Virtual Keyboard Recognition with e-textile Sensors

Eun-Ji Ahn,[1] Sang-Ho Han,[1] Mun-Ho Ryu,[2,3*] and Je-Nam Kim[4]

[1]Department of Healthcare Engineering, Jeonbuk National University, 567 Baekje-daero, Deokjin-gu, Jeonju-si,
Jeollabuk-do 561-756, Republic of Korea
[2]Division of Biomedical Engineering, Jeonbuk National University, 567 Baekje-daero, Deokjin-gu, Jeonju-si,
Jeollabuk-do 561-756, Republic of Korea
[3]Research Center of Healthcare & Welfare Instrument for the Aged, Jeonbuk National University,
567 Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do 561-756, Republic of Korea
[4]CAMTIC Advanced Mechatronics Technology Institute for Commercialization, 67 Yu-Sang-ro, Deokjin-gu,
Jeonju-si, Jeollabuk-do 561-844, Republic of Korea

In this study, we propose a gesture recognition method using e-textile sensors and involving the pressing of numeric keys from "0" to "9". An e-textile sensor comprises a double-layer structure with complementary resistance characteristics, and it is attached to the garment to obtain a resistance signal. For gesture recognition, we tested dynamic time warping (DTW), machine learning, long short-term memory (LSTM), and bidirectional LSTM (BiLSTM). Before applying each machine learning technique, we performed normalization and resized the data to ensure that they are of the same length. A total of 120 iterations were performed for each gesture for a single subject. The results indicate that the lowest gesture classification accuracy for DTW was 74.2%, followed by 78.8 and 91.6% for LSTM and BiLSTM, respectively.

1. Introduction

Gesture recognition technology involves interpreting a user's gestures on the basis of the flexion and extension of fingers, elbows, and knees, and their interaction with objects. Gesture recognition has several applications, including human–computer interface (HCI),[1–7] medicine,[8–11] and virtual reality.[12] Various sensors such as magnetometers, accelerometers, and gyroscopes are used for gesture recognition. Ma *et al.* attached a permanent magnet and a noncontact magnetic sensor to the fingers and wrist to calculate the position and orientation of the magnet and recognize hand posture.[1] Kim *et al.* classified the state of the fingers by installing acceleration, gyroscope, and geomagnetic sensors in data gloves.[4] Lee *et al.* attached an inertial measurement unit to the wrist and recognized gestures on the basis of mouse manipulation.[5]

Flexible and stretchable sensors that can be integrated into clothes and do not interfere with joint motions are required. Recently, e-textile sensors that show resistive characteristics

while maintaining the lightness, flexibility, and stretch ability of textiles have been applied to gesture recognition.[6–12] e-textile sensors detect biological signals and joint motions without compromising comfort. Gesture recognition based on e-textile uses the changes in electrical properties according to the flexion and extension of joints during gesture motions. Han *et al.* recognized six types of mouse gestures for an HCI with dynamic time warping (DTW).[6] Aleotti and Caselli used an e-textile-based data glove for the recognition of hand motion for a virtual reality desktop system.[12]

In this study, we propose a keyboard gesture recognition method using e-textile sensors for an HCI. To perform keyboard gestures, 10 proprietary keyboard buttons simulating numerical keys from "0" to "9" were fabricated and placed on flat surfaces. For keyboard gesture recognition, the use of simple DTW[2,3,6] and computation-intensive deep learning algorithms such as long short-term memory (LSTM)[13,14] and bidirectional LSTM (BiLSTM) has been attempted. These algorithms were applied to 1200 gesture motions, and their accuracies were evaluated.

2.   Materials and Methods

2.1   e-textile sensor and data acquisition

In our previous study,[6] conductive fibers (0.80 mm thickness, EeonTex™ NW170-PI-20, Eeonyx Corp., United States) and stainless steel seals (28 Ω/ft, DEV-11791, Sparkfun Electronics, United States) were used to fabricate an electronic fiber sensor. The fabricated sensor was attached to both sides of a double-sided tape (cat. # 2240, 24 mm width, 2 mm thickness, 3M, United States), yielding each sensor pair in a double-layer structure. The data acquisition system comprised three double-layer e-textile sensors, a current source that supplies constant current to the sensors, a buffer (voltage follower) that converts the output of the sensor into a low impedance voltage, and an analog-to-digital-converter (ADC) to quantize the voltage into a digital value. A microcontroller unit calculates the resistance of the sensor from the digital value and transmits it to a PC at 100 Hz. The PC collects the sensor data and saves them for analysis.

