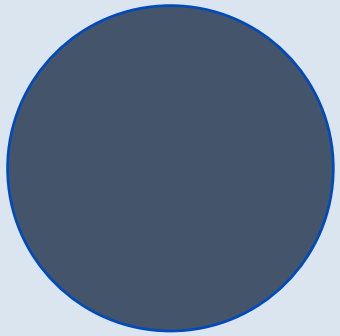
BROOKDALE COMMUNITY COLLEGE

New Jersey Space Grant Consortium

Helping the Visually Impaired with Object Detection



Jack Arteaga

Mentor: Dr. Christopher Ochs

Advisors: Karina Ochs, Ana Teodorescu

Overview

- What is object detection and how can it be used to help the visually impaired?
- Techniques used to implement object detection
- Creating an object detection and avoidance algorithm
- Implementation of algorithm
- Results and conclusion found after implementation

Objective

To determine if off the shelf hardware and object detection techniques can be used to provide live object avoidance of everyday household items to the visually impaired.

Background

Object detection: A computer vision technique that can be used to identify and locate objects in an image or video.

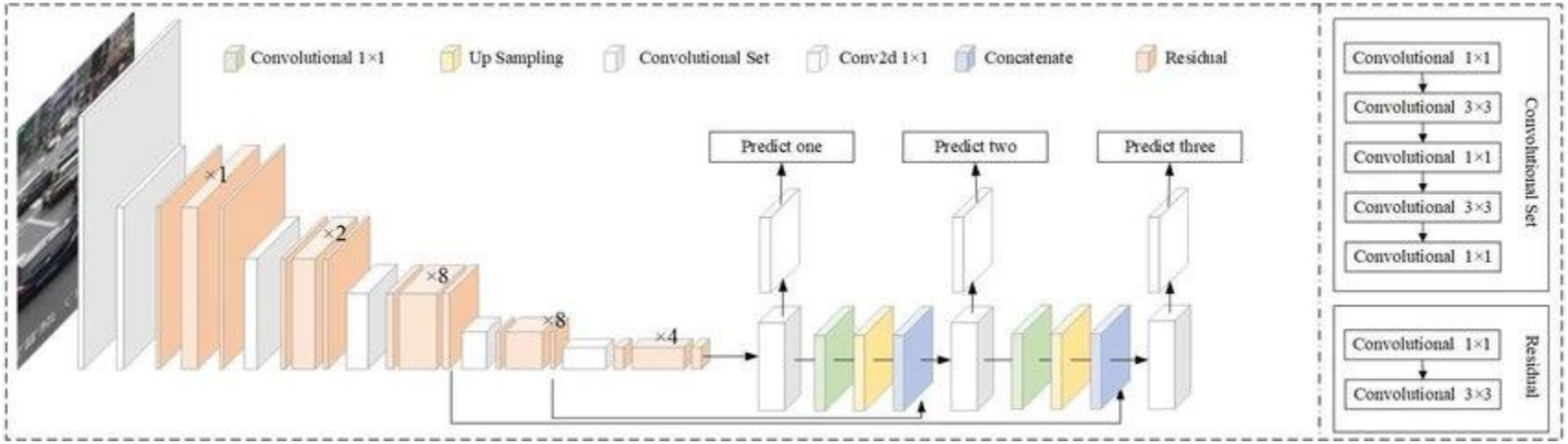
Figure 1: YOLO Multi-Object Detection And Classification.

Background

- An object must be classified and localize in an image in order to be detected

Background

- An object detection model is a collection of mathematical representation of data and an algorithm that can classify objects in an image, typically trained using thousands of images



Background

- Object detection models can be created and trained to fit certain use cases

- Pre-trained models have been trained according to the needs of previous use cases and can be applied to use cases in similar domains

Technology Aiding the Disabled

- Mobile Apps
  - -Be My Eyes
  - -TapTapSee
- Wearable equipment
  - -biped.ai
- QR codes

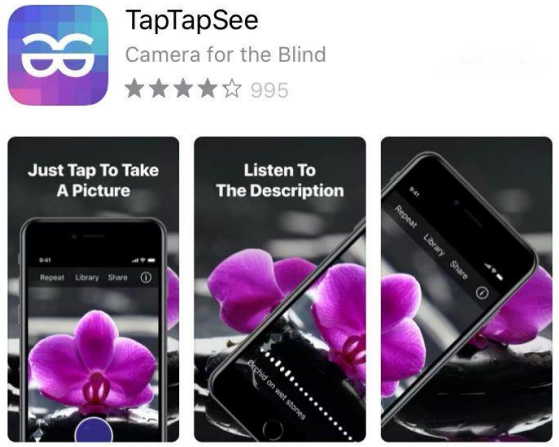


Figure 2: TapTapSee a mobile camera application designed specifically for blind and visually impaired users

