Frontiers in Health Informatics
ISSN-Online: 2676-7104
www.healthinformaticsjournal.com
2024; Vol 13: Issue 6
Open Access

Optimizing CGAN and DNN Architectures for Wearable Systems to Support Students with Developmental Disabilities

Woosoon Jung [1], KyoungOck Park [2], Jeong Tak Ryu [3], Kyuman Jeong [4] ,Yoosoo Oh [5]

[1] Institute of Special Education & Rehabilitation Science, Daegu University

[2] Department of Elementary Special Education, Daegu University

[3] Department of Electronic and Electrical Engineering, Daegu University

[4] School of AI, Daegu University

[5] School of Computer & Information Engineering, Daegu University

---

Cite this paper as: Woosoon Jung , KyoungOck Park , Jeong Tak Ryu , Kyuman Jeong ,Yoosoo Oh (2024) Optimizing CGAN and DNN Architectures for Wearable Systems to Support Students with Developmental Disabilities. Frontiers in Health *Informatics*, *13(6)* 1006-1014

Abstract. In this study, we propose a 6-Degrees of Freedom Inertial Measurement Unit (6-DoF IMU)-based wearable system for recognizing challenging behaviors of students with developmental disabilities. In the proposed system, 6-DoF IMU data is preprocessed and used as input to the Deep Neural Network (DNN) to recognize challenging behaviors of students with developmental disabilities. Building a dataset is one of the biggest challenges in wearable Artificial Intelligence (AI) systems. Since collecting data samples is expensive, there is a limit to the amount that can be collected. In this study, we collect datasets from participants, build a custom dataset, and augment the data using Conditional Generative Adversarial Network (CGAN). We observe the performance change according to the augmentation ratio of the original data, and evaluate the scalability of the developed model by applying data from new participants that have never been shown during the training process. As a result of applying data augmentation techniques to a DNN model that already has high accuracy, a slight decrease in accuracy was observed for the original test set. However, when data from new participants is applied, an accuracy improvement of up to 10% was observed.

Keywords: HAR, Challenging Behavior, Wearable System, DNN, AI, Low-power Design

1. Introduction

Challenging behavior of students with developmental disabilities refers to behaviors that cause physical harm to themselves or others [1]. Challenging behavior is a major obstacle to students with developmental disabilities entering society, so recognizing it can be helpful in supporting and intervening their behaviors in the future. In this study, we propose a wearable system to identify the types and frequency of challenging behaviors. The proposed wearable system consists of a 6-Degrees of Freedom Inertial Measurement Unit (6-DoF IMU) and a low-power processor for constant operation, and applies wearable Artificial Intelligence (AI) to achieve Human Activity Recognition (HAR). HAR can be divided into vision-based research and sensor-based research. Vision-based research has the advantage of requiring only a camera, and various frameworks such as OpenPose [2] and MediaPipe [3] that can be immediately applied have been developed. On the other hand, HAR using IMU has various applications depending on the body part to be analyzed, such as the neck [4], hands [5-7], and waist [8]. In addition, it is also applied to various devices such as orthopedic walker [9], smartphones [10], and earbuds

[11] that are not body parts.

Due to the nature of wearable devices, computing resources are limited, so it is necessary to consider memory usage for implementing on-device AI. Since memory requirements should be considered not only by the size of the AI model (number of parameters) but also by the size of the code, the DNN technique is adopted. NN-based algorithms generally tend to increase in performance with the number of parameters. For this reason, the point where accuracy is saturated is explored for the optimal DNN architecture.

Datasets are a major factor in determining the performance of AI [12]. AI model optimization can proceed only when a dataset with a large number of samples and well-verified quality is prepared first. AI models developed based on public datasets such as MNIST [13] and ImageNet [14] can be objectively evaluated. On the other hand, when using biased data such as this application that utilizes unusual behavioral data, the dataset must be built by hand. Building a dataset is time-consuming and costly. To overcome the limitations, this study adopts data augmentation using CGAN [15]. The performance of the DNN model trained with the collected original dataset and the model trained with the augmented dataset are compared to verify the reliability of the results of data augmentation. Data augmentation is performed up to 5 times the original dataset. In addition, to verify the generalizability of the model trained with the augmented data, evaluation is conducted with data from new participants who have never been shown in training processes.

