On-Device Deep Learning-based Multiple Behavior Detection using IMU Motion Sensors

1st Dong-Eon Kim
Center for ICT and Automotive Convergence
Kyungpook National University
Daegu, South Korea
dongekim@knu.ac.kr

2nd Ngoc-Dau Mai
AI Research Laboratory
VISIONIN
Seoul, South Korea
ngocdaumai@gmail.com

3rd Dong Seog Han*
School of Electronic and Electrical Engineering
Kyungpook National University
Daegu, South Korea
dshan@knu.ac.kr

Abstract*—This study proposes a system for monitoring the behavior of patients using an on-device deep learning-based inertial measurement unit (IMU) motion sensor. The wearable device captures the patient's four active behavior states (walking, eating, falling, and resting) using a three-dimensional accelerometer (ACC) and gyroscope (GYR). Five features, including mean value, standard deviation, median absolute deviation, minimum, and maximum, are applied to each 1-second segmented sample to extract the most significant characteristics from the signals. Four machine-learning approaches, such as support vector machines (SVM), multilayer perceptron neural network (MLP), long short-term memory (LSTM), and convolutional neural networks (CNNs), are used to evaluate the system's viability for different patient behavior identifications. The CNN algorithm showed the highest accuracy in patient behavior classification, surpassing the other algorithms by 92.68%. This algorithm is installed directly on the wearable device due to its exceptional performance, increasing system efficiency, and decreasing data transmission and connection latency. Additionally, a software program installed on the computer helps obtain necessary data from the wearable device through Bluetooth. It enables doctors, nurses, or supervisors to monitor a patient's behavior and other relevant information. The study's analysis results demonstrate the reliability of the device-based deep learning system for patient behavior recognition.

Keywords— Edge machine learning, IMU, Behavior Detection, Deep learning.

I. INTRODUCTION

Falls account for 40% of patient incidents in British Columbia (BCPSLS, 2015), making it crucial for medical practitioners to monitor patients to reduce such hazards [1]. Elderly patients are at higher risk of falling due to declining physical and mental health, mobility and balance issues, and reduced physical strength. The need for more nurse staff and supplies makes it challenging to supervise patients, which can lead to serious safety hazards. Researchers have developed an automated system that recognizes and alerts medical practitioners about potentially dangerous patient behavior to address this issue. Two methods for detecting patient behavior are wearable and non-wearable solutions. Wearable devices collect and interpret patient data using sensors, while non-wearable gadgets like cameras, sensors, and radars observe patient behavior [2], [3], [4].

In this study, we propose a wearable strategy with IMU sensor-based on-device deep-learning behavior recognition system to identify patient behavior. The system records two signal types, accelerometer gyroscope (GYR), to extract five valuable features, including mean, standard deviation, median absolute deviation, minimum, and maximum on the time domain. Four supervised learning algorithms, including SVM, MLP, LSTM, and CNNs, identify patient behavior based on significant values extracted from the feature extraction stage. TinyML is a machine-learning model with very little power, making it suitable for compact, lightweight, and portable embedded devices [5]. On-device machine learning is feasible using TensorFlow Lite [6], which programmers can use to run their models on mobile, embedded, and edge devices. The application makes monitoring and controlling patient behavior easier for nurses and doctors.

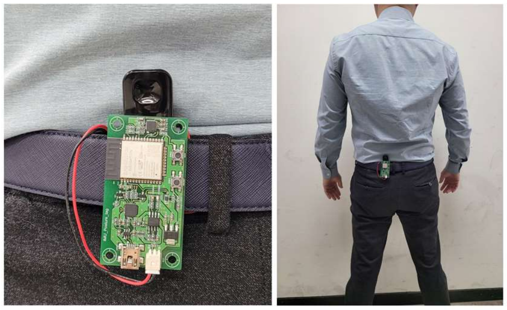
The primary objective of this study is to propose an integrated, wireless, and wearable Inertial Measurement Unit (IMU) device that can continuously and simultaneously monitor a patient's behavior. The paper makes three significant contributions: (1) the design and development of a wearable device that is highly comfortable, stable for long-term attachment, and has wireless connectivity to monitor behavior using the IMU module; (2) the introduction of a remote monitoring software program for computer-based monitoring; and (3) the proposal of a convolutional neural network (CNN)-based patient behavior detection model with superior performance, suitable for use on an embedded device.

II. EXPERIMENTAL FRAMEWORK

A. The Proposed System

The proposed methodology consists of two fundamental components: a smart wearable device and a self-made software application (as illustrated in Fig. 1. a). The LSM9DS1 Inertial Module is an embedded electrical device that integrates a 3D accelerometer (ACC) and 3D gyroscope (GYR) to track the body's orientation and forces. The accelerometer measures acceleration velocity, while the gyroscope detects direction and angular momentum. The sensor tag, which employs Bluetooth Low Energy (BLE) and an ATmega328 chip, can support machine-learning models and handle multiple tasks, such as collecting, analyzing, and classifying data. The wearable device used in the study is depicted in Fig. 1. (b).