2.2   Recognition algorithms

The six e-textile sensor signals captured for the keyboard gesture are shown in Fig. 1. In this figure, the rows and columns correspond to the gesture and sensors, respectively. The left-most column shows the output of the first sensor for the keyboard gesture "0" to "9". In each graph, the vertical axis shows the resistance of the sensor and the horizontal axis shows the gesture duration. The graph indicates that the resistance and gesture duration ranges are different for each sensor and gesture. This is because the motion speeds and patterns of the sensor are different for each gesture. These variations could make gesture recognition difficult if attempted using simple pattern recognition algorithms.

Therefore, the DTW and template matching techniques were applied to consider the variations in data length and pattern characteristics as in our previous study.[6] We explain these
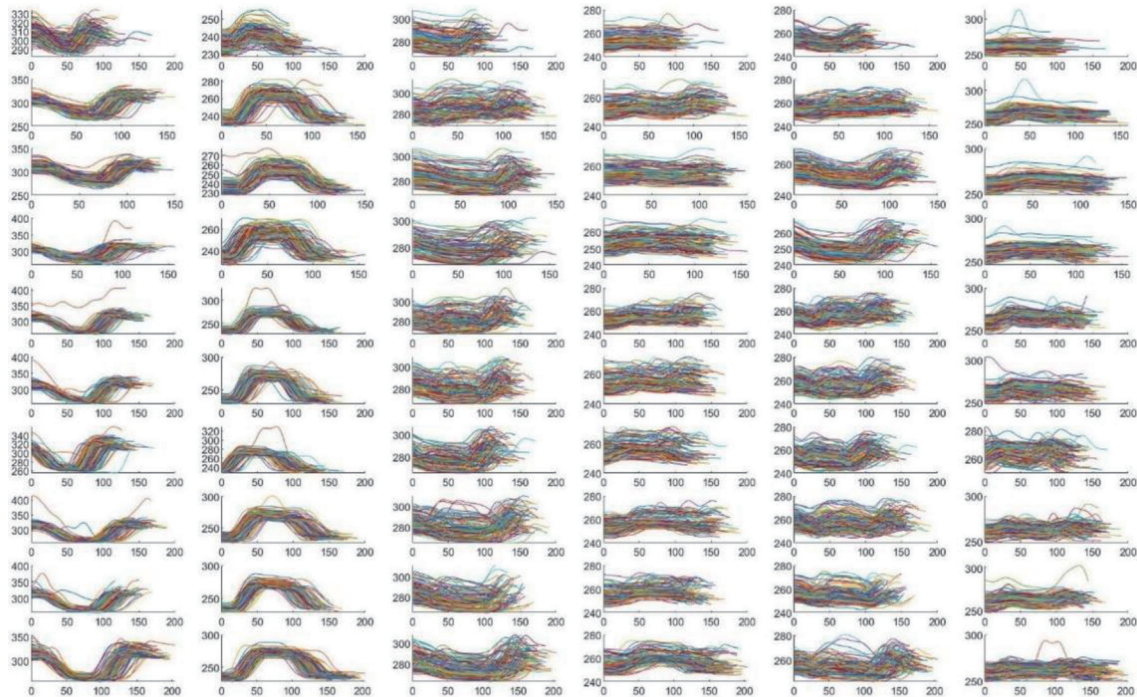
Fig. 1.   (Color online) Sensor raw data (column) for each gesture (row).

techniques in brief here. The DTW technique aligns two time series signals by minimizing the sum of Euclidean distances between the corresponding points. By DTW, we can align two time series signals of different lengths to the same length. Furthermore, the Euclidean distance calculated during the warping process can be used as a measure of similarity between the two time-series signals.[2,3] For the template matching technique, all time series signals for each gesture class are aligned to the same length by DTW and the ensemble is averaged to generate a template gesture time series signal that is compared with ten gesture templates with DTW. The gesture template with the shortest Euclidean distance is determined as the gesture class for the test gesture.

Machine learning algorithms with a more intensive computation were also used to achieve a higher recognition accuracy. To apply machine learning, the acquired data were separated into training, validation, and test data. In this study, 120 keyboard gesture data were obtained from ten keyboards each. The total of 1200 data were separated at rates of 70, 15, and 15% for training, validation, and test, respectively. The hyperparameter was tuned while training with the training data and the accuracy was verified using the validation data. The parameters of the model were adjusted gradually while the accuracy was verified with the validation data. In addition, data preprocessing was performed to improve the learning accuracy before machine learning. For data preprocessing, all data were resized to the same length and normalized in amplitude.