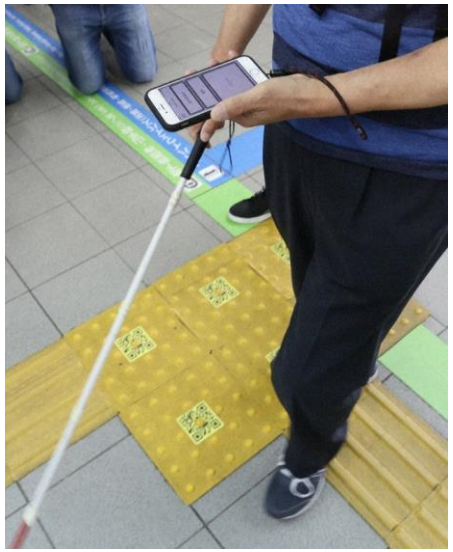
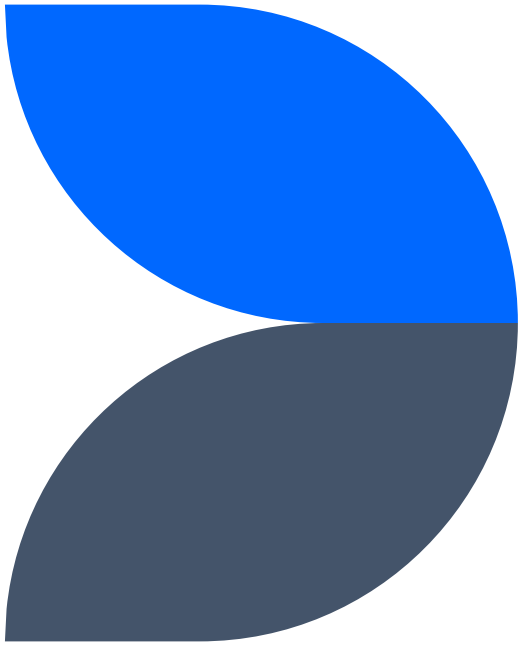


Figure 3: QR codes being read through smartphone to assist the blind navigate JR Shin-Kobe Station

Figure 4: biped smart harness: an AI copilot for persons who are blind or have a visual impairment.

Research

Choosing a Model and Implementing it



Evaluating Approaches

Creating and training a model vs using a pre-trained model

- What are the advantages of each?
- What are the disadvantages of each?
- Which one is better suited for our specific use case?

Evaluating Approaches

Creating and training a model

Pros	Cons
Control of dataset	Large amount of data needed
Can Specify for use case	Time needed to train and test
Ability to modify Algorithm	Model must be retrained for use

Using a pre-trained model

Pros	Cons
Already trained (plug and play)	Cannot specify dataset
Widely used for a general use case	Cannot modify Algorithm
Saves time and resources	

Evaluating Approaches

Based on empirical analysis, a pre-trained model better suited this project

- The objective is to detect everyday household items that may typically be in the way of a visually impaired person in their home
- Pre-trained open-source models exist to detect everyday household items
- Training a model on thousands of household items would be impractical due to data and computing resource requirements

Evaluating Pretrained Models

- The pre-trained models You Only Look Once (YOLO)v3 and Single Shot Detectors (SSD) MobileNet were tested.

- YOLO was more accurate and faster to detect objects in a live video feed than SSD MobileNet, based on testing

Algorithm

- Object avoidance systems must detect objects and calculate their distance from the user
- A six-step algorithm was developed to identify the distance to known objects in a live video stream