The software program is designed to monitor and observe patient behavior. To provide optimal support to doctors and nurses in patient management, the program stores and displays a list containing high-demand patient information on the information panel. Additionally, the user interface appropriately shows behavior status and other essential patient-related data after receiving the information via Bluetooth from the patient's wearable device. The computer



(a) (b)

Fig. 1. (a) Wearable device and (b) Experimental setup.

program for monitoring and recording patient behavior status is presented in Fig. 2.

B. Experiment Protocol

The study enrolled five healthy participants aged between 25 and 30. Each participant utilized a portable monitoring device, which was well-tolerated and did not cause discomfort during use. To record the targeted behavior states, including eating, walking, resting, and falling, each participant was required to execute each behavior state five times within a 5-minute duration.

III. Proposed Method

A. Pre-processing and Feature Extraction

Upon recording, one-second samples with a fixed-length sliding window without overlap are created from the three-axes gyroscope and three-axes acceleration signals. To extract the intrinsic properties of the raw signals, five distinct features are employed in the time domain, including mean value, standard deviation, median absolute deviation, minimum, and maximum. Using the data from the six-axes Inertial Measurement Unit (IMU) sensor, a feature set of 30 features is constructed, utilizing five distinct types of features. The feature set is then converted into a set of feature vectors, normalized between 0 and 1, and presented as input to the learning models, thereby expediting the system computation of the learning models.

B. Classification

Recent years have witnessed the remarkable success of machine learning across various domains, including education, entertainment, and healthcare applications [7]. This study employs four different types of networks, namely SVM, MLP, CNN, and LSTM, to evaluate the system's performance in detecting patient behaviors. SVMs are a popular supervised learning technique for multi-class classification that performs well in high-dimensional domains [8]. The SVM performs two simultaneous tasks in this work: acting as a classifier and being combined with t-SNE (T-distributed stochastic neighbor embedding) for high-dimensional data visualization [9].

The MLP [10] is a feed-forward neural network that learns with backpropagation, consisting of an input layer of neurons, one or more hidden layers processing a set of weighted inputs to generate output using activation functions, and an output layer that predicts the output. A detailed explanation of the proposed MLP architecture is provided in Fig. 3. (c).

CNNs are artificial neural networks widely used for high-dimensional data classification, such as images, videos, and multichannel bio signals [11]. The advantage of CNNs is that

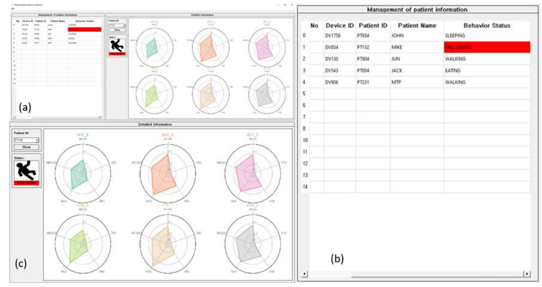


Fig. 2. Software program for multiple patient behavior management and monitoring.

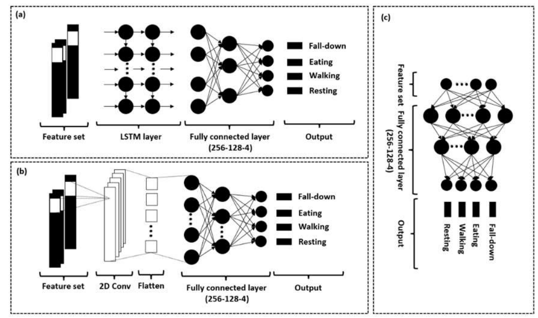


Fig. 3. The proposed (a) LSTM, (b) CNN, and (c) MLP architectures.

they require less manual work or preprocessing techniques. This network features an MLP-like structure, with convolutional layers added before the hidden layers. The convolution layer functions as a filter for extracting essential information from the data. A pooling layer is added to the model to reduce the feature size. This approach speeds up the process and reduces the number of factors. Average pooling and max pooling are also used to average or maximize the output of the pooling layer. The ReLU is used as an activation function in CNN to ensure nonlinearity. The fully linked layer is used to predict the results. A detailed illustration of the proposed CNN architecture is presented in Fig. 3. (b).

