Daily struggles of the visually impaired

- ❌ Depth perception
- ❌ Identifying objects
- ❌ Navigating their way through


#2247251

The Plan

Depth Map Generation

- We use **triangulation** by clicking two images at the same time from slightly different viewpoints.
- This generates a **depth map** indicating the distances of objects from the observer.

Object Detection

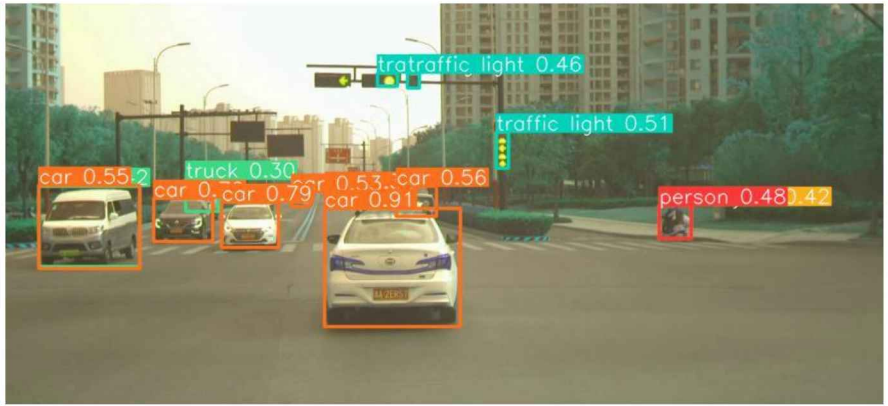
- Simultaneously, our application uses a ML model to **identify objects**.
- It is able to **classify** multiple objects and also **localise** them in the image provided.

Distance to Object Mapping

- Based on the depth map and the objects detected, we inform the observer about the **nearest objects**, their **directions** and **distances**.
- **For Example-**
  Car 20m away at your 11 O'Clock

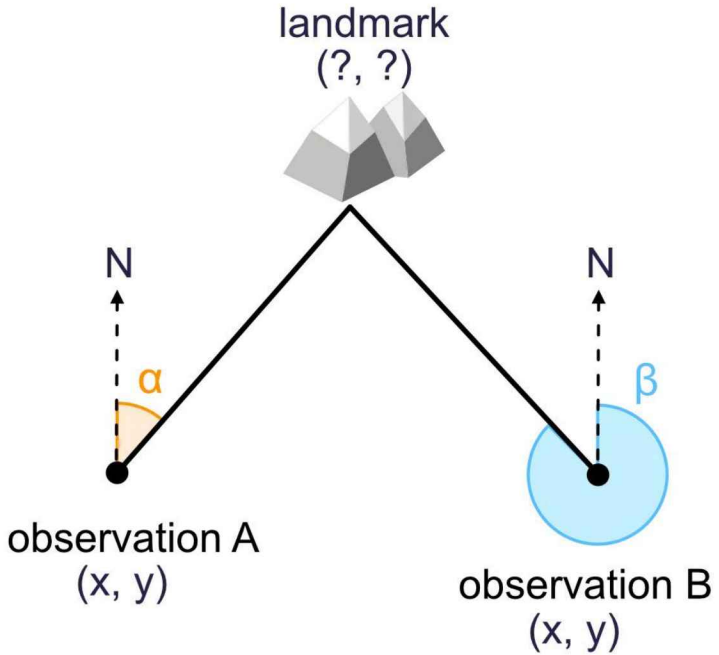
Object Detection

Our objective is to detect objects in the surroundings of the person and classify them. The closer objects will be notified to the user along with the distance and direction.



