Special Collection: Smart Wellness Services for Lifestyles of Health and Sustainability (LOHAS)

INTERNATIONAL JOURNAL OF
ENGINEERING BUSINESS
MANAGEMENT

International Journal of Engineering Business Management
Volume 10: 1–7
© The Author(s) 2018
DOI: 10.1177/1847979018787905
journals.sagepub.com/home/enb
SAGE

An artificial neural network–based fall detection

SunGil Yoo[1] and Dongik Oh[2]

Abstract

With the rise in the elderly population, the importance of care services for the elderly is also increasing. Among the care services, sudden fall detection is one of the most important services that the elderly need. The hip joints are prone to damage when they fall, and most of such injuries can lead to very severe consequences. In recent times, researches on fall detection have been very active. Fall detection by attaching an acceleration sensor to the waist of a person is popular and the detection rate is very high. However, when the fall is detected from a sensor attached to the wrist, which is more convenient as compared to the waist attachment, the detection accuracy is lower. To overcome the problem, in this article, we propose a system that distinguishes falls from the acceleration sensor attached to the wrist using an artificial neural network–based deep learning method. With the proposed method, we could detect the falls with a 100% accuracy in an experiment.

Keywords

Fall detection, artificial neural network, artificial intelligence, deep learning, pattern recognition

Date received: 29 November 2017; accepted: 3 January 2018

Introduction

Population aging is rapidly progressing worldwide. The number of elderly people living alone is increasing as well. A decrease in the number of elderlies living with their children and an increase in the number of those living alone after the death of their spouse are the main causes of the phenomena. Especially in Korea, the suicide rate of the elderly has tripled in the last 10 years. According to the Korea National Statistical Office, the number of the elderly aged 65 or older living alone is expected to increase from 1.19 million in 2012 to 3.34 million in 2035.

In this regard, the medical expenses for the elderly are rapidly increasing. According to the "Annual Health Insurance Statistics for 2016" by the National Health Insurance Corporation and the Health Insurance Review and Assessment Service, the medical expenses for senior citizens aged 65 and older have risen by 13.5% over the previous year and doubled from 2009. The rate of increase in elderly medical care expenditure is 8% in 2012, 9% in 2013, 10.4% in 2014, and 11.4% in 2015.

To deal with such aging-related issues, many studies on elderly care have been conducted, and various services have been suggested and developed. For example, the u-care system is a representative care system developed for the single-person elderly household in Korea. This system has been adopted by many provincial governments of Korea and demonstrated its usefulness. Currently, the service is operational under the name, "Emergency Alert Service for the elderly and severely disabled living alone."[1]

According to the results of various surveys conducted in elderly care services, the most serious fears of the elderly

[1] Department of Medical Science, Soonchunhyang University, Asan, South Korea
[2] Department of Medical IT Engineering, Soonchunhyang University, Asan, South Korea

Corresponding author:
Dongik Oh, Department of Medical IT, Soonchunhyang University, Asan, South Korea.
Email: dohhea@gmail.com

was gas leakage, intrusion by a stranger, fire, fall, and so on.[2] However, according to a research by the Korea Consumer Agency, the actual portion of accidents happen to the elderly was dominantly the fall (63.3%), followed by laceration/vomiting (5.6%), aspiration/choke (4.3%), and burning (3.5%).[3]

Hip fractures of the elderly have 90% mortality if they are not taken care of properly, and the probability of death within 6 months is about 20–30%. Patients who suffer the fall may not be able to move for months. In such cases, the risk of death increases due to the consequential complications such as pressure ulcers, sepsis, and lower thrombus. If they have chronic medical conditions such as hypertension and diabetes, the situation may become worse. Therefore, a prompt and adequate treatment is needed after a fall, and an immediate surgery (within 24 to 72 h) is very important, if necessary.[4]

However, the u-care system currently in use has only the functionalities of fire detection, gas leak detection, and solitary monitoring. It does not provide the fall detection service, which is the most dangerous threat to the elderly living alone. The fall detection functionality is not included in the system mainly because the sensors to detect the fall will increase the system cost. Also, since the elderly need to attach the sensor units on their body, it may give discomfort in the system usage.

