1]);</p> <p>end</p> <p>Step2: Extract sample of enhanced input microphone and add them to the buffer</p> <p>Step3: Compute the spectrogram of the enhanced input microphone</p> <p>Step4: Classify the input microphone into voice orders and voice noise based on thresholding operation</p> <p>1.1 classify voice orders if the following hold</p> <p>begin</p> <p>if Y= B then “voice noise”</p> <p>else if count &lt;= Threshold1 then “voice noise”</p> <p>else if Threshold2 &lt;= 0.5 then “voice noise”</p> <p>else</p> <p>True command</p> <p>end</p> <p>End</p> |

#### VI. Navigation of wheelchair

Navigation of the wheelchair was based on detecting the voice command in the English language such as (“ go”, “ left”, “ right”, “ stop”, that acquired. The equipment of the proposed wheelchair framework comprised of a microcontroller, IC and engines for movement. Algorithm 2 shows the navigation of the wheelchair.

**Algorithm 2: Navigation of wheelchair**

Input

G= True-go  
 S=True-stop  
 L=True-left  
 R=True-right  
 Com port 9600

Output

Specify different voltages on two-wheel drives

Begin

Step1: initialize the serial port

Step2: do until serial port connection close

begin

if input G then

begin

specify 5 volts to pin 11

specify 5 volts to pin 12

end

if input R then

begin

specify 2.5 volts to pin 11

specify 0 volts to pin 12

end

if input L then

begin

specify 0 volts to pin 11

specify 2.5 volts to pin 12

end

if input S then

begin

specify 0 volts to pin 11

specify 0 volts to pin 12

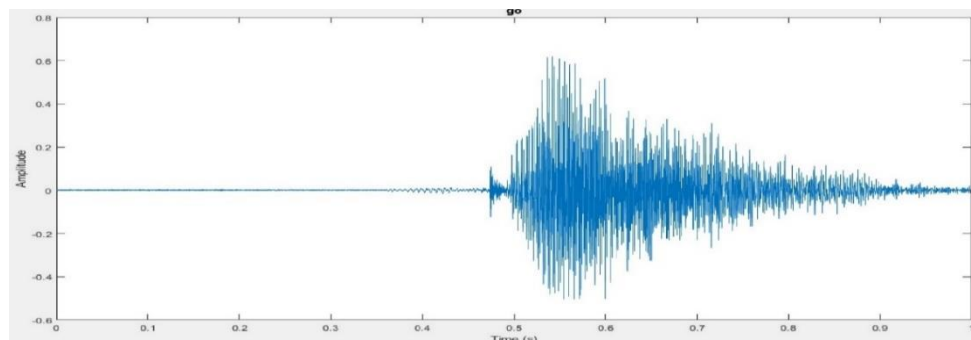
end

end

End

**5. EXPERIMENT RESULTS**

Subsequent to building up the entire framework, the proposed prototype of the wheelchair corresponds to the Arduino UNO board and tested by three humans. An outcome beneath portrays this procedure. After gathering datasets, the preprocessing is the next stage for removing ambient noise from the waveform of the audio file. Figure 3 shows the preprocessing of the off-line original dataset where the amplitude of the original dataset is ranging between (-0.06 to 0.06).



**Figure 3: Off-line of the original audio dataset ('go')**



Figure 4 the amplitude of the original dataset is ranging between (-0.05 to 0.05). Accordingly, ambient noise removed from the off-line original dataset.

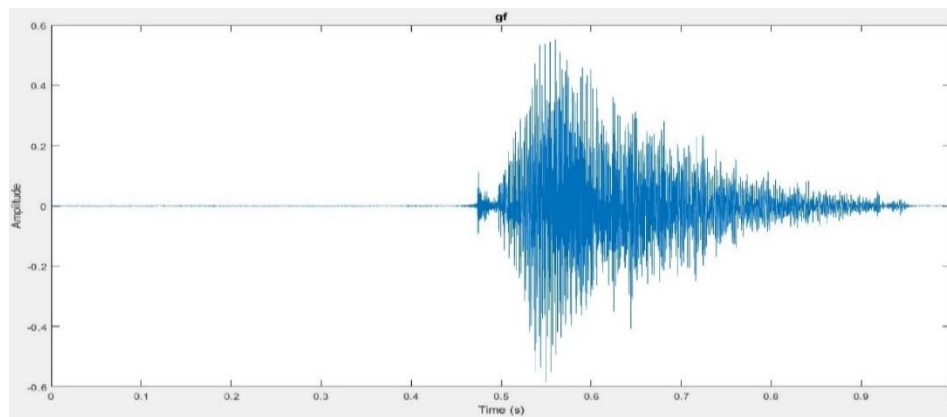


Figure 4: preprocessing of the original audio dataset ('go')