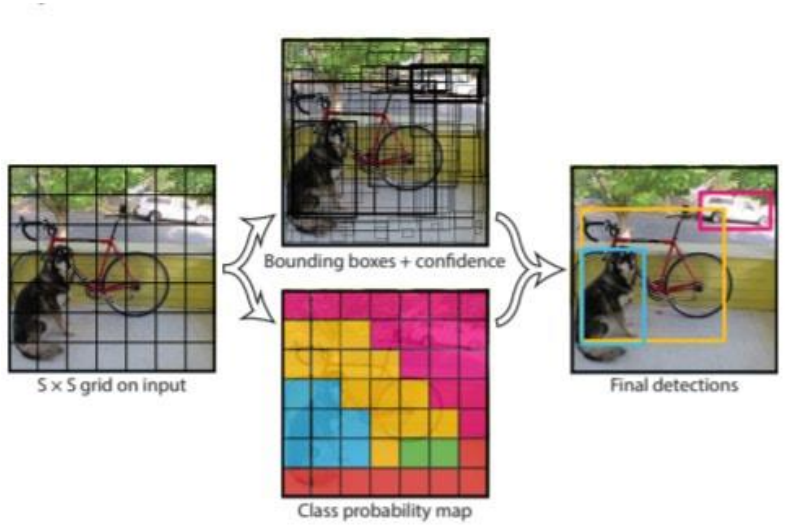


Figure 5: YOLO techniques applied to produce final detection results

Algorithm

1. Process frame to detect objects using YOLOv3 weight files

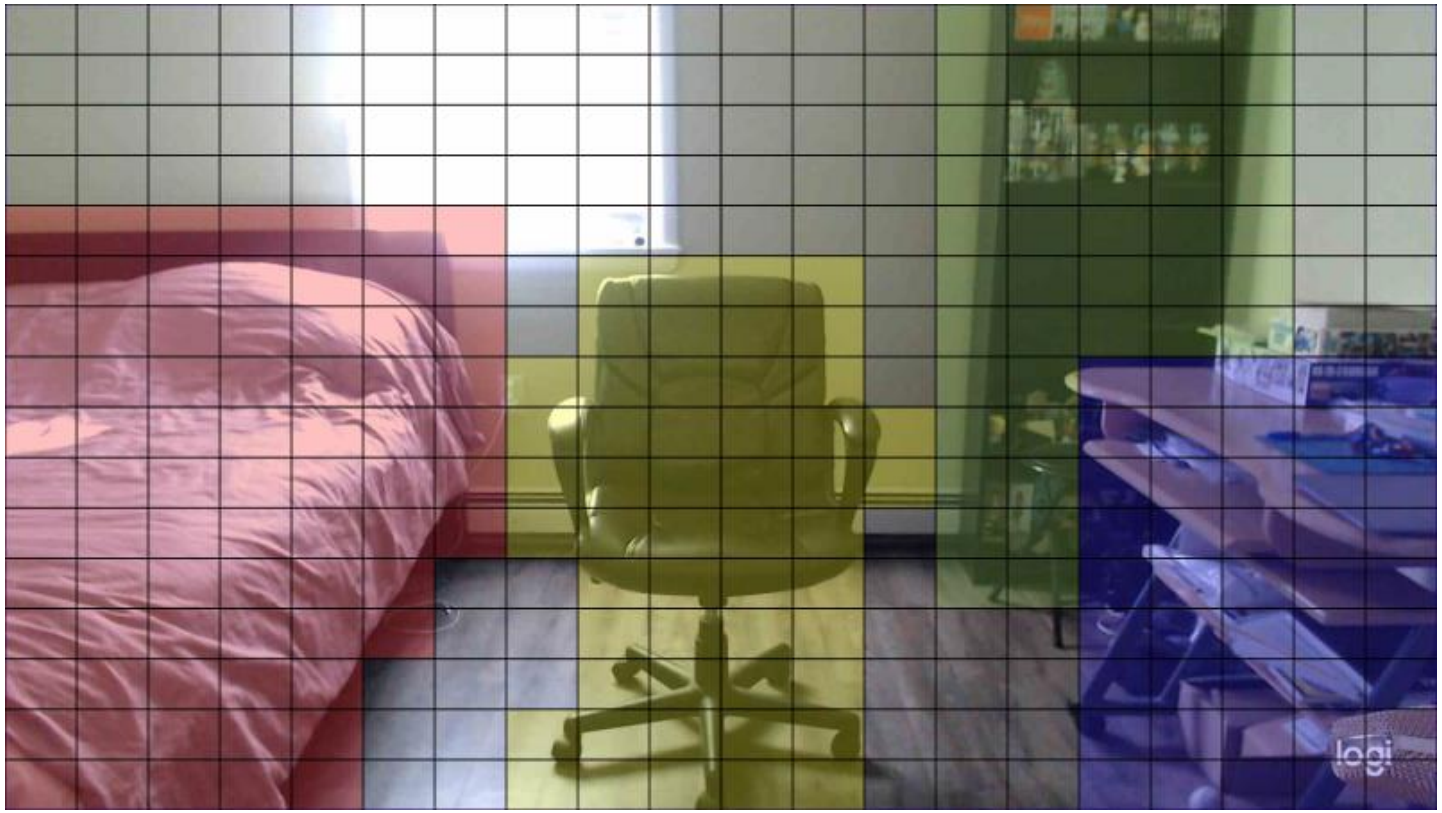


Figure 6

Algorithm

2. Find known objects in object data base

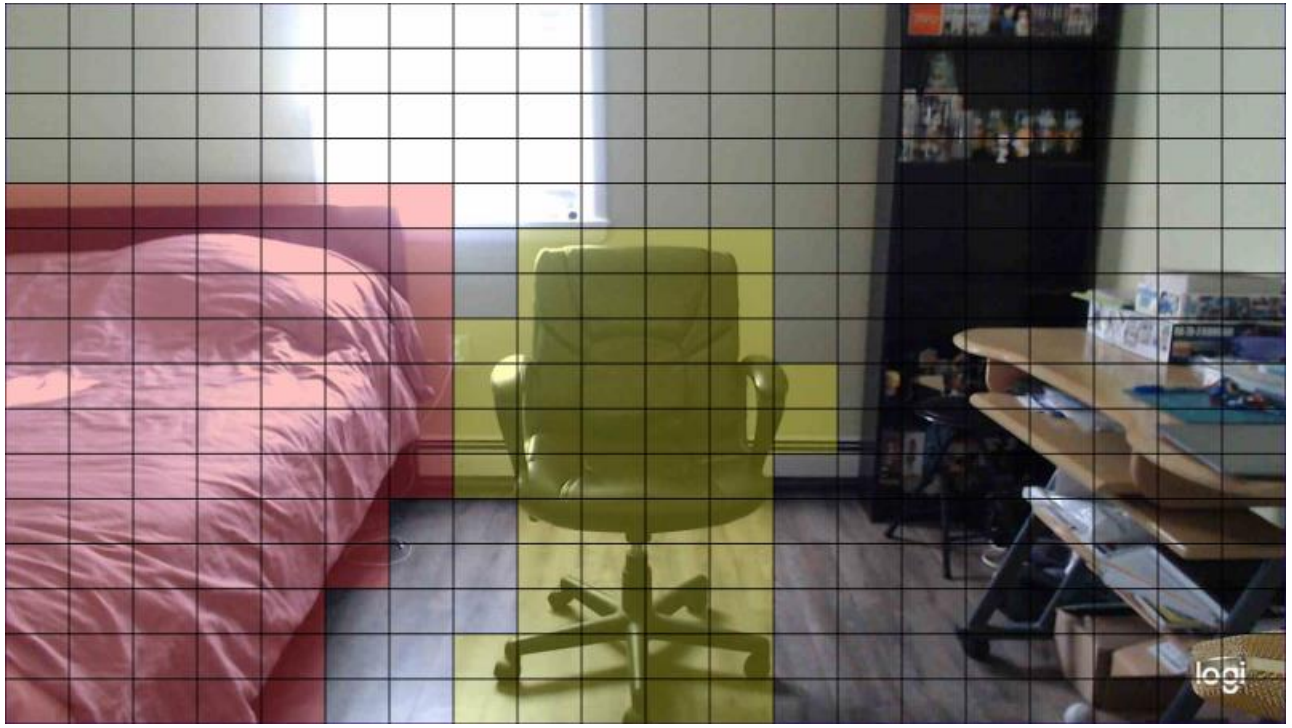


Figure 7

```
57    chair
58    sofa
59    pottedplant
60    bed