Most of the existing studies have distinguished falls by judging whether the acceleration value occurred during a fall exceeded a certain threshold. The most popular method is based on the signal magnitude vector (SMV) value, which eliminates the directionality and uses only the magnitude component of the acceleration. Earlier studies have compared the SMV values at the time of falls with the SMV values of activities of daily living (ADLs) and used the recognition rate for evaluating the performance of the fall detection mechanisms.[5,6]

Previous studies show that the fall recognition rate differs according to the sensor positions. Maarit et al. used the acceleration value from the head, waist, and wrist. The result showed that, with the values from head and waist, the recognition rate was close to 100%. However, with the values from the wrist, the recognition rate was around 64%.[7] For this reason, earlier researches have applied the acceleration sensor mostly to the waist or the chest area and tried to add more sensors (such as a gyroscope) or tried to optimize the software algorithm to improve the recognition rate. However, since the method of determining the fall using the acceleration SMV values of the waist region already has a recognition rate of 90% or more, the meaning of the improvement was marginal. Besides the performance issues, there are two other issues to consider when applying the system in the real-world environment. First, since we need to attach one or more sensors to the human body, the cost of the system will increase. Second, the user needs to bear with the discomfort of wearing the device in some part of the body.

For this reason, many researches have been conducted to detect falls using a smartphone.[8,9] Smartphone is a very popular device, so there may not be extra hardware cost for the fall detection. However, this method also has a problem of discomfort, because the elderly always need to fix a smartphone in some part of their body. For the attachment of the sensor and the convenience of lifestyle to be achieved, recent researches have focused on the use of wrist-bands for fall detection. Nowadays, smart watch is popular and wearing it is natural to most of the people. However, the recognition rate of fall detection using smart-bands is only around 65%.[10]

Our research team has been working on improving the fall recognition rate using the acceleration from the wrist area.[11] We analyzed the frequency characteristics of the SMV values of the acceleration signals to distinguish the fall patterns. We could improve the recognition accuracy to 75%. However, there still is much room for improvement. Because the signal from the wrist has a more complex pattern than the one from the chest or waist, the use of a threshold value or frequency pattern matching could not overcome the fundamental limitation.

With the emergence of the fourth industrial revolution, studies on deep neural networks have been actively conducted. Machine learning techniques using deep neural networks have been applied to various fields, and it demonstrated good results in the areas where "specific rules" fail to solve the problem, such as Go, image recognition, and stock trading.

In this article, we propose a method of analyzing the acceleration signal from the wrist region using neural network techniques. The results of the study showed that the ADLs and the falls (rear) could be distinguished with a 100% accuracy. The ADLs include walking, stepping-down, and lying-down actions. In addition, it was also possible to distinguish the fall from the jump, which has similar acceleration peak value as the fall.

If the fall recognition is possible with the acceleration value generated from the wrist, we may use smart-bands or smart watches for the application, which will lead to a better adoption of the system.

Artificial neural network and deep learning

Artificial neural network

Artificial neural network (ANN) is a computer structure designed by mimicking biological neural network. It is used in various fields that require pattern recognition, data classification, and result prediction because learning is possible through given data. ANN was first proposed in 1943, and the famous multilayer neural network model consisting of input layer, output layer, and hidden layer and back-propagation algorithm had been proposed in the mid-1980s.[12,13] Especially in 2006, Hinton had solved the problems of the deep network such as overfitting and local