Figure 5 the training of all datasets is over with accuracy 97.66 as shown in below

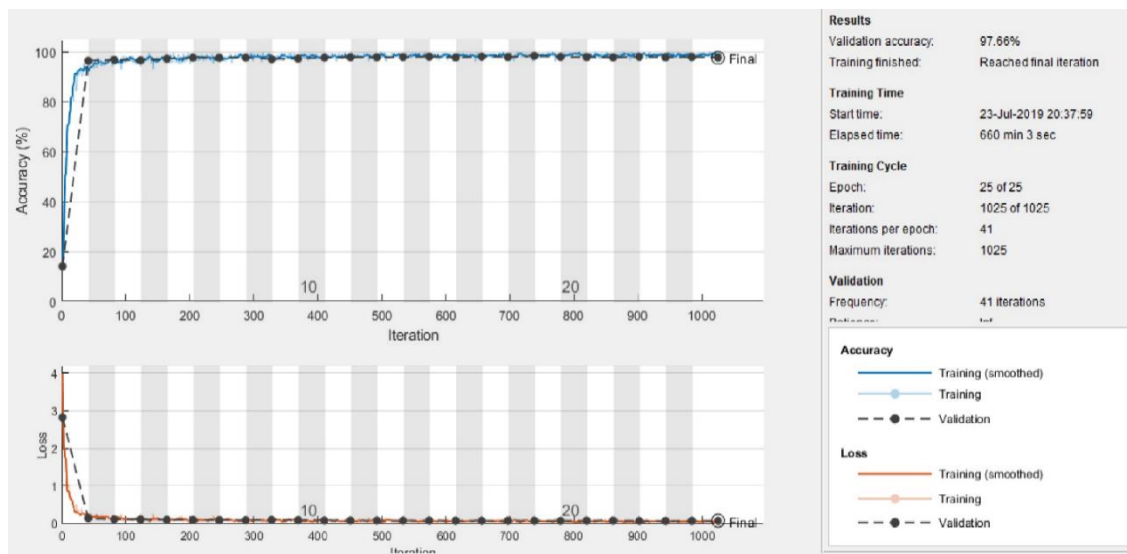


Figure 5: Accuracy of the training dataset

Figure 6 shows the confusion matrix is a table that is frequently used to portray the performance of the model. The column is the true class and the row is predicate class.