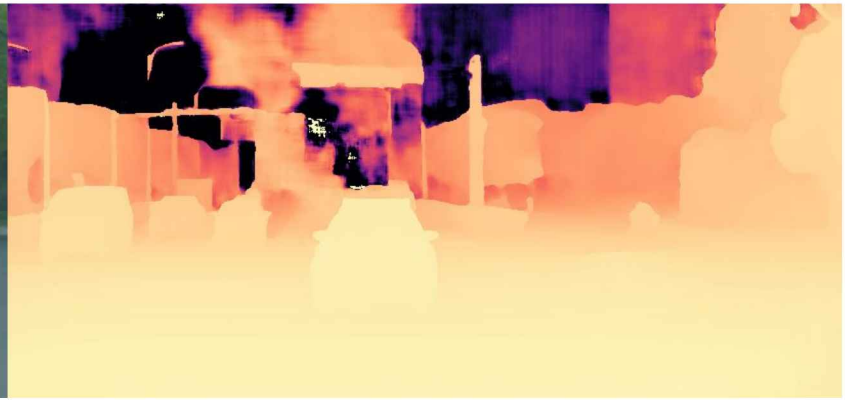
How does depth perception work?

- Our eyes take two images at the same time
- Using triangulation they are able to perceive how far an object is
- Smart phones commonly have multiple slightly displaced cameras which can mimic this.

