A picture containing text, clipart

Description automatically generatedA picture containing text, soup, dish, wheel

Description automatically generated

Cairo University

Faculty of Engineering

Credit Hours System

CCEN480 – GP1

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Visually Impaired Logo Safety Sign 1.2 mm rigid plastic 100x100 :  Amazon.de: DIY & Tools  GP 1 REPORT | Virtual assistant for visually impaired people   |  |  | | --- | --- | | Ahmed Mohamed Ismail | 1180501 | | Mostafa Ashraf Ahmed | 1180406 | | Moaz Mohamed Mohamed | 1180528 | | Nader Youhanna Khalil | 1180477 | |

Table of Contents

[1. Introduction 1](#_Toc119151362)

[2. Surveying 1](#_Toc119151363)

[2.1. VoiceOver 1](#_Toc119151364)

[2.2. TalkBack 1](#_Toc119151365)

[2.3. Microsoft Soundscape 1](#_Toc119151366)

[2.4. Seeing AI 2](#_Toc119151367)

[2.5. Facing Emotions 2](#_Toc119151368)

[2.6. LookTel Money Reader 2](#_Toc119151369)

[3. Features 2](#_Toc119151370)

[4. Research Papers 3](#_Toc119151371)

[5. References 3](#_Toc119151372)

GRADUATION PROJECT REPORT

# Introduction

The project idea is about implementing a virtual assistant for visually impaired/ blind people to help them with their daily tasks and make their lives easier.

The interface will be in the form of an application. The user will communicate with a chatbot through speech.

# Surveying

The following are the already existing applications/technologies that try to achieve the same purpose.

## VoiceOver

It is a mobile application available in App Store that has the following features:

* Enunciate emails or other textual messages aloud
* Use braille for those who have a braille keyboard.
* Describes all the elements on the screen such as app icons.

## TalkBack

Same as VoiceOver however for android users.

## Microsoft Soundscape

Microsoft Soundscape is a product from Microsoft Research that explores the use of innovative audio-based technology to enable people to build a richer awareness of their surroundings

## Seeing AI

A Multipurpose app that permits reading and describing all documents placed under the smartphone camera such as banknotes or product barcodes. [Link](https://www.microsoft.com/en-us/ai/seeing-ai)

## Facing Emotions

It is an app that translates seven major emotions on the human face: anger, fear, disgust, happiness, sadness and surprise.

## LookTel Money Reader

This app allows visually impaired users to accurately count their money.

## Google Assistant

Google Assistant is a virtual assistant that help people in a variety of ways from controlling their home (set temperature, adjust lighting) to helping them with day-to-day operations (sending texts, scheduling an event).

# Features

Our final application will provide the following features/channels/modules:

* Short Text: speaks text as soon as it appears in front of the camera
* Documents: Provides audio guidance to capture a printed page and recognizes the text, along with its original formatting then speaks the document
* Products: audio beeps to help locate barcodes and then scans them to identify products
* Person: Recognizes friends and describes people around you, including their emotions
* Scene: Describes the scene around you
* Currency: Identify currency bills when paying with cash
* Light: Generate an audible tone corresponding to the brightness in your surroundings
* Color: Describes the perceived color
* Handwriting: Reads handwritten text
* Recognize Images in Other Apps: Describe images in your photo gallery and other apps including Mail, Twitter, WhatsApp
* Note Taker: Takes notes for you
* Event Scheduler: Schedules events for you.

# Research Papers

## Cost-effective and collaborative methods to Author Video’s scene Description for Blind People

|  |  |
| --- | --- |
| **Paper Title** | Cost-effective and collaborative methods to Author Video’s scene Description for Blind People |
| **Date of publication** | April 29 – May 5 2022 |
| **Place of publication** | CHI ’22 Extended Abstracts -, New Orleans, LA, USA |
| **Volume number and year** | - |
| **Authors** | Rosiana Natalie |
| **Link To Paper** | [Link](https://web.archive.org/web/20220506043755id_/https:/dl.acm.org/doi/pdf/10.1145/3491101.3503814) |

Summary:

This article discusses the need to provide accessibility of online video content for blind through audio descriptions. Earlier, we used to rely on professionals to author audio descriptions, but their service is costly and not readily available.

