Recent Trends in Intensive Computing
M. Rajesh et al. (Eds.)

doi:10.3233/APC210272

Deep Learning Based Indian Sign Language Words Identification System

P. GOLDA JEYASHEELI [a,1], N. INDUMATHI [b]
[a]*Professor, Department of Computer Science and Engineering, Mepco Schlenk Engineering College, Sivakasi, TN, India*
[b]*Assistant Professor, Department of Computer Science and Engineering, P.S.N.A College of Engineering and Technology, Dindigul, TN, India*

Abstract. In Indian Population there is about 1 percent of the people are deaf and dumb. Deaf and dumb people use gestures to interact with each other. Ordinary humans fail to grasp the significance of gestures, which makes interaction between deaf and mute people hard. In attempt for ordinary citizens to understand the signs, an automated sign language identification system is proposed. A smart wearable hand device is designed by attaching different sensors to the gloves to perform the gestures. Each gesture has unique sensor values and those values are collected as an excel data. The characteristics of movements are extracted and categorized with the aid of a convolutional neural network (CNN). The data from the test set is identified by the CNN according to the classification. The objective of this system is to bridge the interaction gap between people who are deaf or hard of hearing and the rest of society.

Keywords. Convolutional neural network, wearable device, sign language, sensors, gestures

1. Introduction

Deaf and dumb people communicate using physical gestures in Sign language. Hand motions and body expressions are used to express the message. Indians choose a sign language which is known as Indo-Pak sign language, that is a combination of Indian and Pakistani sign languages [1]. Based on the expressed signals, the movements are divided into two categories: The term "manual signals" refers to finger movement, hand orientation, and posture. Mouth action, face expression, and body posture are all important factors to consider as non-manual signals [2]. A sign language identification system is required for non-deaf and mute people to understand deaf and mute people's hand gestures. A smart framework for gesture recognition is being designed for this purpose. The Indian sign language is effectively identified by a smart sign language recognition device in this article. Wearable sensor-based gloves will be used in the smart sign recognition system to decode Indian sign language gestures. On each hand, there are five flex sensors, a pressure sensor, and a Bluetooth module connected to microcontrollers on both sides [3]. The sensor data is gathered in order to identify gestures. The microcontroller is connected to the Bluetooth modules in both hands. After that, the sensor information is sent to the CNN, which categorizes and displays the identified sign term [4].

[1]N. INDUMATHI, Assistant Professor, Department of Computer Science and Engineering, P.S.N.A College of Engineering and Technology, Dindigul, TN, India; Email: indhujegan1996@gmail.com.

2. Related Work

There are two techniques for determining the manual signals which are produced by the deaf and dumb people. The first is a direct approach, while the second is an indirect approach. The gesture is classified using neural networks in a direct approach. Changshui Zhang et.al. [11] suggested a method for converting sign language from videos into a sentence format. For feature extraction, the system uses a deep CNN, and a bidirectional RNN is used to understand the sequence of motions. Ying Xie et al [12] suggested Deaf and dumb people will get benefit from a computer vision-based approach to interpret sign language. This method converts sign language into text, making it easier for the people to interact. The CNN is being used in the model to derive the sign image's spatial features from a series of video clips. And, to train the temporal features, RNN is applied to the image's extracted features [5]. After building, this model illustrates and justifies the actual model to operate on the dataset by using the image dataset to train the network for categorizing the images. Until passing on to the next surface, Inception performs many of the parallel convolutions and combines the resulting feature maps [6].

The gesture is determined by a sensor wearable is being used in an indirect approach, with the neural network obtaining image as input. To determine the letters of American Sign language, the wearable interface uses flex, pressure, and three-axis inertial motion sensors. Sensor data is obtained and evaluated to use an embedded SVM classifier. The acknowledged alphabet is then sent through Bluetooth low-power wireless networking is used to connect to a mobile device [7]. A text-to-speech function has been added to an Android app, which transforms obtained text into loud audio files. A SVM tool used to classify the signs in several different categories [8-10]

3. Prototype Design

The structure of the system is composed of two different modules: one is a wearables interface module and another is sign identification module that identifies deaf and mute people's sign words. These modules will make deaf and mute people's lives easier people to reach out the general public. Figure 1 depicts the proposed system's overall architecture. The smart wearable system is the hardware part in this paper, and it is made up of flex, pressure and Gyroscope sensors that are all linked to the Arduino MEGA microcontroller [13]. They're stitched to the cloth glove to make it more accessible. The master and slave wearables communicate in serial fashion is developed using the HC-05 Bluetooth module. We used flex sensor with 4.5 inch and flex sensor 2.2 inch [14]. On both the hands, the 4.5-inch flexible sensors are used for all fingers except thumb finger. The thumb finger on both the hand uses the 2.2-inch flexible sensors.