2. Proposed IMU-based Wearable System

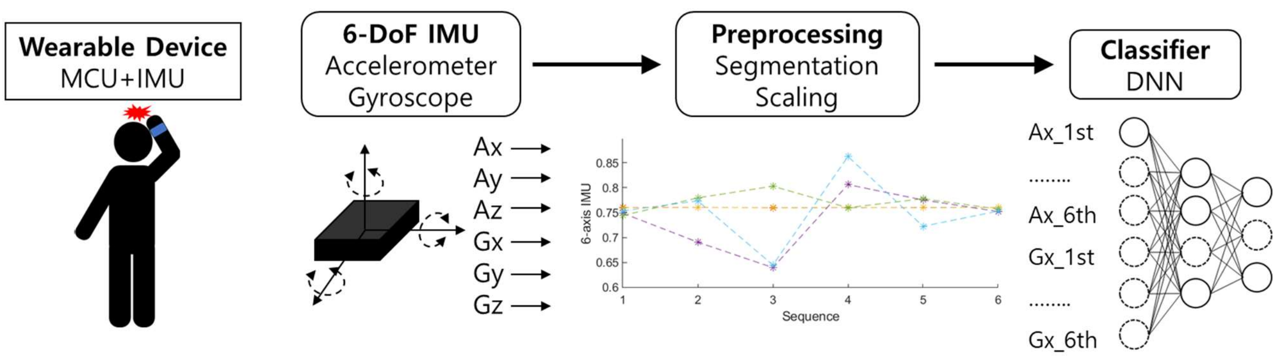
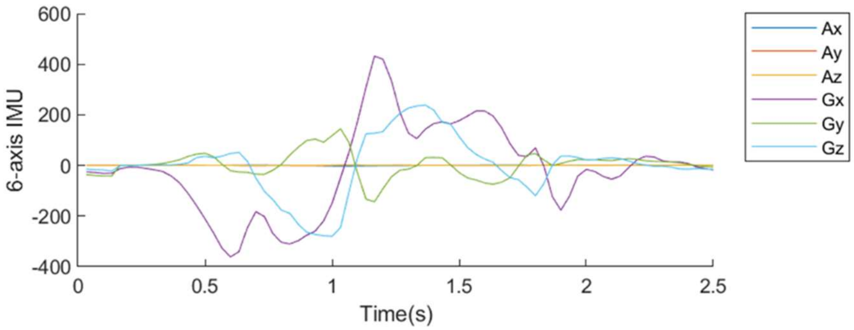


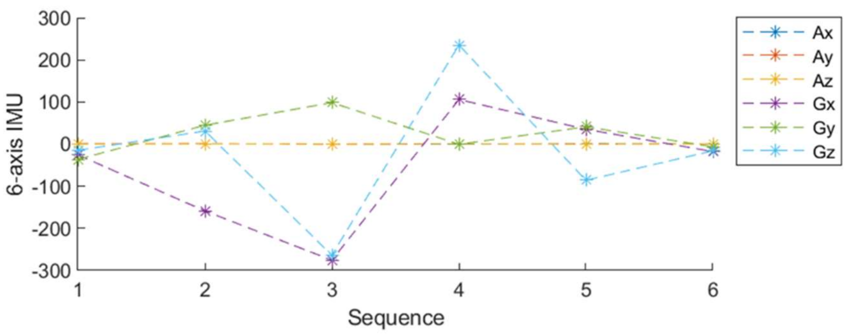
Fig. 1: Wearable system flow diagram

Figure 1 shows the flow diagram of the IMU-based wearable system. The proposed wearable device is attached to the user's wrist and distinguishes the user's self-injurious, aggressive and general behaviors. The wearable device consists of a 6-DoF IMU and a 32-bit low-power processor. Since the average length of the behavioral data collected in this study is approximately 2 seconds, the accelerometer and gyroscope data measured from the IMU are collected at 2-second intervals. The sampling rate of the IMU is 30Hz, and algorithms such as the Kalman filter are not applied considering the computational complexity of the Microcontroller Unit (MCU).