61    diningtable
62    toilet
```

Figure 8

Algorithm

3. Estimate location of object in frame

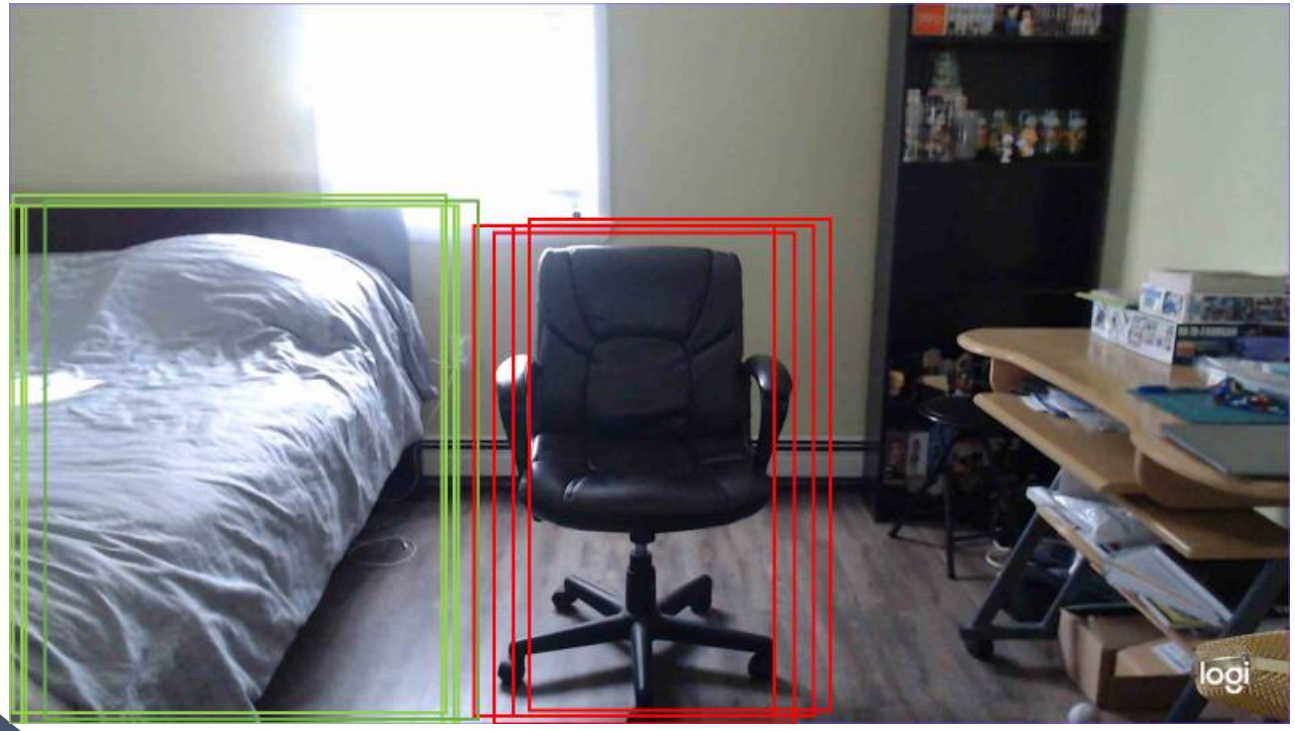


Figure 9

```
Image Shape:  (1080, 1920)
bed Center  x: 328  y: 558
bed Center  x: 324  y: 644
chair Center  x: 953  y: 646
bed Center  x: 290  y: 700
bed Center  x: 325  y: 702
chair Center  x: 953  y: 703
bed Center  x: 330  y: 755
chair Center  x: 951  y: 765
```

