Traffic Flow Detection

Xing Chen, Sicong Chen, Xuemeng Zhao, Yufan Wang

Abstract

In the rapidly growing fields of urban development and traffic management, accurate traffic flow analysis has become the cornerstone of effective planning and decision-making. As urban traffic becomes increasingly complex, traditional vehicle counting and flow analysis methods are becoming increasingly insufficient. The emergence of deep learning and computer vision offers a promising alternative. Leveraging these technologies, especially object detection algorithms, has the potential to revolutionize the way we understand and manage traffic flow.

This project aims to use object detection models to accurately count and analyze traffic flow. In doing so, it attempts to address the urgent need for more reliable, efficient and scalable traffic analysis methods in urban environments. Integrating object detection into traffic management systems not only improves the accuracy of traffic statistics, but also provides real-time data that is crucial for immediate decision-making and long-term urban planning strategies.

1 Background

1.1 Object Detection

Object detection [1] is a fundamental task in computer vision that involves identifying and locating multiple objects within an image or video. The primary goal is to not only classify the objects present but also determine their precise locations by drawing bounding boxes. This task is crucial in various applications [2, 3, 4] across industries, including autonomous vehicles, surveillance, robotics, healthcare, retail and more.

These days, the rapid development of deep learning techniques has greatly promoted the progress of object detection. Generally speaking, the progress of object detection has gone through two historical periods: "traditional object detection period" and "deep learning based detection period". In "deep learning based detection period", many models have significantly evolved, such as RCNN [5], SPPNet [6], Faster RCNN [7], Faster RCNN [8], FPN(Featured Pyramid Networks) [9], YOLO(You Only Look Once) [10].

1.2 Traffic Flow Detection

Traffic flow detection based on object detection involves using computer vision techniques to analyze traffic scenes, identify vehicles, and estimate various parameters related to traffic flow. This approach leverages object detection models to track and monitor vehicles within a road environment, providing valuable insights [11, 12] for traffic management, transportation planning, and intelligent transportation systems.

In conclusion, our project has three main purposes:
1. Apply YOLOv8 on traffic flow detection.
2. Compared the results of proposed model with some baseline models (eg. RCNN, Faster RCNN, YOLOv3).
3. Modified YOLOv8 to improve its performance on detection.

2 Proposed Experimentation/Implementation

2.1 Dataset

In this project, we plan to use UA-DETRAC dataset for training and testing. UA-DETRAC includes 100 challenging videos of real-world traffic scenes with various information in PASCAL VOC format, such as bounding box, trajectory, speed, weather, etc.[13] Vehicles are classified into Car, Van, Bus and Others.

2.2 Method

We will adopt YOLOv8 to the above challenging dataset. YOLOv8 is a SOTA model which is upgraded from previous YOLO versions, supporting a wide range of AI tasks such as object detection. As a result, we will use a pretrained model YOLOv8n or YOLOv8m to train on the UA-DETRAC, and apply our model to videos. Then we will use algorithms such as DeepSort to calculate the traffic flow of these videos. We found that it is hard for YOLO to detect small objects and overlapped objects, which will definitely affect the result of the traffic flow. Therefore, if time and resources are sufficient, we will try to modify the YOLOv8 model.

Possible modification:
1. Add attention, such as CoordAttention
2. Use InceptionNext or ConvNext

3 Feasibility and Limitation

First, despite the fact that UA-DETRAC is an awesome dataset for object detection, only three classes are included and are distributed unevenly. However, vehicle size plays an important role in traffic flow. For example, a truck will occupy more spaces than SUV, so the amount of vehicles we count does not necessarily indicates the road congestion. But as far as we know, other datasets are either too large or too monolithic, so currently, wo do not have a solution for this issue.

Additionally, if the structure of the YOLO model is changed, we might have to train the new model from scratch, which requires a huge amount of computing resources. If it turns out that training the new model is beyond our PC, we will use online services such as CoLab or AWS.

4 Potential Impact

From the most straightforward perspective, traffic flow detection can help optimize traffic management strategies, reducing the amount of time people spend stuck on the road. It can also serve as a foundation for the long-term development of smart cities.

Additionally, there are various application scenarios in other fields. For instance, our model can monitor the status of a company by detecting the traffic flow in front of their parking lot, which could be beneficial for investors. As America is a country on wheels, our model has a wide range of applicable scenarios.

References

[1] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: A survey," *Proceedings of the IEEE*, vol. 111, no. 3, pp. 257–276, 2023.

[2] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhutdinov, R. Zemel, and Y. Bengio, "Show, attend and tell: Neural image caption generation with visual attention," in *arXiv*, 2015.

[3] A. Karpathy and L. Fei-Fei, "Deep visual-semantic alignments for generating image descriptions," in *Computer Vision Pattern Recognition*, 2015.

[4] Q. Wu, C. Shen, P. Wang, A. Dick, and A. v. d. Hengel, "Image captioning and visual question answering based on attributes and external knowledge," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, pp. 1367–1381, 2018.

[5] J. R. R. Uijlings, K. E. A. V. D. Sande, T. Gevers, and A. W. M. Smeulders, "Selective search for object recognition," *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154–171, 2013.

[6] P. Msonda, S. A. Uymaz, and S. S. Karaaa, "Spatial pyramid pooling in deep convolutional networks for automatic tuberculosis diagnosis," *Traitement du Signal*, vol. 37, no. 6, pp. 1075–1084, 2020.

[7] R. Girshick, "Fast r-cnn," in *International Conference on Computer Vision*, 2015.

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[9] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," *IEEE Computer Society*, 2017.

[10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.

[11] T. N. Huy and B. H. Duc, "Traffic flow estimation using deep learning," in *2020 5th International Conference on Green Technology and Sustainable Development (GTSD)*, pp. 180–184, 2020.

[12] C. Chen, B. Liu, S. Wan, P. Qiao, and Q. Pei, "An edge traffic flow detection scheme based on deep learning in an intelligent transportation system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1840–1852, 2021.

[13] L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M.-H. Yang, and S. Lyu, "Ua-detrac: A new benchmark and protocol for multi-object detection and tracking," *Computer Vision and Image Understanding*, vol. 193, p. 102907, 2020.