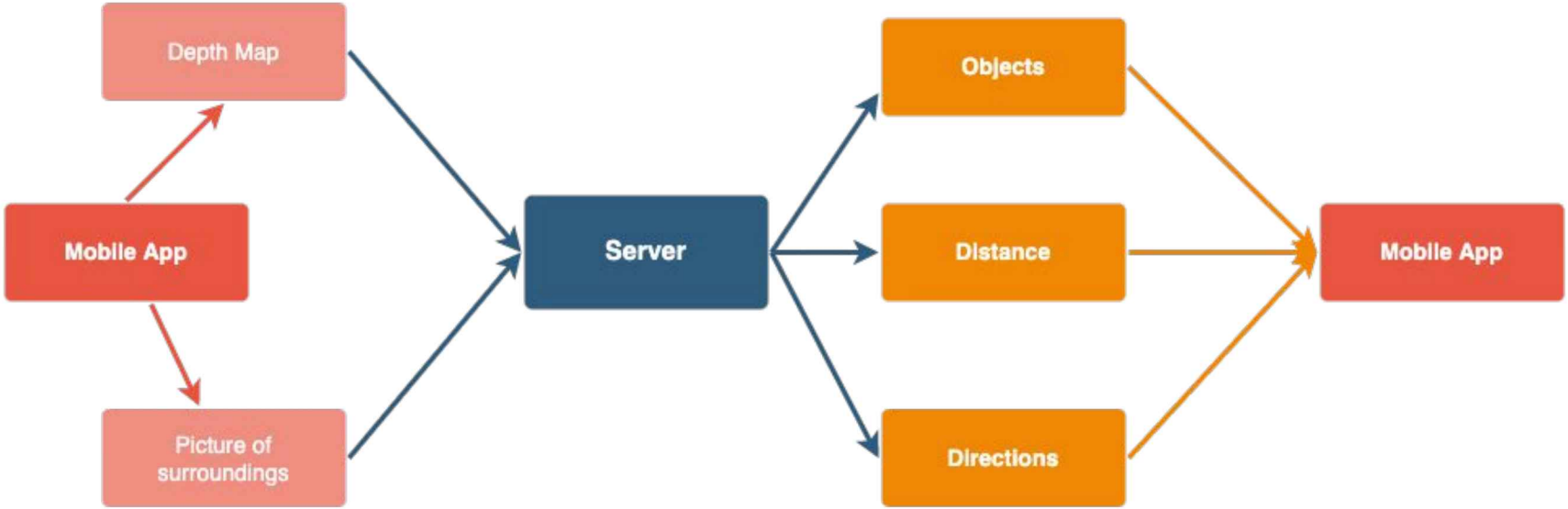


Depth Map Generation

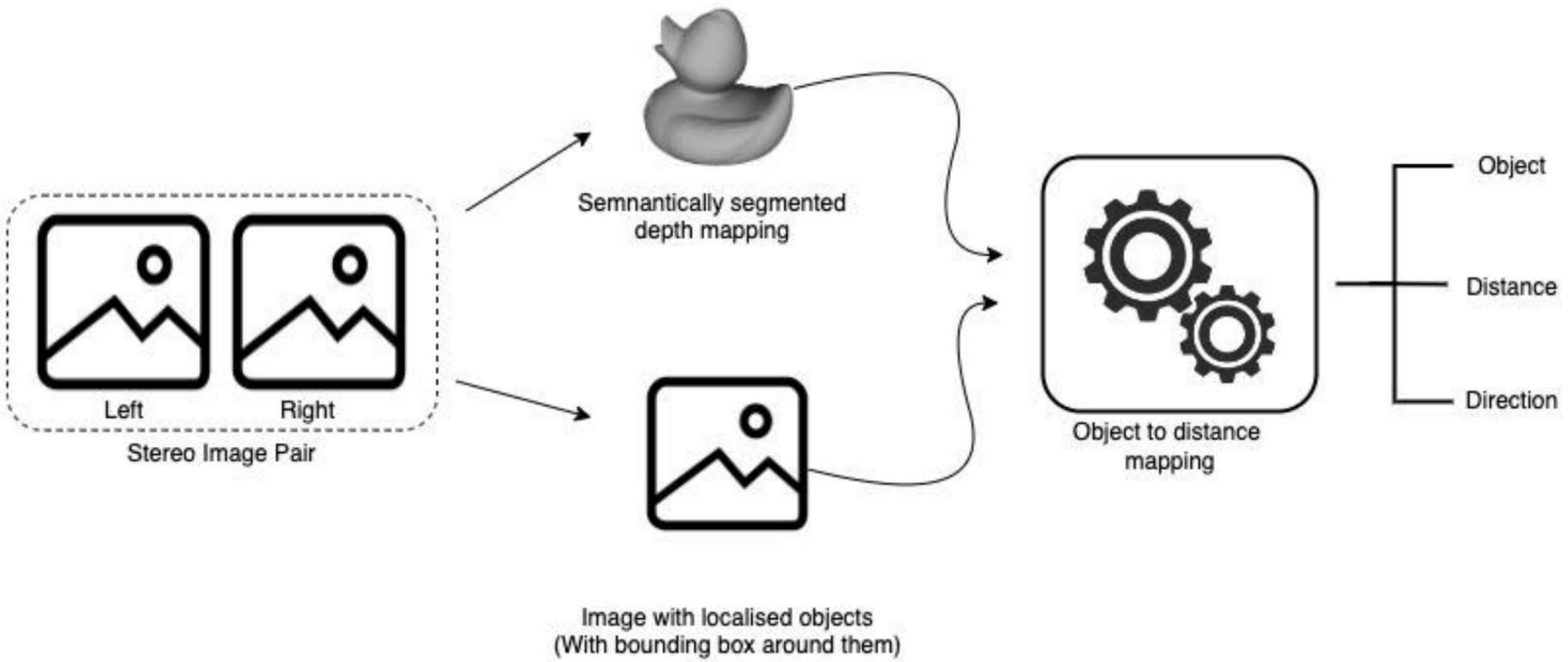
Using two images from slightly displaced but identical cameras depth can be estimated. A depth map is a colour coded representation of the depths of every pixel.