For the machine learning algorithm, LSTM and BiLSTM were used (MATLAB9.6, MathWorks, United States). These deep neural networks show good performance for time-

series classification.[13,14] LSTM is an improved version of the recurrent neural network (RNN); while LSTM processes input and output in sequence, BiLSTM considers the bidirectionality of the LSTM. The LSTM model trains the memory cells of the RNN model concealment layer by adding input, erase, and output gates to erase unnecessary memories and to determine what to remember. The input gate passes the sigmoid function (σ) and hyperbolic tangent function (tanh) to determine the amount of information stored. In the erase gate, σ passes, and the closer the output value is to 0, the more information is deleted. At the output gate, the value that passes σ determines the hidden state. In the hidden state, the cell state passes the tangent tanh, which is computed with the output gate value to filter the value. The values that pass all the gates are directed to the output layer and output as the result. BiLSTM is trained in the same manner as the LSTM model by considering both the forward and backward passes.

The overall network consisted of a sequence input layer with six features, an LSTM layer with 400 hidden units (or a BiLSTM layer with 200 hidden units), a fully connected layer with 10 classes, a softmax layer, and a classification layer in sequence (Fig. 2). The adaptive momentum estimation algorithm was used as a solver. The maximum epochs were set to 430 for LSTM and 200 for BiLSTM, respectively.

2.3   Experimental protocol

Three e-textile sensor pairs (six sensors) were attached at the same positions as in our previous study.[6] The positions on the arm and shoulder indicated large joint angle changes during the gesture motions. The first sensor was attached to the wrist brace, and the center of the sensor was worn on the olecranon. The second sensor was attached to the center of the deltoid medial and the humerus when the rash guard was worn. A third sensor was attached to the rash guard so that one end of the sensor is at the center of the humerus. The rubber band was attached to one end of the sensor to compensate for the problem that the manufactured sensor is not stretched sufficiently according to the degree of bending of the joint.[6]
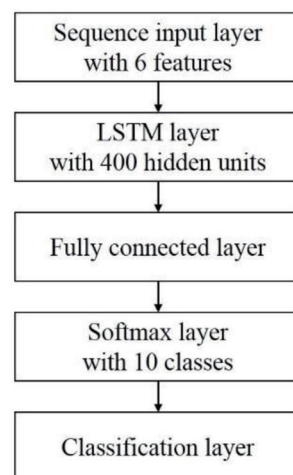

Fig. 2.   Neural network with LSTM model.

To perform keyboard gestures, 10 proprietary keyboard buttons simulating the numerical keyboard from "0" to "9" were fabricated and placed on a flat surface. Each keyboard button was 7×7 cm^2. One push button was connected to the center of the keyboard button for the keyboard pressing motion, and one red LED was connected to the upper center to confirm that the button was pressed. After making a total of 10 keyboard buttons corresponding to "0" to "9," the keyboard was constructed by placing the keyboard buttons at intervals of 15 cm in both directions (Fig. 3).

The subject participated in the experiment while wearing a rash guard with two double-layer sensors and a wrist brace with one double-layer sensor. The subject placed his or her hands at the center of the chest and then pressed the push buttons on the keyboard buttons corresponding to the numbers "0" to "9" with the right hand. The experiment was conducted while sitting on a fixed chair to suppress body motions.

The experimental set consists of two attempts to perform ten gestures once and then ten gestures again. Just before the gesture, the subject pressed a marking switch on his or her left hand and pressed the marking switch again after the gesture was completed. The marking button signal generated at this time was used to segment keyboard gestures. Before the experiment, two sets of practice experiments were conducted to ensure that the subject understood the procedure. A total of 60 sets of experiments were conducted to obtain 1200 keyboard gesture data.

3.   Results and Discussion

After 60 sets of experiments, 120 gesture motion units were obtained for each keyboard gesture. The correct and incorrect classifications for each keyboard gesture were shown in the confusion matrix using the DTW technique (Fig. 4). In the confusion matrix, the rows represent classified gesture classes ("Output Class") and the columns correspond to correct gesture classes ("Target Class"). The classification accuracy and error for each gesture class
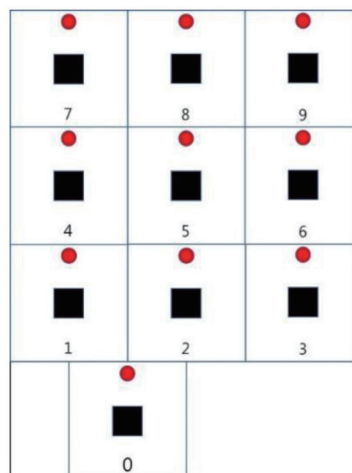

Fig. 3.   (Color online) Proprietary numeric keyboard.

