[1] Sankalp Bisan
[2] S. K. Atre
[3] Satish Sharma

Design of a Real-Time Hand to Text Sign Language Recognition System

Journal of Electrical Systems

Abstract: - Sign language is a very powerful and effective tool for communication among the dumb and deaf. It plays a significant role in ensuring accessibility and inclusion for the individuals with hearing impairment. The proposed work attempts to bridge the communication gap between verbal and non-verbal communicators. In the present work, Indian Sign Language (ISL) is considered for conversion into text form since sign languages are region specific. The designed system uses flex sensors, an inertial measurement unit (IMU) and microcontroller unit as the main processing unit that provides voltage values according to the hand and finger movements. Data as measured using the wearable glove has been analyzed using the Long Short Term Memory (LSTM) algorithm for training the ML model. The in-house system accepts ISL signs as a real time input via flex sensor and IMU those are classified into the appropriate gestures. The classified gesture is then converted into an appropriate label as output in the corresponding legible text format. It is observed that in-house designed system is very simple and reliable for conversion of Indian sign gestures into legible text format.

Keywords: Indian Sign Language (ISL), Flex Sensor, Inertial Measurement Unit (IMU), Machine Learning (ML), Long Short Term Memory (LSTM).

[1] *Department of Electronics and Computer Science, RTMNU Campus, Nagpur – 440 033, Maharashtra, India.
Email: sankalpbisan07@gmail.com

[2] Department of Electronics and Computer Science, RTMNU Campus, Nagpur – 440 033, Maharashtra, India.
Email: atreshivk@gmail.com

[3] Department of Electronics and Computer Science, RTMNU Campus, Nagpur – 440 033, Maharashtra, India.
Email: sharmasat@gmail.com

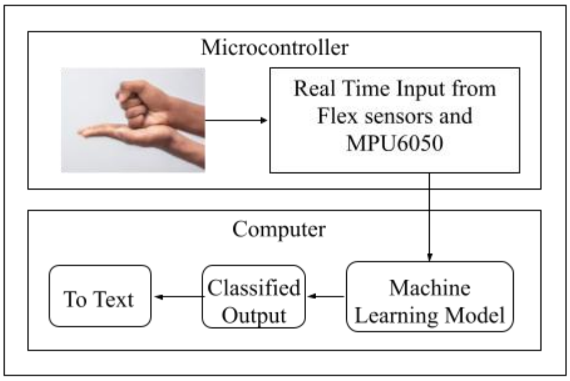
I. INTRODUCTION

Communication is not only limited to the exchanging or sharing of information verbally, but also the information can be exchanged through human emotions, facial expressions, body language and signs performed by hands. Although, verbal communication is far more crucial than other means of communication. But there are people who cannot comprehend verbal communication because they are unable to listen and to speak. Such people are called as Deaf-Mute people and thus they use sign languages which are region specific as their only mode of communication. But the verbally communicating people do not understand sign languages unless they have learnt it. This results in a communication gap between the verbal communicators and Deaf-Mute people. To bridge this gap, a system has been developed which takes the signs performed by the user via hand in real time and aims to convert it into text. This will significantly help the Deaf-Mute people to form relationships, collaborate, grab opportunities and to express their own opinions and feelings to the entire world.

Several researchers have worked in this direction to design an effective communication system. The basic sign gestures like numbers and alphabets can be obtained by using Arduino Nano and comparing its values with predefined values and converting the text into speech format via smartphone via text to speech engine within it [1]. The classification can also be done using some machine learning algorithm like support vector machine to attain high accuracy [2]. Another way of identifying the sign gestures is by using the state estimation method proposed in [3] which aims to track the motion of the hand in 3D space using Raspberry Pi, flex sensor and IMU. The system records the data of hand motion and compares it to a predetermined database. The hand and fingers movement can be determined by the EMG sensor which uses electrical signals generated from the muscles for determining the movements. The classification of these EMG based data along with IMU is done based on LSTM. However, the individual based model outperforms the general model, because the gestures performed vary from person to person [4].

