Pedestrian Detection in Autonomous Vehicles Using AI and Computer Vision

Author: Dinh Hoang (University of Central Arkansas) | Congdinh1801@gmail.com

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 YOLOv5 (You Only Look Once version 5) – a state of the art, fast and accurate model for a 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 nuImages dataset consisting of >14k images
- FLIR dataset was taken by using Forward Looking InfraRed (FLIR) thermal cameras while nuImages dataset was taken by using regular RGB cameras
- Additionally used Efficient Panoptic Segmentation (EfficientPS) and Panoptic DeepLab – state-of-the-art algorithms for panoptic segmentation tasks

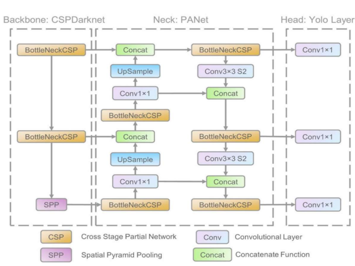

Fig 2. Yolov5 Architecture

3. Inference Results

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

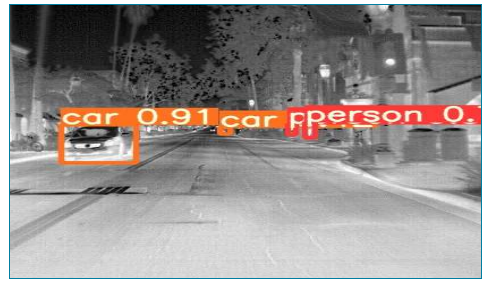

Fig 4. Inference using the pretrained YOLOv5 model on FLIR dataset with inference time of 50.07 FPS

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.trafficcone	static_object.bicycle_rack	vehicle.bus.rigid	Average
YOLOv5	47.5	65.9	79.6	74	62.4	57.3	61.8	60.1	63.58

Table 1. Performance of YOLOv5 on the nuImages validation set in mean Average Precision (mAP) metric

→ YOLOv5 accurately detects "pedestrian.adult" class with high accuracy of 47.5 mAP

Method	road	sidewalk	building	fence	pole	traffic light	traffic sign	vegetation	sky	person	rider	car	truck	bus	train	motorcycle	bicycle	Average
EfficientPS	98.2	79.7	89.2	40.6	59.5	52.8	71.3	90.4	88.3	58.8	56.6	70.9	53.9	74.2	67.7	46.6	50.4	67.59

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

→ EfficientPS accurately performs panoptic segment task on "person" class with high accuracy of 58.7 PQ

5. Conclusion

- YOLOv5 excellently detects pedestrians with high accuracy of >47 mAP and fast inference time of >50 fps
- YOLOv5 achieves state-of-the-art performance 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 that provides me the knowledge foundation for my continuing master education

6. Future Work

- Apply YOLOv5 in Autoware.Auto to test how accurate the model can detect pedestrians in real-time driving application
- Autoware.Auto is a stack open-source software for autonomous driving technology
- Develop my own algorithm based on top of YOLOv5 which will be even faster and more accurate for pedestrian detection as well as other traffic detection

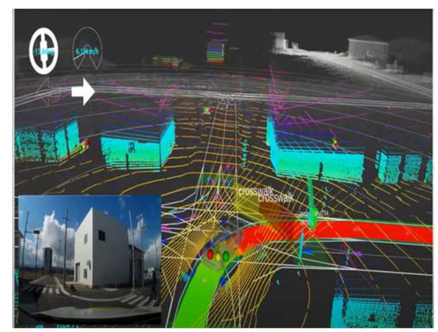

Fig 7. Visualization of real-time environment perception of a car in Autoware.Auto software

7. Acknowledgements

1950872