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](#_Toc119273961)

[2. Surveying 1](#_Toc119273962)

[2.1. VoiceOver 1](#_Toc119273963)

[2.2. TalkBack 1](#_Toc119273964)

[2.3. Microsoft Soundscape 1](#_Toc119273965)

[2.4. Seeing AI 2](#_Toc119273966)

[2.5. Facing Emotions 2](#_Toc119273967)

[2.6. LookTel Money Reader 2](#_Toc119273968)

[2.7. Google Assistant 2](#_Toc119273969)

[3. Features 2](#_Toc119273970)

[4. Research Papers 4](#_Toc119273971)

[4.1. Cost-effective and collaborative methods to Author Video’s scene Description for Blind People 4](#_Toc119273972)

[4.2. Virtual Assistant for blind people 5](#_Toc119273973)

[4.3. A Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired and Blind 1](#_Toc119273974)

[4.4. An insight into smartphone‑based assistive solutions for visually impaired and blind people: issues, challenges, and opportunities 1](#_Toc119273975)

[4.5. Blind- Sight: Object Detection with Voice Feedback 1](#_Toc119273976)

[4.6. Handwriting Recognition using Artificial Intelligence 1](#_Toc119273977)

[4.7. Emotion Detection Algorithm Using Frontal Face Image 1](#_Toc119273978)

[4.8. Currency Detection and Recognition Based on Deep Learning 1](#_Toc119273979)

[4.9. A Survey on Pixel-Based Skin Color Detection Techniques 2](#_Toc119273980)

[4.10. A Review of Various Handwriting Recognition Methods 3](#_Toc119273981)

[4.11. iGenda: An Event Scheduler for Common Users and Centralised Systems 1](#_Toc119273982)

[iGenda: An Event Scheduler for Common Users and Centralised Systems 1](#_Toc119273983)

[4.12. Smart Personal Task Scheduler 1](#_Toc119273984)

[4.13. A Smart Mobile App For Blind Users 1](#_Toc119273985)

[4.14. Smartphone apps for visually impaired persons 2](#_Toc119273986)

[4.15. Voice recognition system: speech-to-text 3](#_Toc119273987)

[4.16. Speech to text Conversion using Deep Learning Neural Net Methods 4](#_Toc119273988)

[4.17. Machine Learning for Handwriting Recognition 5](#_Toc119273989)

[4.18. Smartphone-Based Obstacle Detection for the Visually Impaired 6](#_Toc119273990)

[4.19. Clothing matching for visually impaired persons 6](#_Toc119273991)

[4.20. Real-Time Walk Light Detection with a Mobile Phone 7](#_Toc119273992)

[5. References 7](#_Toc119273993)

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 to help the user in tasks such as color matching for clothes.
* 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.

## A Survey on Pixel-Based Skin Color Detection Techniques

|  |  |
| --- | --- |
| **Paper Title** | A Survey on Pixel-Based Skin Color Detection Techniques |
| **Date of publication** | - |
| **Place of publication** | - |
| **Volume number and year** | - |
| **Authors** | Vladimir Vezhnevets, Vassili Sazonov, Alla Andreeva |
| **Link To Paper** | [Link](https://asset-pdf.scinapse.io/prod/1673002162/1673002162.pdf) |

Summary:

This paper discusses methods for detecting skin. This is extremely helpful in face detection. The paper discusses pixel-based skin detection methods. IT first discusses the different color spaces used. Then, it discusses the existing skin color modelling methods. Finally, it compares all of them together.

## A Review of Various Handwriting Recognition Methods

|  |  |
| --- | --- |
| **Paper Title** | A Review of Various Handwriting Recognition Methods |
| **Date of publication** | 2018 |
| **Place of publication** | International Journal of Applied Engineering Research |
| **Volume number and year** | Volume 13, Number 2 |
| **Authors** | Salma Shofia Rosyda and Tito Waluyo Purboyo |
| **Link To Paper** | [Link](https://www.ripublication.com/ijaer18/ijaerv13n2_44.pdf) |

Summary:

This paper discusses methods for recognizing handwriting. The image is first preprocessed. Then, the preprocessed image is passed to one of the recognition methods. The paper discusses the difference between them. Also, the paper discusses a method for correcting slope and slant as well as segmenting characters. It discusses a total of eight methods.

## iGenda: An Event Scheduler for Common Users and Centralised Systems

|  |  |
| --- | --- |
| **Paper Title** | iGenda: An Event Scheduler for Common Users and Centralised Systems |
| **Date of publication** | 2010 |
| **Place of publication** | Springer, Berlin, Heidelberg |
| **Volume number and year** | Volume 73 |
| **Authors** | Costa, Â., Laredo, J.L., Novais, P., Corchado, J.M., Neves, J. |
| **Link To Paper** | [Link](https://link.springer.com/chapter/10.1007/978-3-642-13161-5_8) |

Summary:

This paper summarizes the need for a scheduler assistant in our daily life as world is walking towards an aged society as a consequence of the increasing rate of longevity in modern cultures and with age comes the fact that memory decreases its efficiency and memory loss starts to surge. That’s why iGenda is used to help the user keep track of their daily routine and activities as well as manage their health by including a Centralized Management System (CMS) on the side of a hospital-like institution.