Figure 10

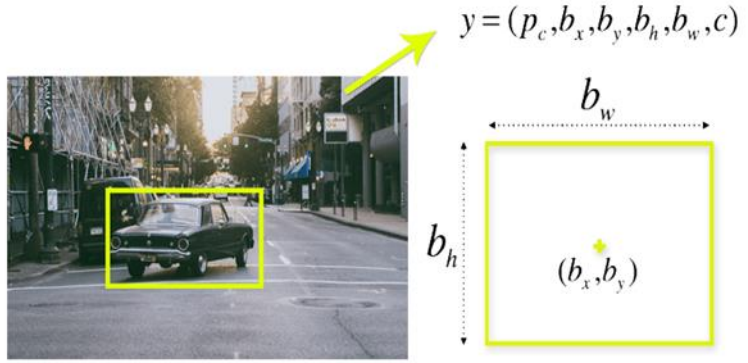


Figure 11

Algorithm

4. Create single bounding box using Non-Maximal Suppression and label object name and confidence

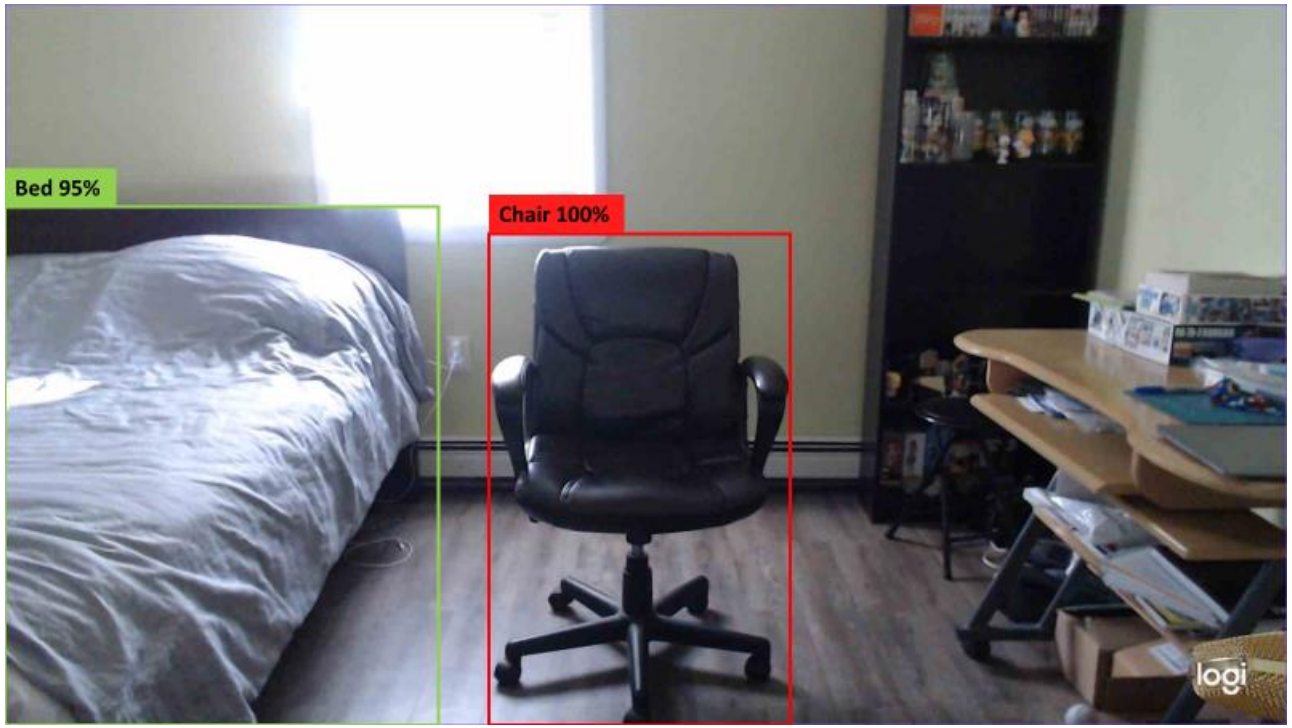


Figure 12

Algorithm

5. If a known measured object is detected, compute distance from camera

Figure 13

Known objects must have a reference image in the database where objects distance and size is predetermined