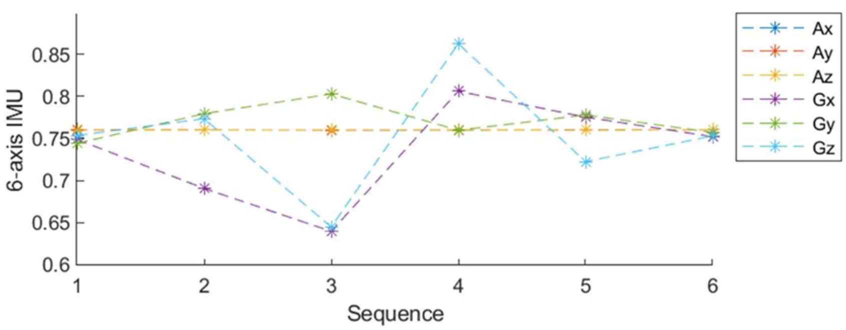
As mentioned in the above chapter, in a computing environment with limited resources, not only the model size but also the code size for implementing the model in the MCU should be considered. Therefore, DNN is adopted for this application. Since DNN has a fixed-size input node, a process for converting time series data into a fixed length is required. For this reason, we applied a segmentation process. Time series data of different lengths are converted into samples of the desired fixed length, and a moving average filter is applied in this process to minimize information loss. In this process, if the number of samples is insufficient, an interpolation method is adopted. Since determining the size of the input layer of the DNN directly affects the size of the AI model, we confirmed through repeated experiments that setting the length per axis of the IMU to 6 maintains sufficient accuracy. Since the IMU used in the study is 6-DoF, the number of input nodes of the DNN is determined to be 36. The final stage of the preprocessing applies maximum and minimum scaling to convert the input distribution into a value between 0 and 1. Figure 2 shows each preprocessing process of the 6-DoF IMU.



(a) Raw data from 6-DoF IMU (Sample rate: 30Hz)



(b) Segmentation (Down-sampling & Moving average filter)



(c) Min-Max Scaler

Fig. 2: Preprocessing flow of 6-DoF IMU data

3. Designing DNN and CGAN for the Proposed System

3.1. Exploration for Optimal DNN Architecture

In order to implement AI on microprocessors with code sizes of tens of kB, this study adopts DNN. DNN is a neural network with two or more hidden layers, and the number of hidden layers is limited to two considering the size of the model. In order to search for a DNN architecture with optimal performance in a limited environment, an experiment is conducted to explore the performance according to the number of nodes in the hidden layer. The PyTorch framework is used, and the learning rate is 0.0075, and the Rectified Linear Unit (ReLU) is applied as the activation function. Table 1 shows the results of the exploration of the DNN architecture.

Table 1: Accuracy exploration results according to DNN model architecture

DNN architecture	36×5×5×3	36×10×10×3	36×15×15×3	36×20×20×3	36×25×25×3	36×30×30×3
Accuracy (%)	95.42	95.83	95.83	96.67	**97.08**	96.25

Experimental results show that the maximum accuracy is saturated when there are two hidden layers and the number of nodes in each layer is 25. The number of parameters in the DNN is calculated as follows (Equation 1):

$$\#\ of\ Parameters = \sum_{l=1}^{L}(n_1 \times n_{l+1} + n_{l+1}) \quad (1)$$

where n_l is the number of nodes in the l-th layer, n_{l+1} is the number of nodes in the next layer, and L is the total number of layers, including the hidden layers and the output layer.

3.2. Designing CGAN for Time Series Data of IMU

The data augmentation technique applied in this study is the CGAN. Since CGAN has the function of generating new data samples based on input conditions, it can generate data that meets specific conditions based on label information. Based on the feature of being able to generate data samples of the desired class, it is suitable for experiments that augment data at a certain rate. Figure 3 shows the CGAN architecture. The input layer of the generator model is determined by the random vector Z and the number of classes, and the output layer is the same as the number of input nodes of the DNN. In this study, three classes of self-injurious, attack, and general behaviors are defined, and the number of input nodes of the DNN is 36. The size of the random vector Z is determined to be 400 through repeated experiments. The generator model consists of a total of five layers, and the number of nodes in the remaining layers excluding the input layer and the output layer increases sequentially by two times. The discriminator model consists of a total of five layers. The input layer is determined by the input node of the DNN and the number of classes, and the output layer consists of one node that determines real/fake. The remaining layers excluding the input layer and the output layer are composed in the opposite order to the generator architecture.