In the present work, a wearable glove with the flex sensors and IMU mounted on it are used along with the STM32F103C8T6 MCU for the processing the data which is solely responsible for taking input from the sensors and then processing of the ML model for sign gesture classification is done on computer.

II. METHODS



A. SYSTEM DESCRIPTION

Fig.1: Block Diagram of the System

Fig.1 shows the block diagram of the system designed in the laboratory. Sign gesture performed by hand using a wearable glove fused with flex sensors and IMU is taken as an input. The Flex sensors account for the bending of fingers whereas MPU6050, the IMU, accounts for the angular momentum of the hand. Flex sensors are used to measure the resistance across the strip. The bending resistances may vary from 45 to 125 Kohms. MPU6050 is a 6-axes motion tracking device, in which 3-axes acts as accelerometer and rest 3-axes act as gyroscope. Thus, it gives a total of 6 values as measured output which defines the shift in angular momentum of the hand from one instance to the another. The input data collected by these sensors is fed to the microcontroller (STM32F103C8T6) for further processing. It performs computations on the measured voltage values received from the sensors and returns its values to the computer in a specified way.

Fig.2: Wearable Glove

Fig.3: Complete Setup of the System

All the electronic components have been mounted onto a PCB so as to make the system portable and light. The sensors are connected to the PCB board via connectors provided by the supplier. Flex sensors are attached to each finger of the rubber glove for recording the bending movement of five fingers and MPU6050 is attached alongside it for capturing hand movements. Fig. 2 shows the wearable glove whereas Fig. 3 shows a screen shot of the complete system for conversion of ISL signs into legible text.

B. SOFTWARE

The design and development of the present system is divided into hardware and software. Flex sensor is procured from the https://quartzcomponents.com site. Most of the electronic components are soldered on the PCB for its convenience and compactness. For software development and, data acquisition and analysis, Arduino IDE and Python language are used. The software part includes collecting data, preprocessing it, training the preprocessed data and deploying the model on STM32F103C8T6. Data is collected and stored into csv file format. In the present work, as the preliminary testing, the dataset consists of two categories of data i.e. for two signs ‘welcome’ and ‘bye-bye’. Data is collected by a single subject and with 30 samples for each sign. Thus, making it a total of 60 csv files. Each csv file contains 50 rows as time points and 11 columns as sensor values. The labeling is done in annotated form, the file name is the label itself [5]. This data is preprocessed and converted into numpy format for more efficient machine learning model training. The model is trained based on the LSTM algorithm with 30 epochs that achieved almost 100% accuracy with the validation data.

III. RESULTS

The collected dataset for two sign gestures is showing variations when their values are plotted into a graph. Although there are slight variations in between them still these signs are differentiable. In time series data, the instance of a particular time point plays a significant role. The values at the initial time point and at the end point depict a significant difference, however it is tough to describe these time points on a graph, which could give a path or pattern of how exactly data is moving/behaving.

The graphical representation of the dataset results in better understanding of the data. To represent all the 30 samples of data into a single graph is challenging and complex at the same time and when the 30 samples were plotted on a graph, the data points were cluttered into each other and were a bit complex to interpret. Therefore, to reduce its complexity, the average of the same cell of all samples has been calculated and the average values for each cell is stored into another csv file. This was calculated for each cell of the data, like, the value in the position [5,5] has been averaged with the values of all csv files on the same position. This similar calculation was completed for each position. This resulted in a new csv file of dimension 50x11. The graph for this csv file represents the data in a concise way. It represents a similar pattern to the pattern formed after plotting all data in a graph.

