International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

Conversion of Sign Language to Text and Speech and Prediction of Gesture

Bharath A Manoj

Abstract: ***People with the inability to speak, use sign language for communication. Ordinary people usually find it very difficult to communicate with mute people due to their lack of understanding of the universal sign language. This paper aims to provide a solution for this very problem through a device that uses an Arduino Uno board, some flex sensors and an Android application to facilitate interaction amongst the users. The flex sensors detect the movements and gestures of the wearer and based on the established conditions for the different values generated, respective messages are sent using a Global System for Mobile (GSM) Module to the user's android device which translates the text message to speech. The GSM module also attempts to create parameters for gesture predictions by sending sensor inputs to a cloud-based server for future reference. The application is ever learning and continues to evolve to be more reliable by examining user behaviours at all times. The use of this device allows mute people to convert sign language to speech, thereby making it significantly easier to talk to others, especially those who do not know sign language. This device empowers mute people and opens them up to previously unattainable opportunities.***

Keywords: ***Particle swarm optimization-back propagation, Sign Languages, Mute people, Glove based device***

I. INTRODUCTION

In a fast-paced world where even normal people find it hard to communicate, imagine the plight of those without the ability to speak.[4] Literally unable to voice their opinions ad thoughts, they pass through life wishing for that which most of the world take for granted. The world is a very different place to a mute person. Communication is one way most of the time, as most people do not understand sign language.[2] All of the day to day activities of a normal person is exponentially more difficult for a mute person. From ordering food at a restaurant, to asking for directions, to even asking a question in class, everything is laborious. [3]

Sign language is a means of communication for mute people using hand gestures. Different hand formations are associated with different letters of the alphabets and shown consequently in order to form meaningful words and phrases. The most common type of sign language is American Sign Language (ASL).

ASL saw its origin in the American School for Deaf in West Hartford in the early 19th century.

Manuscript received on March 15, 2020.
Revised Manuscript received on March 24, 2020.
Manuscript published on March 30, 2020.

* Correspondence Author

Bharath A*, Bachelor's Degree in Computer Science, Mathematics and Electronics.

The school was founded by the Yale graduate Thomas Hopkins Gallaudet who travelled to Europe to learn about deaf pedagogy. He adopted the techniques of the French Institute National de Jeunes Sourds de Paris. Hence ASL is closely related to the French Sign Language (LSF), so much so that about 58% of signs I modern ASL are similar to those in the French Sign Language.

Sign language was invented to allow mute people to communicate with others by using different movements and gestures of the hand each of which have an assigned meaning. However, communication using sign language is difficult as most ordinary people do not learn sign language. "EASY TALK", a glove with flex sensors is being developed to tackle this problem. It implements the American Sign language gestures with the help of these flex sensors [1]. The different gestures are accurately measured and their values and meanings recorded. This information is the sent to the user's android device and using our application, the text is converted into speech using a speaker. The device uses four flex sensors and the information is taken as inputs for Back - propagation neural network for future references. [5]

The rest of the paper is organized as follows. Section 2 gives an overview of the literature review done. Section 3 mentions the proposed system followed by Section 4 which talks about the proposed methodology. Section 5 includes experimental results after which the conclusions and future scope for this research are discussed about.

II. LITERATURE REVIEW

Almost 72 million of the world's population is of deaf/mute people as per the report published by the World Federation of the Deaf [2].Sign language is a fundamental human right for deaf people, from which it is possible to achieve all other human rights [3].There are over 240 sign languages for spoken language around the world [4]. There are a lot of glove-based devices available in the market for different purposes. One of the most important features of glove-based devices is to provide communication for the mute people through identification of gestures and converting it into real-time. Recognition, generation, and translation of sign language is a research area with high potential impact. Sign language processing would reduce the barrier for the mute people while communicating [5]. Google has introduced a wrist band which identifies the gesture shown by the user which is later approved as only theoretically possible [6]. TOSHIBA has also developed an android that helps in identification of sign language [7]. In the existing systems, the device just converts the sign language to its corresponding text or speech. There is no provision of storing the data or prediction of gestures.

One of the best algorithms for prediction is the Particle swarm optimization – back-propagation (PSO-BP). This algorithm is used in the prediction of performance in education management [8]. It is also used in prediction of stay of a patient in a hospital [9]. The PSO-BP algorithm can balance the relationship between energy and economic growth [10].

III. PROPOSED SYSTEM

This proposed system not only detects gestures and converts them to speech, but also predicts the mute person's needs. It connects the flex sensor outputs with the user's needs and tries to predict them using the time of day as a reference. It uses a Particle swarm optimization - backpropagation for its prediction. The system inputs various data in to the neural network such a s time of day, sensor values etc. and trains the neural network to make predictions about the mute person's needs.