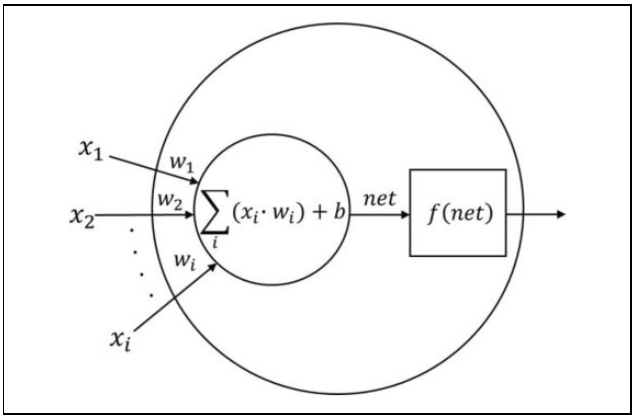

Figure 1. Structure of neurons constituting ANN. ANN: Artificial neural network.

minimum.[14] It is now applied to the problems that we thought in the past unsolvable, such as image recognition, stock trading and the game of Go.

ANN consists of artificial neurons, as shown in Figure 1, which mimics biological neurons. An artificial neuron is a structure that multiplies an input data x by setting an arbitrary weight w, adds it to b, and inputs it to an activation function f.

For the activation function, sigmoid function of equation (1) is traditionally used. Sigmoid is a function that converts an input value to a real number between 0 and 1. However, since the sigmoid function is considered to be the cause of the vanishing gradient issue, various other functions have been proposed. Rectified Linear Unit (ReLU) equation (2) is one of the most used functions.[15] ReLU outputs a value less than 0 as 0, and a value larger than 0 is output as it is.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

$$f = \left\{ \begin{array}{l} (x < 0)\, f(x) = 0 \\ (x \geq 0)\, f(x) = x \end{array} \right\} \tag{2}$$

$$\mathrm{SMV} = \sqrt{A_x^2 + A_y^2 + A_z^2} \tag{3}$$

In a neural network configuration, neurons can be composed of several layers. When a neural network has three types of layers (an input layer, a hidden layer and an output layer), it is called a multilayer neural network. A multilayer neural network with three or more hidden layers is called deep neural network (DNN). In addition, learning deep structure models such as DNNs is called deep learning.[16]

A fall detection mechanism development

Experimental data acquisition

In this article, we used Google TensorFlow for DNN configuration and learning.[17] TensorFlow is a machine learning library that has been released by Google and is currently the most widely used library.

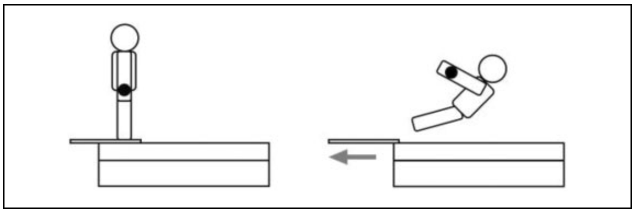

Figure 2. Experimental setting for the fall.

Using the TensorFlow, the neural network implementation can be more convenient, because it provides functions used for machine learning, including the activation function and the initialization function. We used the 1.3.0.1 version of the TensorFlow for Python.

For neural network learning and verification, experiments were conducted to obtain the acceleration values of the fall and non-fall activities.

According to a study on the elderly safety and accidents,[18] slipped (33.7%), tripped (24.0%), and lose one's footing (13.5%) were the most common causes of the fall. We prepared a similar situation of such falls and conducted an experiment. However, we excluded fall types other than slip, because the other types may result in a severe injury to the experimental subjects; the slip is relatively a safe experiment.

The experimental environment was prepared as follows. A floor mat capable of absorbing the fall impact was placed in a double layer, and a sheet was placed on top of the mat. The subject to which the accelerometer was attached on the left wrist was asked to stand on the sheet. By pulling the sheet at high speed, the subject's fall was induced. Figure 2 depicts the experimental setting.

Five men, aged 23 to 35 years old, were selected as subjects and the fall experiment was conducted 15 times for each person.

We have built an accelerometer sensor module that can be attached to the wrist of the participants. The module has a three-axis acceleration sensor and a battery, a Bluetooth chip for communication and an LED for operation confirmation. Figure 3 shows the sensor module.