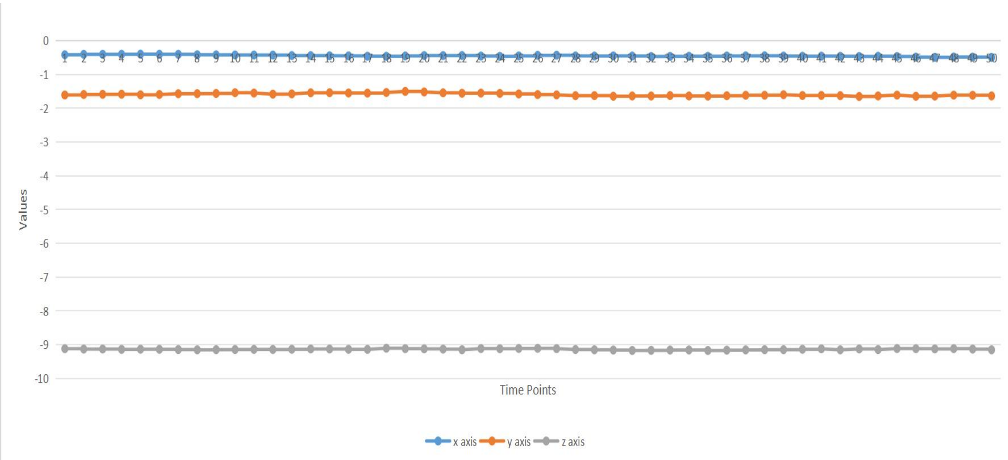


Fig.4: Variation of Accelerometer Sensors' Output with Time for "WELCOME"

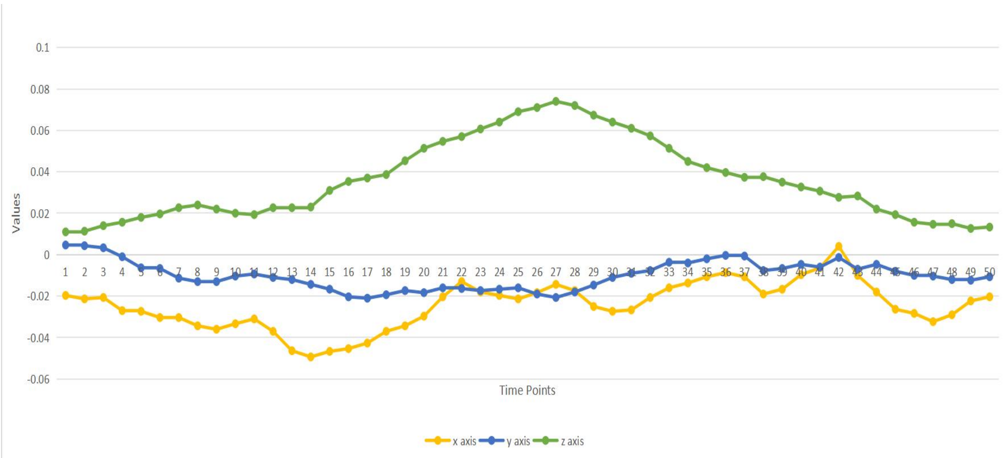


Fig.5: Variation of the Gyroscope Sensors' Output with Time for "WELCOME"'

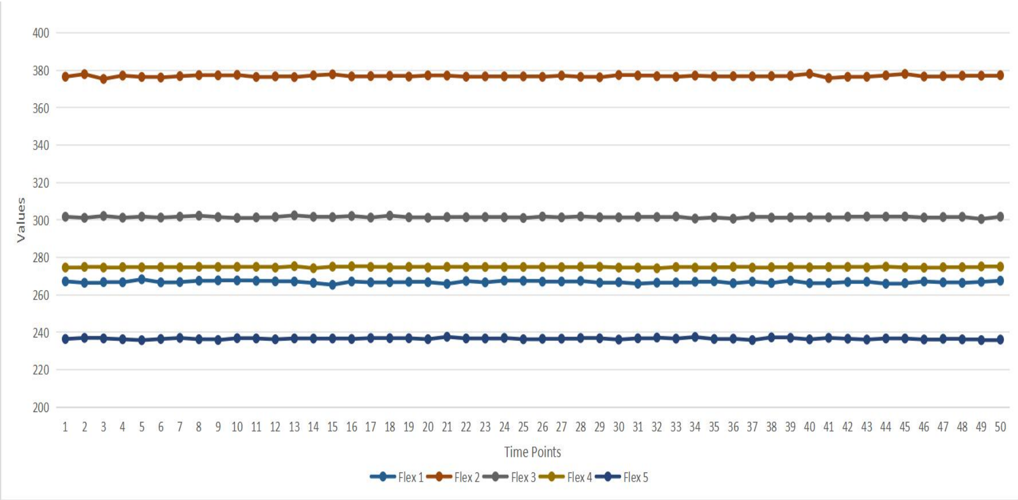


Fig.6: Variation of Flex Sensors' Output with Time for "WELCOME"