Above the palms of both hands, the pressure sensor is mounted. A resistor is linked to the pressure sensor. To ensure that the current flow is stabilized, the Flexi sensor is mounted to a resistor. The above three sensors are interconnected to the microcontroller in order to obtain precise examinations for each sign generated by the smart wearables. Figure 2 depicts the equipment configuration for the wearable's module [15].

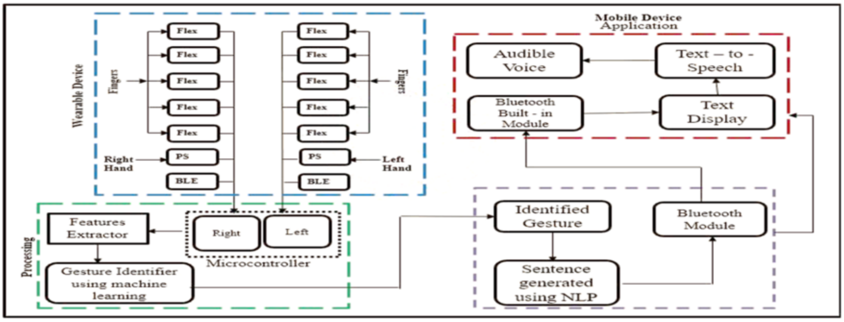

Figure 1. Proposed System Architecture

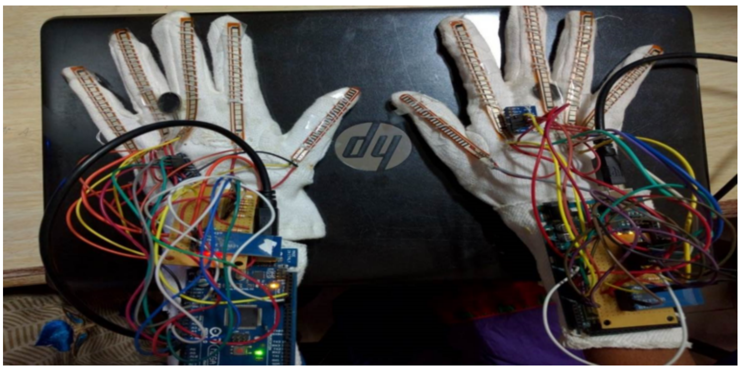

Figure 2. Hardware setup for the wearable device module

Sensor information is gathered from Fifty different individuals. The data collection was generated with all of the sensor data obtained from people who used the system. The sensor examinations are saved in a spreadsheet. The captured sensor information is used in the gesture identification module to classify and identify the gesture. For the categorization of the sign generated by the wearables module, the sensor information is analyzed using a CNN. Figure 3 represents a diagram of the CNN's architecture.

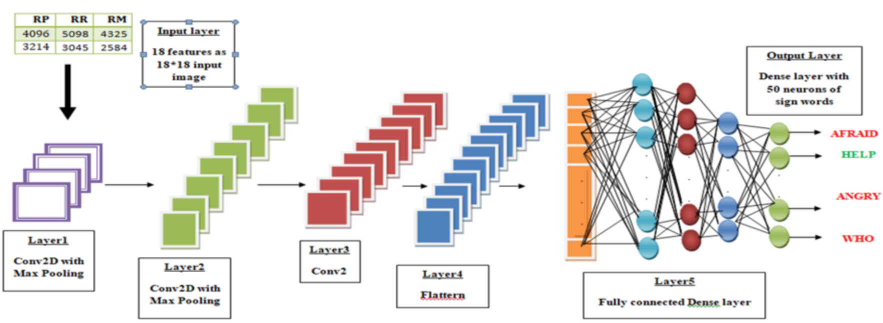

Figure 3. CNN Architecture

In the input layer, eighteen input attributes are passed into the CNN, including the value of all the 3 types of sensors used for all Fifty signs. The output layer would have the same number of neurons as the network's classes. It represents the estimated word as a value of one other word with the value of 0. As a result, the output layer generates Fifty neurons for each identified sign term.

