Smart mouse for quadriplegic patients

Ahmed Ashraf, Abd El-Rahman Omran, Mohamed Hassaan, Ahmed Omar, Khalid Walied
Arab Academy for Science Technology & Maritime Transport
, Egypt, AAH@student.aast.edu, Abdelrahman.omran@student.aast.edu, khaledelkemary@student.aast.edu,

ahmedalshabib210@gmail.com, mohamedhassaan951@gmail.com

Supervisor: Dr. Nashwa Elbendary, Dr., Ammar Mostafa,
nashwa.elbendary@gmail.com, ammar.mostafa@gmail.com

Abstract - In Egypt, more than 2 million people suffer from a mobility impairment some of them have quadriplegia, which usually indicates subsequent damage to the cervical spinal cord, although it may be the result of an injury to the brain or peripheral nerves. The cervical spinal cord is located inside a canal in the vertebrae of the spine, and it connects the brain to the majority of the body's organs, through the nerve fibers running through functional harm includes activities of daily living, such as eating, dressing, bathing, loss of voluntary control over the excretion of urine or faces, loss of control over mobility, and even sexual performance, fertility and the ability to contribute to family life, integration into the labor market and social settings. Also, quadruple patients cannot use technology such as computers, cell phones, and television controls, and other things that need manual control., for example, a computer-primarily relies on a mouse and keyboard for system navigation, navigation, and use apps, so we thought about helping them by designing a motion-controlled mouse.

The head has a set of features that enable the patient to control the system with sound and has a keyboard controlled by eye movement.

Now we made the voice recognition "up"أعلى " & "down" "أسفل" in the Arabic language to help in scrolling up & down, and this can be done by making extractions from the MFCC features and then make binary classification using support vector machine algorithm.

I. Introduction

Causes and risk factors of quadriplegia among the causes of quadriplegia and factors that cause damage to the spinal cord:

Direct injuries, narrowing of the spinal canal and herniated discs between the vertebrae (herniated disc), multiple sclerosis, cancerous tumors, infections, and diseases of blood vessels. The quality of the causative agent and the extent of its response to medical treatment has an impact on the severity and longevity of the damage. But from the moment it arose, the damage has dire consequences, whatever the cause. Complications of quadriplegia Complications result from the interruption of communication between the body's organs and the brain. Respiratory, urinary system complications, and bedsore may lead to respiratory failure, kidney failure, acute, and diffuse. Infections, which are the main causes of morbidity and mortality. Other complications may also occur in the heart and blood vessels, in the digestive system, in the bones, and others. Due to all of this, quadriplegic patients are unable to use technology such as computers, mobile phones, TV controls, and other things that need manual control.

For example, the computer depends mainly on the mouse and keyboard to navigate the system, browse and use applications, so we thought of helping them by designing a mouse that is controlled by the movement of the head and has a set of characteristics that enable the patient to control the system by voice.

II. Methodology & System Design

Our project aims to find solutions to help the people who cannot use their hands for any reason, to help them use the computer with the ways that we worked on. In our project, we worked to make a lot of different features that can help a person that cannot move his/her hand to use computers. We added the speech recognition system, so, the person who can talk will use the computer easily just by using his voice. We use MFCC feature extraction.

The process is divided into four parts.

A. Voice recognition commands speech acquisition is done by a Microphone (headset).

B. The audio files needed for the training and testing are stored locally on the computer storage unit.

C. The data is categorized to "UP" & "Down" in Arabic using MFCC.

D. The system trained on some of the data and used the other for testing.

The goal is to make people with special needs benefit from the current technology of computers, mobile phones, and smart TVs via voice.

The proposed UP & DOWN speaker recognition system is displayed in figure 1. It can be divided into two main components: feature extraction and. Firstly, improvised emotional utterances of the IEMOCAP database are considered to evaluate this system. We extract MFCCs and MFCC-SDC from the speech signals. These obtained feature vectors are then divided into training and test sets. Then, classification is done using two well-known multiclass SVM and logistic recreation approaches which are One-Against-All(OAA), and One-Against-One(OAO). Finally, the decision of the recognition system is specified with accuracy rate using the test set.

