Korean Sign Language Translation using Machine Learning

Angela Caliwag
Department of IT Convergence Engineering
Kumoh National Institute of Technology
Gumi, South Korea
a_caliwag@kumoh.ac.kr

Stephen Ryan Angsanto
Department of Electronics Engineering
Kumoh National Institute of Technology
Gumi, South Korea
angsanto@kumoh.ac.kr

Wansu Lim
Department of Electronics Engineering
Kumoh National Institute of Technology
Gumi, South Korea
wansu.lim@kumoh.ac.kr

Abstract—**Advancement in technology provides a means of doings things that even humankind is incapable of doing. A sign language glove aims to help the hearing-impaired people to eliminate the boundary and communicate with other people even without knowledge in Sign Language. Applying Machine Learning in hand gesture recognition provides a smarter means of conveying information with high accuracy.**

Keywords-Korean Sign Language (KSL); AlexNet transfer learning; Machine Learning (key words)

I. INTRODUCTION

Nowadays, continues innovation and invention enables mankind to live an easier and more comfortable lives. Even people with disabilities are not exceptions as benefactors of the advantages of the advancement in technology. High-tech machines and devices has now the capability of aiding disabled people by providing them means of doing things that they are restricted to do due to their disabilities.

According to the Korean Statistical Information Service, as of 2016, there are 271,843 registered residents with hearing disability and 19,409 registered residents with speech disability. [1] In order communicate with others, people around the hearing-impaired people must learn the Korean Sign Language. However, with the development of a glove that directly translate the hand gestures, communicating with people without the knowledge of the Korean Sign Language.

This paper deals with the design of a sign language glove that uses Machine Learning algorithm to directly translate hand gestures. This study proposes only a one hand-glove that can classify leftward and rightward movement. It aims to benefit hearing-impaired people and serve as a basis for further researches and innovations.

In this paper, we apply two methods: construction of Neural Network from scratch and AlexNet transfer learning.

II. REVIEW OF RELATED LITERATURE

In [2], a dynamic gesture recognition system for the Korean Sign Language (KSL) was proposed. The pair of gloves used by the researchers utilizes a flex-sensor system consisting of fiber optic transducer responsible for measuring the flexing angles of the fingers. It also uses a Real-time 6DOF finger and hand tracking sensor system responsible for measuring the movement of the hand in the x, y and z axis. A fuzzy min-max neural network was used for on-line pattern recognition. The researchers were successful in the classification of almost 85% of the given words.

Another related research conducted in [3] uses an electromyography (EMG) sensor and an inertial measurement (IMU) sensor. An EMG sensor records the electrical activity of muscle tissue in response to a nerve's stimulation of the muscle. It is used in diagnosing muscle or nerve disorders. [4] The IMU sensor, usually used together with gyroscopes and accelerometers, measures linear and angular motion. [5] Instead of using gloves, the researchers used an armband called MYO. The researchers devised three Neural Network (NN) architectures: Convolutional Neural Network (CNN) architecture without dropout, CNN architecture with dropout, and Long Shot Term Memory (LSTM) architecture. The accuracy of each of the NN model in understanding sign language is determined and the researchers concluded that the CNN architecture shows better performance than the LSTM architecture in learning and recognizing KSL.

In [6], the researchers propose a two-way communication system for deafened and speechless people to those who are not familiar with sign languages. The study consists of two systems: hand sign recognition system and speech recognition system. For the hand sign recognition system, a camera is used as an input device and the real-time image processing capability of the Open Source Computer Vision Library (Open CV) for the sign language recognition. This system is proficient in providing the hearing and speech-impaired people to speak out and communicate with others. For the speech recognition system, the researchers used the Hidden Markov Model (HMM) algorithm. The recognized speech is then converted into text. This system provides a means for the hearing and speech-impaired people to understand the outside world.

