1. Introduction

1.1 Background

Japan is an aging society, correspondingly, the number of dementia patients is increasing. It is reported that in 2012, 15% of elderly people aged 65 and over have dementia. That is to say, about 4.46 million elderly people have dementia. Fall prevention sensors are often used to support the care of dementia patients at night. There are three major types of fall prevention sensors. The first type of sensor is a mat type. The sensor is placed over a mattress and detects the pressure applied to the mat by the patient lies down. The second one is a clip type, which is the easiest and least expensive option. The clip is attached to the garment of the patient and connected to the sensor switch. Whenthe patient leaves the bed, the sensor switch turns on and the sensor is activated. The third type of sensor uses infrared light. When the infrared light from the emitter to the receiving unit is blocked by the body movement of a patient, the sensor is activated. This type of sensor is flexible but expensive compared to other types of sensors. All of the above sensors are designed only to detect when the patient leaves the bed. Therefore, in principle, they all have the same problem. When a caregiver responds to a sensor call, it is often too late to support the patient’s movements as they leave the bed. We are studying the ways to predict the behavior of patients, rather than waiting for them to leave the bed. In order to realize such a sensor system, we focus on the sleeping positions. Since turning over in one’s sleep occurs between REM sleep and non-REM sleep [1], we think that predicting the arousal level is possible by monitoring positions while sleeping. In this study, we focused on the frequency of change in sleeping posture, in order to verify which postures closely related to awakening timings. In consideration of the privacy of the care recipient, we have studied and developed a method to detect changes in posture while sleeping by deep learning technology using data obtained from a sheet-type pressure-sensitive sensor.

1.2 Related Work

In order to predict bed-leaving behavior at night, a method of categorizing turning motions has been created using the data gathered from sleeping postures obtained by the seat type pressure-sensitive sensor [2]. Building on this research, we used special sensors with high detection sensitivity, compared various methods of machine learning, and confirmed the highest detection accuracy with posture estimation by a support vector machine. Next, we used a seat type pressure-sensitive sensor that was less expensive and usable as a general bed sheet.

# corresponding author: Noriaki Kuwahara (E-mail: nkuwahar@kit.ac.jp)
Based on the body pressure data obtained by using this sensor, we anticipated the sleeping posture using deep learning, a method which has been drawing attention in recent years. It is reported that the learning model constructed using CNN (Convolution Neural Network), a deep learning technique, has higher accuracy than the conventional image recognition method [3]. However, the accuracy of results obtained by CNN was worse than results using a support vector machine [4].

1.3 Purpose

The sensor we used in the previous study had low sensitivity, making detections of the upper and lower arms especially difficult. For this reason, it was difficult to distinguish between Log and Yearner postures, typical sleeping postures shown in Fig. 1. Soldier and Starfish were also difficult to distinguish. When trying to detect the patient turning over in bed, it is unnecessary to distinguish between these postures. In this research, we decided to identify six different sleeping postures as shown below.

1. Foetus Left
2. Foetus Right
3. Log or Yearner Left
4. Log or Yearner Right
5. Soldier or Starfish
6. Freefall

CNN and AE (Auto Encoder) are typical methods in deep learning technology and have few application examples in fields other than image and speech recognition; however, it is widely anticipated that they will be effectively applied elsewhere. Therefore, in this study, we aim to examine the method for predicting how someone leaves the bed by comparing the accuracy of classification in sleep positions using these two methods.

2. Method

2.1 Convolution Neural Network

Many methods of deep learning have been proposed. Various approaches are also being considered in the field of image recognition. Currently, CNN is considered to be the most successful. CNN is an extension of a classical multilayer perceptor, based on findings in the structure of the visual cortex. In concrete terms, it consists of a convolution layer responsible for local feature extraction of the image and a pooling layer (subsampling layer), these layers are then repeated. Since the parameters of the convolution filter are shared at all places in the image, the number of parameters is greatly reduced compared to a full connected network. Furthermore, by interleaving the pool layer, the number of parameters can be further reduced, and at the same time invariance can be added to the translation of the input, which is essential for general object recognition.

2.2 Auto Encoder

The auto encoder was supervised while learning using the same data for the input layer and the output layer in the 3-layer neural network. This could be called a special case of back propagation. Since learning is performed by back propagation, it becomes a nonlinear optimization problem. Activation functions of the intermediate layer and output layer can be arbitrarily selected. When the teacher data is a real number and there is no value range, the activation layer function is often chosen to be the identity map (that is, it does not change anything). If the identity mapping is also selected for the activation function of the intermediate layer, the result nearly agrees with the principal component analysis.

2.3 Experimental Setup

In this study, body pressure data was measured using a pressure sensor system produced by Tsuchiya Co., Ltd. The textile-type pressure sensor

![Fig. 1 Typical Sleeping Postures](image-url)
made of conductive fibers was soft and comfortable, even when placed in a bed. The size of the pressure sensitive area was 1600 mm × 800 mm, and the seat had 80 × 40 sensing points. The sensor’s output was as high as 12 frame rates of data per second. In this experiment, the pressure sensor was placed on the bed, then the mattress cover was placed on it and the sensor was fixed as shown in Fig. 2.

