# Traffic Accident Detection and Classification in Videos based on Deep Network Features

Zain Ul Arifeen, Jang-Eui Hong, Bo-Seok Seo, Jae-Won Suh
*College of Electrical and Computer Engineering*
*Chungbuk National University*
Cheongju, Korea
zain@cbnu.ac.kr, jehong@cbnu.ac.kr, boseok@cbnu.ac.kr, sjwon@cbnu.ac.kr

***Abstract*—The number of vehicles on the road has increased significantly, which pose a number of challenges to cope with for traffic management. Especially road accidents need instant attention to reduce the loss of people life and their property. In this paper, we propose a traffic accident detection and classification framework, which automatically detects accident in traffic videos using deep networks features and also classify that accident into car-car and car-bike collisions. The proposed framework works in two phases. Accident anomaly detection: We explore three convolution neural networks (CNN's) named GoogLeNet, AlexNet and VGGNet, where deep features are extracted using these networks and a one class support vector machine (OCSVM) is trained on each network deep features, which are used to detect accident anomalies in a outlier fashion. Accident anomaly classification: where a multi-class SVM model is trained using the features of the detected accident frames and is used to classify accident into car-car and car-bike collisions. The experimental results on UCF-Crime road accident video sequences show that the proposed approach achieves high accuracy on both traffic accident anomaly detection and classification.**

***Index Terms*—GoogLeNet, AlexNet, VGGNet, Accident detection, Accident classification, One class svm, Multi-class svm**

## I. INTRODUCTION

Recently efforts have been put in installing the smart city projects across the globe. The smart city projects are meant to serve a more secure sense of life to the people [1]. In the so called smart cities, a major challenge is the information management concerning the transportation department. Due to increase in population and economic progress the number of vehicles on the roads are increasing rapidly. Consequently there is an increased burden on road traffic management in handling the road accidents, providing instant attention to the emergency events concerning the life and property of people. According to WHO (world health organization) among the people who lose their life's every year in different events (approximately 1.35 million people die in road traffic crashes) a significant amount is killed world wide in road traffic accidents and these accidents cost most countries 3 percent of their GDP [2]. However in order to process the road traffic information to identify an abnormal event we have to rely on the human observation. For a human it is quite impossible to observe and recognize all abnormal events with out missing any one of it in such a huge amount of surveillance cameras video scenes. In order to automate the traffic monitoring, we need to develop automatic methods for accident detection.

Over the past few years the research community has been focusing on development of the automated traffic anomaly detection methods using the computer vision and pattern recognition techniques [3]–[5]. However, the presented methods suffer from limitations dealing with real world scenarios [6], [7]. Computer vision based traffic anomaly detection is a difficult and challenging problem in many aspects, which include: variations in light conditions, poor quality of videos, weather conditions, variety of scenes (like highway, market place) etc. Recently deep learning techniques have been explored in many computer vision tasks like image classification [8], segmentation [9] etc. The performance of deep learning in representing the visual data is becoming a success because of the rich features which they extract using the convolution layers. This paper is organized as follows. The proposed method for accident detection and classification is described in section II. We describe the dataset used and the experimental results in section III. A conclusion is drawn in section IV.

## II. PROPOSED APPROACH

In this section the proposed framework is elaborated as shown in "Fig. 1". We describe how our proposed framework recognized vehicle accident. Our main goal is to provide simple but yet prompt model for solving the challenge of traffic accident detection which work effective and efficient and deliver the important data to the concerned authority in quick time. The key tasks includes in the proposed framework is as follow:

- Features extraction
- Accident detection
- Accident classification

### *A. Feature extraction*

We use three pre-trained deep networks called as GoogLeNet, AlexNet and VGGNet to extract features of the video frames but the features extracted through GoogLeNet network and then trained on OCSVM outperforms the other two networks.The process of feature extraction for all three networks is same. We will explain the process of GoogLeNet as feature extractor: Note that we only use the convolution

This work was supported by the National Research Foundation of Korea grant funded by the Korea government (MSIT) (No. 2022R1A5A8026986).

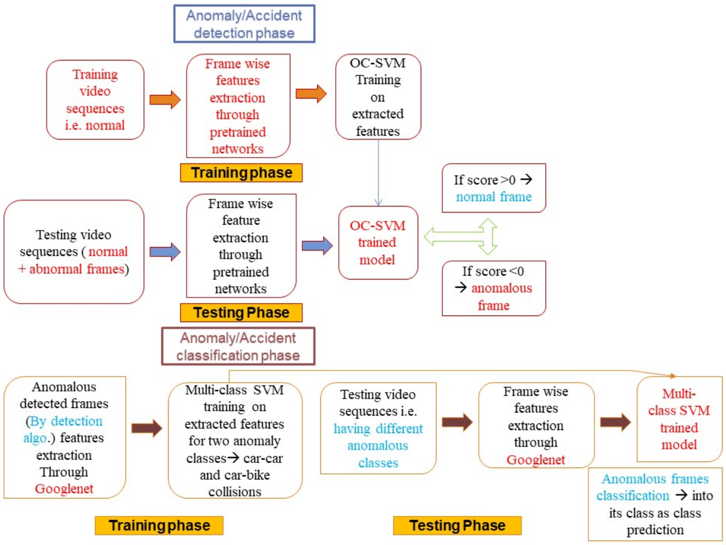


Fig. 1. Proposed accident detection and classification framework.