Acceleration values measured at the wrist are transmitted to the Android device via Bluetooth communication, and the Android device transmits the acceleration value to the server via Wi-Fi communication.

We measured the sensor signal when the subject began to fall, until the fall completed. The acceleration variation of the sensor's x, y, z axes and the SMV values of the signal were calculated and recorded using equation (3). The sampling rate of the acceleration sensor was set to 50 Hz.

The ADLs to compare with the falls are walking, running, going downstairs, sitting, lying down, and jumping. Five persons were measured five times each for the above-mentioned activities. Acceleration values and SMV values

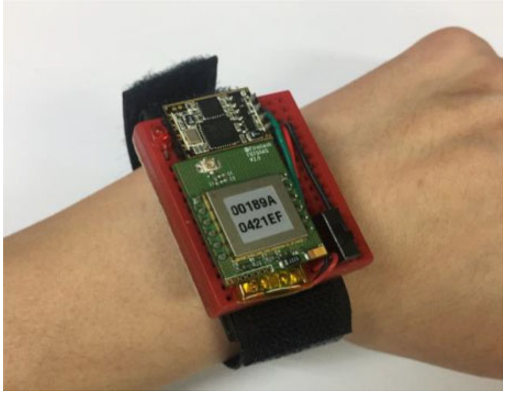

Figure 3. Sensor module.

Table 1. Description of experiments.

Motion	Explanation
Walking	Walking on a flat floor of around 30 m
Running	Running on a flat floor of around 30 m
Going downstairs	Go downstairs for 10 s
Sitting	Sitting and standing: Repeat for 5 counts
Lying	Laying and standing: Repeat for 5 counts
Standing jump	Standing jump: Repeat for 5 counts
Fall	Falling motion (rear) 15 counts

of x, y, z axes are recorded. The sampling rate of the sensor is set to 50 Hz. Table 1 describes the details of the experiments.

From the acquired experimental data, 60 fall data and 60 non-fall data (total 120 data) were randomly selected and used for neural network learning and testing.

Data preprocessing

When data is transmitted to the designed neural network, the size of input data should be the same as the number of input layer variables (in our model, the input layer size is 525). However, since the duration of each action including the fall is different, we needed to make the data size the same. To do so, among the recorded acceleration data, those related to the standing posture were removed. In the case of walking or going downstairs, because the same pattern is repeated, data is cut out to have 4 to 5 repeated patterns only (Figure 4).

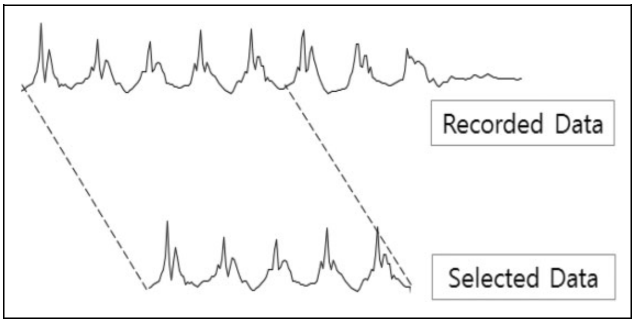

Figure 4. Example of data selections (down the stairs).

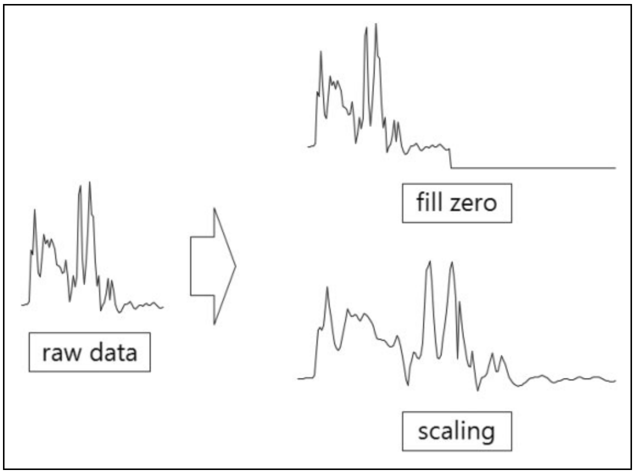

Figure 5. Example of scaling and fill zero processing (fall data).