III. METHODOLOGY

A. Hardware

The researchers created a breadboard-based testbed composed of Adafruit Precision NXP 9-DOF (Degree of Freedom) Breakout Board and the Adafruit Feather nRF52 Bluefruit LE. The Adafruit Precision NXP 9-DOF Breakout Board, capable of motion and orientation sensing, is a combination of FXOS87000 3-Axis accelerometer and FXAS21002 3-Axis gyroscope. Adafruit Feather nRF52 Bluefruit LE is a Bluetooth Low Energy board with a native-Bluetooth chip, Nordic's nRF52832, which used as a main microcontroller.
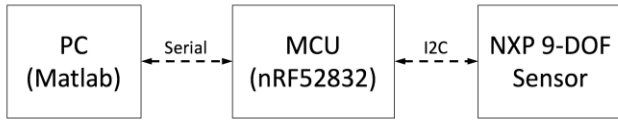

Fig. 1. Single gyro-accelerometer Block Diagram

Figure 1 shows the block diagram of the single gyro-accelerometer hardware prototype. The prototype was assembled on a breadboard for quick setup and testing. Although the nRF52 Bluefruit LE has a build-in Bluetooth chip, Serial Communication was used to send the sensor's data to Matlab during testing.

B. Software

The researchers use two NN architectures: a NN constructed from scratch and a predefined CNN, AlexNet. The gesture recognition time, learning time and accuracy of the two architectures are compared and analyzed.
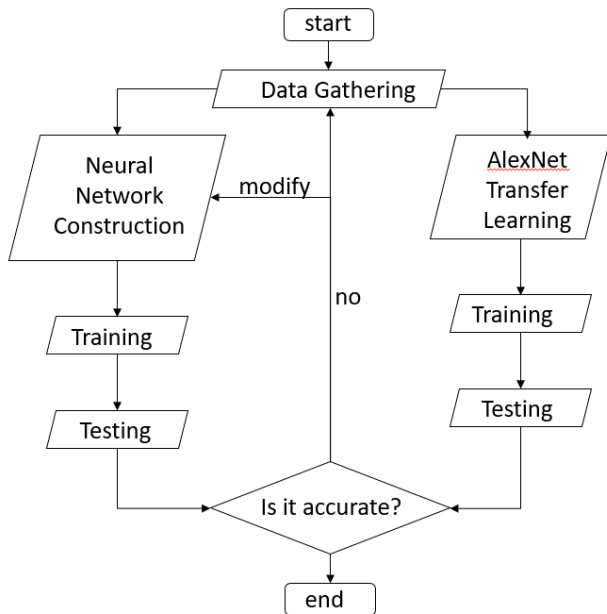

Fig. 2. Conceptual Framework

The conceptual framework of this research is shown in Fig. 2. The first step done was to gather and analyze data. After some analysis and interpretation, a neural network was constructed, trained and then tested. Using the same data, a predefined CNN, AlexNet was trained using transfer learning. The results of both methods were observed and compared.

To simplify gathering of data, a base point was set. That is, the sensor goes back close, if not exactly, to the initial point after a gesture was made as shown in figure 3. Thus, to be able to be recognized as one complete hand gesture, the sensor has to be positioned back to the base where it can start another gesture again. For further research, it would be better if the machine recognizes a single gesture without going back to the initial position. Since more than a thousand samples must be obtained, going back to an initial position is more convenient especially when data collection is wished to be done continuously.
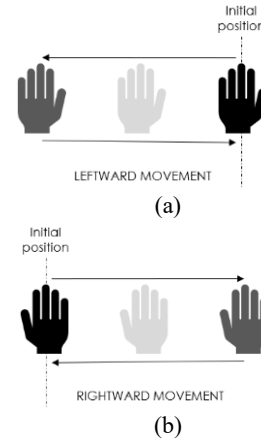

Fig. 3. (a) Leftward movement (b) Rightward movement

The device was programmed such that a plot, as shown in Fig. 4, displays the changes in values of the sensor for every gesture in real-time. In the sample graph, two sets of curves can be observed. The Fig. 4a resulted when a leftward movement was made while Fig. 4b resulted when a rightward movement was made.

The graph at the top of the window shows the values of gyro-meter over time while one at the bottom shows the values of the accelerometer over time. Three lines can be observed in each graph, each representing the values in x, y and z axes direction. The x-axis represents the sample time. That is, the temporary storage can hold up to 100 pocket data.