Fig. 4 and Fig. 5 represent the pattern of the x, y and z axes of the accelerometer and gyroscope, respectively for the "WELCOME" gesture. The accelerometer shows gradual fluctuations in the output signal values along each axis. This shows that the data collected for "WELCOME" gesture at each x, y, z axis is almost similar for all the samples. Thus, it can be significantly distinguished from other gesture signs. Whereas, the output signals at the gyroscope shows noticeable fluctuations. These fluctuations are within an extremely small range which represents the values collected for this sign does not include the angular movement of the hand as "WELCOME" gesture is typically a horizontal movement of hand from a distance from the chest towards chest with palm of hand facing the sky. Also, this sign does not involve any flex (finger) movements, therefore the values of all the flex sensors remain constant in the graph as shown in Fig. 6.

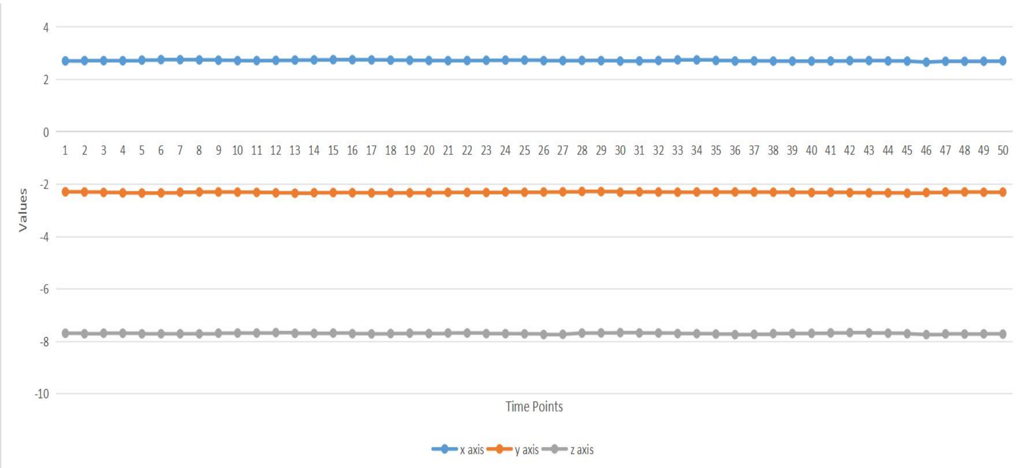


Fig.7: Variation of Accelerometer Sensors' Output with Time for "BYE-BYE"

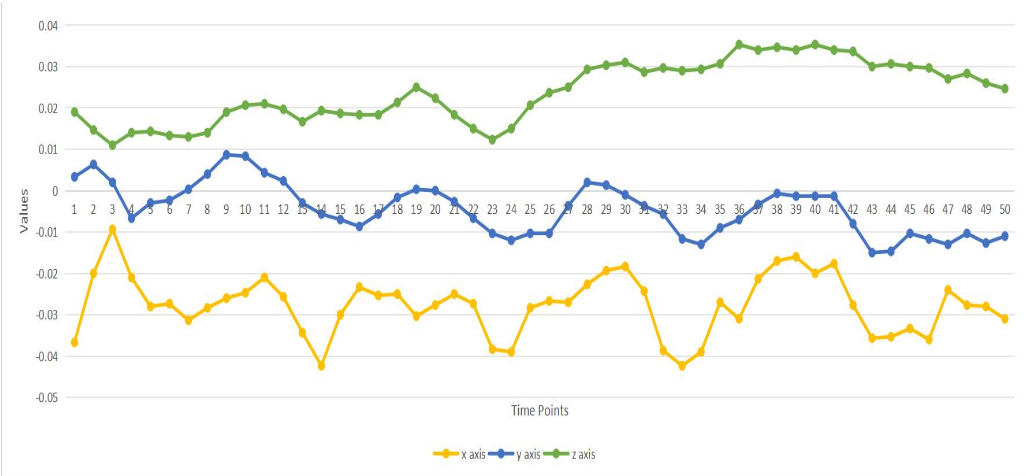


Fig.8: Variation of Gyroscope Sensors' Output with Time for "BYE-BYE'

