Person Anomaly Detection based on Autoencoder with Obrid-Sensor

Taiki Sunakawa^a, Yuki Horikawa^a, Atsushi Matsubara^a, Seiji Nishifujii^a, Shota Nakashima^{a*}

^aYamaguchi University, 2-16-1, Tokiwadai, Ube, Yamaguchi 755-8611, Japan

*Corresponding Author: s-naka@yamaguchi-u.ac.jp

Abstract

This paper explores a novel method of fall detection assuming elderly people which can be trained easily by using AutoEncoder. The classifier has accuracy is 98.7%, which is 2.1 points higher than conventional method. In this method, Obrid-Sensor acquire brightness information. Moreover, the information based to detect whether a person is in a falling state with protecting privacy. On the other hand, the conventional method uses a classifier built by Support Vector Machine for fall detection. However it is necessary to prepare the data of the falling state as well as the standing state for training. In the proposed method, 78% less required training data than the conventional method, and only use the data of standing state for training.

Keywords: Obrid-Sensor, Privacy protection, Abnormal detection, Auto-encoder, convolution neural networks.

1. Introduction

The number of elderly people is increasing, it result in a seriously aging society⁽¹⁾. Elderly people have a risk of falling during their daily life due to dulling of their body movements, which lead to serious accidents⁽²⁾. For this reason, they need to be watched over in their daily life. However, it is difficult for a caregiver to watch over them at any time when they live alone or are in a private room. In order to reduce the burden on the caregiver, it is expected to develop a system to watch over the person. In conventional watching over systems, detection using cameras and sensors such as acceleration sensors, mat sensors, and ultrasonic sensors have been proposed⁽³⁻⁶⁾. Nevertheless, the camera possibly violate privacy in a private room⁽⁷⁾. In addition, the accelerometer needs to be a wearable device, which there is a risk of removed. The others installed sensor, such as mat sensors, are difficult to introduce because the sensor is needed to install all over the floor. Thus in our previous study, we developed a system using a Obrid-Sensor to watch over the elderly people based on the acquired brightness distribution⁽⁸⁾. the sensor can acquire brightness information like a camera, but it can protect privacy due to compression to one dimension. In our previous study, we proposed a method to detect falls by Support Vector Machine (SVM)⁽⁹⁾. Although, it is necessary to prepare enormous number of data of the falling state as well as the standing state for training. At the same time, it is necessary to reproduce the situation similar to the case of a fall which is a big burden for elderly people. Thus by AutoEncoder, it can detect abnormal state using the model

building from the data of only normal state⁽¹⁰⁾. As an example, the AutoEncoder with convolutional layer is used for concrete defect detecting⁽¹¹⁾. By applying this method, this paper proposes a detection method that can watch over dangerous status such as falling state using Obrid-Sensor.

2. Obrid-Sensor

2.1. Theoretical-Structure

A image of a Obrid-Sensor is shown in Fig. 1. The Obrid-Sensor is the optical sensor, which consists of line sensor with light receiving devices arranged in a row and a type of plano-convex lens called cylindrical lens. The line sensor is used such as scanners and 2D barcode readers. Cylindrical lens has a lens effect in one direction and does not have lens effect in the opposite direction. The theoretical structure of the Obrid-Sensor is shown in Fig. 2(a). The top view and side view of the sensor are shown in Fig. 2(b) and Fig. 2(c). As shown in Fig. 2(b), the lens is regarded as a plano-convex lens that has lens effect in this direction. Thus, vertical lights from each points U, C, L of the target space enter to the corresponding points S^U , S^C , S^L on the line sensor. Therefore, the sensor acquire vertical brightness information in the target space. As shown in Fig. 2(c), the lens does not have the lens effect because it is the shape of flat glass in this direction. Thus, lights from each points C^L , C, C^R of the target space enter to one point S^C on the line sensor. Therefore, horizontal light irradiated from the target space are integrated and the brightness that corresponds to integral value is acquired. The system will detect the state of the subject without violating privacy by Obrid-Sensor acquired brightness distribution waveform.