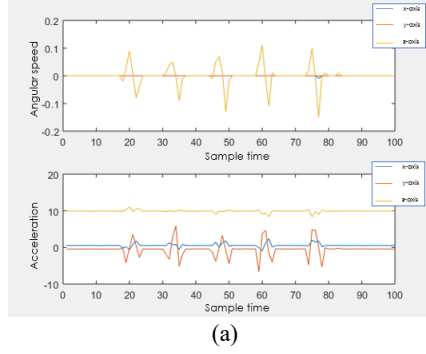## IV. RESULTS AND DISCUSSION

1) 200 datasets

200 datasets were gathered: 100 leftward gesture and 100 rightward gestures. The data for each gesture is stored in a matrix which can then be converted into an image to be used as an input to AlexNet. 120 of the data gestures are used as

training data and the remaining 80 data gestures are used as testing data.

a) New Neural Network (NNN)

Table 1
Number of Correctly Classified Gestures using the NNN

Data type	No. of data	Correctly labelled
Training	80	80
New data	20	19

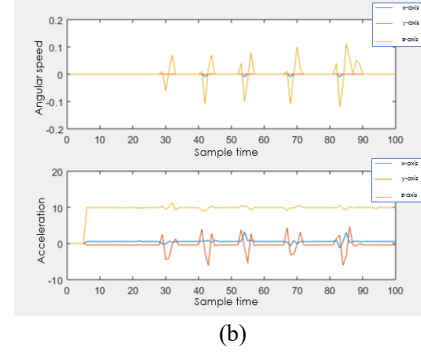(a)   (b)

Figure 4 (a) graph for leftward gesture (b) graph for rightward gesture

As shown in Table 1, the NNN is successful in correctly classify a total of 99 gestures out of 100 gestures using only 120 datasets as training data.

b) AlexNet

Table 2
Number of Correctly Classified Gestures using the AlexNet

Dataset	No. of data	Correctly labelled
Training	80	80
New data	20	11

As shown in Table 2, the AlexNet is successful in correctly classify a total of 99 gestures out of 100 gestures using only 120 datasets as training data.

c) NNN vs. AlexNet

Using AlexNet in training with few datasets is found to have less reliability and accuracy than the NNN

2) 2000 datasets

Due to the result of the testing with 200 datasets, the researchers gathered another set of data. 2000 datasets were gathered: 1000 leftward gesture and 1000 rightward gesture. 1200 of the data gestures are used as training data while the remaining 800 of the data gestures are used as testing data. Similar to what was done using 200 datasets, another new set of data – 20 new data, were gathered and tested. The result of the experiments showed that training the AlexNet with more data sets increases the accuracy in classifying new sets of data (classifying 17 out of 20 gestures). The numerical data of each gesture is stored in a matrix. Thus, the longer the gesture is performed, the more columns the matrix will contain.

V. CONCLUSION

In this paper, we present a machine learning-based sign language glove that classifies a leftward and rightward hand gesture. The performance of a Neural Network constructed from scratch is compared to the performance of a pre-trained AlexNet when it comes to classification of the hand gestures. From the results of the experiment we conclude that, for simple gestures, constructing a Neural Network from scratch is more efficient as it is designed for a specific purpose and only contains what is necessary. On the other hand, for complex gestures, the pretrained AlexNet is more powerful but requires thousands of training data give higher accuracy.

REFERENCES

[1] Korean Statistical Information Service, *Number of the Registered Disabled-by Year, Types of Disability, Disability rating and Gender,* 2016. [Online] Available: http://kosis.kr/eng/statisticsList/statisticsList_01List.jsp?vwcd=MT_E TITLE&parentId=D#SubCont

[2] Kim J.S., Jang W., Bien Z., "A Dynamic Gesture Recognition System for the korean Sign Language (KSL)," *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics,* vol. 26, no. 2, pp. 354-359, Apr. 1996, [Online].

[3] Shin, S., Baek Y., Lee J., Eun Y., Son S.H., "Korean Sign Language Recognition Using EMG and IMU Sensors Based on Group-Dependent NN Models," presented at Computational Intelligence *SSCI), 2017 IEEE Symposium Series on, Honolulu, HI, USA, USA, 27 Nov.1-1 Dec. 2017.

[4] Johns Hopkins Medicine, *Electromyography (EMG),* Available: https://www.healthline.com/health/electromyography#purpose. Accessed on: Jan. 30, 2017.

[5] Xsens, *IMU Inertial Measurement Unit,* Available: https://www.xsens.com/tags/imu/. Accessed on: Jan. 30, 2017.

[6] Nath G., S A. V., "Embedded Sign Language Interpreter System for Deaf and Dumb People," presented at Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017 Internation Conference on, Coimbatore, India, India, 17-18 March 2014