Distance Calculation

$$\text{Focal } Length = \frac{(widthInPixels \ * measuredDistance)}{actualWidth}$$

$$Distance = \frac{(actualWidth \ * focalLength)}{widthInPixels}$$

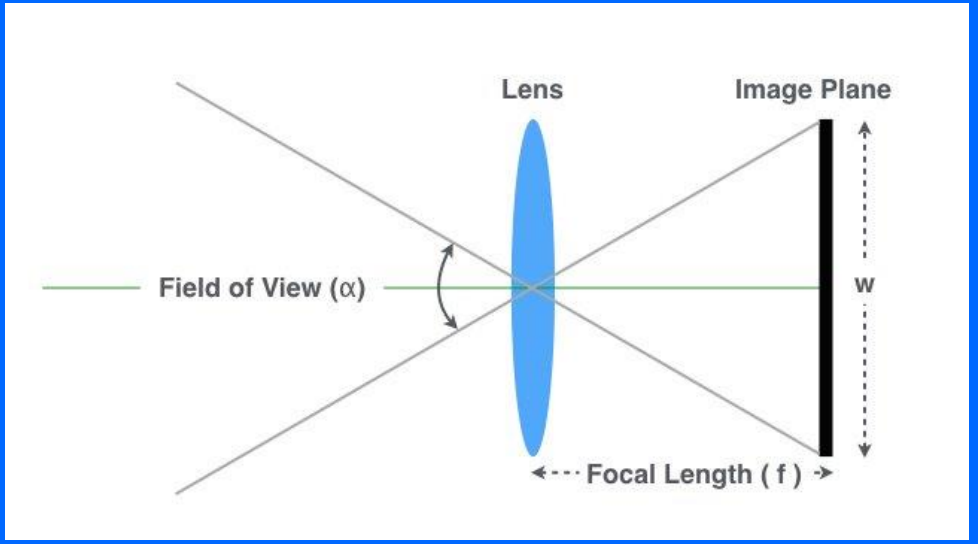


Figure 14: Relationship between focal length, Field of View ands Sensor Size

Algorithm

6. Determine if user is within distance threshold and trigger an alert if they are

Figure 15

Distance Calculation

Focal Length:

$$1293 \text{ pixels} = \frac{(431\ pixels * 72\ inches)}{24\ inches}$$

Distance:

$$36\ inches = \frac{(24\ inches\ * 1293\ pixels)}{862\ pixels}$$

System Implementation

- Using Python programming language, OpenCV (a Computer Vision library) captures frames in a video feed

- Once captured, each frame is processed using the pre-trained YOLOv3 algorithm to detect objects. Only objects with a confidence score greater than 60% will display a bounding box.