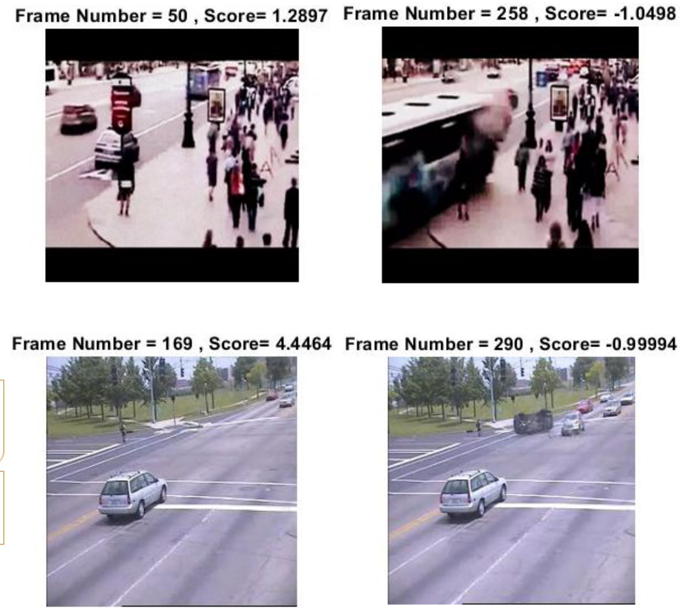


Fig. 2. Left side figures show normal frames (no anomaly) with the positive scores provided by the trained model. Right side figures show anomalous frames (an anomaly occurs in the form of an accident) and the model provides negative scores.

layers of the network to extract features. Where these features are the output of the activations function in the last pooling layer of the GoogLeNet (pool5-drop-7 $\times$ 7-s1). The size of the extracted features for each frame is $1 \times 1024$.

First the training video sequences are re-sized to match the input size of the network i.e. $224 \times 224 \times 3$, then the features of training sequences at the frame level are extracted from the last pooling layer (as mentioned above). We arrange the extracted features as an array of feature vectors where each row gives the features of a video frame. Note that all these features are extracted using the video frames of normal class videos (i.e. the frames having no traffic anomaly).The extracted features are used to train a classifier model. The trained model is then used to classify a traffic video frame as anomalous or non-anomalous. In this work, we use a variant of the support vector machine (SVM) presented in literature as one class SVM (OCSVM).

### B. Accident detection

The extracted features of the video frames from the normal traffic sequences using the pre-trained networks are used to train the OCSVM classifier with labels. The OCSVM learns a decision boundary based on these normal class features. During testing phase the extracted features from the video frames of the test sequences are given to the trained OCSVM classifier, which decides about the test frame as anomalous or normal. The OCSVM works in an outlier detection fashion by calculating the anomaly score of the frame based on the features. If the score of the frame falls in normal class category it is classified as normal and if it falls in abnormal class then it is classified as abnormal (which is an indication that an accident occurs) as depicted in “Fig. 2”. Note: After experiment we observe that features extracted through GoogLeNet network and then trained on OCSVM outperforms the other two networks.

### C. Accident classification

For the frames, which are classified as anomalous we further perform classification into two categories i.e. car-car or car-bike collision. In order to achieve the anomaly classification, at first we extract deep features of each category training video frames using GoogLeNet because of the good performance at accident detection phase. A multi-class support vector machine is trained on the extracted features with category labels. The SVM learns the decision boundary based on the features and the provided labels. At the testing stage, based on the features extracted from the testing video frames, the trained model provides category labels classifying an anomalous frame into one of the categories. We show example video frames in “Fig. 3” which are assigned the correct category.

## III. Experimental Evaluation

Experiments were performed on Intel(R) Core(TM) i7 CPU @ 1.80GHz with 8GB RAM. The programs were written in Matlab 2019b.

### A. Dataset Used

We evaluate the performance of the proposed accident detection and classification framework using the UCF-Crime dataset, which contains real-world anomalies. UCF-Crime dataset is a publicly available benchmark for anomaly detection collected by UCF (University of Central Florida) center for research in computer vision which covers 13 real world anomalies [10]. Among the 13 different types of anomalies in


Fig. 3. Anomalous frames classification by the proposed approach into car-car collision (Label 1) and bike-car collision (Label -1).

the dataset, we perform evaluation of the proposed approach on road accidents anomalous videos. There is total 150 videos containing road accident anomalies. The videos contain traffic accidents involving vehicles, pedestrians or cyclists.

