REAL-TIME MOBILITY ASSISTANCE FOR THE LEGALLY BLIND

Raunak Khaitan, Dr. Mohammad Habibur Rahman

Dept. of Biomedical Engineering and Computer Science, BioRobotics Lab, College of Engineering and Applied Sciences, UW-Milwaukee

UNIVERSITY of WISCONSIN UWMILWAUKEE

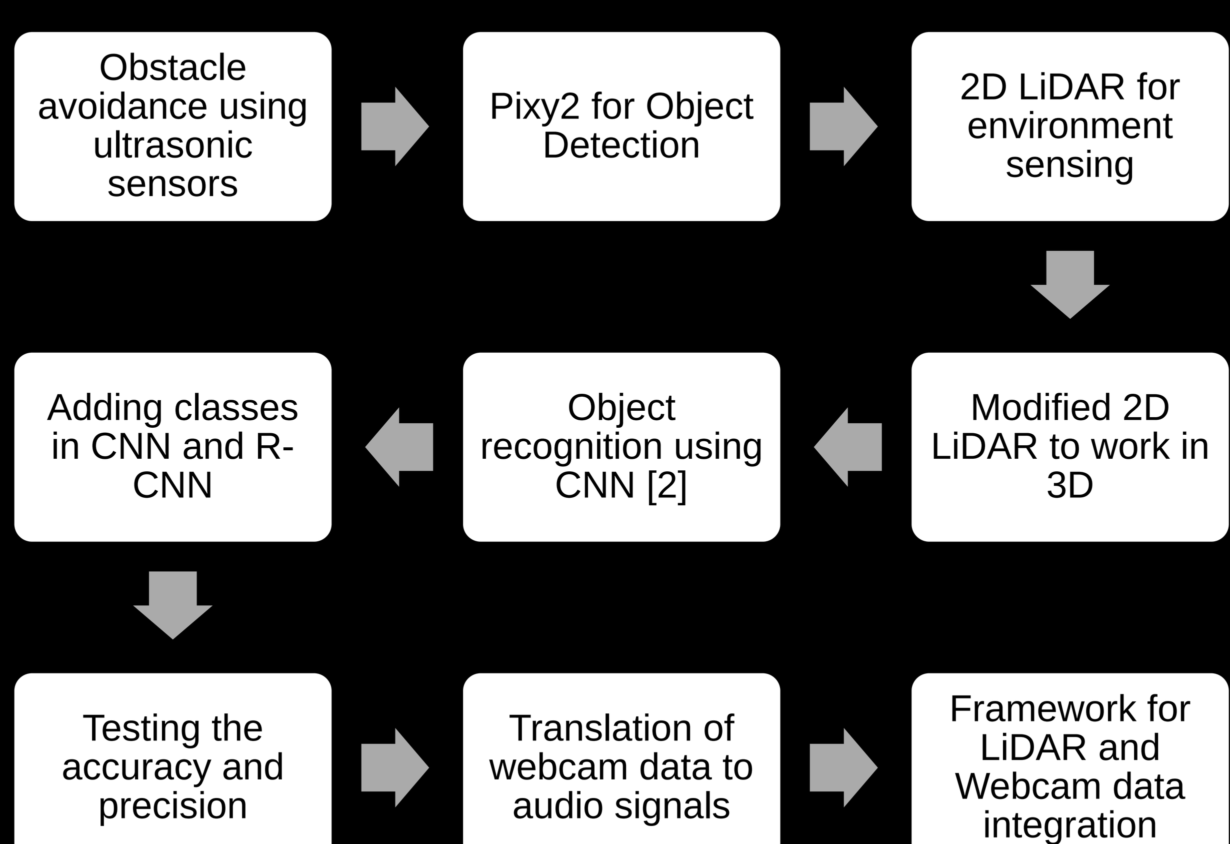
ABSTRACT

This project is powered by machine learning and deep learning infrastructure that is aimed to ease the way legally blind people navigate on the streets and make them feel independent just like anyone else. The project is being developed and tested on a rover that contains a lidar that maps the surrounding, ultrasonic sensors that give feedback on obstacles in forward and reverse motion and a camera that identifies and classifies objects. The rover will be able to differentiate between stationary and moving objects. The entire information will be converted into an audio message that the person can use in real-time on the streets. Future work will also include a braille system which will help both deaf and blind people. After successful completion of the entire technology, the rover will be scaled down to a wearable smart glass that will be both trendy and highly useful to those in need. The vision is to make the disabled "differently-abled."