One major limitation of DNNs is that they only provide sparse temporal modeling since they employ a fixed-size sliding window of data. In contrast, recurrent neural networks (RNNs) incorporate cycles that feed network activations from a previous time step into the network as inputs to affect predictions at the current time step. The LSTM network is described as a customized RNN for modeling temporal sequences and their long-range relationships compared to conventional RNNs [12]. The recurrent hidden layers of the LSTM consist of memory blocks with memory cells having self-connections and special multiplicative units called gates for controlling the information flow. The network's secular state is regulated in memory cells with input and output gates. The proposed LSTM design in this work is presented in detail in Fig. 3. (a).

IV. Results and Discussion

Fig. 4 exhibits four discrete behavioral states and unprocessed signals captured by IMU sensors. These signals were initially divided into non-overlapping 1-second segments, to which five feature types were applied to create a meaningful set of features before normalization. T-SNE dimensionality reduction technique was used to visualize the data. Fig. 5 presents the outcomes of the combined T-SNE and

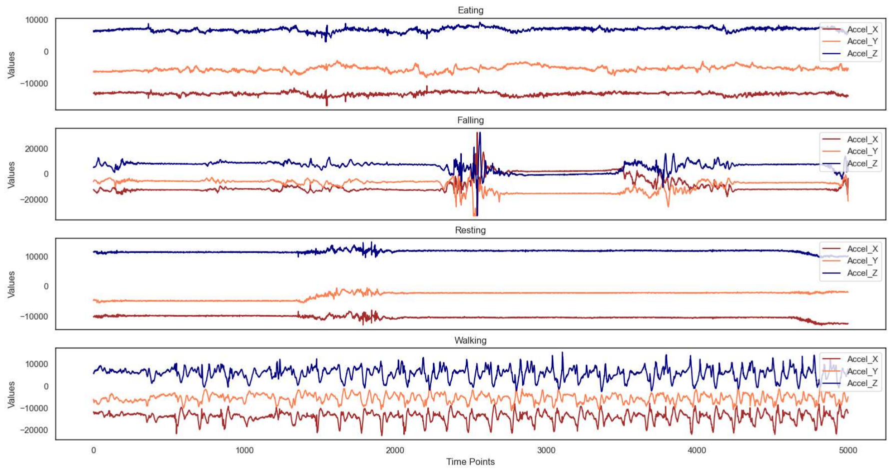


Fig. 4. Visualization of the three-axis accelerometer (ACC) with a complementary filter applied.

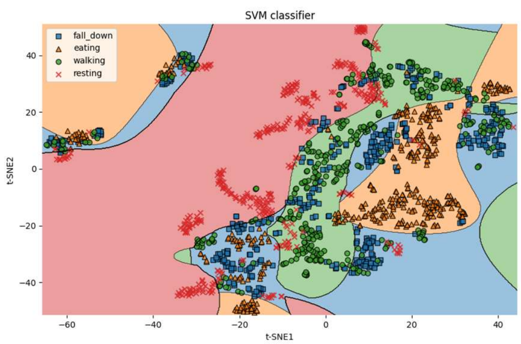


Fig. 5. Visualization for the feature set using t-SNE and SVM.

SVM classifier application for the feature set dimensionality reduction and visualization. The study used fivefold cross-validation to train and validate SVM, MLP, CNN, and LSTM models. For each subgroup, five subsets of feature sets were created, one of which was used as a test set, and the other four were used to train models. The performance of these models, using information from five different feature types, was compared, and CNN outperformed the other models in behavioral classification, achieving an accuracy level of up to 92.68%.

In contrast, LSTM, MLP, and SVM achieved detection performances of 90.12%, 88.95%, and 83.37%, respectively. Therefore, the CNN model was customized to classify patient behaviors on our device, as it exhibited superior classification accuracy. TensorFlow Lite was utilized to enable the machine-learning model to work on embedded and edge devices. The software developed captured important information through the Bluetooth protocol, including sensor ID, behavior status, and extracted attributes. This user-friendly interface assists healthcare professionals in monitoring a patient's behavior and relevant information. Fig. 6 shows the training and validation loss curves, indicating an increase in model accuracy and a decrease in loss with each epoch , and the performance of the CNN model, measured by the confusion matrix on the test set, was presented in Fig. 7. The confusion matrix depicted in Fig. 7 reveals that the system achieves the highest accuracy in detecting the "eating" state, followed by the "resting" and "fall-down" states, respectively.

TABLE I. COMPARISON OF THE FOUR PROPOSED MODELS IN PATIENT BEHAVIOR CLASSIFICATION

	Proposed Models			
	SVM	***MLP***	***LSTM***	***CNN***
Accuracy (%)	83.37	88.95	90.12	92.68

In contrast, the system's performance detecting the "walking" state is notably inferior to other actions. This phenomenon can be attributed to the fact that the extracted features have a lesser influence on the "walking" state than other movements. This finding has also been illustrated visually in Fig. 5.