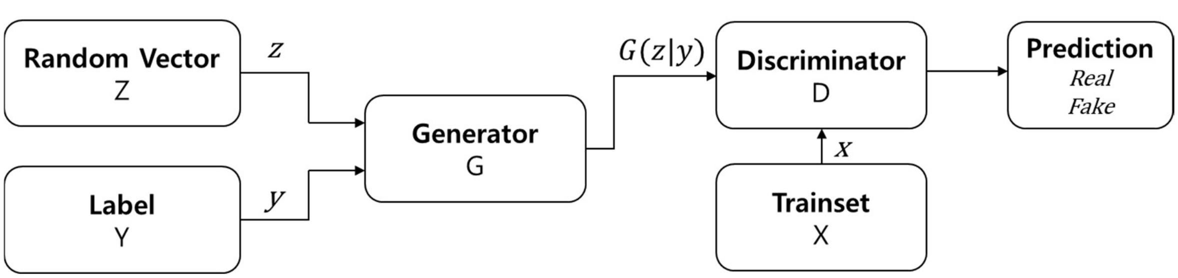
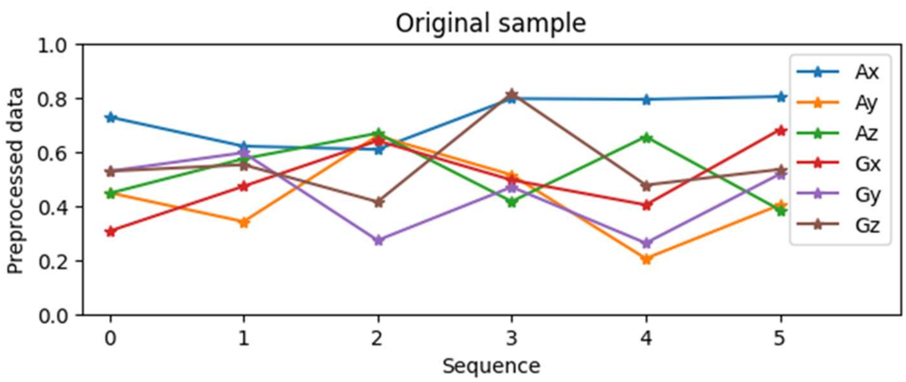
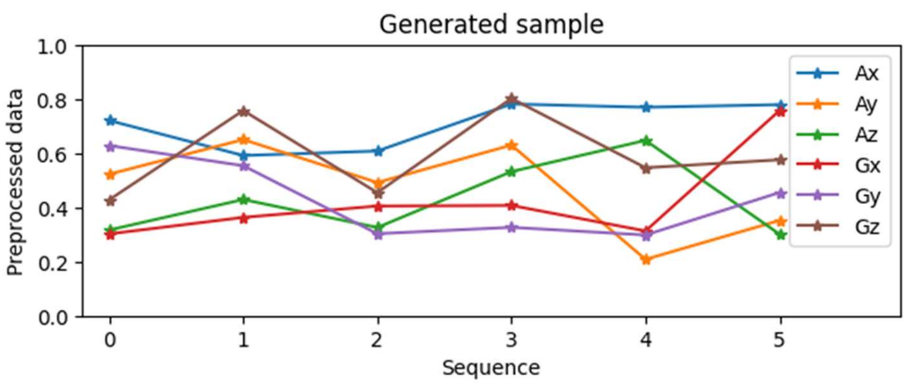


Fig. 3: CGAN architecture

Figure 4 shows the original data sample after the preprocessing process and the data sample generated using the CGAN technique.



(a) Example of original data sample



(b) Example of data sample generated by CGAN

Fig. 4: Example of original data sample and data sample generated by CGAN

4. Result

4.1. Experimental Setup

The wearable device used in the study consists of a 6-DoF IMU and a 32-bit low-power processor based on Cortex-M4. The Printed Circuit Board (PCB) is about 3×3cm in size and is attached to the user's wrist (Figure 5). The current consumption of the entire PCB is about 8mA or less, and it operates for more than 12 hours with a 20×30mm sized lithium polymer battery.