Fig. 1. Image of the Obrid-Sensor

2.2. Feature Values Acquisition

In this study, fall detection is based on the features of subject from distribution brightness waveform. To extract the feature because the waveform has features of background too, the background subtraction method is used. The background subtraction method acquire subject because calculate the difference of background data where there is no subject and the current data(12). The brightness information acquired from the Obrid-Sensor is shown in Fig. 3(a), and the image of the target space is shown in Fig. 3(b). In Fig

Fig. 2. Theoretical structure of Obrid-Sensor

Fig. 3. Background subtraction method

3(a), subtracted waveform is made by subtraction of background waveform from acquired waveform. This subtracted waveform is pointing features of subject.

3. Proposed Method

3.1. Autoencoder

AutoEncoder is a neural network of unsupervised learning which is trained to reconstruct that are close to input data. It is composed of an encoder to extracts features, and a decoder to reconstructs data of the same dimension as the input layer from the extracted features. the encoder and decoder are symmetric. By training the encoder and the decoder using only the data of the normal state, the model that able to correctly reconstruct the features of the normal state is built. While, AutoEncoder is not able to reconstruct well with anomaly data which does not using to training. Therefore this method detect anomaly status by reconstruction error. In particular, the error between the input and the reconstructed output becomes small when the input data is normal state. in the case input data is anomaly, the error becomes large because the reconstruction is not correct. This error is used for anomaly detection. In the proposed method, the model is training by inputting the distribution brightness waveform of standing subject from an Obrid-Sensor. The method is detection the standing state as a normal state and the falling state as an abnormal state.

3.2. Proposed Model

In this method, the model compose of convolutional and full connected layers for the network of AutoEncoder, which is a common model structure used in image processing⁽¹³⁾. The reason for using this model is that the waveform of the brightness distribution of a Obrid-Sensor has features similar to images. Thus, the method apply image processing such that convolution neural network. Fig. 4(a) and Fig. 4(b) show the brightness information of standing state and the images of the target space, and Fig. 5(a) and Fig. 5(b) show the brightness information of falling state and the images. In the figure, it can be seen that the brightness information is wide when the subject is standing and narrow when it is falling in this image. In this way, the features of the standing position are acquired from relations with neighboring pixels. In addition, this waveform often slides depending on the depth of the standing place. In the target space of the sensor, the waveform slides upward when the subject moves away from the sensor, and slides downward when the object moves closer. For this reason, a simple full connected layer can not learn the features between pixels. Therefore, a convolution layer can be used to learn generic situation for the training data. Table 1 shows an overview of the AutoEncoder architecture used in this study and Fig. 6 shows a schematic. Comparing the waveforms from 0 pixel to 150 pixel in Figs. 4 and 5, it can be seen that the waveforms are different in a wide range. Thus, we set a relatively wide kernel size and stride, and compress the waveforms to 3 dimensions to acquire a wide range of features. the decoder, which is a contrast to the encoder, reconstructs data to the same dimension as the input layer. The proposed method calculates the anomaly score using

the Residuals Sum of Squared (RSS) between the reconstructed output data and the input data from a Obrid-Sensor. Based on the anomaly score of normal and abnormal state, a threshold is defined and detect anomaly.

4. Evaluation Experiment

4.1. Methodology

Using the proposed method, brightness information of standing and falling states was acquired using a Obrid-Sensor in an experimental environment. The model was trained from the brightness information and fall detection

Fig. 4. Obtained information of standing person

Fig. 5. Obtained information of falling person

Table 1. Overview of the AutoEncoder architecture

Name	Kernel Size/Stride	Output Size
Conv 1	7/5	50
Conv 2	5/3	16
Full connected 3		3
Full connected 4		16
Conv 5	5/3	50
Conv 6	7/5	256

Fig. 6. Network structure of Autoencoder

was evaluated.