V. CONCLUSION

This research paper proposes an on-device deep learning-based system that employs IMU motion sensors to track the behaviors of several patients. The system extracts the most important features of six-axis inertial signals using five features on the temporal domain for each 1s-segmented frame. Four different machine-learning approaches, namely Support Vector Machine (SVM), Multilayer Perceptron (MLP), Short-Term Long Memory (LSTM), and Convolutional Neural Network (CNN) were tested to examine the effectiveness of the proposed system in spotting various patient behaviors such

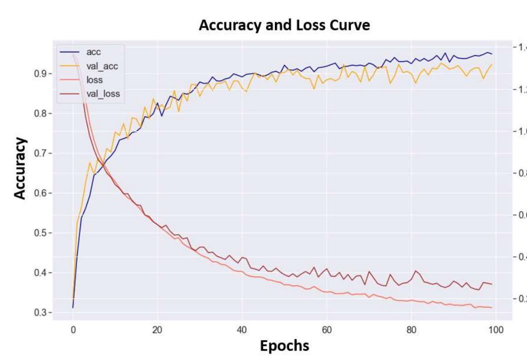


Fig. 6. Accuracy and loss curves of the CNN model.

as walking, eating, falling, and resting patterns. The results show that the CNN approach was the most accurate, achieving an accuracy of 92.68% in categorizing patient behaviors. Consequently, the CNN model was deployed on the wearable device using TensorFlow Lite.

Moreover, the software installed on the computer enables users to obtain critical data via Bluetooth protocol from the device, monitor patient behavior, and issue alerts for prompt response to unforeseen patient problems. The evaluation findings of this study demonstrate the applicability of the deep learning-based on-device system for predicting patient behavior using inertial signals. Future plans include developing a system that integrates driving service technology related to patient transportation based on this behavior information and expanding the system to detect vehicle occupant fall accidents, which can be utilized in the safety system during patient transportation.

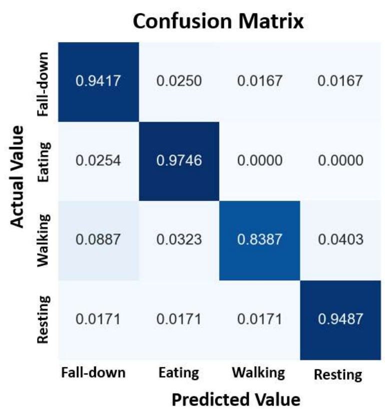


Fig. 7. Confusion Matrix of the proposed CNN model.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1A6A1A03043144).

REFERENCES

[1] https://pressbooks.bccampus.ca/clinicalproceduresforsaferpatientcaretrubscn/chapter/3-11-fall-prevention/, last accessed 2022/07/07.

[2] Cippitelli, Enea, et al. "Radar and RGB-depth sensors for fall detection: A review." IEEE Sensors Journal 17.12 (2017): 3585-3604.

[3] Ahmed, Nurzaman, Hafizur Rahman, and Md I. Hussain. "A comparison of 802.11 ah and 802.15. 4 for IoT." Ict Express 2.3 (2016): 100-102.

[4] Liu, B., Yan, Z., & Chen, C. W. (2016). Medium access control for wireless body area networks with QoS provisioning and energy efficient design. IEEE transactions on mobile computing, 16(2), 422-434.

[5] Samie, Farzad, et al. "Computation offloading and resource allocation for low-power IoT edge devices." 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT). IEEE, 2016.

[6] Warden, Pete, and Daniel Situnayake. Tinyml: Machine learning with tensorflow lite on ar-duino and ultra-low-power microcontrollers. O'Reilly Media, 2019.

[7] Mohammadi, Mehdi, et al. "Deep learning for IoT big data and streaming analytics: A survey." IEEE Communications Surveys & Tutorials 20.4 (2018): 2923-2960.

[8] Noble, William S. "What is a support vector machine?." Nature biotechnology 24.12 (2006): 1565-1567.

[9] Van der Maaten, Laurens, and Geoffrey Hinton. "Visualizing data using t-SNE." Journal of machine learning research 9.11 (2008).

[10] Riedmiller, Martin. "Advanced supervised learning in multi-layer perceptrons—from back-propagation to adaptive learning algorithms." Computer Standards & Interfaces 16.3 (1994): 265-278.

[11] Lee, S. Y., Hung, Y. W., Chang, Y. T., Lin, C. C., & Shieh, G. S. (2021). RISC-V CNN coprocessor for real-time epilepsy detection in wearable application. IEEE transactions on biomedical circuits and systems, 15(4), 679-691.

[12] Yu, Yong, et al. "A review of recurrent neural networks: LSTM cells and network architec-tures." Neural computation 31.7 (2019): 1235-1270.