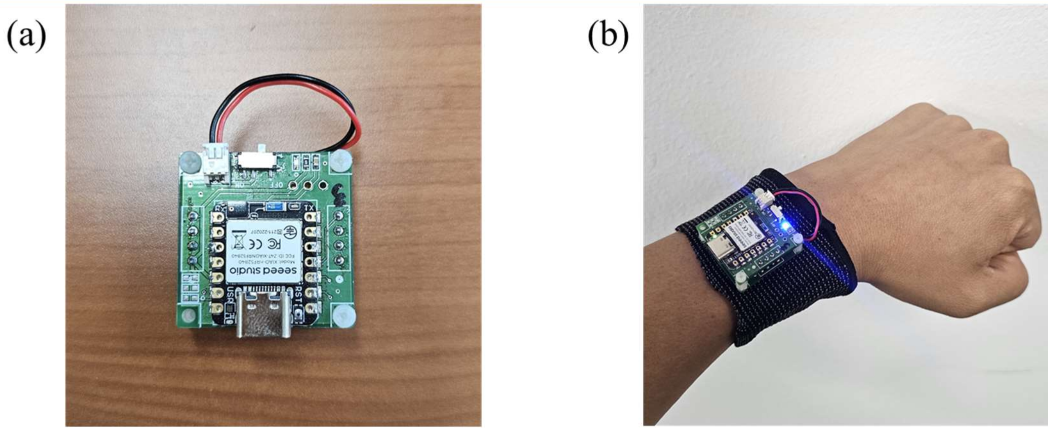


Fig. 5: (a) Wearable device consisting of 6-DoF IMU and low-power processor, (b) Example of wearing the device

Figure 6 shows three defined classes. Self-injurious, aggressive, and general behaviors are defined, and in the case of general behavior, five behaviors are mixed: 'walking', 'writing', 'reading', 'standing', and 'computer work'. The data collected from the IMU with a sampling rate of 30 Hz have an average length of about 2 seconds. 50 samples per class were collected from four participants, constructing a dataset with a total of 600 data samples, 3 classes, and 36 features. Of these, 240 samples, corresponding to 40%, were used as the test set, and the remaining 360 samples were used as the training set. The experiment consists of two experiments: one to evaluate a model trained with the original training set and the other one to evaluate a model trained with the augmented training set. The same test set (240 samples) is used in both cases.

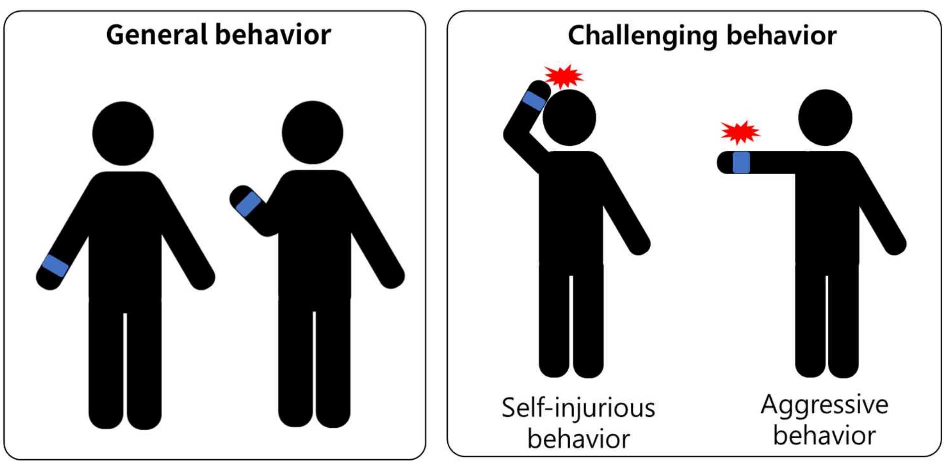


Fig. 6: Example of defined classes (General behavior, self-injurious behavior and aggressive behavior)

4.2. Accuracy Changes According to Data Augmentation Ratio

We conducted an experiment to verify whether the data shortage environment can be overcome by applying CGAN. In order to observe the change in accuracy according to the number of samples generated through the data augmentation technique, a training set was generated up to 5 times the size of the original dataset. As mentioned in Chapter 3, the highest accuracy was shown in the DNN with a 36×25×25×3 architecture based on the original dataset, so the search range of the DNN architecture is limited to a maximum of 25 nodes in the hidden layer. Of the 600 samples in the original dataset, 240, which is 40%, are used as a test set (80 per class, 240 in total). All samples generated through CGAN are used only in the training set (120 per class, 360 in total). Table 2 shows the performance change according to the augmentation ratio of CGAN. Since all augmented training sets contain the original training set, a 1× augmentation has a total of 720 samples.