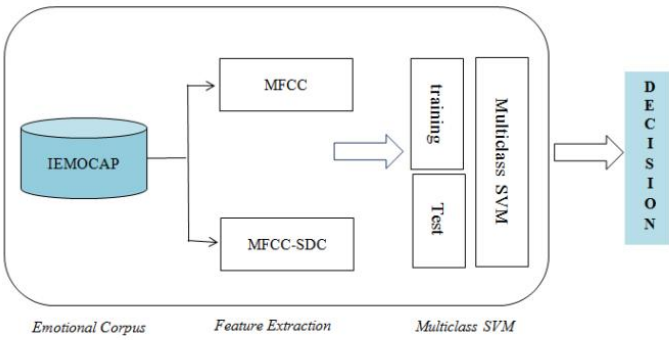


Fig. 1

Choosing a suitable feature that represents useful information increases the precision of the recognition system. In this section, we describe the procedure of MFCC feature extraction.

Mel Frequency Cepstral Coefficients Mel Frequency Cepstral Coefficients (MFCC) is one of the commonly used techniques of feature extraction. MFCC coefficients are based on human hearing perceptions that cannot perceive frequencies over 1Khz. After the frame blocking and windowing step, the FFT is computed, and the power coefficients are filtered by a triangular bandpass filter bank, also known as Mel-scale. They have been utilized to collect phonetically significant aspects of voice signals. Filters spaced linearly at low frequencies below 1000Hz, and filters spaced logarithmically above 1000Hz are available in MFCC. As a result, the Mels for a given linear frequency f may be calculated using the following approximate formula:

$$Mel(f) = 2595 \times \log_{10}(1 + f / 700) \quad (1)$$

The full extraction procedure of the Mel-frequency cepstral coefficient is described in figure 2.

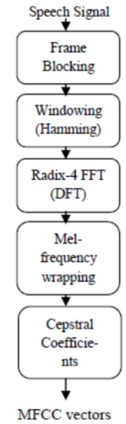


Fig. 2

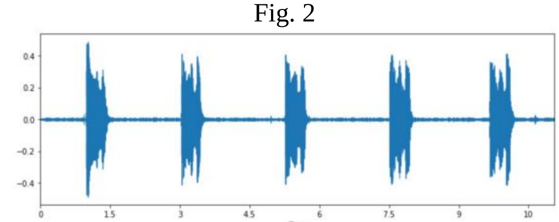


Fig. 3. Audio amplitude for up sample.

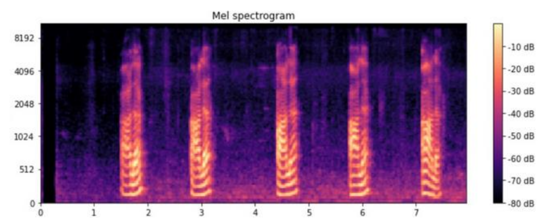


Fig. 4. Mel spectrogram for up sample.

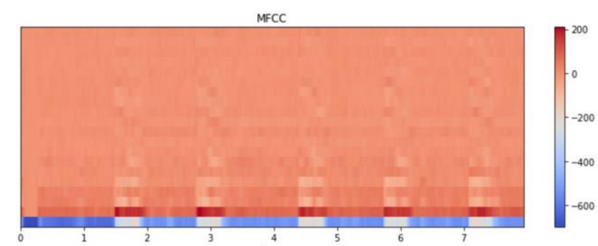


Fig. 5. MFCC coefficients for up sample.

Then the output is passed to the classifiers. We use two classifiers to compare between them and look for the best result.

A. Logistic Regression

It is a predictive model analysis technique where the target variables (output) are discrete values for a given set of features or input (X). For example, whether the word is up (1) or down (0). It is a powerful and simple classification algorithm in machine learning borrowed from statistics algorithms. It is one of the most common machine learning algorithms used for binary classification. It predicts the probability of occurrence of a binary outcome using a logit function (sigmoid function). It's a type of linear regression that uses the log function to predict outcome probabilities. It's a mathematical function with the property of being able to translate any real value to a number between 0 and 1 in the shape of the letter "S" as in Fig. 1. [5] Where sigmoid function is denoted as (Y), (e) is a constant number, and (z) is (Y old) $Y = 1 / 1+e^{-z}$. *(2)*

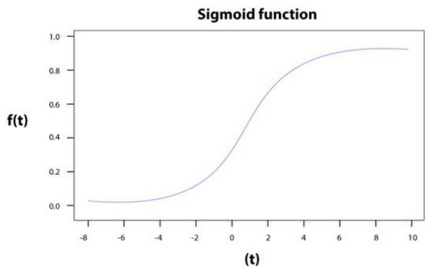


Fig. 6 Sigmoid function

B. Support vector machine (SVM)

It is an algorithm that can be used for classification at most and regression, and it is determined as a supervised machine learning algorithm.