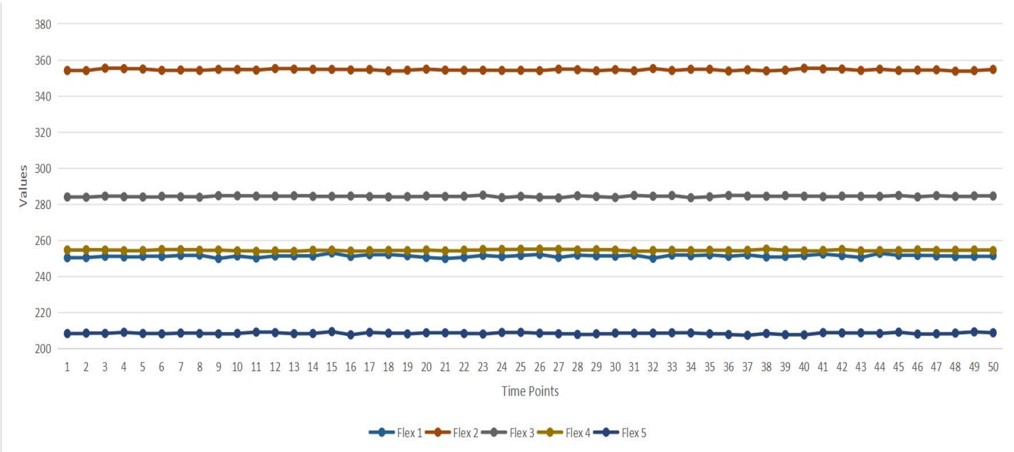


Fig.9: Variation of Flex Sensors' Output with Time for 'Bye-Bye'

The figures 7, 8 and 9 show variation of the output signals at the accelerometer, gyroscope and flex sensors, respectively for the sign "BYE-BYE". The accelerometer output signal values are not varying much as expected and almost follow the straight line with different values as compared to the "WELCOME" gesture. The gyroscope output signal values are slightly in a higher range than that of "WELCOME" gesture values and it has formed a pattern, which depicts that the repeated action has occurred which sounds convincing because the "BYE-BYE" gesture is just similar to what everyone performs to signal indicating departure. This action includes the movement of the hand and its action is just like waving a hand. Therefore, "BYE-BYE" sign gesture is also distinguishable with other gestures as per the data collected.

