

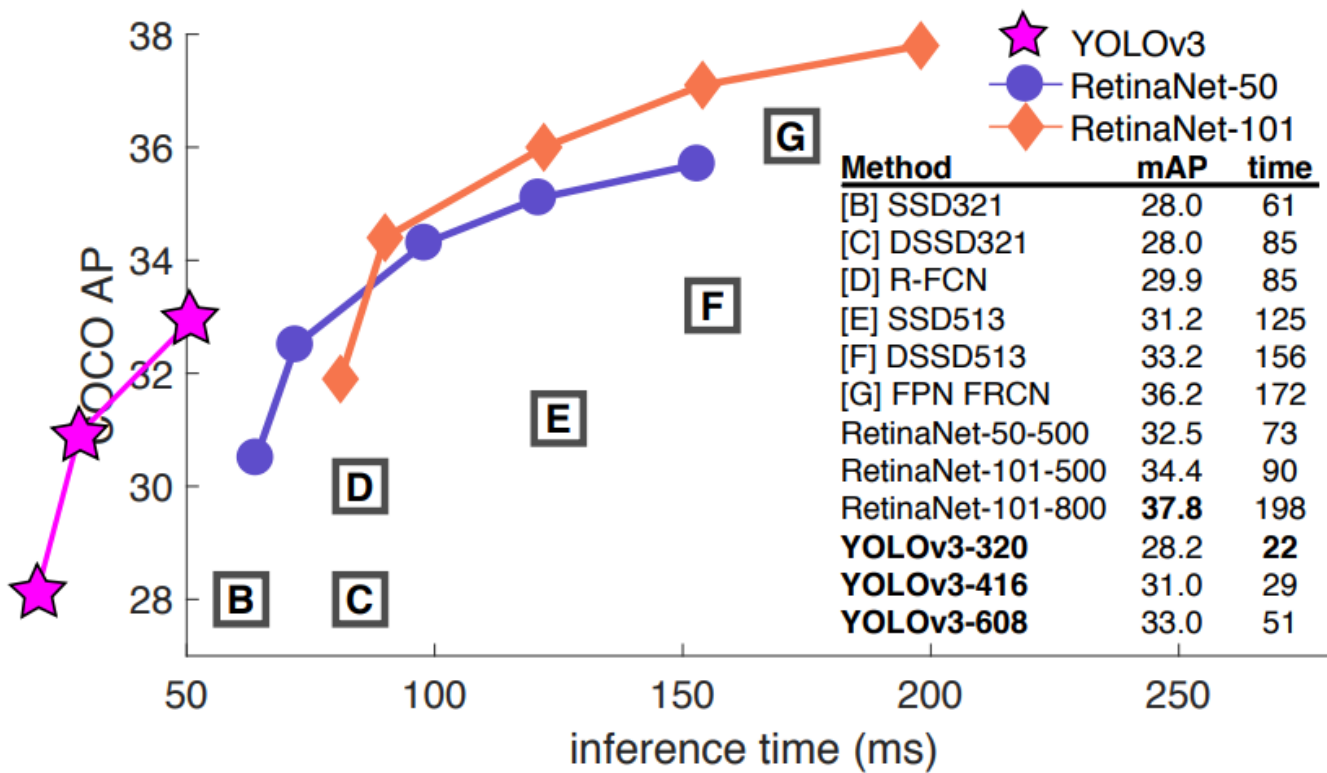
Abstract

Introduction

최근 자율 주행에 관한 연구가 활발히 일어나고 있다. 신호등 탐지 및 인식 기술은 자율 주행에 있어서 핵심 기술 중 하나이다. 신호등 탐지 및 인식 기술은 신호등의 신호를 인식하고 해당 차량의 주행 가능 여부를 판단하는 기술이다. 차량에 부착된 카메라를 통해 얻어낸 영상에서 신호등을 탐지하고 인식한다.

기존의 신호등 인식에 관련해서 신호등의 종류와 관계 없이 신호등 인식 혹은 신호등의 색상을 통해 어떤 신호 인지 인식하는 연구들만이 있었다. 한국에는 중앙버스전용차로가 있어 기존 차량들과의 신호 혼동을 줄이기 위해 일반적인 신호등이 아닌 버스모양 신호등이 있다. 기존의 연구들을 통해 신호등을 인식한다면 일반 차량이 버스 모양 신호등을 인식하고 주행하는 경우 혹은 중앙버스전용차로를 이용하는 버스의 경우에는 일반 신호등을 인식하고 주행하는 경우가 발생한다.

본 논문은 상위와 같은 문제를 해결하기 위해 딥 러닝 기술 적용하여 일반 신호등과 버스 모양 신호등을 구별하여 인식하는 모델을 제안한다.



신호등의 탐지 및 인식에는 영상 속 개체를 실시간으로 검출할 수 있을 정도로 빠르고 정확한 모델인 YOLO(You Only Look Once)v3 프레임워크를 기반으로 만들어진 모델들을 사용하며, 쓰인 각 모델들의 성능을 비교 분석한다.

Related work

객체 인식 분야에서 CNN(Convolutional Neural Network)을 기반으로 하는 인공지능망 네트워크 모델들이 발전되어 왔다. 대표적으로 AlexNet[5], GoogLeNet[6], R-CNN[7], Fast R-CNN[8], Faster R-CNN[9], SSD[10], YOLO[1, 2, 3]가 있다.

여러 모델 중에 YOLO 네트워크는 빠른 처리 속도와 높은 정확도를 가지고 있기에 실시간 영상을 처리하기에 적합하다.

YOLO는 객체 인식 과정에서 입력 이미지를 그리드로 나누어서 Bounding Box의 좌표와 어떤 Class가 어느 정도의 Confidence를 가지는지 한번에 확률적으로 추정하는 Regression 문제로 단순화시켜 수행 속도를 크게 높인 기술이다.

Method

일반 신호등과 버스 신호등의 검출 성능을 비교하기 위해 크게 세가지 모델을 사용했다. 첫째는 YOLOv3의 기본 모델로 75개의 Convolutional layer를 가진다. Downsampling을 Convolutional layer로 하는 특징을 가진다. 둘째로 YOLOv3-spp 모델은 COCO Dataset에서 YOLOv3 보다 더 좋은 성능을 보이는 모델로 76개의 Convolutional layer를 가진다. Convolutional layer와 Max-pooling layer에서 가장 좋은 feature를 얻어냄으로 downsampling을 하는 특징을 갖는다. 셋째로 YOLOv3-tiny 모델은 13개의 Convolutional layer를 가지며 Max-pooling layer를 통해서만 downsampling을 한다.





Cropping



학습 데이터로 두가지 종류의 데이터를 사용한다. 한 종류는 2288x1080 해상도를 가졌고 다른 하나는 그 이미지를 자른 1024x1024 해상도를 가졌다.

Result

Conclusion

References

- [1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
- [2] Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7263-7271).
- [3] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.
- [4] Huang, Z., & Wang, J. (2019). DC-SPP-YOLO: Dense Connection and Spatial Pyramid Pooling Based YOLO for Object Detection. arXiv preprint arXiv:1903.08589.
- [5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
- [6] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).
- [7] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).
- [8] Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

- [9] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems* (pp. 91-99).
- [10] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21-37). Springer, Cham.