2.4 Participants

There were seven participants in this experiment, ranging from 21 to 27 years of age. Every participant was healthy. Each participant performed the above mentioned sleeping positions in their own manners. There were three trials in each position and, as a result, we obtained 126 samples (21 samples per position × 6 positions) in total. First, we measured the initial pressure value and then, we instructed the participants to lie on the bed in each of the nine postures. We then at last measured the pressure after checking that the posture was stable.

2.5 Data Processing

The data set for machine learning in this research was the pressure data on which the following preprocessing was performed. The pressure sensor can continuously record the pressure value of 12 frames per second. In the measurement of this experiment, when the posture of the participants stabilized, the time was recorded. One frame at that time was cut out and used as one body pressure data. In this experiment, in order to increase the number of data samples, two data frames were extracted and used for one posture (1st frame and 15th frame) in response to the fluctuation of periodic data from the pressure sensor. Since the measured pressure value included the pressure of the cover and the pillow, noise, pretreatment for removing them was performed. After converting the pressure data to a pressure distribution image, it was used as input to the discriminator. One body pressure data has 3200 (40×80) pressure values. By standardizing the pressure value with 0-255, it was regarded as a pixel value and a grayscale pressure distribution image was created, as shown in Fig. 3.

2.5.1 CNN configuration

Table 1 CNN configuration

| Input Image | 40 pixel × 80 pixel × gray scale |
|-------------|----------------------------------|
| Convolution Layer | Filter (5 pixel × 5 pixel × 16 types) |
| Pooling Layer | 2 × 2 filter |
| Convolution Layer | Filter (3 pixel × 3 pixel × 32 types) |
| Pooling Layer | 2 × 2 filter |
| Convolution Layer | Filter (3 pixel × 3 pixel × 64 types) |
| Pooling Layer | 2 × 2 filter |
| Flatten Layer | 3200 (5 × 10 × 64) |
| Dense Layer | 512 |
| Dense Layer | 6 |
| Active Function | Softmax |

(a) Sheet-type Pressure Sensitive Sensor

(b) Bed Equipped with Sensor System

Fig. 2 Experimental Setup

(a) Sheet-type Pressure Sensitive Sensor

(b) Bed Equipped with Sensor System

Fig. 2 Experimental Setup

Fig. 3 Data of the Pressure

Fig. 3 Data of the Pressure

Table 1 CNN configuration

Journal of Fiber Science and Technology (JFST), Vol.73, No.12 (2017) 345
2.6 Performance Evaluation

We constructed CNN or a Stacked AE classifier using the machine learning framework "Keras" [5] and identified sleeping postures. The language “Python” was used for programming. CNN configuration was shown in Table 1. Loss function was “Cross-Entropy,” and the optimizer was “Adam.” The stacked AE configuration was shown in Table 2.

3. Results and discussion

All data were randomly divided into three equal parts and set as data sets A, B, and C, respectively. As the learning data to test data ratio was 2:1, as shown below, learning and testing was tried three times, and the average was regarded as the discrimination rate. We performed validation tests using test data, which was not used for learning, and verified the general-purpose performance of classifiers by CNN and Stacked AE.

Test A. Learning data = B + C, test data = A
Test B. Learning data = A + C, test data = B
Test C. Learning data = A + B, test data = C

CNN:
Test A. 0.85
Test B. 0.96
Test C. 0.95
Average: 0.92

Stacked AE:
Test A. 0.81
Test B. 0.91
Test C. 0.92
Average: 0.88

In this research, the CNN recognition rate exceeded AE rates for all test data; the average recognition rate was 88% for AE and 92% for CNN. Processing of feature extraction is different between Stacked AE and CNN. We thought that it was possible to extract minor features of the images by using convolution, which led to an improvement in the recognition rate of CNN. Regarding the difference of discrimination rate by data set, data set A contained a lot of face down and supine postures, when compared with B and C. Therefore, when learning was performed using the data sets B and C, we believe that the recognition rate decreased because the classifier could not be learned sufficiently. Also, the image for face down and supine lying postures were difficult to discriminate by sight. For this reason, we thought that preparing a data set at the stage of preprocessing that clearly discriminates lying postures of face down and supine leads to better recognition ratio for Stacked AE and CNN.

4. Conclusion

As a result, with regard to any test data, the classification accuracy rate of CNN exceeded that of Stacked AE by around 4%. We found that the feature extraction method of CNN (combination of convolution and pooling layers) was more suitable for the data used in this experiment than the feature extraction by Stacked AE. However, we do not know by merely looking at the result which part of the sleeping posture the CNN classifier captures.
Therefore, it is necessary to observe how CNN learned the classifier by visualizing the output of the intermediate layer.

References

1. T. Kogure, *Journal of the Society of Biomechanisms*, vol. 29, no. 4, pp. 189–193 (2005), (in Japanese)
2. Mineharu, *ICIEV*, June (2015).
3. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton, *Advances in neural information processing systems*, pp. 1097–1105 (2012).
4. Kengo Wada, Noriaki Kuwahara, Kazunari Morimoto, *LNCS*, vol. 10286, pp 188–197 (2017).
5. https://keras.io/ (referred on 2 Feb. 2017)