IV. PROPOSED METHOD

The methodology followed for this work is represented in Figure 1.

When the gestures are shown by the user by wearing the gloves, the flex sensors in the gloves produce sensor values. These values are sent to the Arduino for the identification of the shown gesture.

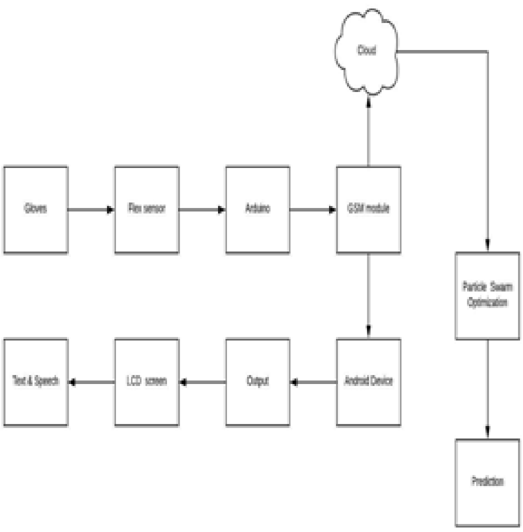

Fig 1. Block diagram

The GSM module, which is connected to the Arduino, sends the corresponding message to the users android device as well as LCD display. It also sends the sensor values to the ThingSpeak cloud from which we can retrieve the values for prediction.

With the help of the block diagram in Fig 1, a circuit of the same was made as shown in Fig 2. Four flex sensors are used in this where one end of each flex sensor is connected to A0, A1, A2, A3 of the Arduino uno board respectively and to the 5V power supply. A 10kΩ resistor is also connected to it. The other end of each flex sensor is connected to the ground. These flex sensors are fixed on the four fingers of the gloves. The Rx and Tx of GSM module is connected to Tx and Rx of the Arduino respectively. Both the GSM and Arduino should be grounded. D4, D5, D6, D7 of LCD should be connected to pin 5,4,3 and 2 of Arduino respectively. Additionally, wire a 10K potentiometer to the 5V power supply.

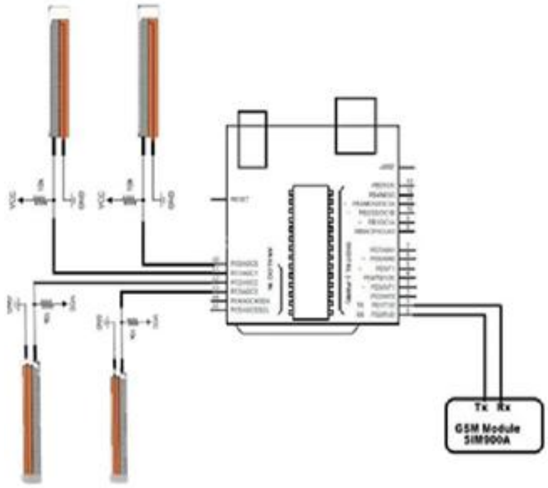

Fig.2.Circuit diagram

A. Particle-swarm optimization

Particle swarm optimization is an optimization method derived from the social behavior of bird flocking. Each entity stores coordinates in itself based on which the most apt solution is reached. The values are altered for each permutation. In our device, the artificial neural network is applied to the values generated from the flex sensors. The device then uses these values along with the information stored in the cloud to simulate a predictive model for the next day.

B. Back-propagation neural network

The back-propagation algorithm is used as a piece of layered feed for the artificial neural network. It has many layers with monitored learning frameworks in light of gradient descent learning rule. The algorithm is given situations of the sources of the inputs and outputs we need the framework to enlist and the errors that occur afterwards is identified. The purpose of this algorithm is to diminish these errors unto the point where the artificial neural network takes in the training information.

The corresponding weights w_{ji} is multiplied with the sum of inputs x_i for generating activation function.

$A_j(\overline{x},\overline{w})=\sum_{i=0}^{n} x_i w_{ji}$

Sigmoidal function is utilized as the result function in this work.

$O_j(\overline{x},\overline{y}) = \frac{1}{1+e^{A_i(\overline{x},\overline{w})}}$

Since the error correction between legitimate and predicted results, the error depends on the weights and we have to change the weights so as to reduce the error. We can show the error function for the result of every neuron:

$E_j(\overline{x},\overline{w},d = (O_j(\overline{x},\overline{w} - d_j)^2$

Then it is discovered how error value is affected by inputs, outputs and weights.