In this paper, they introduce four threads to create tools that are both time and cost-effective in providing good-quality audio descriptions which are:

* The development and evaluation of a mixed-ability collaboration authoring tool
* The formative study to uncover the feedback pattern from the reviewer
* The evaluation and generation of real-time supports for novice authors to write AD
* The design, development and evaluation of a system that demonstrates the utility of semi-automatically authoring AD.

Audio description (AD) verbally explains visual events that are not audible to blind users.

This research paper works toward making every video that has been and will be created in the future accessible for blind people using CV and NLP.

## Virtual Assistant for blind people

|  |  |
| --- | --- |
| **Paper Title** | Virtual Assistant for blind people |
| **Date of publication** | May 2021 |
| **Place of publication** | International Journal of advance scientific research and engineering trends. |
| **Volume number and year** | Volume 6 || Issue 5 || May 2021 || ISSN (Online) 2456-0774 |
| **Authors** | Avanish Vijaybahadur Yadav, Sanket Saheb Verma, Deepak Dinesh Singh |
| **Link To Paper** | [Link](http://ijasret.com/VolumeArticles/FullTextPDF/831_36.VIRTUAL_ASSISTANT_FOR_BLIND_PEOPLE.pdf) |

Summary:

This project proposes to use AI, ML, Image, and Text recognition to assist people who are blind or visually impaired. The concept is realized using and Android mobile App that includes features such as voice assistant, image recognition, currency recognition, e-book and chatbot.

It is a visual-based project consisting of few main components such as a camera, raspberry Pi, Sensors, Microphones and vibrators mountain together.

## A Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired and Blind

|  |  |
| --- | --- |
| **Paper Title** | A Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired and Blind |
| **Date of publication** | 9 March 2022 |
| **Place of publication** | Applied Science |
| **Volume number and year** | - |
| **Authors** | Aaron Raymond See, Bien Grenier Sasing and Welsey Daniel Advincula |
| **Link To Paper** | [Link](https://drive.google.com/drive/folders/1mQ-Gul-ri8XNKHMReG3P8dKI1zqSP_ev?usp=sharing) |

Summary:

In this research, they made use of a mobile phone with a depth camera function for obstacle avoidance and object recognition. It includes a mobile app that is controlled using voice and gesture controls to assist in navigation

. The proposed system gathers depth values from 23 coordinate points that are analyzed to determine whether an obstacle is present in the head area, torso area, or ground area, or is a full body obstacle. In order to provide a reliable warning system, the research detects outdoor objects within a distance of 1.6 m. Subsequently, the object detection function includes a unique interactable feature that enables interaction with the user and the device in finding indoor objects by providing an audio and vibration feedback, and users were able to locate their desired objects more than 80% of the time.

## An insight into smartphone‑based assistive solutions for visually impaired and blind people: issues, challenges, and opportunities

|  |  |
| --- | --- |
| **Paper Title** | An insight into smartphone‑based assistive solutions for visually impaired and blind people: issues, challenges and opportunities |
| **Date of publication** | 4 July 2020 |
| **Place of publication** | Springer-Verlag GmbH Germany, part of Springer Nature 2020 |
| **Volume number and year** | - |
| **Authors** | Akif Khan, Shah Khusro |
| **Link To Paper** | [Link](https://link.springer.com/content/pdf/10.1007/s10209-020-00733-8.pdf) |

Summary:

The paper reviewed research avenues in smartphone-based assistive technologies for blind people, highlighted the need for technological advancements, accessibility-inclusive interface paradigm, and collaboration between medical specialists, computer professionals, usability experts and domain users to realize the potential of ICT-based interventions for blind people.

It analyzes a comprehensive review of the issues and challenges for visually impaired and blind people with the aim to highlight the benefits and limitations of the existing techniques and technologies.

