CSU Bakersfield
School of Natural Sciences, Mathematics, and Engineering

HELPING THE VISUALLY IMPAIRED USING ARTIFICIAL INTELLIGENCE OBJECT DETECTION WHILE REDUCING LIGHTING DISTORTIONS

By: Jose Figueroa, Frankie Sanchez, Enrique Tapia, and Eduardo Ramos Soriano

Project Summary

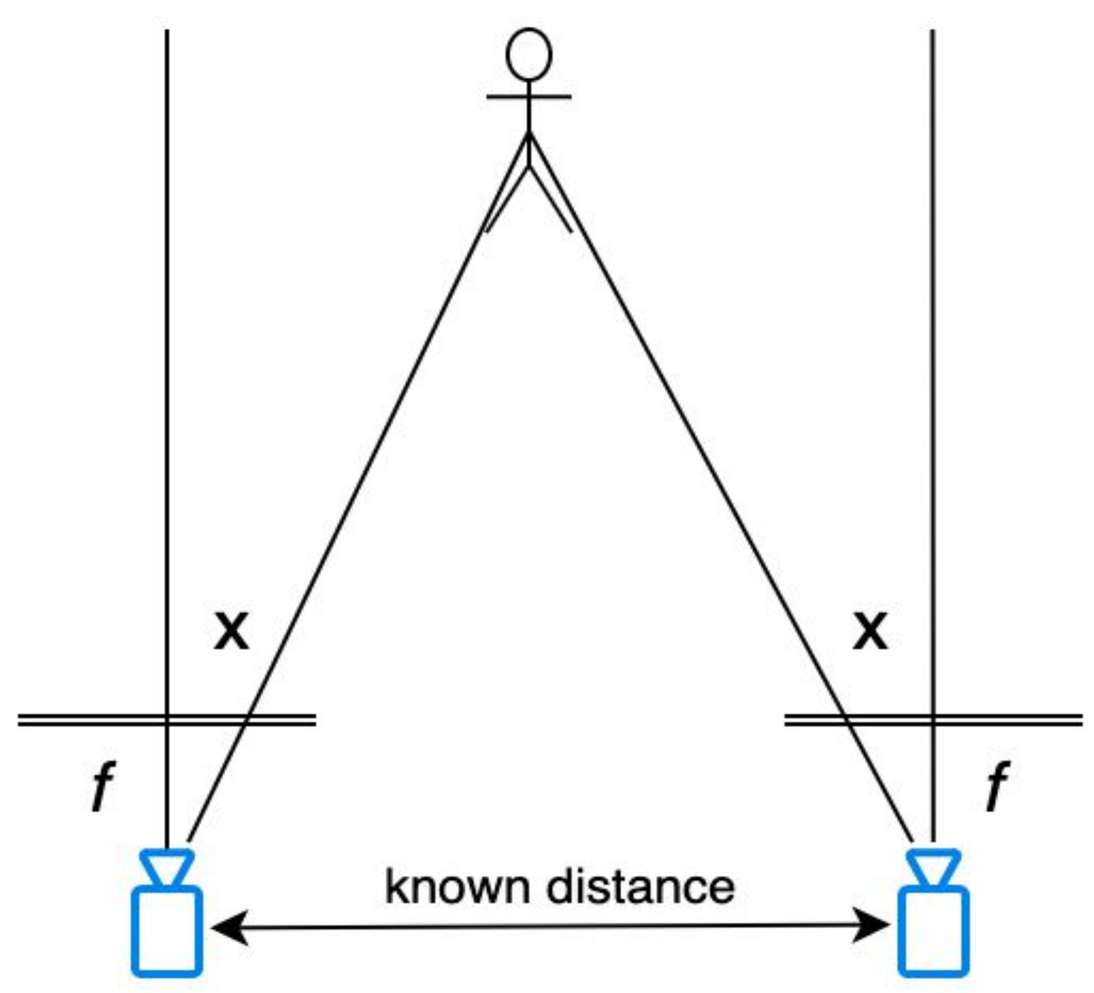
Our project aims at helping CSUB students who are visually impaired by using modern technology to assist them in avoiding objects that can obstruct them as they are walking across campus. To accomplish this, we developed an algorithm that works within the YOLOv5 architecture that helps supports our main goal, as well as implementing various hardware components to use with our algorithm.

