Category-based Memory Bank Design for Traffic Surveillance in Context R-CNN

M. Takahashi, K. Iino, H. Watanabe
Waseda University
I. Morinaga, S. Enomoto, X. Shi, A. Sakamoto, T. Eda
NTT Software Innovation Center

***Abstract*— Traffic surveillance cameras recognize a large number of objects. In this configuration, since vehicles appear in the same lanes, past footage is useful information. Context R-CNN has been proposed to store past videos in a memory bank and use them for recognition. In this paper, we propose a method to improve recognition performance by selecting objects to be stored in memory banks. We found that the proposed method can improve mAP by 0.37 points on average compared to the conventional method.**

I. Introduction

Object detection in traffic surveillance is an important task for measuring and tracking the number of vehicles and improving the accuracy of anomaly detection. In addition, detection in fixed cameras can be difficult because objects are obscured by weather or lighting conditions, or objects are only partially visible in the image.

Context R-CNN [1] improves detection performance by storing context information of nearby images from the same camera in memory banks supplemented with attention block. The strategy puts emphasis on long-term context and stores the highest scoring context in each frame from all frames from the same camera within a pre-defined time period. However, in traffic data, many objects appear in the image and similar objects exist across multiple frames, so this strategy leads to storing only similar context information in the memory bank, which makes it difficult to store context information about objects that appear infrequently.

In this work, we focus on the number of objects for each category stored in memory banks. The same number of context information is stored for all objects.

II. Proposed Method

We employ a method of selecting the contexts to be stored in the memory bank not by frame, but by object category. Figure 1 shows an example of the difference between the conventional and proposed methods in retrieving context information from the same frame.

After collecting all object context information from all frames of the same camera within a predefined time period, we store the same number of contexts in all categories, in order of the highest score for each object category. By obtaining context information for all categories, the contexts of objects that appear less frequently can also be stored in the memory bank.

III. Experimental Result

The results of storing the same number of contexts using the CityCam dataset [2] with the conventional and proposed methods are shown in Figure 2. The proposed method can improve mAP by 0.37 points on average, and some categories improved performance by up to 2.17% mAP due to the number of stored contexts being approximately double of that of the conventional method.

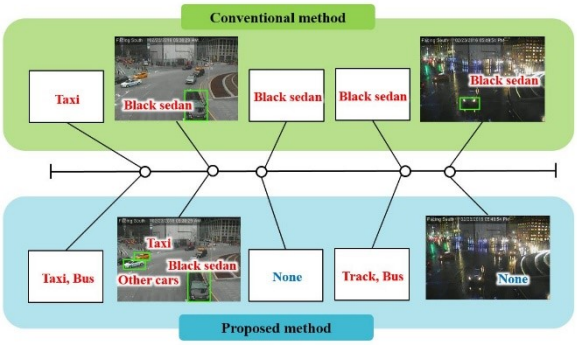

Fig. 1. Difference in context information stored in memory banks between the conventional method and the proposed method.

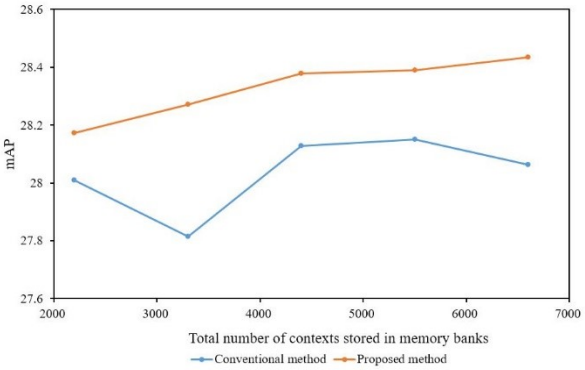

Fig. 2. Results of mAP for the conventional and proposed methods for each total number of contexts stored in memory banks.

One of the reasons that existing methods reduce detection accuracy is that they store false positive context information in memory banks. However, the proposed method can store multiple contexts that can be detected from clearly visible objects, so there is less risk of storing false context information.

IV. Conclusion

In this work, we proposed a method for storing contexts in memory banks by category. The proposed method can utilize memory banks for all categories regardless of the number of objects in a frame or the number of occurrences per category, thus improving detection performance.

References

[1] S. Beery, G. Wu, V. Rathod, R. Votel, and J. Huang, "Context r-cnn: Long term temporal context for per-camera object detection," CVPR, pp. 13072–13082, Jun. 2020.

[2] S. Zhang, G. Wu, J. P. Costeira, and J. M. Moura, "Understanding Traffic Density from Large-Scale Web Camera Data," CVPR, pp. 4264-4273, Jul. 2017.