Fig. 4. (Color online) Confusion matrix for 10 keyboard gestures.

correspond to the bottom row. The classification accuracy and error for all keyboard gestures correspond to the bottom-right cell. Gestures that press the number "0" keyboard button had the highest accuracy of 99.2%, while gestures that press the number "7" keyboard button had the lowest accuracy of 33.2%. The average accuracy for all keyboard gestures ("0" to "9") was 74.2%.

Ten keyboard gestures were trained and verified by LSTM as shown in Fig. 5. Figure 5(a) shows the training accuracy (blue) and validation accuracy (black) according to the number of learning iterations. Figure 5(b) shows the loss according to the number of learning iterations. The accuracy of training using the LSTM model was 100% and the final validation accuracy was 73.33%; the maximum validation accuracy was 73.89%. The final minibatch loss was 0.0024 and the final validation loss was 1.3044. After the training using the LSTM model, the test data were tested and showed a recognition accuracy of 78.89%.

The result of training using the BiLSTM model, which learns forward and backward passes, is shown in Fig. 6. The training accuracy was 91.67% and the final and maximum validation accuracy was 91.67% [Fig. 6(a)]. The final minibatch loss was 0.4612 and the final validation loss was 0.3107 [Fig. 6(b)]. After testing with the trained BiLSTM model, the accuracy was shown to be 91.67%.