After the process, the longest data of all actions was going downstairs with 165 acceleration values for each of the x, y, z axes. Based on this, all data were unified to have 175 values. If the number of recorded data was smaller than the standard, the missing part was filled with zero or scaled using one-dimensional linear interpolation. Figure 5 shows the result of preprocessed SMV values of falls by this method. The method of filling empty spaces with zero is the commonly used technique in natural language processing, and the scaling method is routinely used in image processing.[19,20] We then compared the recognition rates of these two types of data.

Of the total of 120 preprocessed data, data was randomly selected and divided into 90 training sets and 30 test sets. Using the training set data, a model that can distinguish between the fall and non-fall activities was created, and the test set data was fed to the model to verify if it classifies the data as intended.

Optimizing training model

We used a neural network constructed as shown in Figure 6. It is made of a three-layer concealment of 500-500-2000, which is the same as the DNN used by Hinton.[14] Hinton showed a good performance of recognizing 28x 28 pixels images using a hidden layer of this kind of structure.

The number of variables in the input layer is set to 175×3. This is because the sample size of the experiment is 175, and the acceleration data has values of x, y, z axes per sample. Or, we may use 175 calculated SMV values for the network input layer. The output layer consists of two neurons – [0, 1] for the fall and [1, 0] for the non-fall. In the output stage, the activation function was set as a sigmoid, so that two-neurons each output a real number between 0 and 1. In other layers, sigmoid or ReLU function was used as the

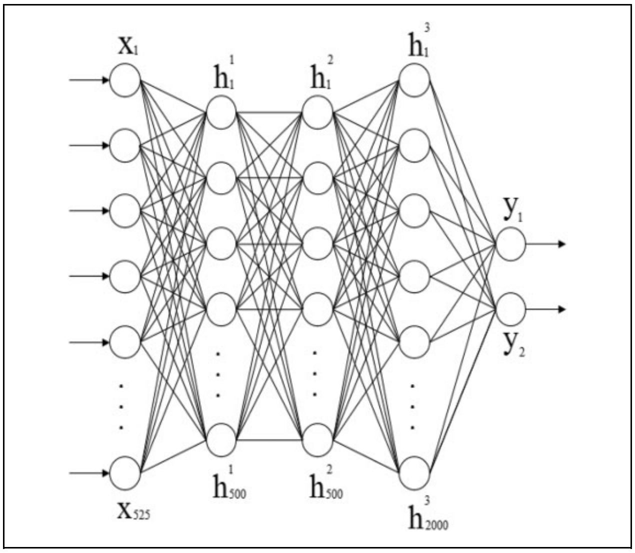

Figure 6. Structure of the neural network.

activation function. We will compare the performance of activation functions in section "Comparisons and results."

The network can be used for learning based on two types of data. The first is 175 tuples of *x*, *y*, *z* values, and the second is 175 SMV values. We will compare the recognition accuracy between the cases of *x*, *y*, *z* tuples and the SMV values in section "Comparisons and results" as well.

The neural network sets the weight value w_i and *b* (bias) value of each neuron in the entire network according to the training data produced by the learning process (see Figure 1). In this process, the difference between the actual data and the learning result is minimized. In this model, when data is input, the learning process proceeds to output a value close to [0, 1] when the input data is a fall, and a value close to [1, 0] when the data is a non-fall.

After rendering data to the model, we compare the values of two neurons in the output layer. If the output value of the first neuron is closer to 1 (if it is like [1, 0]), then it is not a fall, and if the value of the second neuron is closer to 1 (like [0, 1]), it is classified as a fall.

