歩行者の姿勢ヒートマップを用いた
監視カメラ映像からの異常行動検出に関する初期検討

Temuroglu Onur*、川西 康友、出口 大輔、平山 高嗣、井手 一郎、村瀬 洋(名古屋大学),
岩崎 麻友、塚田 篤(住友電気工業株式会社)

A Preliminary Study on Pedestrian Abnormal Behavior Recognition based on Pose Heatmap from Surveillance Camera Images

Onur Temuroglu, Yasutomo Kawanishi, Daisuke Deguchi, Hirayama Takatsugu, Ichiro Ide, Hiroshi Murase (Nagoya University),
Mayuu Iwasaki, Atsushi Tsukada (Sumitomo Electric Industries, Ltd.)

1. Introduction

Surveillance cameras are expected to become an essential part of ITS, as they can be used to capture and relay relevant information to drivers almost instantaneously through Infrastructure-to-Vehicle communication. Their elevated viewpoint makes it possible to detect anomalies that may not be visible to in-vehicle cameras, such as abnormally behaving pedestrians. In this research, we are tackling the problem of detecting abnormally behaving pedestrians, especially drunk walking ones, by treating them as anomalies.

AutoEncoder variants are a popular way to tackle the problem of anomaly detection. It extracts the features of input, and then reconstructs the input data from said features. Anomaly detection with AutoEncoder works by training the network on normal data which restricts it in a way that it will not be able to reconstruct abnormal data with high accuracy on test time. Hence, by comparing the reconstruction result to the original input and calculating the difference between the two, we can classify the data as normal or abnormal.

There are multiple works tackling the problem of pedestrian abnormal behavior detection. However, most of these works use sequences of images directly, which contain a variety of features that are unnecessary and possibly even detrimental to the detection of abnormal behavior. As such, in this paper, we propose using sequences of pedestrian pose heatmaps instead for abnormal behavior detection leading to more accurate classification.

2. Anomaly Detection Using Pose Heatmap Sequence

In this paper, we use OpenPose [1] to extract pedestrian pose. Since the outputs are easily affected by occlusions, instead of using the final keypoint output we use heatmap for our input type as they are more robust to data loss. Even in the case of keypoints being missing from the final output, heatmap of probability of body parts existing will relatively keep the keypoint information.

To use heatmaps as the input for AutoEncoder, we pass the output for each frame through a few preprocessing steps. First, we remove keypoints deemed unnecessary, resulting in 12 out of 25 available keypoints being used. Then, we resize and pad the frames to 96×96 pixels and use them to generate 30 frame long sequence.

For detection, we use the output of an AutoEncoder and calculate the reconstruction error as the indicator of anomalous behavior.

Table 1. Balanced accuracy of classification results

Method	20 Epoch	100 Epoch
Proposed (Pose)	**0.891**	0.878
Comparison (Image)	0.589	0.578

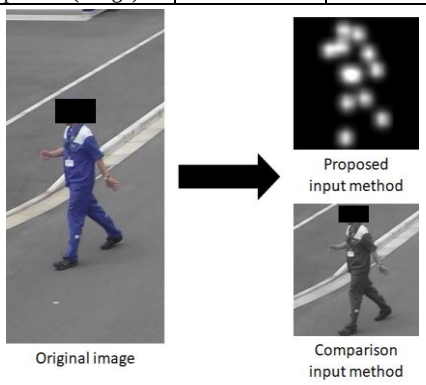

Figure 1. Example input frames

3. Experiment

We conducted an experiment to confirm the effectiveness of using sequences of pose heatmaps in abnormal behavior detection. The data we used for the experiment consists of 4,124 normal sequences for training, 1,147 normal sequences for validation, and 2,952 sequences for testing with 1,952 of them normal and the remaining 1,000 abnormal. We also prepared sequences of actual images of pedestrians cropped from surveillance camera images. We trained two instances of the same AutoEncoder network for 100 epochs with each type of data and compared the results.

We used sum of absolute error to calculate reconstruction difference, and classified the data on 100 different threshold intervals, taking the best result for each.

4. Conclusion

In this paper, we proposed a method that uses sequences of pose heatmaps to determine if an input pedestrian movement sequence is abnormal or not. We conducted an experiment and confirmed the effectiveness of the proposed method.

Acknowledgment: Parts of this research were supported by MEXT, Grants-in-Aid for Scientific Research.

References

(1) Z. Cao, et al., Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition, pp.1302-1310, July 2017.