OBJECTIVES

- Making navigation easier for the legally blind population, both indoors and outdoors.
- Analyzing the challenges in street navigation by pedestrians including terrain difference, surface quality, etc. using yolo architecture [1].
- To train deep neural network models for object recognition with capabilities to distinguish roads and pedestrian walkways, traffic lights, stationary and moving objects, etc.
- Decoding the numerical data into human understandable audio format.
- Implementing all these technologies on a rover initially for testing purposes.

METHODOLOGY



5 direction mechanism: L,LFD,C,RFD,R
Bit 0 represents clear or no obstacle.
Bit 1 represents obstacle.

Example: 00100 = obstacle at front
01100 = obstacle at front and left front diagonally.

```
begin str=01100
begin str=11111
begin str=11111
begin str=11111
begin str=00100
begin str=00000
begin str=01110
```

Fig 1: Ultrasonic Sensor Output in 5 bit sequence detecting obstacles.

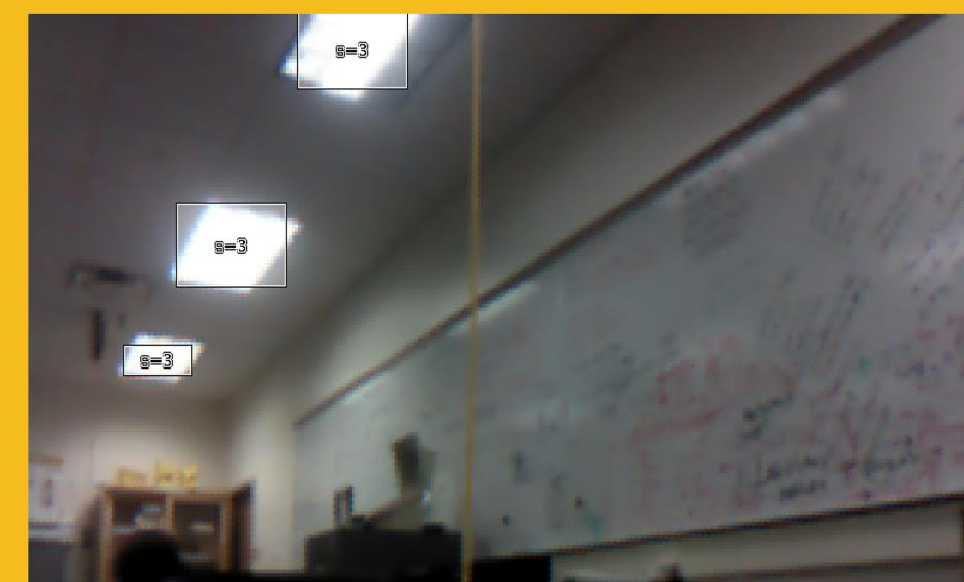


Fig 2: Pixy2 Camera detecting the ceiling lights as trained

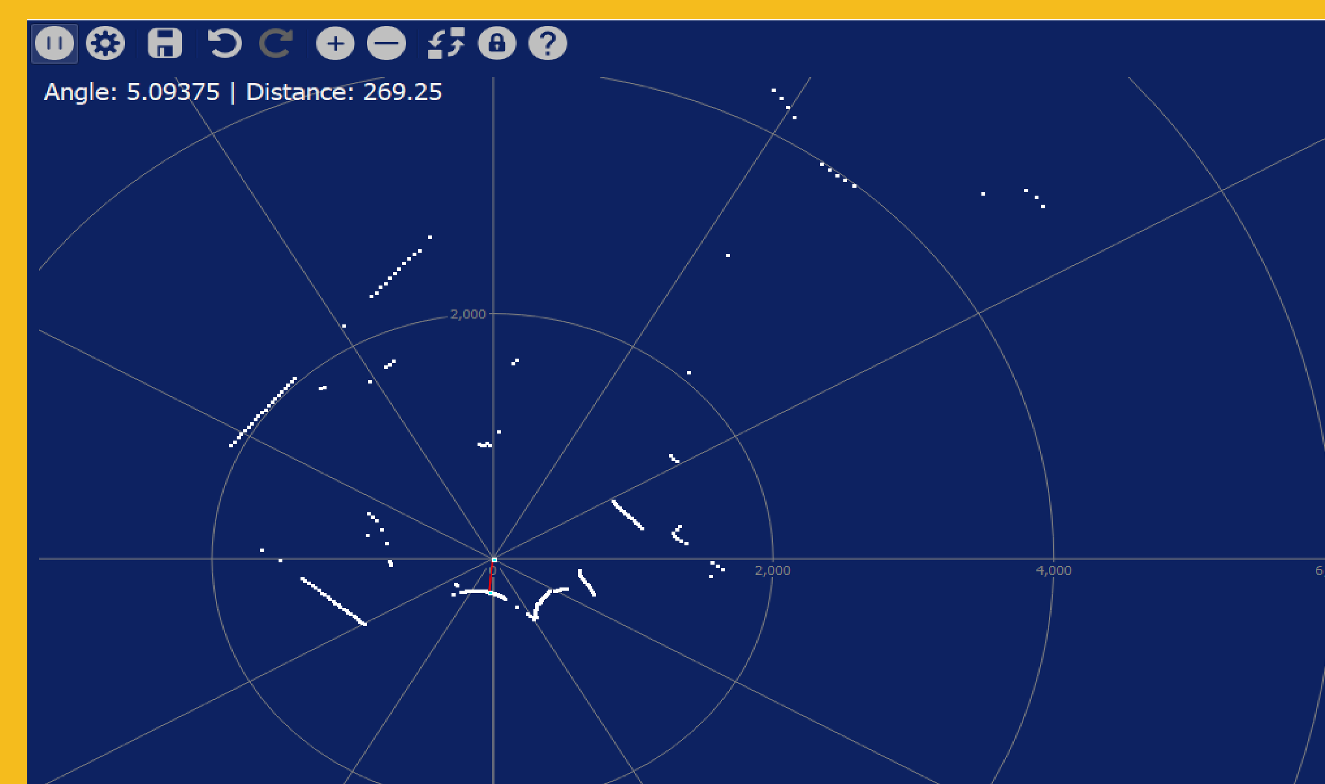


Fig 3: LiDAR sensor mapping the environment in two dimensional. Prints the angle and distance on top left corner.

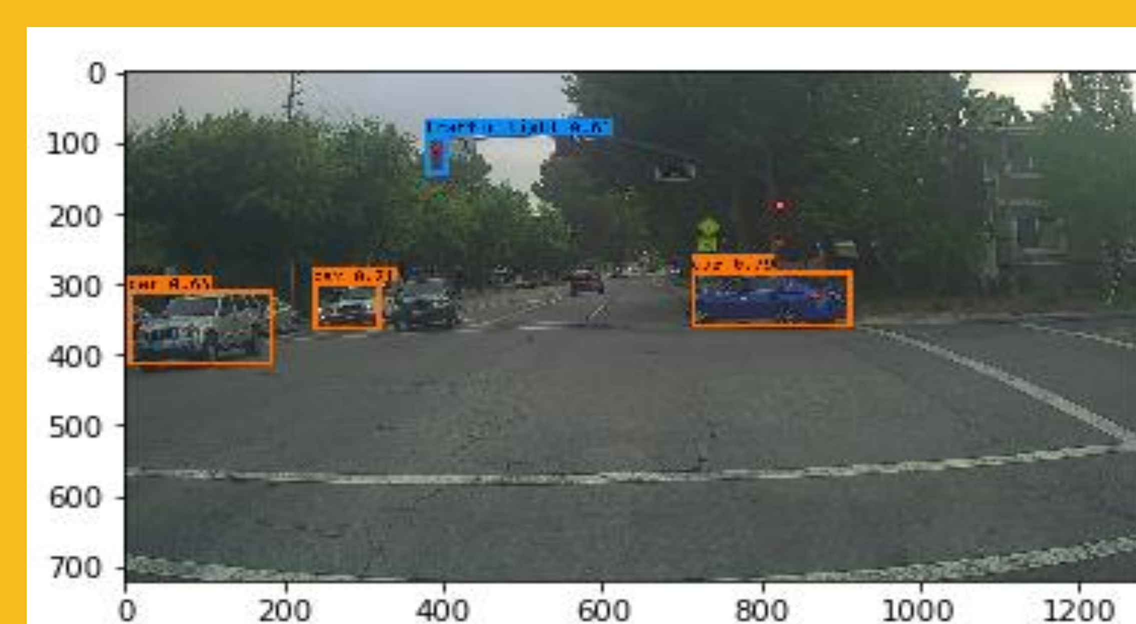


Fig 4: CNN based output of a picture that detects the cars in the streets along with the traffic lights. Note that it detects the traffic light as a whole but not particularities like if the signal on or off for pedestrian crossing. It also misses out on detecting a few cars [1].