If the cost is minimized, we may be certain that the learning model is optimized for the given data set, namely, the smaller the cost, the more accurate the training is. However, even if training data sets are identified with 100% accuracy, the cost may not necessarily be the minimum. Depending on the given data, even if the cost is not the minimum, the classification may give 100% accuracy. Therefore, the cost determination process still has a room for improvement. In this article, we used gradient descent algorithm which is one of the most popular algorithm among various optimization (cost minimization) methods.

Activation functions and initializers

When we begin the training and initializing of the w_i and *b* values of each neuron, there is an initializer that may give better performance than entering a random value. Xavier Initializer (2010), proposed by Glorot and Bengio, shows good performance by selecting initial values within a reasonable range, and it is a widely used initializer.[21] Also, sigmoid functions with random number initialization is another alternative that has traditionally been widely used.

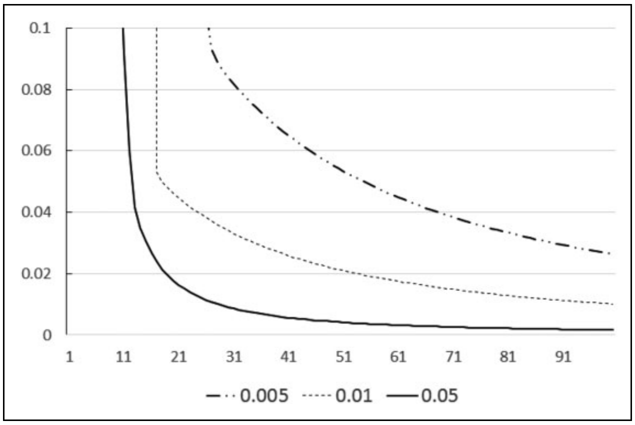

Figure 7. Comparing cost according to learning rate.

In the training process, a learning rate value must be set appropriately. If the value is set too large, the cost value will diverge and the training will not be successful. If the value is set too low, the training will be delayed and the total training time will increase.

To find the optimal learning rate, we compared the changes in cost by training 100 epochs for each learning rate. From the experiment, we found that the cost was minimized when it was set to 0.05. If it was set to 0.1, the cost value was divergent, and the learning failed. If it was 0.001, the learning proceeded slowly, and the cost was relatively high. Figure 7 shows the change of cost value regarding the change in the learning rate using the ReLU and Xavier Initializer. In the case of sigmoid function and random value initialization, learning was best done at the training rate of 0.5.

Comparisons and results

Comparison targets

The choice of activation function and the initializer as well as the input data are the main variants that affect the performance of the system. In this article, two types of acceleration data (three-axis raw value vs. SMV value), two ways of compensating the missing data parts (fill zero vs. scaling), and two types of activation/initialization functions (Sigmoid/Random vs. ReLU/Xavier) are used. Therefore, there are eight choices to be compared for performance as specified in Table 2.

Performance comparisons

We created classification models by conducting 50 epoch trainings for each of the eight cases in Table 2. For each

Table 2. Eight cases to compare performance.

Group number	Activation function/initializer	Data type	Missing data compensation
①	Sigmoid/Random	Raw value	Fill zero
②	Sigmoid/Random	Raw value	Scale
③	Sigmoid/Random	SMV	Fill zero
④	Sigmoid/Random	SMV	Scaling
⑤	ReLU/Xavier	Raw value	Fill zero
⑥	ReLU/Xavier	Raw value	Scale
⑦	ReLU/Xavier	SMV	Fill zero
⑧	ReLU/Xavier	SMV	Scale

case, we created 30 models to see the average performance of the case. This is because even if we have the same training data, we may get different models from the training.

The average cost, training time, and accuracy of the fall/non-fall detection, which are the performance comparison matrix of this study, were measured for each type. The results of the eight cases mentioned in Table 2 are given in Tables 3 and 4.

