

Pedestrian Detection in Autonomous Vehicle Using AI and Computer Vision

Author: Dinh Hoang (University of Central Arkansas)

Mentor: Dr. Brendan Tran Morris

Research Experiences for Undergraduates (REU) Summer 2021 at University of Nevada, Las Vegas

1. Background and Introduction

- During 2014 to 2018, a total of 387 fatalities and 376 fatal pedestrian crashes occurred on Nevada roadways [7]
- The hours of 6:00 PM and 11:59 PM had the greatest number of fatal pedestrian crashes [7]
- Autonomous vehicles with advanced safety technology can help reduce pedestrian deaths
- Autonomous vehicles must understand their environment through various sensors such as RGB & FLIR camera, LiDAR, Radar, etc.
- This work was focusing on computer vision & deep learning for environmental perception and pedestrian detection with the use of RGB & IR thermal camera (IR for better nighttime detection)



Fig 1. Autonomous vehicles must understand their environment

2. Methodology

- Used YOLO v5 (You Only Look One version 5) – a state of the art, fast and accurate model for real time object detection system
- Some advantages of using YOLOv5:
 - Easy installation
 - Fast training
 - Flexible inference media
 - Intuitive file layout
- Training YOLOv5 on FLIR thermal dataset consisting of >93k images and nulumages dataset consisting of >14k images
- FLIR dataset were collected by using Forward Looking InfraRed (FLIR) thermal camera while nulumages dataset were collected by using regular RGB camera
- Additionally used EfficientPS and Panoptic DeepLab – state-of-the-art algorithms for panoptic segmentation tasks

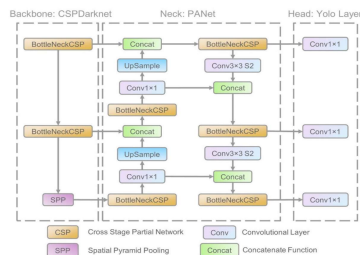


Fig 2. YOLOv5 Architecture

3. Inference Results



Fig 3. Inference using the pretrained YOLOv5 model on nulumages dataset with inference time of 54.55 FPS



Fig 4. Inference using the pretrained YOLOv5 model on FLIR dataset with inference time of 50.07 FPS



Fig 5. Inference using the pretrained EfficientPS model on Cityscapes dataset with inference time of 0.61 FPS



Fig 6. Inference using the pretrained Panoptic DeepLab model on Cityscapes dataset with inference time of 0.29 FPS

4. Performances

Method	pedestrian.adult	vehicle.bicycle	vehicle.car	vehicle.motorcycle	vehicle.truck	movable_object.barrier	movable_object.trafficcone	static_object.bicycle_rack	vehicle.bus.rigid	vehicle.construction	Average
YOLOv5	0.475	0.659	0.796	0.74	0.624	0.608	0.573	0.618	0.601	0.555	0.625

Table 1. Performance of YOLOv5 on the nulumages validation set (in mAP metric)

Method	road	sidewalk	building	wall	fence	pole	traffic light	traffic sign	vegetation	terrain	sky	person	rider	car	truck	bus	train	motorcycle	bicycle	Average
EfficientPS	98.2	79.7	89.2	34.7	40.6	59.5	52.8	71.3	90.4	40.1	88.3	58.8	56.6	70.9	53.9	74.2	67.7	46.6	50.4	64.4

Table 2. Performance of EfficientPS on the Cityscapes validation set (in PQ metric)

5. Conclusion

- YOLOv5 excellently detects pedestrian with high accuracy of >60 mAP and fast inference time of >50 fps
- YOLOv5 achieves the state-of-the-art for object detection
- EfficientPS and Panoptic DeepLab have high accuracy but low inference time
- This project helps me learn more about AI, deep learning and computer vision
- This work is a useful research experience which provides me the knowledge foundation for my continuing master education

6. Future Work

- Apply YOLOv5 in Autware.Auto to test how accurate the model for pedestrian detection in real time driving application
- Autware.Auto is a stack open-source software for autonomous driving simulation and real time car application controlling
- Develop my own algorithm based on top of YOLOv5 which is even faster and more accurate for pedestrian detection as well as other traffic detection

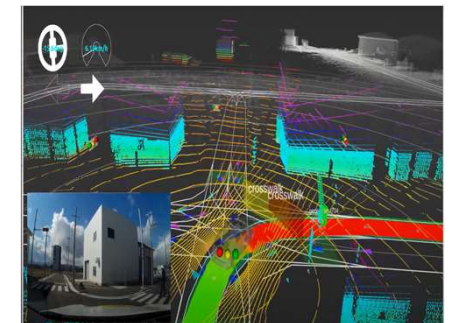


Fig 7. Environment perception in Autware.Auto software

7. Acknowledgements

