Real Time Pedestrian Detection using YOLOv5 and Jetson Nano

Author: Samuel Sokalzuk
Mentor: Dr. Sarah Harris
Research Experiences for Undergraduates (REU) Summer 2021 at University of Nevada, Las Vegas

1. Background and Introduction

- Pedestrian fatalities have increased across the U.S. and in Nevada:
  - U.S. increase: 54% (from 4,300 to 6,600 from 2010-2020) [1]
  - Nevada increase: 28% (from 2019-2021) [1]
- Steps made toward making vehicles autonomous will help decrease this number.
- The purpose of this project is to make a low-cost hardware-software device, the Pedestrian Detection System (PDS), that count the number of pedestrians at an intersection and transmits this data to the RTC (Regional Transportation Commission) to aid in traffic light regulation or to share that number with a driver using an app to alert when pedestrians are present.

2. Methodology

- Designed and built a Pedestrian Detection System the YOLOv5 neural network and the GPU-based Jetson Nano.
  - **Jetson Nano** - GPU-based microcontroller (128-core Maxwell) optimized for running multiple neural networks in parallel with the free software Jetpack.
  - **YOLOv5** (You only look once) is a free neural network that predicts bounding boxes and class probabilities from images in one evaluation.
- Trained YOLOv5 using the publicly available Cityscapes dataset [2] with three different YOLOv5 models on the classes car, bicycle, person, motorcycle, rider, bus, and truck. Parameters are the numbers of weights and biases in a model.
  - YOLOv5n: nano model that has 1.9 million parameters.
  - YOLOv5s: small model that has 7.2 million parameters.
  - YOLOv5m: medium model that has 21.2 million parameters.

Table 1. Cost of Software and Hardware

Hardware/Software	Price
Jetson Nano	$99
256 GB Micro SD Card	$12
Rasberry Pi Camera 2	$25
ESP32 (Wifi/Bluetooth)	$22.50
Total	**$158**

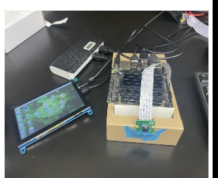

Fig 1. Pedestrian Detection System(PDS)

3. Inference Results

Fig 3.Label Key

Fig 2. Unannotated Image(Cityscapes).

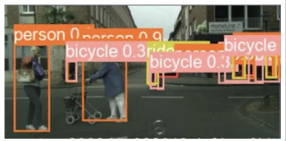

Fig 4. Results of Yolov5m model

Fig 5. Results of YOLOv5s model

Fig 7. Preliminary Live Detection Results using PDS with model YOLOv5n.

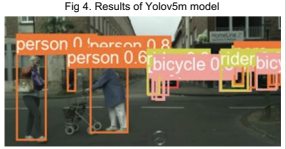

Fig 6. Results of YOLOv5n model

4. Model Performances

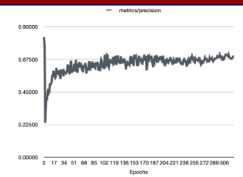

Fig 8. Precision of YOLOv5m model.

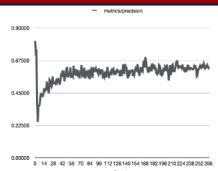

Fig 9.Precision of YOLOv5s model.

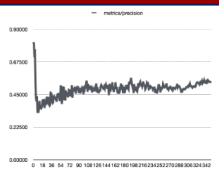

Fig 10. Precision of YOLOv5n model.

Table 2. Precision mAP(mean average precision metric) and Inference Speed on Cityscapes dataset

Model	Time	Car	Bicycle	Person	Motorcycle	Rider	Bus	Truck	Precision(Average)
YOLOv5m	2.0ms	0.732	0.622	0.586	0.55	0.694	0.76	0.664	**0.658**
YOLOv5s	0.8ms	0.688	0.583	0.561	0.517	0.715	0.845	0.8	**0.673**
YOLOv5n	0.4ms	0.565	0.398	0.418	0.418	0.549	0.452	0.447	**0.464**

5. Conclusion

- YOLOv5m has the highest accuracy when it comes to identifying pedestrians at 0.586 mAP but has the slowest inference time of 2.0 ms.
- YOLOv5s has the second highest accuracy at identifying pedestrians at 0.561 mAP but is almost twice as fast as YOLOv5m at 0.8ms.
- YOLOv5n has the lowest accuracy at 0.418 mAP but has the fastest inference time at 0.4 ms.
- YOLOv5s balances speed and accuracy better than models YOLOv5m and YOLOv5n.

6. Future Work

- Run additional tests on YOLOv5 models on live video with the Jetson Nano with the Raspberry Pi Camera in the Las Vegas area.
- Have the pedestrian count acquired through models on the Jetson Nano be displayed to an app.
- Make the Pedestrian Detection System more accurate and power-efficient by modifying the algorithm or hardware.

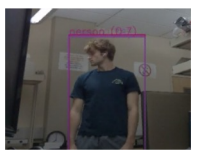

Fig 11.Screenshot of Yolov5s model live on Jetson Nano with Raspberry Pi Camera.

7. References

[1]. Petraglia, Elizabeth. @*GHSAHQ Pedestrian Traffic Fatalities by State*

8. Acknowledgements

NVIDIA
1950872
CITYSCAPES DATASET
[2]