In the cases using sigmoid activation function and random initialization, SMV data showed the highest fall/non-fall detection accuracy when processed with zero-fill (case ③, 99.89%). In addition, there was a statistically significant difference between this case compared to the cases ② and ④ (Tukey HSD tests, p <0.001). For the cost value, there were no statistically significant differences among the four cases. For the training time, case ③ had the shortest time value. Also, there was a statistically significant difference between this case and cases ① and ② (Tukey HSD tests, $p = 0.025 <0.05$, $p < 0.001$). Overall, case ③ gave the best result for the performance matrices of the cases given in Table 3.

In the cases where ReLU activation function and Xavier initialization were used (Table 4), case ⑦showed the highest performance in all three matrices. There was a statistically significant difference in this case as compared to cases ⑤, ⑥ and ⑧ for training time (Tukey HSD tests, p <0.001, $p < 0.001$, $p = 0.045 <0.05$). There was also a statistically significant difference in this case as compared to cases ⑤ and ⑥ for the cost value (Tukey HSD tests, $p < 0.001$). For the fall/non-fall detection, all four cases had 100% detection accuracy.

From the results in Tables 3 and 4, we could conclude that case ⑦ is the best choice for the fall detection model. We could get 100% fall/non-fall detection rate and the shortest training time.

Table 3. The evaluation results of models using sigmoid function + random initialization.

	Raw value		SMV value	
	Fill zero ①	Scale ②	Fill zero ③	Scale ④
Cost	1.64E-07	4.94E-07	4.65E-07	1.33E-06
Spent time (s)	5.4375	5.5941	5.1057	5.3200
Training set accuracy (%)	100.00	100.00	100.00	100.00
Test set accuracy (%)	98.67	96.89	99.89	97.11
Test set sensitivity (%)	97.78	94.00	99.78	98.89
Test set specificity (%)	99.56	99.33	100.00	95.33

Table 4. The evaluation results of models using ReLU + Xavier initialization.

	Raw value		SMV	
	Fill zero ⑤	Scale ⑥	Fill zero ⑦	Scale ⑧
Cost	7.18E-03	7.26E-03	3.77E-03	4.15E-03
Spent time (s)	6.0065	5.7584	5.2650	5.5774
Training set accuracy (%)	100.00	100.00	100.00	100.00
Test set accuracy (%)	100.00	100.00	100.00	100.00
Test set sensitivity (%)	100.00	100.00	100.00	100.00
Test set specificity (%)	100.00	100.00	100.00	100.00

Conclusions

In this study, we measured the acceleration data from the wrist and used them to detect the fall. We used an artificial neural network technique to determine the fall. The data was used for training the model, and various learning and optimization techniques were applied and tested. By comparing the use of different preprocessing techniques and choice of activation functions and initializers, we found that the use of ReLU activation function with Xavier initializer while filling the missing data with 0 is the best choice for fall detection.

When SMV data is used, we gather the training time is reduced because the input size is 1/3. Also, it seems that a more efficient training is achieved when zero-fill is applied than the scaling. When the input is zero, we think, the neuron seems to be diluted during the model building process (optimization), so that the processing could be more simplified.

In the previous fall detection studies, the fall was recognized using the acceleration sensor from the waist or the chest, and the recognition rate was over 95%. However, when the acceleration sensor from the wrist was used, the recognition rate was about 75%. The proposed artificial neural network method of this study was able to recognize falls with a100% accuracy using the acceleration from the wrist. This is a huge improvement to the conventional fall detection mechanism. With the wrist-band type devices, we can also cut down the system cost (we may use existing smart-watch or band) and provide comfort while wearing.

In this article, the experiment was conducted only for rear-falls because of the safety of the experimental subjects.

We may introduce dummies to the experiment so that many different types of falls can be tested and detected. Also, many other activities other than the fall should be able to be classified so that the living pattern of the elderlies can be recognized and used for better elderly services in the future.