Ten gesture recognition accuracies obtained using the three methods are shown in Table 1. As shown in the table, the overall keyboard accuracy of DTW was 74.2%. When the gesture
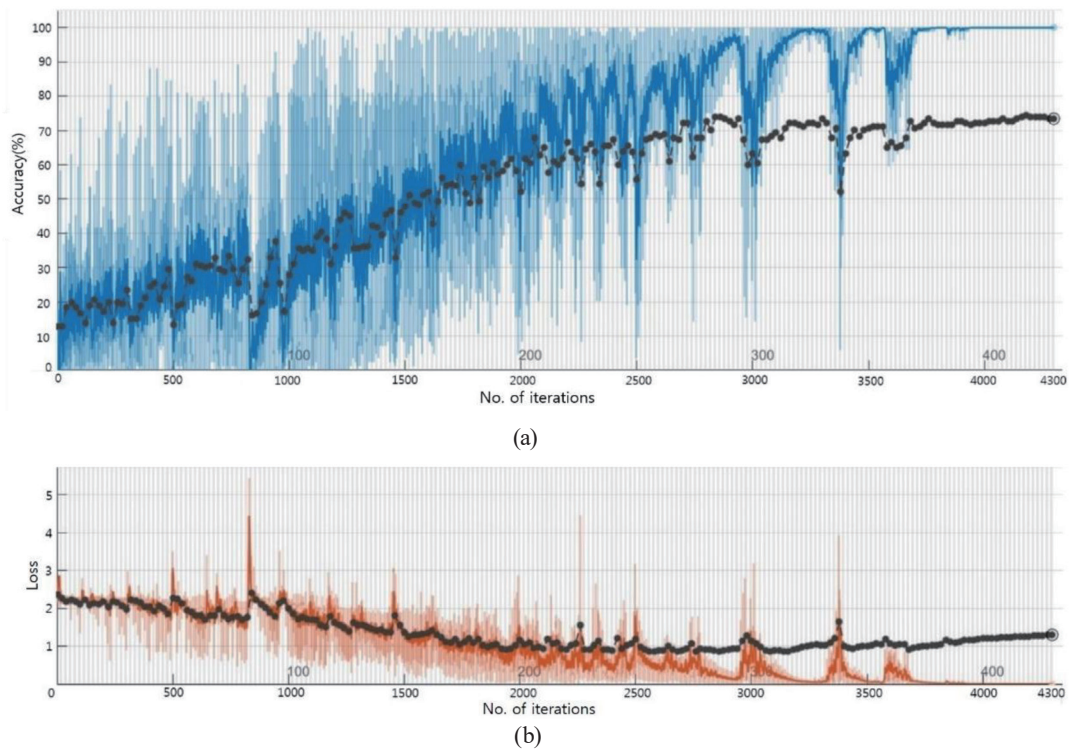
Fig. 5.   (Color online) Result of training using LSTM model. (a) Train accuracy (blue) and validation accuracy (black). (b) Minibatch loss (red) and validation loss (black).
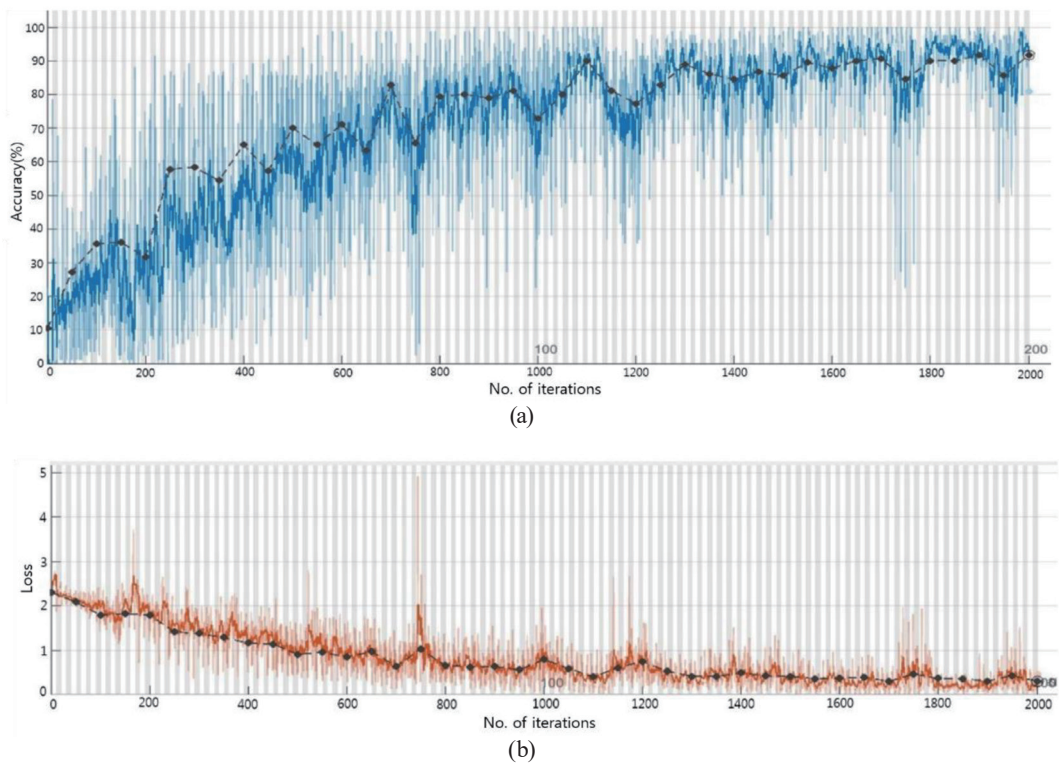

Fig. 6.   (Color online) Result of training using BiLSTM model. (a) Train accuracy (blue) and validation accuracy (black). (b) Minibatch loss (red) and validation loss (black).

Table 1
Recognition accuracy for each gesture.

Gesture	DTW	LSTM	BiLSTM
0	99.2	90.9	98.3
1	84.2	64.7	80.5
2	61.7	80.0	94.1
3	74.2	84.6	90.0
4	78.3	64.7	82.5
5	72.5	74.0	89.1
6	86.7	84.8	92.5
7	33.3	50.0	86.6
8	70.0	95.0	95.8
9	81.7	90.9	95.8
Overall	74.2	78.9	91.7

was recognized by the trained LSTM model, the accuracy was 78.89%, which is slightly higher than that of the DTW technique. The gesture recognition accuracy obtained using BiLSTM was the highest (91.67%). In particular, the recognition rates for gesture "7" were determined to be 33.3 and 50.0% in DTW and LSTM, respectively, but 86.6% in BiLSTM.

4. Conclusions

We proposed a method of recognizing keyboard gestures for an HCI using e-textile sensors. Each e-textile sensor comprised a double-layer structure showing complementary resistance characteristics. Constant current sources, ADCs, MCUs, and PCs were used to acquire and save numeric keyboard gesture motion data from "0" to "9". For gesture recognition, the DTW technique, which is known to show excellent performance with a small amount of computation for dynamic signals, was used. In addition, the computation-intensive deep neural networks LSTM (a higher version of RNN) and BiLSTM (bidirectional version of LSTM) were used.