System Implementation

Chair Detected at 100% confidence

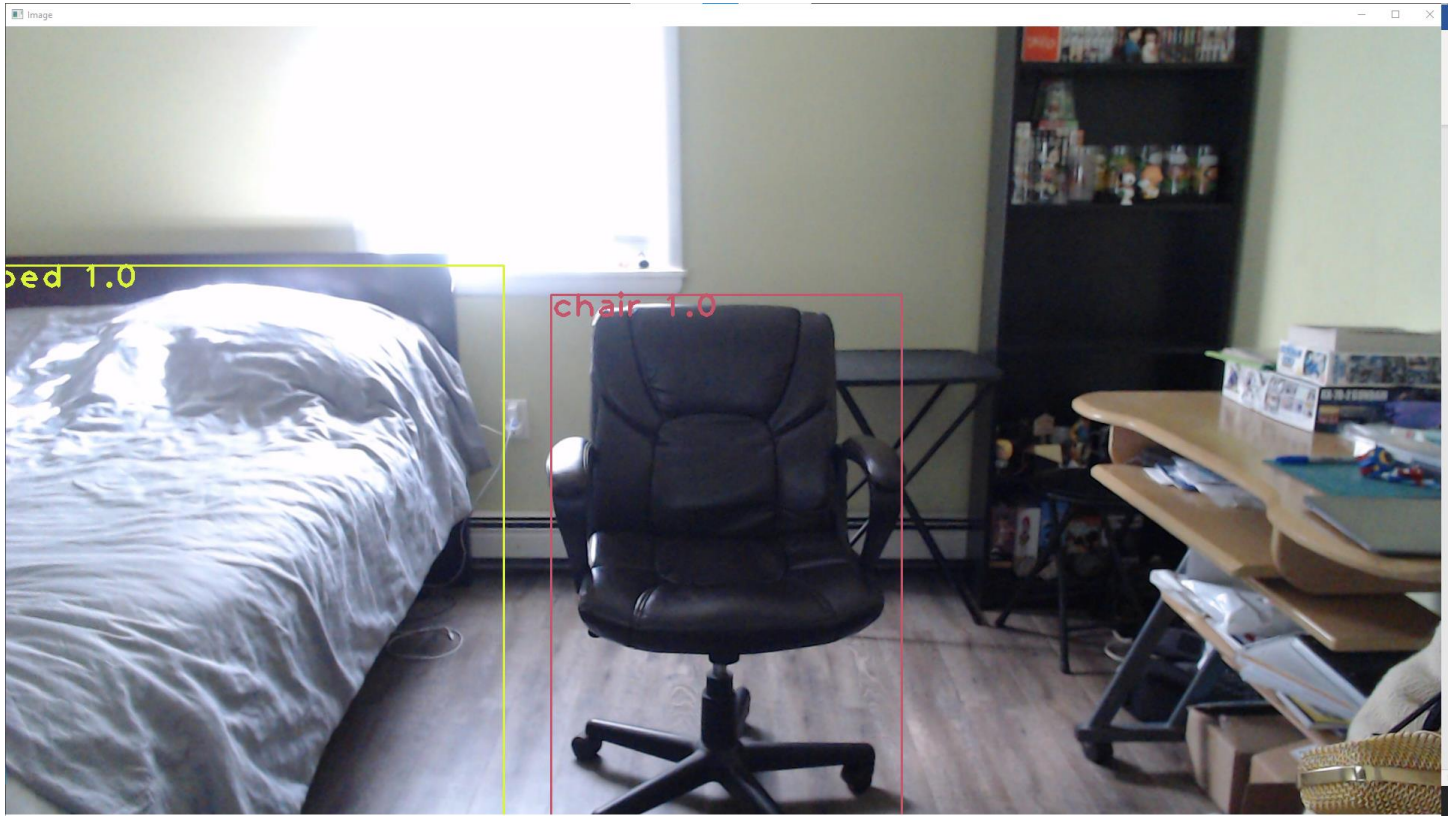


Figure 16

System Implementation

- Known objects will have their current distance to the camera calculated

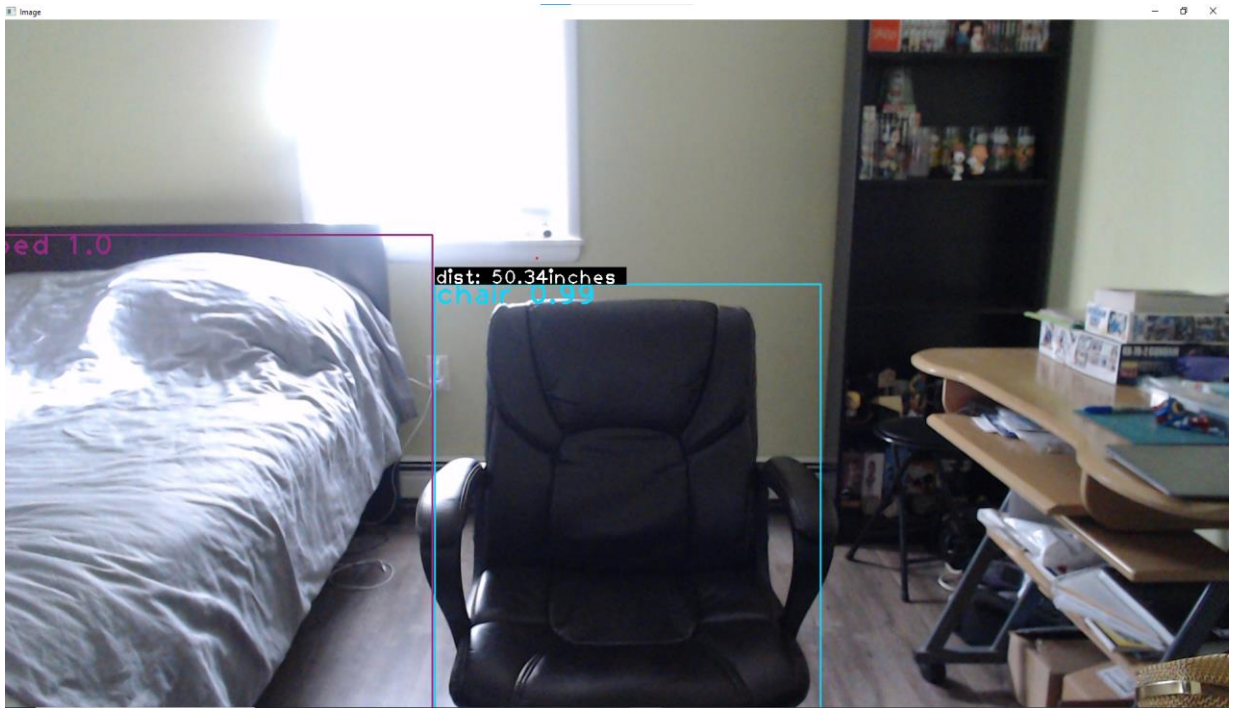


Figure 17

System Implementation

- If known objects come within a threshold of 48 inches (4ft), an alert is triggered indicating an object is ahead

