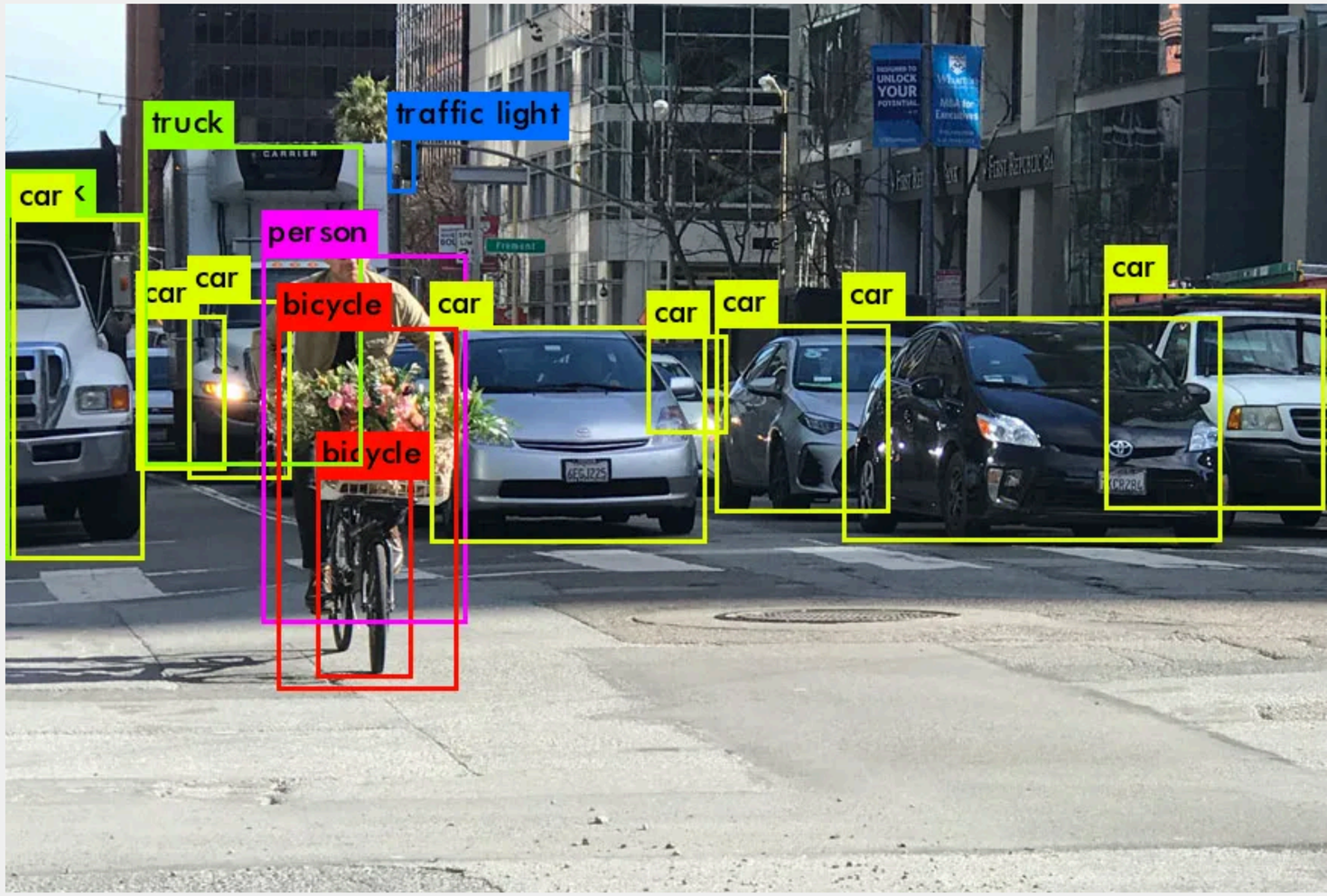
Assistive Vision System



GROUP 5
Type 1: Engineering Project

- Siyi Gao
- Yufei Huang
- Xuedong Pan
- Tongle Yao

Defining Problems

Problem 1

Real-time applications for visually impaired are rare and require human assistance

Problem 2

Models with good performance are too large to be applied on edge computing devices.

Problem 3

Existing models perform poorly in complex environments or low-light conditions,

Project Goals

1. Describing all objects in the user's front, including their class and dimensions.
2. Alerting users about potential obstacles such as potholes and curbs on the way.
3. Notifying users of traffic signals and other important visual cues.
4. Maintaining high recognition accuracy in low-quality and low-light environments.
5. Achieving real-time processing on small, low-power computing platforms.

Ideas to solve the problems

1. **Model Compression Techniques：**

- Utilizing YOLO v4 for object detection, replacing the CSPDarknet53 backbone with lighter architectures like MobileNet.

- Investigate efficient YOLO variants such as YOLOv8-nano

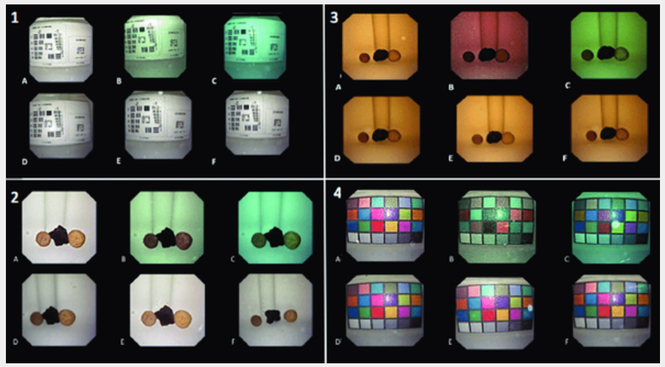
2. **Monocular and Dual Camera**

- Investigate the MonoDepth algorithm, which uses a single camera to estimate depth

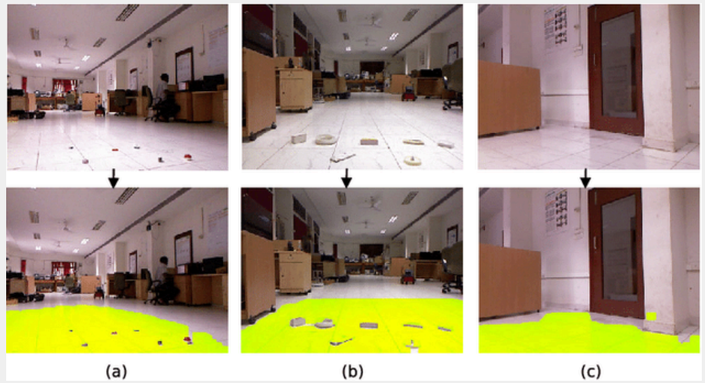
- Compare accuracy with dual camera system using the Semi-Global Matching (SGM) algorithm

Datasets

BDD100K datase



Images in various lighting conditions



Images of low-lying obstacles

$

Employ data augmentation technique

Related works

- Object Detection: Investigated YOLO developments (Redmon et al. 2016; Bochkovskiy et al. 2020) for foundational object detection techniques.
- Distance Estimation Approaches:
  - Monocular Vision: MonoDepth by Godard et al. (2017) for single-camera depth assessment.
  - Stereo Vision: Semi-Global Matching (SGM) algorithm by Hirschmüller (2008) for accurate dual-camera depth analysis.
- Efficient Computing Models: Explored "MobileNets" by Howard et al. (2017) for implementation on mobile and embedded systems.

Reference

- Redmon, J., et al. (2016). "You Only Look Once: Unified, Real-Time Object Detection." In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp. 779-788.
- Bochkovskiy, A., et al. (2020). "YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv preprint arXiv:2004.10934.
- Godard, C., et al. (2017). "Unsupervised Monocular Depth Estimation with Left-Right Consistency." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 270-279.
- Hirschmüller, H. (2008). "Stereo Processing by Semi-Global Matching and Mutual Information." IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2), pp. 328-341.
- Howard, A. G., et al. (2017). "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." arXiv preprint arXiv:1704.04861.

Thank you