We use the number of features denoted as (n) to plot each data item from a dataset and differentiate between the data to two classes up and down in Arabic and to divide between the two data then we use the hyper-plane to put the new data points easily to the correct category by choosing the support vectors in each class as in Fig. 7.

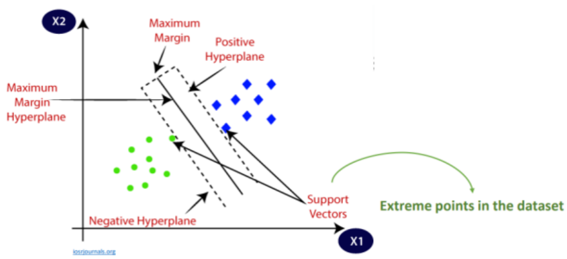


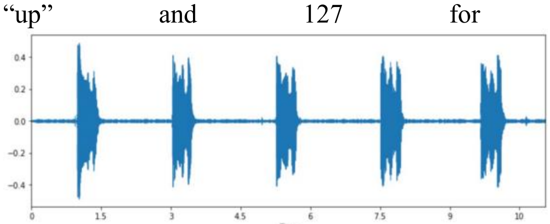
Fig. 7. SVM

III. DATASET

We have collected our dataset from a group of our family members and friends, and then we recorded it manually and labeling it manually.

Firstly, the total data received were 200 ‘’up’’ and 200 for ‘’down’’ and the suitable data valid to work on was only 113 for ‘’up’’ and 165 for ‘’down’’.

Then the total amount of data for testing was only 38 for ‘’down’’ and 28 for ‘’up’’, while the training data was 85 for “up” and 127 for “down”.

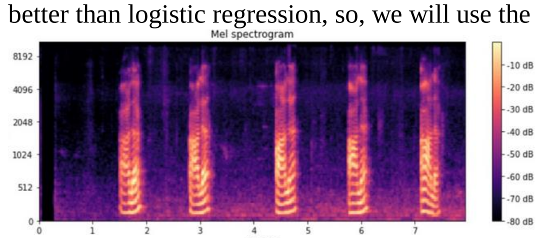


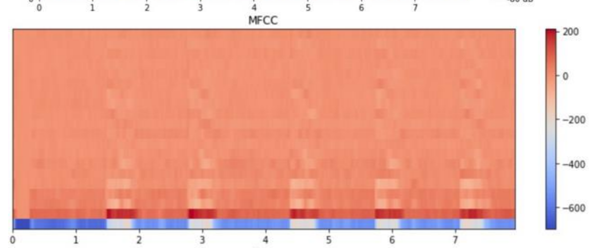
IV. RESULTS AND DESICCATION

After we get features extraction of MFCC we computed the delta of MFCC for up and down, then we integrated with the MFCC.

Then we added it to the classifier logistic regression and support vector machine.

We used the logistic regression and added it to the confusion matrix to get the result for it to be [[50 0] [4 30]], the test classification report 0.93 and then 1.00, the recall was 1.00 0.88, f1-score 0.96 then 0.94, support was 50 then 34 and we found that the test accuracy was 0.9523809523809523, while for that of the support vector machine, We added it to the confusion matrix to get the result for it to be [[50 0] [1 33]], the test classification report 0.98 and then 1.00, the recall was 1.00 then 0.97, f1-score 0.99 in both steps, support was 50 then 34, and we found that test accuracy was 0.9880952380952381. So, we can conclude that the SVM is better than logistic regression, so, we will use the SVM.





V. CONCLUSION AND FUTURE WORK

In this project, the basics of developing a VRC system have been discussed. The proposed methodology has been implemented primarily based on MFCC analysis and Logistic regression & Support vector machine (SVM) classifiers. The core achievement of this project is that the higher rate of accuracy is about 98% of the SVM than Logistic regression with a rate of 95%. The proposed work has achieved a high rate of accuracy only using a modicum of training samples and 38 + 28 = 66 samples dictates the reliability. We look forward to adding more features to the system to recognize more words like a press, shut down and restart, etc... all that to help quadriplegic patients.

VI. ACKNOWLEDGMENT

For this project, we would like to acknowledge faculty supervisors for our project, Dr. Nashwa Elbendary & Dr. Ammar Mostafa Hassan for all the forms of help that they provide to the team of this project.