$\Delta W_{ji} = -\eta \frac{\partial E}{\partial W_{ji}}$

$\frac{\partial E}{\partial O_j} = 2(\,O_j - d_j\,)$

Now, to discover how much the result relies on upon the activation and the weights, we calculate the following:

$\frac{\partial O_j}{\partial W_{ji}} = \frac{\partial O_j}{\partial A_j}\frac{\partial A_j}{\partial W_{ji}} = O_j(1-O_j)x_i$

Retrieval Number: F9502038620/2020©BEIESP
DOI:10.35940/ijrte.F9502.038620
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The variation in regards to the weights will be

$\Delta W_{ji} = -2\eta(O_j-d_j)O_j(1-O_j)x_i$

If we are to alter the weights of a previous layer, we expect first to record how the error relies not on the weights, yet in the findings in the previous layer i.e. outing replacing W by x as shown in the following equation.

$$\Delta V_{ik} = -\eta\frac{\partial E}{\partial V_{ik}} = -\eta\frac{\partial E}{\partial x_i}\frac{\partial x_i}{\partial V_{ik}}$$

Where

$$\frac{\partial E}{\partial W_{ji}} = 2(O_j-d_j)O_j(1-O_j)W_{ji}$$

$$\frac{\partial x_i}{\partial V_{ik}} = x_i(1-x_i)V_{ik}$$

C. Flex sensors

The device uses flex sensors to detect gestures. The sensors are linked to the gloves of the mute person. The movements of the flex sensor in different angles give the respective numerical values. The different flex angles change the value of the resistors for its unusual un-flexed value of 10k. This work uses four flex sensors and the values are inserted into the neural network. The configuration of the flex sensors has two methods: firstly, calculate the change in voltage as per the bend and secondly, the conversion of the voltage into current so as to generate the value of current according to the flex.

D. Arduino Uno

Arduino Uno is a micro-controller board which has 14 digital (input/output) pins. 6 of these pins can be used as paw output. It also has 6 analog pins, a power jack and a USB connection. It can be connected to a computer using a USB cable, or by an AC-DC adapter of battery. The source code for Arduino Uno is added used the Arduino Software (IDE) and uploaded t the micro-controller. A serial monitor is used to see the values of the flex sensors. The micro-controller is connected with flex sensors, Sim800A GSM module and an LCD display. The Arduino has various pins to which all its components are connected and the source code is generated based on such pins to which each component is connected.

E. SIM 900A

This project uses a SIM900A GSM module to send SMSs to the receiver's mobile device. This GSM module is also used to send values to the cloud. SMS control, data transfer and remote control is done using the GSM module. We use the ThingSpeak cloud to acquire values from our device. The ThingSpeak cloud is also where the values are plotted. It can be used in different situations, eg if the pulse sensor values exceed its normal range, a message can be sent to the mobile device of the mute person, his doctor or his relative.

F. Android application

This project involves the development of an android application which translates text to speech. On receipt of a text message by the user, the application will display the message and it will be put out through the speakers of the user's device. The user may replay the message by pressing the "TEXT TO SPEECH" button. The application is given permission to access the contact name and the text messages so as to screen messages from different sources as per the user's need. The only goal of the application is to display and covert the message to text whenever one is received from the GSM module. The user can listen to every new message by pressing the button and replay old messages by pressing he same button again.

E. MATLAB

MATLAB can be used for various purposes like recognizing patterns, detecting gestures etc. In this project the MATLAB neural network tool is used to implement training and predicting flex sensor values. It has the provision to implement various algorithms and methods such as back-propagation neural network, particle swarm optimization etc. The neural network tool can be used for training the data and simulate the results and the input, output and training sample values can be easily inserted. The variation between the predicted values and the actual values can be understood by posing them accurately.

V. EXPERIMENTAL RESULTS

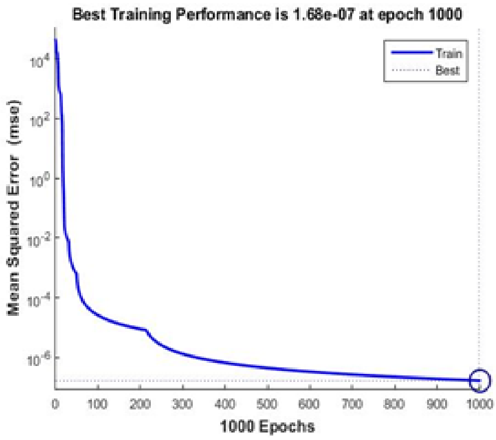

Fig 3. Plotted graph of mean square error of the data passed to the network

The above table shows the set of values estimated from the neural network based on the inputs. It was then compared to the expected sensor values of the same. The difference between these two values provided the mean square error, also known as mean square deviation. The mean square deviation values which are close to zero are good estimators.