In the case of using 1200 gesture motion data to test 10 keyboard gestures, the accuracies of keyboard gesture detection by DTW, LSTM, and BiLSTM were 74.2, 78.89, and 91.67%, respectively. Gesture recognition technique DTW showed the lowest accuracy (74.2%); gesture recognition through LSTM showed a slightly improved accuracy (78.89%) over DTW. Keyboard gesture recognition through BiLSTM showed a significantly improved recognition rate compared with other methods with an accuracy of 91.67%. The minibatch loss for BiLSTM was 0.4612, which is higher than that of LSTM (0.0024). However, the validation losses of the BiLSTM and LSTM were 0.33107 and 1.3044, respectively. Since the BiLSTM model has a lower validation loss than the LSTM model, the hyperparameter was adjusted to show an excellent network.

References

1   Y. Ma, J. Yang, Z. H. Mao, W. Jia, C. Li, and M. Sun: IEEE Trans. Magnetics. **47** (2011) 970.
2   S. Celebi, A. S. Aydin, T. T. Temiz, and T. Arici: Proc. VISAPP (2013) 620.
3   S. Seto, W. Zhang, and Y. Zhou: arXiv preprint (2015) arXiv:1512.06747.
4   K. W. Kim, M. S. Lee, B. R. Soon, M. H. Ryu, and J. N. Kim:Technol. Health Care **24** (2016) S223.
5   M. S. Lee, K. W. Kim, M. H. Ryu, and J. N. Kim: Sens. Mater. **28** (2016) 655
6   S. H. Han, E. J. Ahn, M. H. Ryu, and J. N. Kim: Sens. Mater.**31** (2019) 1387.
7   M. Bobin, H. Amroun, S. Coquillart, F. Bimbard, and M. Ammi: Proc. IEEE Int. Conf. Systems, Man, and Cybernetics (IEEE, 2016) 2192.
8   S. H. Han, M. H. Ryu, and J. N. Kim: Sens. Mater. **30** (2018) 1787
9   P. T. Gibbs and H. Asada: J. Neuroeng. Rehabil. **2** (2005) 7.
10 A. Tognetti, F. Lorussi, G. D. Mura, N. Carbonaro, M. Pacelli, R. Paradiso, and D. De Rossi: J. Neuroeng. Rehabil. **11** (2014) 56.
11 L. Santos, N. Carbonaro, A. Tognetti, J. L. González, E. de la Fuente, J. C. Fraile, and J. Pérez-Turiel: Technologies **6** (2018) 8.
12 J. Aleotti and S. Caselli: Proc. VRIPHYS (2008).
13 Y. Chen, K. Zhong, J. Zhang, Q. Sun, and X. Zhao: Proc. ICAITA (2016).
14 F. J. Ordonez and D. Roggen: Sensors **16** (2016) 115.

About the Authors

Eun-Ji Ahn received her BS degree from Jeonbuk National University, South Korea, in 2018. She is currently pursuing her MS degree in the Department of Healthcare Engineering, Jeonbuk National University, South Korea. Her research interests include biomedical, neural network, electronic, and control engineering.

Sang-Ho Han received his BS degree in biomedical engineering from Jeonbuk National University, South Korea, in 2009. From 2010 to 2016, he worked as a researcher at R&D centers for several companies. He is currently an MS student in the Department of Healthcare Engineering at Jeonbuk National University. His current research interests are in Internet of Things and embedded systems.

Mun-Ho Ryu received his BS degree in control and instrumentation engineering and his PhD degree in biomedical engineering from Seoul National University, South Korea, in 1990 and 2004, respectively. He worked for Daewoo Heavy Industry Machine Tool Division from 1990 to 2000. He also worked for Biomed Lab Co., Ltd. from 2000 to 2005. He has been a professor at Jeonbuk National University since 2005. His research interests are in human motion tracking, inertial navigation, telerehabilitation, rehabilitation, and embedded systems.

Je-Nam Kim received his BS degree in biomedical engineering and his MS and PhD degrees in healthcare engineering from Jeonbuk National University, South Korea, in 2009, 2011, and 2014, respectively. Since 2014, he has been with CAMTIC Advanced Mechatronics Technology Institute, South Korea. His research interests include embedded software, embedded hardware, and application contents.