Data Set

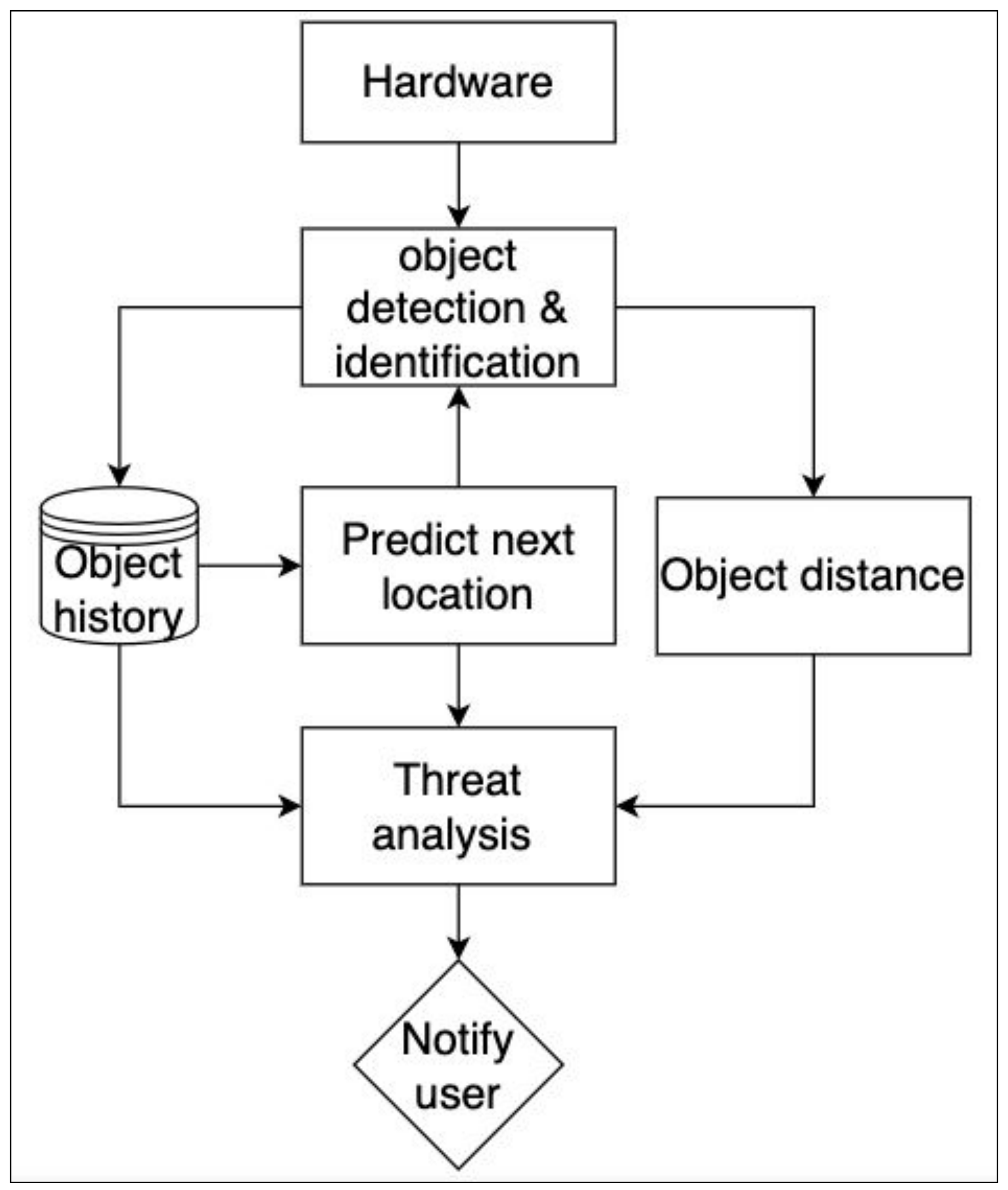
The data set is comprised of objects found on the CSUB campus that have the potential for causing harm to the visually impaired. Our data set includes objects such as vehicles, stairs, traffic signs, and much more. To help improve accuracy, the artificial intelligence system was trained with pictures taken during the day and in the late afternoon. For each image captures we made multiple copies and made changes to them such as flipping, mirroring, and placing a filter over the images. This allowed our object detection algorithm to have a better chance of identifying objects with greater accuracy.

How it Works

YOLOv5 is a state-of-the-art object detection algorithm that has become the main method of object detection in computer vision, a field of artificial intelligence. In our project, we created and trained a custom data set that's comprised of objects located at CSUB. The algorithm will detect objects on campus in real-time and alert the user when approaching them. Furthermore, our algorithm will calculate the distance of each object and run a threat analysis to determine if the user needs to be notified via a text-to-speech notification.



Simplified diagram of distance detection



Simplified overview of our project

The Algorithm

Our algorithm is responsible for handling the various hardware components along with processing and analyzing each object that is detected. For each object that is detected the algorithm assigns it a unique identifier and stores its location history. The algorithm then analyzes the object's location history and applies the Kalman filter to predict the location of the object in the next frame. It will then capture the object's distance and run a threat analysis to determine whether or not to notify the user.

Voice Recognition/TTS

The main objective of voice recognition and text-to-speech was to create an interactive and hands-free experience for the user. The voice recognition will allow the user to ask our project to look for specific objects by invoking a wake-work to activate the system. The text-to-speech gives the user feedback and notifies the user of any potential hazards that may cause harm.

Hardware

Our project supports the combination of various hardware configurations such as the use of one or two cameras and an ultrasonic sensor. By using a two-camera setup our algorithm can calculate the distance of objects from the user. Additionally, the two-camera setup will also help combat lighting distortions. The use of an ultrasonic sensor is optional but can be used to detect the distance of objects directly in front of the user.

Conclusion

Throughout this project we have gained experience in collaborating, implementing features, coding standards, adapting and revising a complex artificial intelligence project.