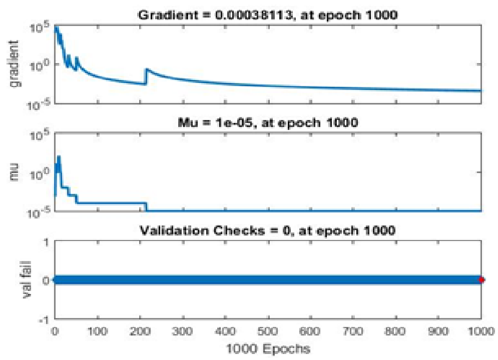

Fig 4. Plotting the gradient values, mu and validation fail

The gradient points to the direction in which there is high rate of increase for the considering function. 'mu, is the control parameter for the prediction model that we used. Validation checks helps to terminate the study by neural network.

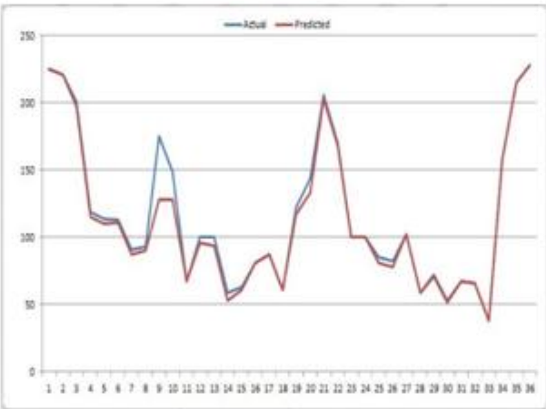

Fig 5. Plotting actual data along with predicted data

The predicted value is plotted along with the actual value. The x axis represents the hours of the day and y axis represents the sensor values.

VI. CONCLUSION

This project associates the different needs of mute people to the values of flex sensors in order to predict their needs using back propagation neural network. With the use of this device mute people can vastly enhance their ability to communicate to others and improve their lives. The neural network model adjusts their weights and ensures that the predictions are accurate. Future projects can be carried out to use the predicted values from this model to the mute person's android application so as to facilitate communication between them and others. This project not only converts the American Sign language to speech in real time, it also involves the prediction of the mute person's needs. This allows for greater clarity on what exactly a mute person needs at all times of the day with much greater efficiency. Future enhancements can be done so as to make the predictions available on the mute person's mobile device, allowing them to evaluate their accuracy. These evaluations can be used as training inputs to further improve the future predictions. Overall this device reduces the gap between mute people and others, allowing them to express themselves more and be involved in the world in a way that they have previously not been able to do.

FUTURE SCOPE

There are many more works that can be carried out as an extension of this project. This system predicts the need of the mute person but future systems may be developed that could communicate to the mute person's mobile device, allowing the system to learn the needs of the user, thereby provisioning the development of recommendatory systems as they have the relevant data related to the mute person that can easily be learned thought he neural network model.

REFERENCES

1. Mr Jose L. Hernandez-Rebollar" A New Instrumented Approach For Translating American Sign Language Into Sound And Text"
2. Ms Maryam Rehman" American Sign Language Translation through Sensory Glove; SignSpeak"
3. Ms Jenny Nilsson" WORLD FEDERATION OF THE DEAF"
4. Mr Taner Arsan and Oğuz Ülgen" SIGN LANGUAGE CONVERTER"
5. Mr Oscar Koller" Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective"
6. http://gizmodo.com/electronic-wristbands-translatesign-language-into-smar-1594190782
7. http://gizmodo.com/toshibas-eerie-sign-languagerobot-will-silently-stare-1643265457
8. Mr Hongyi Li" PSO-BP Neural Network Grade Prediction Model Based on Bagging Ensemble Learning"
9. Mr Abhijit Suresh" Particle swarm optimization over back propagation neural network for length of stay prediction"
10. Mr Rencai Gao" Application of BP Neural Network Optimization Based on PSO Algorithm in Energy Development"

AUTHORS PROFILE

Bharath A Manoj was born in February 1997 and has a Bachelor's Degree in Computer Science, Mathematics and Electronics. He is passionate in the field of UI Designing.

Retrieval Number: F9502038620/2020©BEIESP
DOI:10.35940/ijrte.F9502.038620
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication