Title	A study on Anomaly Detection in Surveillance videos
Author(s)	顧, 超逸
Citation	
Issue Date	2020-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/16395
Rights	
Description	Supervisor: 丁 洛榮, 先端科学技術研究科, 修士(情報科学)

A study on Anomaly Detection in Surveillance videos

1810061 Gu Chaoyi

Anomaly detection is a task to detect abnormal and normal actions of people in terms of surveillance videos. Anomaly detection could play an important role in different areas. For example, it could release the problems of lacking the labor force in nursing/daycare facilities since it could detect abnormal actions of the elderly people and/or children to keep their securities and comfort. It could also detect abnormal actions of people in public space to keep public space safe. If abnormal actions are detected, the enforcement agencies will be informed. All of the work can be finished by the anomaly detection system, so lots of human force will be saved. Therefore, the research of anomaly detection is essential, necessary and promising.

The concept of anomaly detection has been proposed in the last century. Due to the limitations of technologies of computer science and sensor, the development of anomaly detection is slow until entering 21 century. Because of the development of science, especially the rapid development of computer science, anomaly detection develops rapidly. More and more attention has been drawn by researchers to the anomaly detection field, and the obtained achievements are remarkable. However, it is inescapable that researchers must define abnormal and normal actions regardless of methods used either traditional or based on machine learning. That is a subjective task as the boundary between abnormal and normal actions are not clearly defined, and it is difficult to define the boundary between them. Another limitation is data labeling. Labeling data is a task that requires lots of human effort, especially when supervised learning is applied to detect abnormal actions. Since the input data consists of image frames, all the frames have to be labeled before training the model. In order to overcome these two main limitations, we propose to apply the Multiple Instance Learning (MIL) for anomaly detection. There are two merits in MIL; Firstly, MIL is a category of weakly supervised learning. The input of MIL is a video-level label instead of a frame-level label. That would save lots of human labor. Second, it is unnecessary for human experts to define the boundary between abnormal and normal actions. Our input for MIL is a video. There is no need to label the start-time and end-time of abnormal actions. Thus, the computer learns to define the boundary between actions by itself.

Moreover, the main contributions of this research are proposing a new model based on a baseline model (deep MIL ranking model) and improving the performance of the baseline model. Before changing the inner structure of the baseline model, we optimized the parameter settings of the Fully Connected Neural Network (FCNN) at the end of the baseline model in order

to obtain a better performance. After optimizing the parameter setting of FCNN, we apply Bi-directional Long Short-Term Memory (Bi-LSTM) model between pre-trained C3D model and FCNN in the baseline model to improve the performance further. In order to avoid overfitting, we optimized parameter settings of the Bi-directional LSTM module and provided the best performance in all of the models tested in this thesis.

In order to evaluate performance, ROC, AUC, F1-Measure, and Recall are used. Comparing to the baseline model, experimental results show that our model could improve AUC from 73% to 79%. F1-Measure increases from 9.1%. Recall is improved from 0.55 to 0.665. The main reason leading to this performance improvement is that since the temporal features between adjacent video segments provide valuable information for classification in FCNN, a Bi-LSTM could extract temporal features from adjacent video segments in more efficient than LSTM. Thus, our model performs better.