Leveraging Computer Vision to Build a Label Scanner for the Visually Impaired

Rijul Mahajan

2.2 Billion

Globally, at least 2.2 billion people have a near or distance vision impairment. [WHO]

Background

- The majority of people with vision impairment and blindness are over the age of 50 years; however, vision loss can affect people of all ages. [WHO]
- By 2050, these figures are projected to double to approximately 2.01 million people who are blind, 6.95 million people with VI (visual impairment), and 16.4 million with VI due to uncorrected refractive error [Varma, 2016].
- The annual global costs of productivity losses associated with vision impairment from uncorrected myopia and presbyopia alone were estimated to be US$ 244 billion and US$ 25.4 billion, respectively.

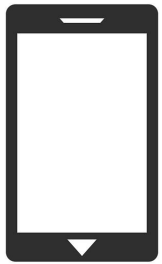
The Problem

- Adults with vision impairment often have lower rates of workforce participation and productivity and higher rates of depression and anxiety. In the case of older adults, vision impairment can contribute to social isolation, difficulty walking, a higher risk of falls and fractures, and a greater likelihood of early entry into nursing or care homes.
- Most devices and software are geared towards fully able people, meaning there is usually a large accessibility gap for the visually impaired, who form a sizable part of the population at almost 3.2 million, not counting the 1.02 million that are officially blind [CDC, 2020].

Our Investigation

Reading existing research papers on solutions for helping visually impaired people

Interviewing people from charities like Vision Loss Resources, and VI students

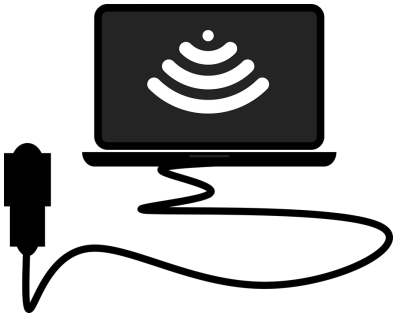


Using and examining existing technology for VI people

The Solution

- We propose a tool for identifying, reading, and answering questions about medical bottles.
- We identified that was one of the main areas that visually impaired people struggled with, and had no specialised solution. There were many general purpose OCR solutions, but we felt this could be a useful area to focus on.

Methodology



1. Image stitching: There must be a way to use a camera to capture the labels on the bottles and put them together in a machine-readable way. Usually this wouldn't be an issue, but we must account for the curvature of the bottle which makes it harder to put together a concrete image, and for people who may have shaky hands or may be pointing it at an angle.
2. Text Reading: Once the images have been stitched together, the text on it must be read and put together so that it can be understood and interpreted by the system.

Methodology, Pt. 2

3. FDA API: After the text has been read and stored, the plan is to look for an “NDC Number”, which is a unique code on each medical bottle that is assigned to it by the FDA. Once this number has been extracted and verified, the FDA has set up an online API called OpenFDA that can be used to search and return many results. One parameter that the API takes is the NDC number from above, and it returns a wide variety of information about the medicine back – everything from the generic names, recall enforcement reports to adverse events. The plan is to use this information in conjunction with the conversational agent to help the user.