Acknowledgement

This work was supported by the Soonchunhyang University Research Fund.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

SunGil Yoo http://orcid.org/0000-0003-3678-9624

References

1. *2017 Emergency Alert Service for the elderly · severely disabled living alone project*. Korea: Ministry of Health and Welfare Korea, 2017.
2. Min-Soo K, Saizmaa T and Hee-Cheol K. A questionnaire study for the elderly's awareness and notification of accidents in smart home environments. *Korea Multi Soc Acad Pub* 2006; 2006(1): 99–104.
3. Korea Consumer Protection Board. *Survey on Safety of Elderly People at Home*. Korea: Korea Consumer Protection Board, 2003.
4. Hyoung-Keun O.Intertrochanteric fracture: how to improve the surgical outcomes?. *J Koren Orthop Assoc* 2015; 50(3): 192–201.
5. Karantonis D, Narayanan M, Mathie M, et al. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Trans Inf Technol Biomed* 2006; 10(1): 156–167.
6. Fabio B, Becker C, Cappello A, et al. Evaluation of accelerometer-based fall detection algorithms on real-world falls. *PLoS ONE* 2012; 7(5): e37062.
7. Maarit K, Konttila A, Lindgren P, et al. Comparison of low-complexity fall detection algorithms for body attached accelerometers. *Gait Post* 2008; 28(2): 285–291.
8. Tong Z, Wang J, Liu P, et al. Fall detection by embedding an accelerometer in cellphone and using KFD algorithm. *Int J Comput Sci Netw Sec* 2006; 6(10): 277–284.
9. Jiangpeng D, Bai X, Yang Z, et al. PerFallD: a pervasive fall detection system using mobile phones. In: *2010 8th IEEE international conference on pervasive computing and communications workshops (PERCOM Workshops)*, Mannheim, Germany, 2010, pp. 292–297.
10. Degen T, Jaeckel H, Rufer M, et al. Speedy: a fall detector in a wristwatch. In: *7th IEEE international symposium wearable computers (ISWC)*, 2003, pp. 184–189.
11. Suhwan L, Oh D and Nam Y. Enhancement of fall-detection rate using frequency spectrum pattern matching. *J Int Comput Ser (JICS)* 2017; 18(3): 11–17.
12. Warren SM and Pitts W. A logical calculus of the idea immanent in nervous activity. *B Math Biophy* 1943; 5(4): 115–133.
13. David ER, Geoffrey EH and Ronald JW. Learning representations by back-propagating errors. *Nature* 1986; 323(9): 533–536.
14. Geoffery EH, Osindero S and Teh YW. A fast learning algorithm for deep belief nets. *Neural Comput* 2006; 18(7): 1527–1554.
15. Vinod N and Hinton GE. Rectified linear units improve restricted boltzmann machines. In: *ICML'10 Proceedings of the 27th International Conference on International Conference on Machine Learning*, Haifa, Israel, 2010, pp. 807–814.
16. Yann L, Bengio Y and Hinton G. Deep learning. *Nature* 2015; 521: 436–444.
17. Martín A, Agarwal A, Barham P, et al. Tensor flow: large-scale machine learning on heterogeneous systems. 2015. https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45166.pdf (accessed 4 July 2018).
18. Yoo Y. Falls and functional levels associated with falls in older people living in the community. *J Korea Gerontol Nurs* 2010; 12(1): 40–50.
19. Baotian H, Lu Z, Li H, et al. Convolutional neural network architectures for matching natural language sentences. In: *NIPS'14 Proceedings of the 27th international conference on neural information processing systems*, Vol. 2, Montreal, Canada, 2014, pp. 2042–2050.
20. Nanne VN and Postma E. Learning scale-variant and scale-invariant features for deep image classification. *Pattern Recogn* 2017; 61: 583–592.
21. Glorot X and Yoshua B. Understanding the difficulty of training deep feedforward neural networks. In: *Proceedings of the thirteenth international conference on artificial intelligence and statistics (AISTATS 2010)*, Vol. 9, Sardinia, Italy, 2010, pp. 249–256.