### *B. Results*

In this section we present the performance of the proposed approach using qualitative and quantitative ways. We provide the qualitative performance evaluation of the proposed approach in the form of visual results shown in “Fig. 2” and “Fig. 3”. The quantitative evaluation for accident detection is done through a common evaluation metric of anomaly detection task called Receiver operating curve (ROC) and then the performance in terms of the ROC is summarized by area under the curve (AUC).

In “Table. I”, we provide the frame level accident detection AUC from resulting ROC’s for UCF-Crime road accidents videos. The quantitative anomaly classification results obtained by the proposed approach are provided in “Fig. 4” in the form of a confusion matrix.

## IV. Conclusion

In this research, we present a traffic accident detection system which automatically detect accidents in traffic surveillance videos and also classify that accident into a class among car-car or car-bike collision. The proposed approach consists of two phases. In first phase three deep networks named as GoogLeNet, AlexNet, VGGNet and are used to extract features from the training videos frames. These video frames are of normal class and a one class SVM is trained on normal class features, which is used to classify the testing video frames into normal (no accident) and anomalous (accident occurs) class. In second phase a multi class SVM is trained on the features of anomalous frames from classes of car-car or car-bike anomalies. The trained SVM model is found to classify the anomalous accident frames with high accuracy. The experiments are performed on UCF-Crime accident videos. The proposed approach is found capable of classifying the normal and accident frames then classifying these accident frames into their respective classes.

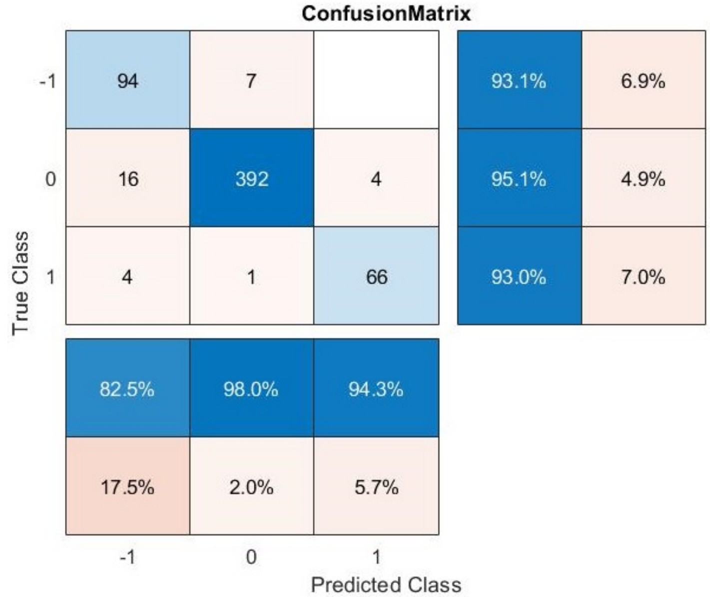


Fig. 4. The label 1 denotes car-car collision, label -1 denotes the bike-car collision and the label 0 denotes the normal frames.

TABLE I
Frame level AUC (%) performance comparison based on used pre-trained network features .

| Network features | AUC on UCF-Crime dataset |
|---|---|
| AlexNet | 80.57 |
| VGGNet | 89 |
| GoogLeNet | **97.27** |

## References

[1] V. Kostakos, T. Ojala, and T. Juntunen, “Traffic in the smart city: Exploring city-wide sensing for traffic control center augmentation,” *IEEE Internet Computing*, vol. 17, no. 6, pp. 22–29, 2013.
[2] https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.
[3] W. Sultani, C. Chen, and M. Shah, “Real-world anomaly detection in surveillance videos,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6479–6488.
[4] S. Bai, Z. He, Y. Lei, W. Wu, C. Zhu, M. Sun, and J. Yan, “Traffic anomaly detection via perspective map based on spatial-temporal information matrix.” in *CVPR Workshops*, 2019, pp. 117–124.
[5] J. Zhao, Z. Yi, S. Pan, Y. Zhao, Z. Zhao, F. Su, and B. Zhuang, “Unsupervised traffic anomaly detection using trajectories.” in *CVPR Workshops*, 2019, pp. 133–140.
[6] F. Mehboob, M. Abbas, R. Jiang, A. Rauf, S. A. Khan, and S. Rehman, “Trajectory based vehicle counting and anomalous event visualization in smart cities,” *Cluster Computing*, vol. 21, no. 1, pp. 443–452, 2018.
[7] Z. Hui, X. Yaohua, M. Lu, and F. Jiansheng, “Vision-based real-time traffic accident detection,” in *Proceeding of the 11th World Congress on Intelligent Control and Automation*. IEEE, 2014, pp. 1035–1038.
[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
[9] S. Hong, H. Noh, and B. Han, “Decoupled deep neural network for semi-supervised semantic segmentation,” in *Advances in neural information processing systems*, 2015, pp. 1495–1503.
[10] UCF-Crime dataset at: https://visionlab.uncc.edu/download/summary/60-data/477-ucf-anomaly-detection-dataset.