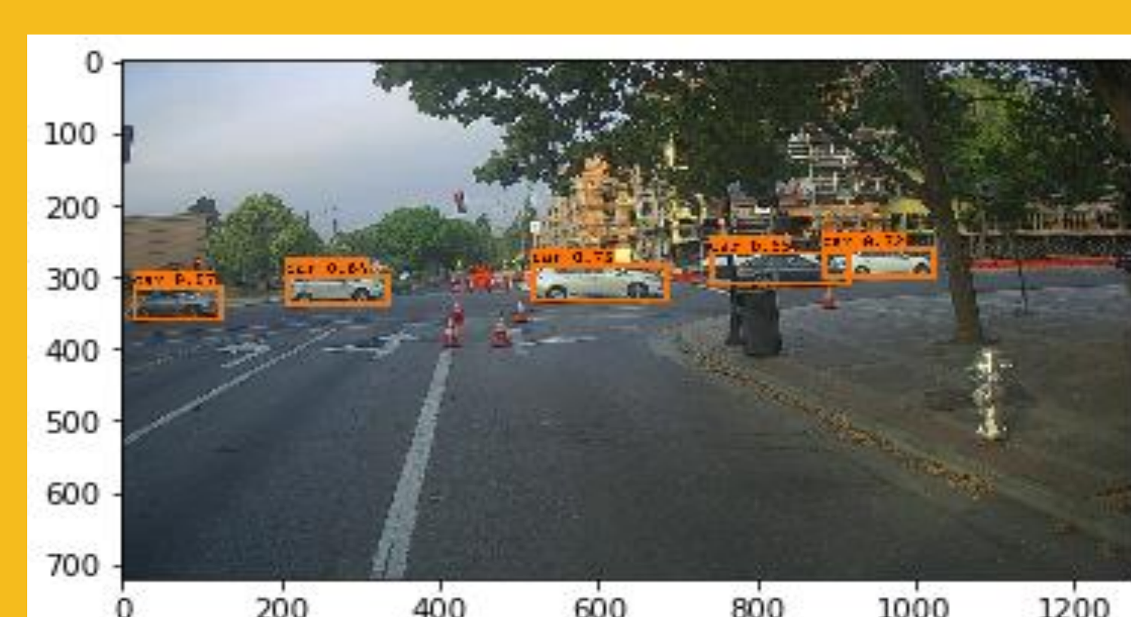


Fig 5: CNN based output of a picture that detects the cars in the streets. Note that this time it detects all the cars in its vision. For pedestrian walkways, challenges increase because it needs to identify the trees, trash cans, etc. Also, the cones in the zebra crossing needs to be recognized [1].

RESULTS

Method	Positive aspects	Negative aspects
Pixy2 Camera	• Accurately detects the objects trained to look for. • No training required.	• Limited objects to < 10. • Not feasible in outer environment.
2D LiDAR Sensor	• Effective room mapping with precise angle & distance measurement. • No training required.	• Works in 2D, not 3D. • Identifies obstacles but can't recognize them.
Ultrasonic Sensor	• Accurate measurement of obstacles in 180º. • Sufficient accuracy for indoor operations. • No training required.	• Standalone sensor works in 2D, not 3D. • Identifies obstacles but can't recognize them.
CNN	• Recognizes objects accurately & with reasonable confidence percentage. • Can be modified to recognize any type of object.	• Requires a lot of training and testing before deployment.

CONCLUSION

The rover detects obstacle in forward and reverse motion and stop or changes path accordingly, the CNN model recognizes the objects it is trained for [2]. The farther the object from the lens, lesser accuracy is observed and vice-versa.

FUTURE WORK

- Increasing accuracy of CNN model and adding layers to recognize particularities like the status of traffic signal, speed of moving object and it's direction, etc.
- Translating the recognized objects into audio signals and tuning its time precision.
- Modifying CNN's to detect stationary and moving objects with speed and location.
- Designing a framework to integrate LiDAR point cloud data and webcam data and ensuring its parallel accuracy.

BIBLIOGRAPHY

1. **YOLO Architecture:**
   - Redmon, Josheph, et al. "You Only Look Once: Unified, Real-Time Object Detection." *ArXiv.org*, 9 May 2016, arxiv.org/abs/1506.02640.
   - Redmon, Joseph, and Ali Farhadi. "YOLO9000: Better, Faster, Stronger." *ArXiv.org*, 25 Dec. 2016, arxiv.org/abs/1612.08242.
2. **Deep Learning:**
   -A. Ng, Y. Bensouda, and K. Katanforoosh, "Deep Learning by deeplearning.ai," *Coursera*. [Online]. Available: https://www.coursera.org/specializations/deep-learning.