Table 2: Accuracy of the model trained on the augmented dataset

	No. of samples in the training set (Augmentation ratio)					
No. of hidden nodes	360 (Original)	720 (1×)	1,080 (2×)	1,440 (3×)	1,800 (4×)	2,160 (5×)
5	95.42	93.33	92.08	95.83	91.67	94.58
10	95.83	95.42	95.83	96.67	94.17	95.00
15	95.83	96.67	95.42	95.83	95.42	95.00
20	96.67	96.67	96.25	95.83	94.58	95.83
25	**97.08**	96.25	95.42	95.83	94.58	95.00

As a result of evaluating with the same test set, the accuracy of the model trained with data generated by CGAN is rather low overall. In the case of 25 hidden nodes, the accuracy decreases from a minimum of 0.83% to a maximum of 2.5% depending on the augmentation ratio. Despite the accuracy decrease, it shows a high accuracy of about 95% or more in most cases.

An experiment was conducted to evaluate the scalability of the model trained with the data set augmented by CGAN. An additional experiment was conducted by configuring the test set with data collected from two new participants who had not been shown before. 50 samples were collected per class, and the results of the experiment with a total of 300 samples are as shown in the Table 3.

Table 3: Accuracy of the model trained on the augmented dataset (Evaluated using only the data from new participants)

	No. of samples in the training set (Augmentation ratio)					
No. of hidden nodes	360 (Original)	720 (1×)	1,080 (2×)	1,440 (3×)	1,800 (4×)	2,160 (5×)
5	78.33	80.00	64.33	82.33	78.33	81.00
10	83.67	77.33	81.33	83.33	80.67	82.33
15	71.67	83.00	79.67	81.67	86.00	83.67
20	78.33	83.67	79.00	83.00	88.33	79.67
25	79.00	83.00	80.67	85.33	84.00	**89.33**

Experimental results show that when evaluating data from new participants with a model trained with data augmented by CGAN, the accuracy is up to 10% higher than that of a model trained using only the original dataset in the section with 25 hidden nodes. The higher the data augmentation ratio and the more complex the model, the greater scalability.

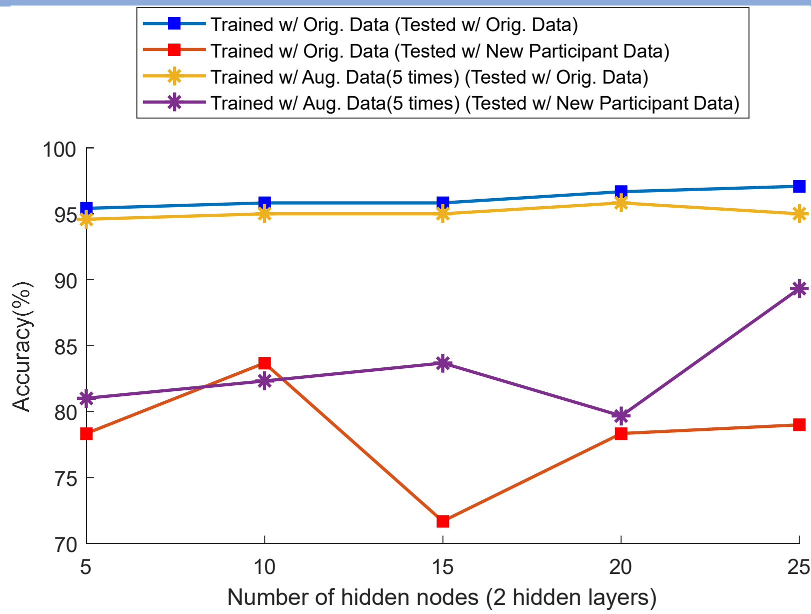


Fig. 7: Accuracy of the model trained on the original dataset and augmented dataset

Figure 7 shows the performance change when training with the original dataset and the augmented dataset (5 times). When there are 25 hidden nodes, when the participants' data is reflected in the training process (blue and yellow lines), the accuracy of the model applying the original dataset is about 2.5% higher. On the other hand, when evaluating with only the data of new participants (red and purple lines), the data of the model applying the augmented dataset is 10.33% higher.