## Blind- Sight: Object Detection with Voice Feedback

|  |  |
| --- | --- |
| **Paper Title** | Blind- Sight: Object Detection with Voice Feedback |
| **Date of publication** | March-April-2021 |
| **Place of publication** | International Journal of Scientific Research & Engineering Trends |
| **Volume number and year** | Volume 7, Issue 2, March-April-2021, ISSN (Online): 2395-566X |
| **Authors** | A. Annapoorani, Nerosha Senthil Kumar, Dr. V. Vidhya |
| **Link To Paper** | [Link](https://ijsret.com/wp-content/uploads/2021/03/IJSRET_V7_issue2_211.pdf) |

Summary:

. Image classification techniques are used to identify the features of the image and categorize them into their appropriate class. The text description of the recognized object will be sent to the Google Text-to-Speech API using the gTTS package. Voice feedback on the 1st frame of each second will be scheduled as an output to help the visually impaired hear what they cannot see.

The following Modules are implemented: Image Capture, Feature Extraction, Object Classification and Speech synthesis

## Handwriting Recognition using Artificial Intelligence

|  |  |
| --- | --- |
| **Paper Title** | Handwriting Recognition using Artificial Intelligence |
| **Date of publication** | 2020 |
| **Place of publication** | (IJACSA) International Journal of Advanced Computer Science and Applications |
| **Volume number and year** | Volume 11, No. 7 |
| **Authors** | Sara Aqab, Muhammad Usman Tariq |
| **Link To Paper** | [Link](https://pdfs.semanticscholar.org/2590/ccf9445b96ef6ec17de8adae603f420517e2.pdf) |

Summary:

An artificial neural network is used to recognise handwritten digits and characters and convert them to an electronic format or a machine-encoded form. A neural network is used because it simulates how the human brain works when reading handwriting in a more simplified form. This paper focuses on artificial intelligence networks, machine learning, Hidden Markov Model, and the Support Vector Machine.

The modules used are image acquisition and digitization, preprocessing, segmentation, feature extraction, and recognition.

## Emotion Detection Algorithm Using Frontal Face Image

|  |  |
| --- | --- |
| **Paper Title** | Emotion Detection Algorithm Using Frontal Face Image |
| **Date of publication** | 2nd June 2005 |
| **Place of publication** | Institute of Control, Robotics and Systems |
| **Volume number and year** | - |
| **Authors** | Moon Hwan Kim, Young Hoon Joo, and Jin Bae Park |
| **Link To Paper** | [Link](https://koreascience.kr/article/CFKO200533239341594.pdf) |

Summary:

This paper proposes an emotion detection algorithm using a frontal facial image. There are three stages: image processing, facial features extraction and emotion detection. In image processing stage, the face region and facial component is extracted by using fuzzy color filter, virtual face model, and histogram analysis method. The features for emotion detection are extracted from facial component in facial feature extraction stage. In emotion detection stage, the fuzzy classifier is adopted to recognize emotion from extracted features.

The modules used are image processing, facial features extraction and emotion detection.

## Currency Detection and Recognition Based on Deep Learning

|  |  |
| --- | --- |
| **Paper Title** | Currency Detection and Recognition Based on Deep Learning |
| **Date of publication** | 2018 |
| **Place of publication** | IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). |
| **Volume number and year** | - |
| **Authors** | Qian Zhang and Wei Qi Yan |
| **Link To Paper** | doi:10.1109/avss.2018.8639124 |

Summary:

This paper proposes a method for currency detection. It uses CNN as a feature extractor under the framework of Single Shot Multi Box Detector (SSD) model. The image of the currency is first filtered and sent to the MLP layer for currency classification, finally the currency recognition is done.

The modules used here are image filtering, feature extraction and classification.

# References

[13 Must-have Apps for Blind or Visually Impaired People in 2022](https://www.inclusivecitymaker.com/apps-blind-visually-impaired-people/)

[26 Best Apps for the Visually Impaired](https://www.everydaysight.com/best-apps-for-visually-impaired/)