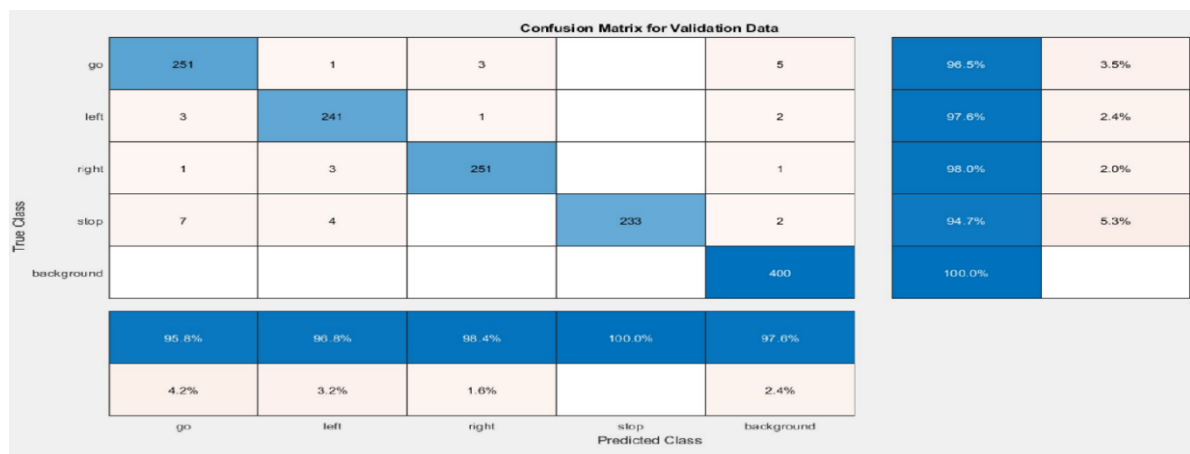
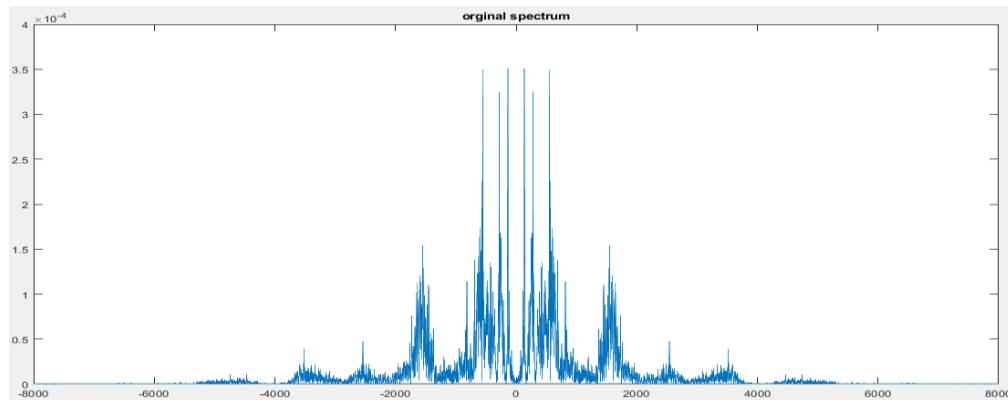
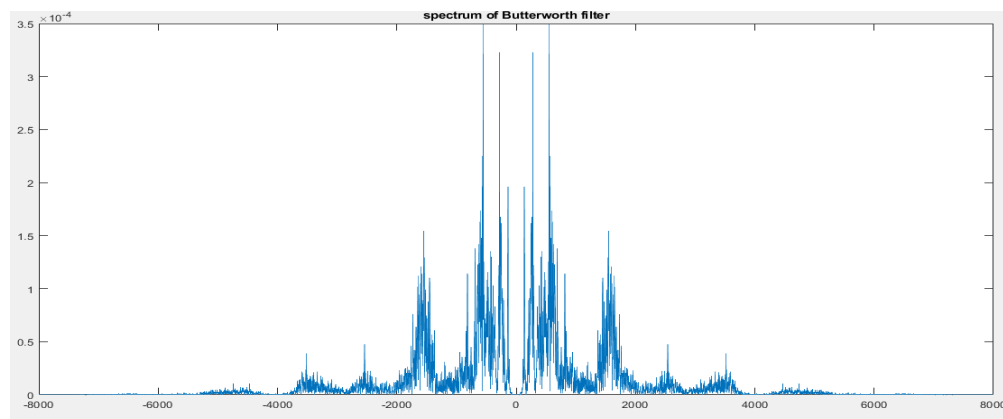


Figure 6: Confusion matrix

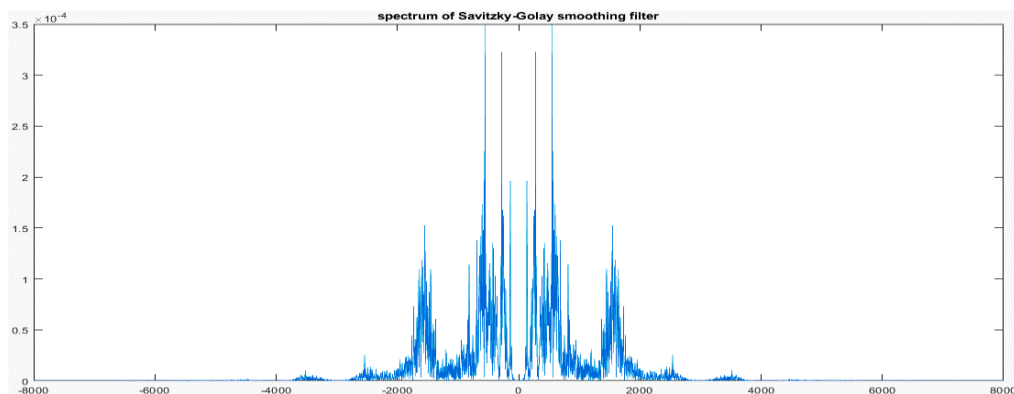
Figure 7, Figure 8 and Figure 9 illustrate preprocessing of human1 voice via MIC, by two filters that were Butterworth filter and Savitzky-Golay filter respectively. In the first filter, the ambient noise was removed from the voice waveform. In the second filter, the red square shows reduce the power spectrum of the human voice from 2500Hz to 3700Hz. Accordingly, the ambient noise and low level of noise were removed from the original voice waveform of the human1 in such a way without effects on the original voice waveform of the original signal.