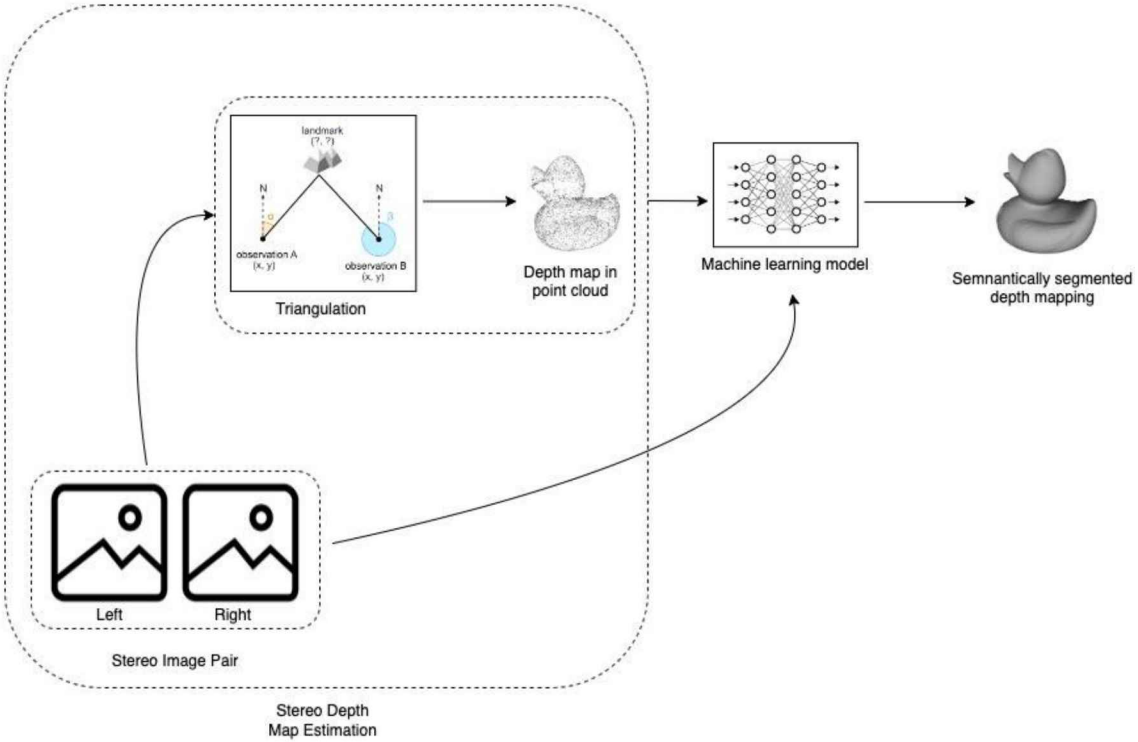
Mobile App Architecture



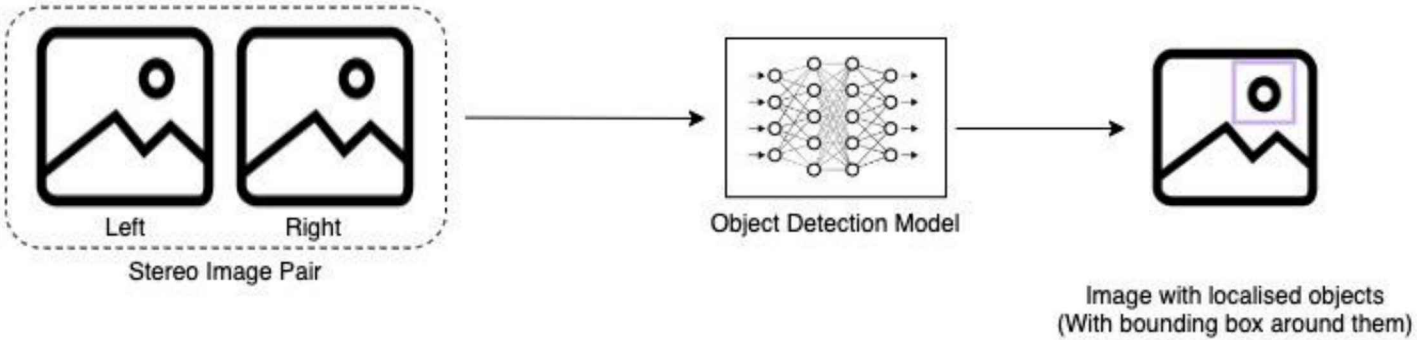
Machine Learning Model Architecture



M1 - Depth Map Estimation



M2 - Object Detection

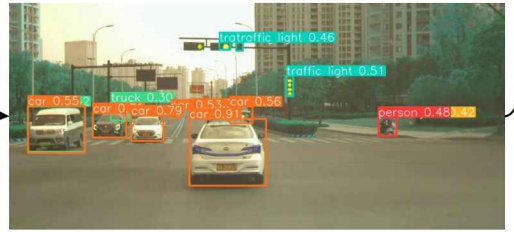


An Example

Left

Right

Depth Map



Object Detection

- A car is 17m at 11`o clock
- A car is 39m at 10`o clock
- A traffic light is 49m at 12`o clock
- A person is 45m at 2`o clock
- …. and a few more