Learning Behavior Recognition and Analysis by Using 3D Convolutional Neural Networks

Rui Zhang[1,2] and Baorong Ni[1]

[1] Graduate School of Engineering, Fukuoka Institute of Technology,3-30-1, Wajiro-higashi, Higashi-ku, Fukuoka 811-0295, Japan
[2] School of Optical Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 20094, China

***Index Terms*— Learning behavior, Human behavior recognition, 3D convolutional neural networks**

1. Introduction

The brain has two modes of thinking: the focus mode and the diffuse mode[1]. Under the focus mode, we pay the most attention and have the highest learning efficiency. Nowadays people's attention is distracted by too many things. In order to deal with the problem, a time management method--the pomodoro technique was proposed. However, since work time and rest time are malleable, everyone is different. This paper focuses on the recognition and analysis of individual learning behavior. By comparing learning efficiency, the best tomato clock can be obtained. As a preliminary study, kth-dataset which has certain similarities with learning behavior was used to investigate the feasibility of 3D CNN and confirm the effectiveness of this method.

2. System by using 3D CNN

On the basis of human behavior recognition technology. Because 3D convolution neural network(CNN) can better capture the time and space feature information in video, 3D convolution is carried out in CNN convolution layer to capture the distinguishing features in spatial and temporal dimensions[2]. In this study, an individual learning behavior recognition and analysis system based on 3D CNN was designed.

In tensorflow platform, a 3D CNN framework was built with keras, including convolutional layer, pooling layer and full connection layer. Due to the current lack of learning behavior database, an existing kth-dataset was used to test my 3D CNN framework which consists of six types of human actions: walking, jogging, running, boxing, hand waving and hand clapping in four different scenarios as shown in Fig.1. The library of openCV was used to segment the target video, and to extract the main behavioral characteristics.

3. Results and discussion

For each behavior that needs training, its long-term behavior information was extracted as its high-level behavior characteristics. The motion information is long enough to contain much more information than CNN's cube of input frames. By adjusting the structural parameters for testing, the accuracy of the structure was improved as shown in Fig.2. In this study, different batch sizes and epochs were used to train our structure. The best train accuracy was 83.51% and the best test accuracy was 69.17% when the batch size was 4 and the epoch was 100.

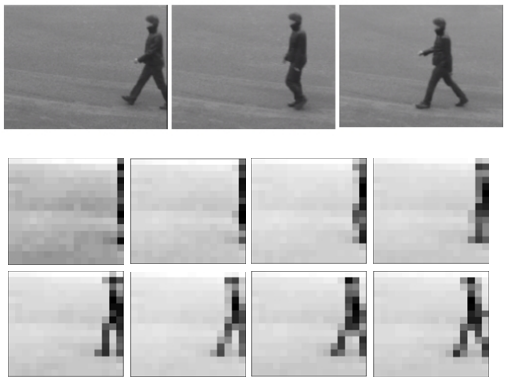

Fig. 1. Walking behavior segmentation.

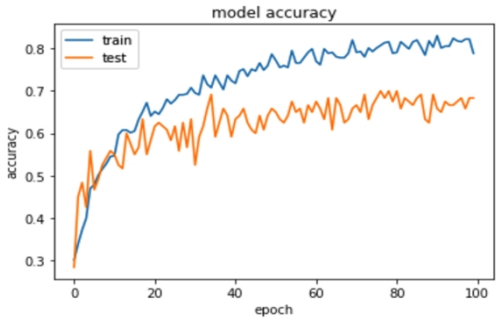

Fig. 2. The model accuracy

4. Conclusion

Under this structure, the training accuracy was up to 83.51%, but the test accuracy was only 69.17%. In the future, a more complete database of learning behavior will be established, including learning, reading, experimenting, napping, chatting and playing. In addition, in the following studies, the number of structural layers should be appropriately increased and parameters should be adjusted to optimize the results. Through the monitoring everybody's daily time management, recognition and analysis of these behaviors, we think that the most suitable time management which leads to an improvement of learning efficiency can be specified individually.

REFERENCES

[1] Barbara Oakley.A Mind for Numbers[M].Createspace Independent Pub:America,2014.
[2] Ji S , Xu W , Yang M , et al. 3D Convolutional Neural Networks for Human Action Recognition[C]// International Conference on Machine Learning. DBLP, 2010.