GR-78 AI for Social Good: Assisting the Elderly in Reading Prescription Medications

Abstract

According to the Centers for Disease Control (CDC), roughly 12 million people in the United States above the age of 40 are visually impaired. In adults 18 years and older, visual impairment is one of the top 10 disabilities, and can have an enormous effect on one's independence and quality of life. Assistive technology through artificial intelligence can support the visually impaired in everyday life functions. Our application is specifically designed for reading text such as small print on prescription labels utilizing AI technologies of computer vision and natural language processing. After the user takes a picture of the text with the app, the text is extracted from the image, converted to speech, and played to the user in audio format. Overall, this app is a benefit to society through increasing the quality of life for the visually impaired.

Introduction

Our app is an AI project that assists visually impaired individuals. Americans are living longer, and our aging population is growing. We wanted to develop a system that assisted the aging population to live an independent lifestyle longer. Our goal is to improve the quality of life through using our app that takes text from an image and converts it to audio. Computer vision is an AI technology that allows a computer program to deduct meaningful information from an image. Utilizing computer vision technology through Google Cloud Vision API, the app extracts the text from the image. Natural language processing in AI is the ability for the computer program to understand human language. Then, using natural language processing through Google Cloud Text to Speech API, the app processes the extracted text into audio.

Research Questions

How will the application simplify prescription label identification for the visually impaired?

What features will simplify the usability of the application for the visually impaired?

What are future potential improvements that could be made to simplify the usability of the application that will cater toward the visually impaired?

Materials and Methods

The application was developed in Python using the Kivy framework. We utilized Google Cloud for computer vision and natural language processing. Draw.io was used to create images for screens.

The development steps were allocated into the following categories:

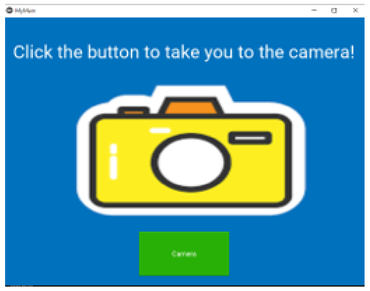
- General framework development (i.e., screen navigation, usability, and user experience features) using Kivy
- Camera functionality using Python and Kivy
- Image (.png) to text (.txt) file using Google Cloud Vision API which helps to integrate vision detection through optical character recognition into the application. This is performed through the feature TextDetection which extracts text from detection within an image.
- Text (.txt) to audio (.mp3) file that uses Cloud Text To Speech API using DeepMind's WaveNet and Google's Neural Networks. The Text to Speech API has 30 voices in many languages. DeepMind's WaveNet generates raw audio from deep neutral networks.

Overview

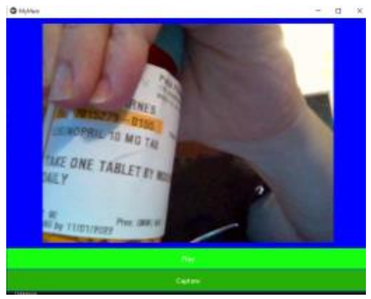
Implementing AI technology, the app is simple but powerful for the user.

The user enters the app and the welcome screen is displayed. The user is prompted to press the green button labeled "Click Here to Enter".

Another window appears that informs the user to "Click the button to take you to the camera!".



Once the camera is loaded on the third screen, the user clicks "Play" to engage the camera. Pressing "Capture", takes the image and saves the image file.



The text is then extracted from image, converted into text, and the text is converted to audio. On the fourth screen, the user presses the audio button, and the captured text is read aloud to the user.

Conclusions

This project makes an original creative contribution to the discipline of artificial intelligence through creative use of computer vision and natural language processing to serve a specific population and cater to their needs. It seeks to use AI technology to assist those who are visually impaired in reading prescription bottles. It has a large potential to be impactful within the field through its dissemination into the academic community. Through academic research, the effectiveness of this application can be studied and determined. If proven to make a positive impact in the intended community, this application can then be improved upon for use in the public.

The intellectual merit of this application spans wider than the current focus group for this project. Through the production of this application, there are many different studies that could be conducted to assist in advancing the knowledge and understanding of communities with low vision, reading comprehension, or health literacy. This could include understanding better how to assist persons that are not native English speakers or those who have a higher need for verbal communication over written communication.

Future research to test this product on visually impaired individuals would be the next step to determine the ease of use and necessity of the application and identify potential improvements for the application. The large text and bright colors on each screen and button were constructed to make the content easy to decipher by application users. Testing the features among a group could provide ideas to improve the application. Some improvements may include adjusting the button size, text size, altering the colors, and any other potential enhancements that could increase the ease of use. Another action that could advance the application's use is to add a function for language translation. One more action for future improvement could be to add audio features through each step of the application that allow users to listen to each prompt along the way.

Acknowledgments

Dr. Ying Xie

Contact Information

sallen16@students.kennesaw.edu
jbarne54@students.kennesaw.edu
lpope19@students.kennesaw.edu
chill165@students.kennesaw.edu

Scan to access our website

References

https://www.cdc.gov/grand-rounds/pp/2017/20170919-senior-aging.html
https://www.cdc.gov/visionhealth/basics/ced/fastfacts.htm
https://www.ibm.com/topics/computer-vision
https://monkeylearn.com/blog/nlp-ai/
https://cloud.google.com/vision/docs/ocr
https://deepmind.com/blog/article/wavenet-generative-model-raw-audio
https://cran.r-project.org/web/packages/googleLanguageR/vignettes/text-to-speech.html

KENNESAW STATE UNIVERSITY
COLLEGE OF COMPUTING AND SOFTWARE ENGINEERING

Stacie Allen, Jessica Barnes, Chenelle Hill, Lauren Pope
Dr. Ying Xie