We should also mention our participating partners, the Arab Academy for science, technology & maritime transport (AAST), the Arab organization for industrialization, and the Elaraby group.

VII. References

[1] Manuscript Templates for Conference Proceedings, IEEE. http://www.ieee.org/conferences_events/conferences/publishing/templates.html

[2] M. King, B. Zhu, and S. Tang, "Optimal path planning," *Mobile Robots*, vol. 8, no. 2, pp. 520-531, March 2001.

[3] H. Simpson, *Dumb Robots*, 3rd ed., Springfield: IOS Press, 2004, pp.6-9.

[4] M. King and B. Zhu, "Gaming strategies," in Path Planning to the West, vol. II, S. Tang and M. King, Eds. Xian: Jiaoda Press, 1998, pp. 158-176.

[5] B. Simpson, et al, "Title of paper goes here if known," unpublished.

[6] Sahidullah, Md.; Saha, Goutam (May 2012). "Design, analysis and experimental evaluation of block-based transformation in MFCC computation for speaker recognition". Speech Communication. 54 (4): 543–565. doi:10.1016/j.specom.2011.11.004

[7] S.B. Davis, and P. Mermelstein (1980), "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," in IEEE Transactions on Acoustics, Speech, and Signal Processing, 28(4), pp. 357366.

[8] European Telecommunications Standards Institute (2003), Speech Processing, Transmission and Quality Aspects (STQ); Distributed speech recognition; Front-end feature extraction algorithm; Compression algorithms. Technical standard ES 201 108, v1.1.3.

[9] T. Ganchev, N. Fakotakis, and G. Kokkinakis (2005), "Comparative evaluation of various MFCC implementations on the speaker verification task Archived 2011-07-17 at the Wayback Machine," in 10th International Conference on Speech and Computer (SPECOM 2005), Vol. 1, pp. 191–194.

[10] Tolles, Juliana; Meurer, William J (2016). "Logistic Regression Relating Patient Characteristics to Outcomes". JAMA. 316 (5): 533–4. doi:10.1001/jama.2016.7653. ISSN 0098-7484. OCLC 6823603312. PMID 27483067

[11] Cortes, Corinna; Vapnik, Vladimir N. (1995). "Support-vector networks" (PDF). Machine Learning. 20 (3): 273–297. CiteSeerX 10.1.1.15.9362. doi:10.1007/BF00994018. S2CID 206787478.

[12] Press, William H.; Teukolsky, Saul A.; Vetterling, William T.; Flannery, Brian P. (2007). "Section 16.5. Support Vector Machines". Numerical Recipes: The Art of Scientific Computing (3rd ed.). New York: Cambridge University Press. ISBN 978-0-521-88068-8. Archived from the original on 2011-08-11.

[13] Cuingnet, Rémi; Rosso, Charlotte; Chupin, Marie; Lehéricy, Stéphane; Dormont, Didier; Benali, Habib; Samson, Yves; and Colliot, Olivier; "Spatial regularization of SVM for the detection of diffusion alterations associated with stroke outcome", *Medical Image Analysis*, 2011, 15 (5): 729–737.

[14] Gaonkar, Bilwaj; Davatzikos, Christos; "Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification".

[15] Pradhan, Sameer S., et al. "Shallow semantic parsing using support vector machines." Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics: HLT-NAACL 2004. 2004.

[16] Smola, Alex J.; Schölkopf, Bernhard (2004). "A tutorial on support vector regression"(PDF). *Statistics and Computing*. **14** (3): 199–222. CiteSeerX 10.1.1.41.1452. doi:10.1023/B:STCO.0000035301.49549.88. S2CID 15475. Archived (PDF) from the original on 2012-01-31.

[17] Suykens, Johan A. K.; Vandewalle, Joos P. L.; "Least squares support vector machine classifiers", *Neural Processing Letters*, vol. 9, no. 3, Jun. 1999, pp. 293–300.

[18] P. Mermelstein (1976), "Distance measures for speech recognition, psychological and instrumental," in *Pattern Recognition and Artificial Intelligence*, C. H. Chen, Ed., pp. 374–388. Academic, New York.

[19] Meinard Müller (2007). *Information Retrieval for Music and Motion*. Springer. p. 65. ISBN 978-3-540-74047-6.

[20] L. C. W. Pols (1966), "Spectral Analysis and Identification of Dutch Vowels in Monosyllabic Words," Doctoral dissertation, Free University, Amsterdam, The Netherlands