4. Execution

The measured sensor data serve as bend categories, indicating how each term bends depending on its sensor data. To prevent differences between sensor readings, the wearable system generates normalized values. The flex bend values determine how movements are categorized. CNN are used to assess the system's identification. Although this system contains 50 signs with ordinal attributes, classification of the signs centred on the data from their sensors is necessary. To be able to accurately accommodate changes in hand movement, the gesture is identified with a duration interval of 10 seconds. CNN is loaded with random weights and implements numerical data conversion into image format for processing. After the image conversion, the images are stored as the image data store and perform convolution operations on the images. The validation of the training and testing set will be evaluated and the recognized words are displayed as the class labels based on the prediction of the test data. The image gets reshaped in the size of 18*18 for making all the images in the same size for validation. The class labels are displayed based on the prediction.

5. Results and Discussions

The network's accuracy is determined by analyzing the samples of both training and testing are shown in Figure 4. The data samples are collected at random, so further training would be required to boost accuracy. For better accuracy, a certain ratio of the dataset is retained. In this project, the training and testing ratio is defined as 60:40 to achieve higher accuracy. The 50 signs will be identified by the machine.

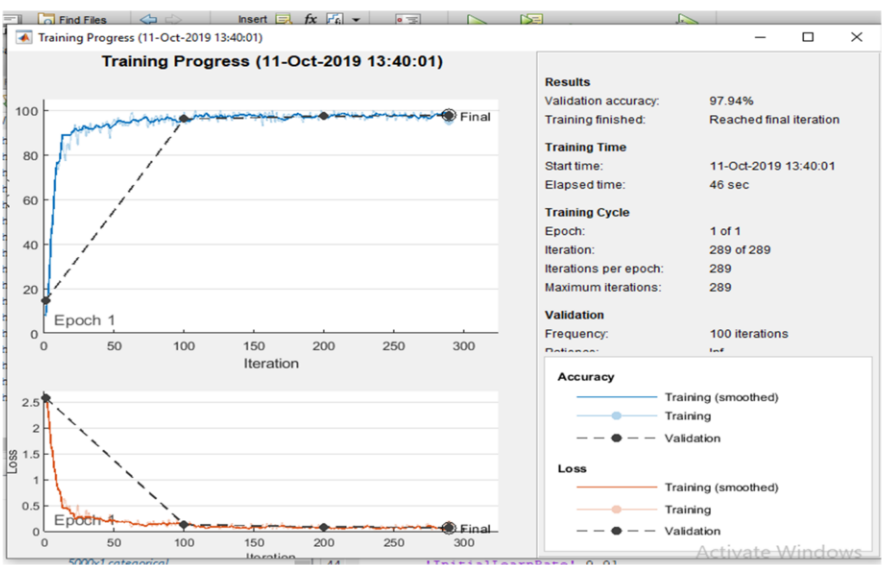

Figure 4. Training progress of CNN

There are 50 different samples in each sign. In total, there are 2500 samples. The accuracy of the proposed system's convolutional neural network is shown in Table 1.

Table 1. Convolutional Neural Network Accuracy

Trial Cases	Total no of Samples	Training samples	Testing Samples	Accuracy of CNN
Trial case 1	2500	1500	1000	89%
Trial case 2	2500	2000	500	**97.8%**

In contrast to current systems, the proposed system has improved accuracy. The accuracy of the proposed system is compared to that of the current system in Table 2.

Table 2. Accuracy comparison between other existing systems and proposed system.

Author	Used Methodology	Achieved Accuracy
Su Min Lee [5]	Sensor fixed gloves	92.5%
Gwang Soo Hongb [6]	Sensor Gloves	93.9%
Pamela Godoy-Trujill [10]	Sensor gloves	93.3%
Nishith A. Kotak [11]	Sensor gloves	91.5%
R Jafari [14]	Wrist-Worn, SEMG	91.2%
Proposed system	Sensor gloves	97.8%

6. Conclusion

A smart glove device is proposed in this project for categorizing Indian sign words. This smart glove captures the gesture created by deaf and dumb people. For identifying each sign word, the sensor values of the respective signs are processed and categorised using a CNN. This network helps to identify the word with the aid of training data. For each value of the 3 types of sensors, the network is trained with a collection of 18 neurons as inputs. The word is recognised based on the training results.