**Figure 7: Original signal of human**



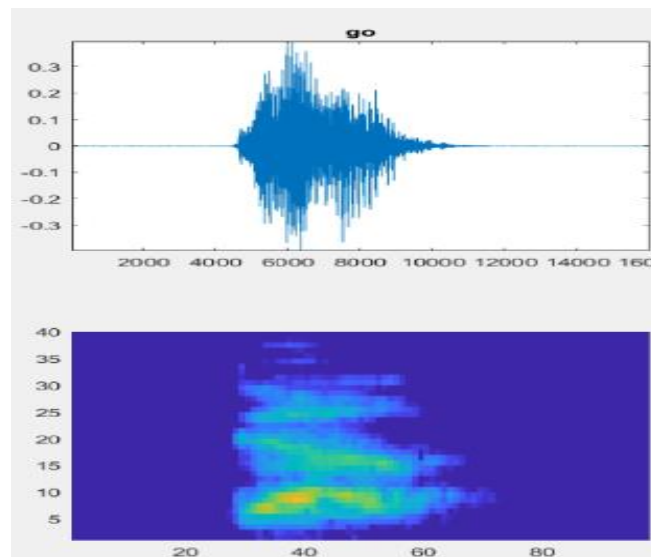
**Figure 8: Spectrum of the first filter**



**Figure 9: Spectrum of the second filter**

Feature extraction is the essential step during the transformation of the dataset of speech sound to a stream of a feature vector that involved only critical information for recognition of a given speech. MFCC is very great for representing speech signals in speech processing. Figure 10 shows the results of the feature extraction of voice 'go'.





**Figure 10: Spectrum of audio dataset 'go'**

Table I shows the Precision of the development of a wheelchair when the patient talks across the microphone by using one of the commands (go, stop, left and right) All accuracy of the commands is calculated by using the equation (2).

$$\text{Precision} = ((\text{true movement} + \text{true detect}) / (\text{false detect} + \text{false movement} + \text{true detect} + \text{true movement})) * \% \quad (2)$$

true detect = number of words detected through classification rate =20 Hz

false detect = reverse the true detect

true movement = number of words that achieved for the navigation of the wheelchair

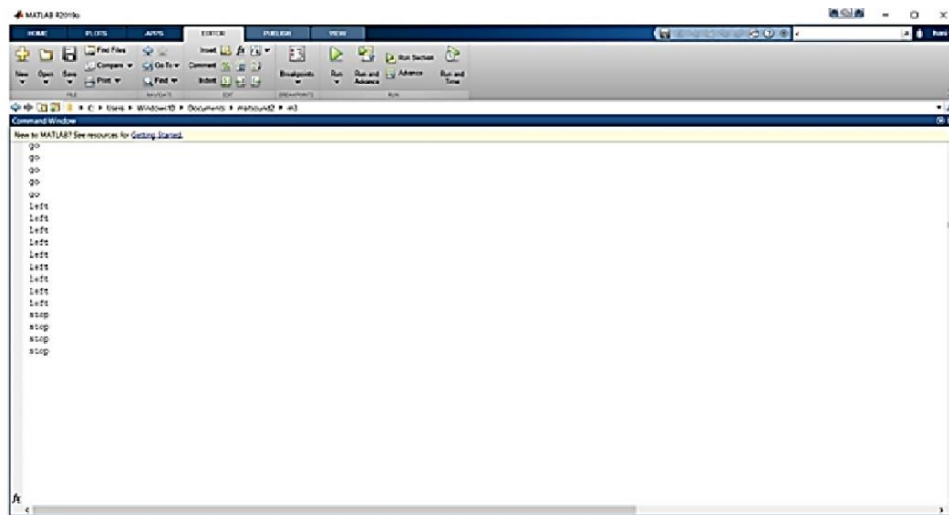
false movement= reverse the true movement

Example, the true detect = 30, true movement = 30, false detect = 7 and false movement = 0. By Applying equation (3) the precision result =83 %.

**TABLE I: Accuracy of wheelchair movement**

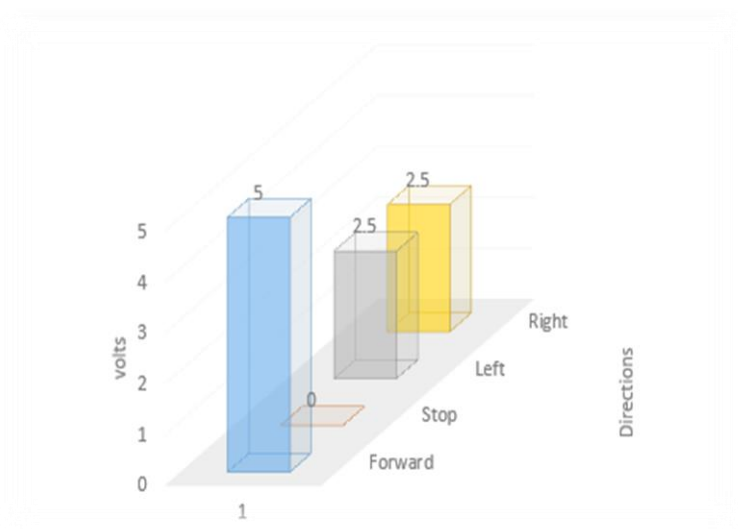
| Commend      | Movement | Accuracy % |
|--------------|----------|------------|
| <b>Go</b>    | Forward  | 90         |
| <b>Left</b>  | Left     | 90         |
| <b>Right</b> | Right    | 92         |
| <b>Stop</b>  | Stop     | 93         |

Figure 11 delineates the execution result in the MATLAB environment, detection direction with supplying different voltages for engines depending on PWM technique



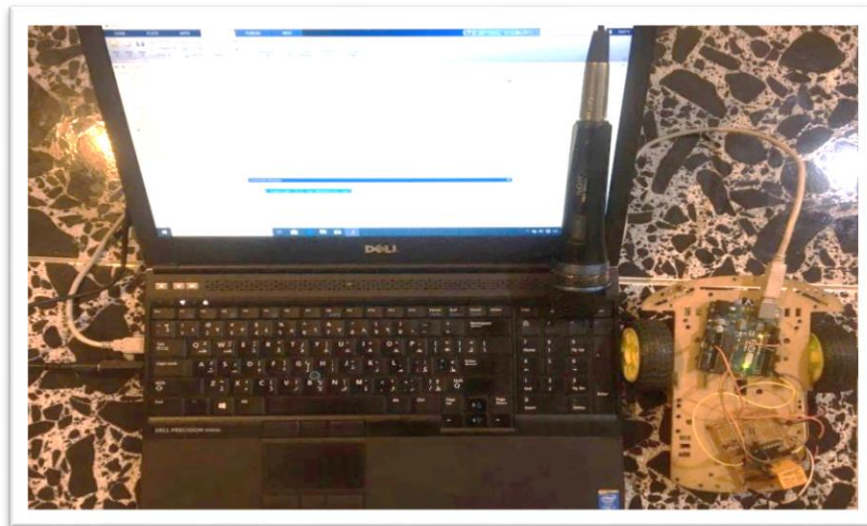
**Figure 11: Command window of MATLAB is detecting voice orders**

Figure 12 explains the accuracy chart of the relationship between proposed directions and volts supplied by Arduino UNO. These accuracies depend on voltages between (0-5) volts. In other words, when the orders are “go” the controlling voltage PWM supplying 5 volts to wheel drives. In the same way, when orders “left” , “right” PWM supplying 2.5 volts. In case of orders “stop” PWM is supplying zero volt to wheel drives.



**Figure 12: Relationship between volts and directions**

Figure 13 shows the wheelchair prototype that comprises, electronic circuit, of Arduino Uno and two motors, indicate M1 and M2. It ought to be referenced that all equipment parts of the prototype work with a continuous current of 5 volts.



**Figure 13: Prototype of wheelchair**

## 6. CONCLUSIONS

The point of our proposed system is to propose a prototype of the wheelchair for helping old, amputee and disabled people. This proposition utilizes perceive the continuously English discourse and this prototype relies upon two important steps of preprocessing to get accurate of specific audio orders and discriminative model, preprocessed genuine dataset of off-line English voice to reject background noise and uniforms of the voice dataset. Prepared a filtered of the genuine dataset by utilizing profound learning and the approval exactness result and the high accuracy are forward heading. Our prototype was hearty against foundation commotion. Voice directions of the patient by means of a receiver that deciphered into a signal through an Arduino, it maintained by PWM to make various volts reliant on voice orders.

## REFERENCES

- [1] I. McLoughlin, Applied Speech and Audio Processing, 1st edition, Cambridge University Press, 2009.
- [2] Y. A. Karmakar, K. A. Suthar, Y. A. Karmakar, N. Pandya, Survey paper on Different Speech Recognition Algorithm: Challenges and Techniques, Int. J. Comput. Appl., 175 (2017) 31-36.
- [3] R. A. Kalantri, D. K. Chitre, Automatic Wheelchair using Gesture Recognition, Int. J. Eng. Adv. Technol., 2 (2013) 146-150.
- [4] C. G. Patil, S. K. Gharge, S. V. Modhave, Y. S. Angal, Design of Wheelchair Using Finger Operation With Image Processing Algorithm, Int. J. Res. Eng. Technol, 3 (2014) 232-237. <https://doi.org/10.15623/IJRET.2014.0302039>
- [5] B. Buvanswari, T. K Reddy, Eye Scrutinized Wheelchair for People Affected with Tetraplegia, Int. J. Comp. Sci., Eng. Inf. Technol., 5 (2015) 15-24.
- [6] H. S. Hasan, Multimedia Based Computer Human Interface, MSc. Thesis, University of Technology, Iraq. 2016.
- [7] A. Hartman, V. K. Nandikolla, Human-Machine Interface for a Smart Wheelchair, J. Robot.. 2019. <https://doi.org/10.1155/2019/4837058>
- [8] Ch. Chen, Handbook of Pattern Recognition and Computer Vision, 5th edition, World scientific publishing Co. pte. Ltd. USA. , 2015.
- [9] I. McLoughlin, Applied Speech and Audio Processing, 2nd edition, Cambridge University Press, (2016) 174-176.
- [10] S. Albawi, T. A. Mohammed, S. Al-Zawi, Understanding of a Convolutional Neural Network, 2017 Int. Conf. Eng. Technol., Antalya, Turkey, (2017) 1-6. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
- [11] D. P. Kingma, L. J. Ba, Adam: A method for stochastic optimazion, proceedings of the 3rd International Conference on Learning Representations, San Diego, 2015.

- [12] A. Yashwanth, S. Shammer, R. Sairam, G. Chamundeeswari, A novel approach for indoor-outdoor scene classification using transfer learning, *Int. J. Adv. Res., Ideas Innov. Technol.*, 5 (2019) 1756-1762.
- [13] K. U. Shajeesh, K. S. Sachin, D. Pravena, K.P. Soman, Speech Enhancement based on Savitzky–Golay Smoothing Filter, *Int. J. Comp. App.*, 57 (2012) 39-44.
- [14] Z. Katibeh, Techniques for Real-time Multi-person Face Tracking for Human-robot Dialogue. M.Sc. thesis, Blekinge Institute of Technology, School of Engineering, 2013.
- [15] D. Te'eni, J. M. Carey, P. Zhang. *Human Computer Interaction: Developing Effective Organizational Information Systems*, 2nd edition, Wiley, 2007.
- [16] M. L. Martin, B. Carro, A. S. Esguevillas, J. Lloret, Network traffic classifier with convolutional and recurrent neural networks for Internet of Things, *IEEE access*, 5 (2017) 18042-18050. <https://doi.org/10.1109/ACCESS.2017.2747560>