Conclusion

In this study, a small-sized DNN was applied to perform HAR on a wearable device. A very small-sized custom dataset with 200 samples per class was augmented with the CGAN technique to increase the training set by up to 5 times. As a result, when the training set was augmented by 5 times in a DNN with two hidden layers with 25 nodes each, the accuracy decreased by 2.08% compared to the case where the original dataset was used. On the other hand, for a new participant who had never been seen in the training set, an accuracy increase of 10.33% was confirmed. When the participant's data was reflected in the training process, the accuracy of the model trained with the augmented dataset decreased from 97.08% to 95.00%, but it still showed high accuracy. On the other hand, when evaluating only the data of a completely new participant, an accuracy increase of 10.33% was confirmed, confirming that the data augmentation using CGAN has high scalability. This allows for a bit more flexibility in dealing with overfitting issues that can occur when using very small datasets.

Acknowledgements

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5C2A07091326).

References

[1] E. Emerson, "Challenging Behaviour: Analysis and Intervention in People with Severe Intellectual Disabilities," Cambridge University Press, 1995.

[2] Z. Cao, G. Hidalgo, T. Simon, S. -E. Wei and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 1, pp. 172-186, 1 Jan. 2021, doi: 10.1109/TPAMI.2019.2929257.

[3] F. Zhang, V. Bazarevsky, A. Vakunov, et al., "MediaPipe Hands: On-device Real-time Hand Tracking," *arXiv preprint arXiv:2006.10214*, 2020.R. Dewri, and N. Chakraborti. Simulating recrystallization through cellular automata and genetic algorithms. *Modelling Simul. Mater. Sci. Eng*. 2005, **13** (3): 173-183.

[4] I. -C. Severin, "Head Posture Monitor Based On 3 IMU Sensors: Consideration Toward Healthcare Application," 2020 International Conference on e-Health and Bioengineering (EHB), Iasi, Romania, 2020, pp. 1-4, doi: 10.1109/EHB50910.2020.9280106.

[5] X. Chu, J. Liu and S. Shimamoto, "A Sensor-Based Hand Gesture Recognition System for Japanese Sign Language," 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), Nara, Japan, 2021, pp. 311-312, doi: 10.1109/LifeTech52111.2021.9391981.

[6] P. Shenoy, A. Gupta and V. S.K.M., "Design and Validation of an IMU Based Full Hand Kinematic Measurement System," in IEEE Access, vol. 10, pp. 93812-93830, 2022, doi: 10.1109/ACCESS.2022.3203186.

[7] C. K. Jha, Y. Shukla, R. Mukherjee, P. Rathva, M. Joshi and D. Jain, "A Glove-Based Virtual Hand Rehabilitation System for Patients With Post-Traumatic Hand Injuries," in IEEE Transactions on Biomedical Engineering, vol. 71, no. 7, pp. 2033-2041, July 2024, doi: 10.1109/TBME.2024.3360888.

[8] M. Shang et al., "Otago Exercises Monitoring for Older Adults by a Single IMU and Hierarchical Machine Learning Models," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 32, pp. 462-471, 2024, doi: 10.1109/TNSRE.2024.3355299.

[9] M. Huang and E. Clancy, "Smart Walker: an IMU-Based Device for Patient Activity Logging and Fall Detection," 2022 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, USA, 2022, pp. 1-4, doi: 10.1109/SPMB55497.2022.10014869.

[10] A. Dubey and J. Zacharias, "IMU Data Based HAR Using Hybrid Model of CNN & Stacked LSTM," 2024 International Conference on Advancements in Power, Communication and Intelligent Systems (APCI), KANNUR, India, 2024, pp. 1-6, doi: 10.1109/APCI61480.2024.10616429.

[11] V. Mollyn, R. Arakawa, M. Goel, C. Harrison, and K. Ahuja, "IMUPoser: Full-Body Pose Estimation using IMUs in Phones, Watches, and Earbuds," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2023.

[12] D. C. Marcu and C. Grava, "The Importance of Data Quality in Training a Deep Convolutional Neural Network," 2023 17th International Conference on Engineering of Modern Electric Systems (EMES), Oradea, Romania, 2023, pp. 1-4, doi: 10.1109/EMES58375.2023.10171785.

[13] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[14] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 248–255.

[15] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," arXiv preprint, arXiv:1411.1784, 2014. Available: https://arxiv.org/abs/1411.1784.