References

[1] Mittal, A., Kumar, P., Roy, P. P., Balasubramanian, R., & Chaudhuri, B. B. (2019). A Modified LSTM Model for Continuous Sign Language Recognition Using Leap Motion. *IEEE Sensors Journal*, *19*(16), 7056–7063. https://doi.org/10.1109/JSEN.2019.2909837

[2] Cui, R., Liu, H., & Zhang, C. (2019). A Deep Neural Framework for Continuous Sign Language Recognition by Iterative Training. *IEEE Transactions on Multimedia*, *21*(7), 1880–1891. https://doi.org/10.1109/TMM.2018.2889563

[3] Kaur, J., & Rama Krishna, C. (2019). An efficient indian sign language recognition system using sift descriptor. *International Journal of Engineering and Advanced Technology*, *8*(6), 1456–1461. https://doi.org/10.35940/ijeat.F8124.088619

[4] Bantupalli, K., & Xie, Y. (2019). American Sign Language Recognition using Deep Learning and Computer Vision. In *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018* (pp. 4896–4899). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/BigData.2018.8622141

[5] Lee, B. G., & Lee, S. M. (2018). Smart Wearable Hand Device for Sign Language Interpretation System with Sensors Fusion. *IEEE Sensors Journal*, *18*(3), 1224–1232. https://doi.org/10.1109/JSEN.2017.2779466

[6] Kim, J. H., Hong, G. S., Kim, B. G., & Dogra, D. P. (2018). deepGesture: Deep learning-based gesture recognition scheme using motion sensors. *Displays*, *55*, 38–45. https://doi.org/10.1016/j.displa.2018.08.001

[7] Bhujbal, V. P., & Warhade, K. K. (2019). Hand Sign Recognition Based Communication System for Speech Disable People. In *Proceedings of the 2nd International Conference on Intelligent Computing and Control Systems, ICICCS 2018* (pp. 348–352). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICCONS.2018.8663054

[8] Shaheen, H., & Mehmood, T. (2018). Talking Gloves: Low-Cost Gesture Recognition System for Sign Language Translation. In *2018 IEEE Region 10 Symposium, Tensymp 2018* (pp. 219–224). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/TENCONSpring.2018.8692069

[9] Jadhav, D., & Tripathy, A. (2018). Gesture Aided Speech for Deaf and Mute. In 2018 3rd International Conference for Convergence in Technology, I2CT 2018. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/I2CT.2018.8529411

[10] Rosero-Montalvo, P. D., Godoy-Trujillo, P., Flores-Bosmediano, E., Carrascal-Garcia, J., Otero-Potosi, S., Benitez-Pereira, H., & Peluffo-Ordonez, D. H. (2018). Sign Language Recognition Based on Intelligent Glove Using Machine Learning Techniques. In *2018 IEEE 3rd Ecuador Technical Chapters Meeting, ETCM 2018*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ETCM.2018.8580268

[11] Jani, A. B., Kotak, N. A., & Roy, A. K. (2018). Sensor Based Hand Gesture Recognition System for English Alphabets Used in Sign Language of Deaf-Mute People. In *Proceedings of IEEE Sensors* (Vol. 2018-October). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICSENS.2018.8589574

[12] Oliveira, T., Escudeiro, P., Escudeiro, N., Rocha, E., & Barbosa, F. M. (2019). Automatic sign language translation to improve communication. In *IEEE Global Engineering Education Conference, EDUCON* (Vol. April-2019, pp. 937–942). IEEE Computer Society. https://doi.org/10.1109/EDUCON.2019.8725244

[13] Galka, J., Masior, M., Zaborski, M., & Barczewska, K. (2016). Inertial Motion Sensing Glove for Sign Language Gesture Acquisition and Recognition. *IEEE Sensors Journal*, *16*(16), 6310–6316. https://doi.org/10.1109/JSEN.2016.2583542

[14] Wu, J., Sun, L., & Jafari, R. (2016). A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors. *IEEE Journal of Biomedical and Health Informatics*, *20*(5), 1281–1290. https://doi.org/10.1109/JBHI.2016.2598302

[15] Hsu, Y. L., Chu, C. L., Tsai, Y. J., & Wang, J. S. (2015). An inertial pen with dynamic time warping recognizer for handwriting and gesture recognition. *IEEE Sensors Journal*, *15*(1), 154–163. https://doi.org/10.1109/JSEN.2014.2339843