The experimental environment was set to satisfy the minimum size of a private room defined by the Ministry of Health, Labor and Welfare of Japan, considering use in a nursing home. Fig. 7(a) and Fig. 7(b) show the top and side views of the experimental environment. The height of the sensor was 0.93 m and the angle of inclination was 20 deg so that the sensor measure the head that is 172 cm for mean of Japanese at the farthest point⁽¹⁴⁾. The brightness of the environment was 335 lx with fluorescent light, and the aperture of the sensor lens was set to f/3.125 to adjust the brightness to the environment. The subjects were five adult males of different heights, physiques and clothing. Each subject took 100 standing state as upright and 20 falling state as lying on the ground at random locations. Thus as shown Table 2, the total data was 500 standing state and 100 falling state. Moreover, the data was divided into 90% training data and 10% test data. In order to train the model, data of standing state was prepared 450 data to train the model and 50 data to test. In addition, 100 test data of the falling state were prepared to evaluate fall detection. For learning, the number of epochs was 1000 and the mini-batch size to 256. Anomaly scores of RSS were calculated from the trained model with test data. In the evaluation, the

falling state was defined as positive and the standing state as negative. Besides, the data of falling state is classified as positive, it called the true positive. the data of standing state is classified as negative, it called the true negative. the data of standing state is classified as positive, it called the false positive. Finally, the data of falling state is classified as negative, it called the false negative. To avoid overlooking dangerous situations, a threshold value of 0.42 was used, which is the smallest value for false negatives. Furthermore, the detection is evaluated by these functions, such as, accuracy, precision, recall, and F-score⁽¹⁵⁾.

4.2. Results and Discussion

Fig. 8(a) shows one example of the test data in the standing state and reconstructed output data, and Fig. 8(b) shows one example in the falling state and output data. In the standing

state, the output data almost follows the input test data and the reconstruction is correct. In this case, the anomaly score is 0.15. Next, in the falling state, the output data and the input data are far apart, and reconstruction has failed. Therefore, the anomaly score at this case is 8.99, which is much higher than the threshold value of 0.42. Fig. 9(a) shows the distribution of the anomaly score in the standing state, and Fig. 9(b) shows the distribution in the falling state. In addition, the vertical dashed line is the threshold of 0.42. It shows that the test data are most correctly classified in the standing state, and there are no cases below the threshold in the falling state, thus all the test data are correctly classified. Table 3 shows the results of the detection using the test data. Moreover, the evaluation function calculated from the results is shown in Table 4. The results of the detection by SVM proposed in the previous study were accuracy is 96.6%, precision is 93.5%, and F-score is 96.7%. In comparison, accuracy was 98.7%,

Fig. 7. Experimental environment

Table 2. Breakdown of data used to training

State	Training Data	Test Data	Total
Standing	450	50	500
Falling		100	100

(a) Data of standing person

(b) Data of falling person

Fig. 8. Exemplary test data and output
]of proposed method

(a) Score of standing person

(b) Score of falling person

Fig. 9. Histogram of anomaly score

Table 3. Classification results

	State	Output	
		Standing	Falling
Target	Standing	48	2
	Falling	0	100

Table 4. Performance

Accuracy [%]	Precision [%]	Recall [%]	F1-Score [%]
98.7	98.0	100.0	99.0

it is 2.1 points higher than previous, precision was 98.0%, it is 4.5 points higher, and the F-score was 99.5%, it is 2.8 points higher. Therefore, the classification is more accurate than in previous studies, even though only standing data were used. On the other hands, there is a disadvantage which is difficult for improving model than SVM. The disadvantage is because in SVM tune some parameter but in Autoencoder do not have such parameter. However, the method has advantage not only that is higher precision but also does not generate false positives. It is because the threshold is defined unlike SVM to the recall is 100%. For these reason, the method is suitable for fall detection.