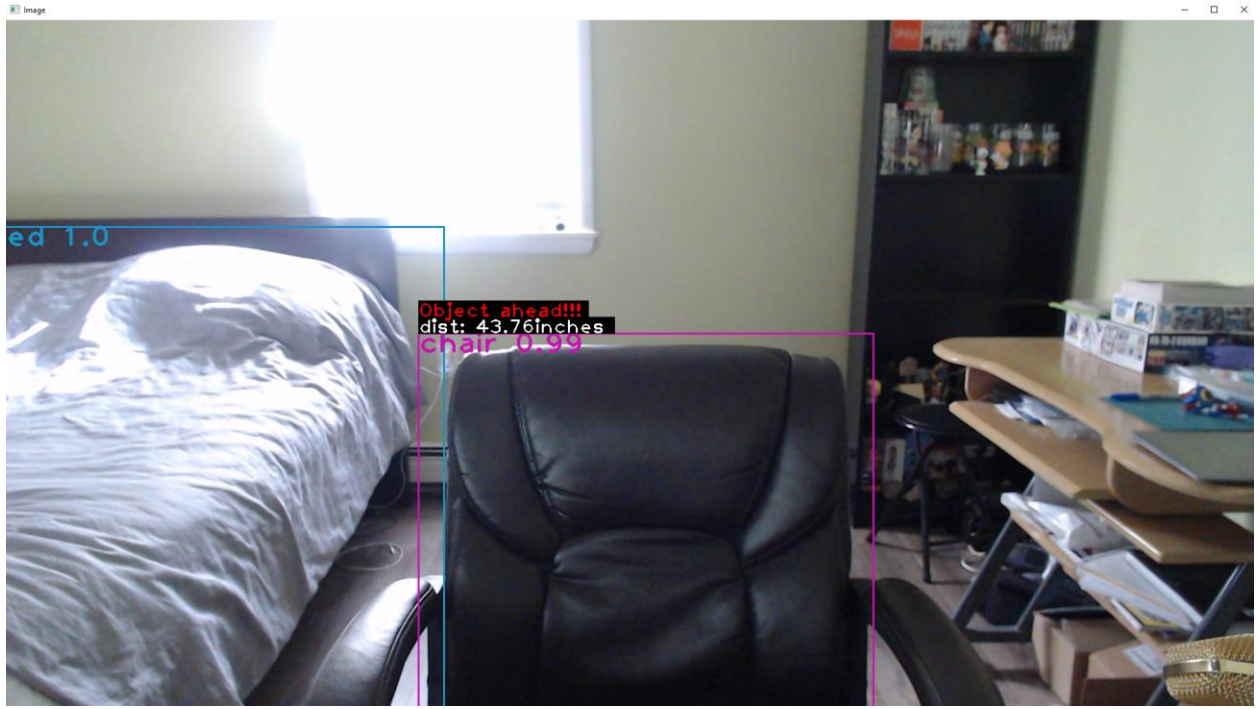


Figure 18

Results

An off the shelf, pre-trained object detection model can be used to build an object avoidance system.

The system can:

- Process live video footage and detect objects
- Calculate the size of and distance to known objects
- Trigger an alert to avoid objects once a certain distance threshold is met

Limitations

Testing and evaluation of the system uncovered several issues that affect the outcome of the object avoidance process.

Limitations include:

- Must have a database of known objects previously measured
- Use of only one camera, restricting depth and spatial capture of footage
- Image quality due to camera resolution, angles, and lighting
- Dependent on accuracy of YOLOv3 algorithm and model

Future Work

- Determine an object's position in 3D space using techniques such as triangulation
- Investigate more specialized hardware such as stereo vision cameras and lidar devices
- Expand number of objects in database that the system can detect and calculate distance to
- Deeper quantitative evaluation to test the benefits of training a custom model
- Conduct tests of the system with users (people with visual impairments)

Conclusion

- Developed a collision avoidance algorithm and system using a pre-trained object detection model
- Objects can be detected and their distance from the camera can be computed
- Objects need to be more consistently and accurately detected by the system before performing user tests
- Further evaluation is required to assess the utility of the system when implemented in the real world

Acknowledgments

This work is Supported by:

I would like to thank Dr. Christopher Ochs Professor, Karina Ochs, and Professor Ana Teodorescu for their continued assistance and support throughout this project.

References

- Mihajlovic, I. (2021, September 24). *Everything you ever wanted to know about computer vision. here's a look why it's so awesome.* Medium. Retrieved March 29, 2022, from https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e
- Naser, et al. "Yolo Object Detection with Opencv." PyImageSearch, 5 Feb. 2022, https://pyimagesearch.com/2018/11/12/yolo-object-detection-with-opencv/.
- "Using UPC Bar Codes and Quick Response (QR) Codes for Accessible Identification." The American Foundation for the Blind, https://www.afb.org/blindness-and-low-vision/using-technology/accessible-identification-systems-people-who-are-blind-1#:~:text=Using%20QR%20Codes%20for%20Identification&text=QR%20codes%20can%20relay%20a,tag%20items%20for%20later%20identification.
- "Introduction to Yolo Algorithm for Object Detection." Section, https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/#:~:text=YOLO%20algorithm%20employs%20convolutional%20neural,in%20a%20single%20algorithm%20run.
- Mallick, Satya. "Approximate Focal Length for Webcams and Cell Phone Cameras: Learnopencv #." LearnOpenCV, 19 Apr. 2021, https://learnopencv.com/approximate-focal-length-for-webcams-and-cell-phone-cameras/.
- "Biped Unveils an AI Copilot for Blind and Visually Impaired People at CES 2022." Designboom, 11 Jan. 2022, https://www.designboom.com/technology/biped-ai-copilot-blind-visually-impaired-people-ces-2022-01-10-2022/.
- "Distance Estimation Using Single Camera Opencv." YouTube, YouTube, 21 Nov. 2020, https://www.youtube.com/watch?v=zzJfAw3ASzY&t=341s.
- "Japan Firms Devise Tech Solutions to Help Visually Impaired Safely Navigate Stations." The Mainichi, https://mainichi.jp/english/articles/20200912/p2a/00m/0na/007000c.