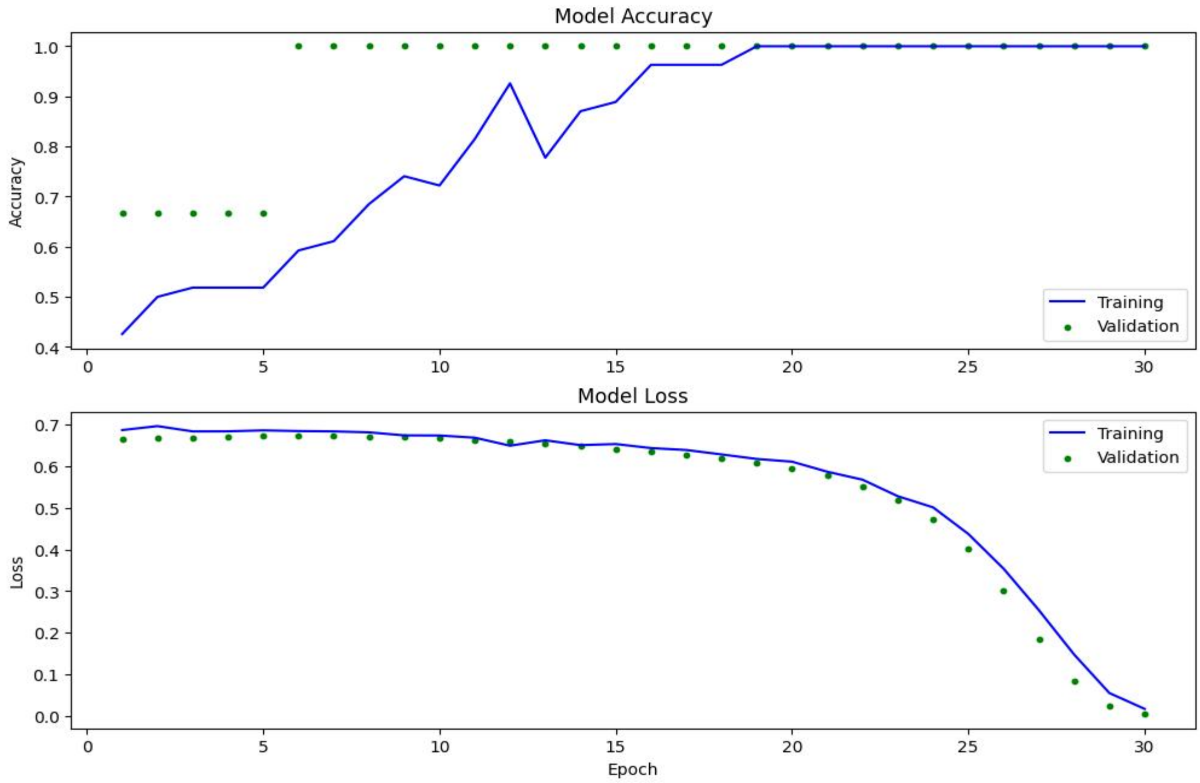


Fig.10: Loss and Accuracy Graph of ML model.

Fig. 10 shows that the model achieved a profound accuracy of 100% after 10-12 epochs. The model has only been trained with 60 samples of data collected from a single person. The ML model is trained based on the LSTM algorithm with 30 numbers of epochs.

IV. CONCLUSION

The work presented formulates the sign gestures and converts the two hand gesture signs, “WELCOME” and “BYE-BYE” into text format. The present work is able to distinguish between the “WELCOME” and “BYE-BYE” sign gestures in real time with the machine learning model running on the computer itself. Although, the classification of the gestures is more promising when the model is tested with the person’s gesture whose data has been collected for training ML model. More data can be collected from various users in order to get more promising results which will certainly add the flexibility trait to the present work.

REFERENCES

[1] Y. Jhunjhunwala, P. Shah, P. Patil, and J. Waykule, “Sign Language to Speech Conversion using Arduino”, *Int. J. Recent Innov. Engg. Res., vol. 2, pp. 37-44, 2017*

[2] M. M. Chandra, S. Rajkumar, and L. S. Kumar, “Sign Languages to Speech Conversion Prototype using the SVM Classifier”, *TENCON 2019 - 2019 IEEE Region 10 Conference, pp. 1803-1807, Kochi, India, 2019*

[3] K. A. Bhaskaran, A. G. Nair, K. D. Ram, K. Ananthanarayanan, and H. R. Nandi Vardhan, “Smart gloves for hand gesture recognition: Sign language to Speech Conversion system”, *International Conference on Robotics and Automation for Humanitarian Applications (RAHA), pp. 1-6, Amritapuri, India, 2016*

[4] S. Shin, Y. Baek, J. Lee, Y. Eun, and S. H. Son, “Korean Sign Language Recognition using EMG and IMU Sensors based on Group-dependent NN Models”, *IEEE Symposium Series on Computational Intelligence(SSCI), pp. 1-7, Honolulu, USA, 2017*

[5] E. Abraham, A. Nayak, and A. Iqbal, “Real-Time Translation of Indian Sign Language using LSTM,” *2019 Global Conference for Advancement in Technology (GCAT). IEEE, Oct. 2019*

[6] S. Y. Heera, M. K. Murthy, V. S. Sravanti, and S. Salvi, “Talking hands - An Indian Sign Language to Speech Translating Gloves”, *International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), pp. 746-751, Bengaluru, India, 2017*

[7] K. Suri and R. Gupta, “Convolutional Neural Network Array for Sign Language Recognition using Wearable IMUs”, *6th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 483-488, Noida, India, 2020*