5. Conclusions

This paper proposes a method of abnormal detection such as falling using classifier which is trained by only normal data. In this method, the abnormal is detected the error between the input data and reconstructed data by AutoEncoder. Experimental results show that the proposed method has a high accuracy and is more accurate than the previous method. By using the proposed method, it is easier to introduce the system to watch over danger than the previous study, and the system becomes easier to use in the real environment.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP21K12797.

References

- (1) United Nations : “World Population Prospects 2019: Highlights”, Department of Economic and Social Affairs, 2019
- (2) S. P. Baker, and A. H. Harvey : “Fall injuries in the elderly”, *Clinics in Geriatric Medicine*, Vol. 1, Issue 3, pp. 501-512, 1985
- (3) H. Seki and Y. Hori : “Detection of Abnormal Action Using Image Sequence for Monitoring System of Aged People”, *IEEJ Transactions on Industry Applications*, Vol.122, No.2, pp. 182-188, 2002
- (4) G. C. Chen, C. N. Huang, C. Y. Chiang, C. J. Hsieh, C. T. Chan : “A Reliable Fall Detection System Based on Wearable Sensor and Signal Magnitude Area for Elderly Residents”, *International Conference on Smart Homes and Health Telematics*. Springer, pp. 267-270, 2010
- (5) L. Klack, C. Möllering, M. Ziefle, and T. Schmitz-Rode: “Future Care Floor: A Sensitive Floor for Movement Monitoring and Fall Detection in Home Environments”, *International Conference on Wireless Mobile Communication and Healthcare*. Springer, pp. 211-218, 2010
- (6) T. Hori, Y. Nishida, H. Aizawa, S. Murakami, H. Mizoguchi : “Sensor network for supporting elderly care home”, *SENSORS*, IEEE., pp. 575-578, 2014
- (7) T. Sugihara, T. Fujinami, and K. Nakagawa. : “Effects of Cameras and Monitors on Caregivers' Work Stress in the Group Home”, *IEICE Technical Report; IEICE Tech. Rep*, Vol. 107, No. 555, pp. 57-62, 2008
- (8) S. Nakashima, S. Mu, S. Okabe, K. Tanaka, Y. Wakasa, Y. Kitazono, and S. Serikawa : “Restroom Human Detection Using One-Dimensional Brightness Distribution Sensor”, *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing 2012, Studies in Computational Intelligence* (Springer), Vol. 492, pp. 1–10, 2013
- (9) Y. Horikawa, D. Hamasuna, A. Matsubara, and S. Nakashima : “Subject posture recognition by Support

Vector Machine Using Obrid-Sensor”, International Symposium on Computer, Consumer and Control, pp. 553-556, 2020

- (10)G. E. Hinton, R. R. Salakhutdinov : “Reducing the Dimensionality of Data with Neural Networks”, Science, Vol. 313, Issue 5786, pp. 504-507, 2006
- (11)J. K. Chow, Z. Su, J. Wu, P. S. Tan, X. Mao, Y. H. Wang : “Anomaly detection of defects on concrete structures with the convolutional autoencoder”, Advanced Engineering Informatics, Vol. 45, pp. 101-105, 2020
- (12), E. Komagal and A. Vinodhini: “real time background subtraction techniques for detection of moving objects in video surveillance system”, in 2012 International Conference on Computing, Communication and Applications, pp. 1-5, 2012
- (13)K, Simonyan, A, Zisserman : “Very Deep Convolutional Networks for Large-Scale Image Recognition”, eprint arXiv:1409.1556, 2014
- (14)Statistics Bureau of Japan : “Physique of Youths by Age”, Japan Statistical Yearbook 2015, p. 676, 2015
- (15)A. Baratloo, M. Hossein, A. Negida, G. E. Ashal : “Simple Definition and Calculation of Accuracy, Sensitivity and Specificity” Annals of Emergency Medicine3, pp. 48-49, 2015