## Smart Personal Task Scheduler

|  |  |
| --- | --- |
| **Paper Title** | Smart Personal Task Scheduler |
| **Date of publication** | 2010 |
| **Place of publication** | Springer, Berlin, Heidelberg |
| **Volume number and year** | Volume 265 |
| **Authors** | Sai Swaroop Krishna, N., Krishna Tej, A., Suchithra, M. |
| **Link To Paper** | https://doi.org/10.1007/978-981-16-6482-3\_44 |

Summary:

The paper proposes the development of an efficient task scheduler. What makes it different from other virtual assistants such as Google Assistant, Alexa, Siri, is the fact that it takes into consideration world factors like traffic. This can be done by tracking the user current location of the user while managing the user schedule and reminding the user about an event.

The software uses technologies like Dialogflow and Google Maps APIs

## A Smart Mobile App For Blind Users

|  |  |
| --- | --- |
| **Paper Title** | A Smart Mobile App for Blind Users |
| **Date of publication** | July - August 2016 |
| **Place of publication** | International Journal of Computer Techniques |
| **Volume number and year** | Volume 3 Issue 4, July - Aug 2016 |
| **Authors** | Ms. Kalpanagayathri M., Ms. Sangeetha Lakshmi G |
| **Link To Paper** | [Link](https://www.academia.edu/37509222/A_SMART_MOBILE_APPS_FOR_BLIND_USER) |

Summary:

This paper introduces the concept of “Low vision Mobile App Portal”, which provides a way to access mobile apps specifically designed for visually impaired users. It describes some aspects of the design and shows preliminary results. The paper discusses what Universal Design means and questions whether everyday mobile apps are really designed to be used by everyone, shedding light on accessibility options of visually impaired people.

## Smartphone apps for visually impaired persons

|  |  |
| --- | --- |
| **Paper Title** | Smartphone apps for visually impaired persons |
| **Date of publication** | January 2019 |
| **Place of publication** | Kerala Journal of Ophthalmology |
| **Volume number and year** | - |
| **Authors** | John Davis Akkara, Anju Kuriakose |
| **Link To Paper** | [Link](https://www.researchgate.net/publication/338363439_Smartphone_apps_for_visually_impaired_persons) |

Summary:

This paper highlights the fact that while many low-vision aids products may be available in the market for visually impaired people to use, they are not generally used as one would expect. Other than the factors of affordability and accessibility, social stigma and denial are other factors. However, a smartphone is often available and the ability to use a smartphone can empower the visually impaired. Most of these apps do not require any extra hardware and many are free. There are apps designed to read out loud from the smartphone menus, voice control, read out books, and even read out text from photographs clicked by the phone. Some apps can identify colors, objects, and money

## Voice recognition system: speech-to-text

|  |  |
| --- | --- |
| **Paper Title** | Voice recognition system: speech-to-text |
| **Date of publication** | November 2015 |
| **Place of publication** | Journal of Applied and Fundamental Sciences |
| **Volume number and year** | - |
| **Authors** | Pranab Das, Vijay Prasad |
| **Link To Paper** | [Link](https://www.researchgate.net/publication/304651244_VOICE_RECOGNITION_SYSTEM_SPEECH-TO-TEXT) |

Summary:

This paper discusses voice recognition systems. The system consists of two components , first component is for processing acoustic signal which is captured by a microphone and second component is to interpret the processed signal, then mapping of the signal to words. Model for each letter is built using Hidden Markov Model(HMM). Feature extraction is done using Mel Frequency Cepstral Coefficients(MFCC). Feature training of the dataset is done using vector quantization and Feature testing of the dataset is done using viterbi algorithm.

## Speech to text Conversion using Deep Learning Neural Net Methods

|  |  |
| --- | --- |
| **Paper Title** | Speech to text Conversion using Deep Learning Neural Net Methods |
| **Date of publication** | 2021 |
| **Place of publication** | Turkish Journal of Computer and Mathematics Education |
| **Volume number and year** | Vol.12 No.05 (2021), 2037-2042 |
| **Authors** | Babu Pandipati, Dr. R.Rraveen Sam |
| **Link To Paper** | [Link](https://turcomat.org/index.php/turkbilmat/article/download/12651/9133/22435) |

Summary:

This paper intends to explore the various methods for conversion of speech-to-text that can be utilized in an email system that is based on voice. This method is built on the interactive voice response. The goal is to research and evaluate the different methods that are used in STT conversions, and find the most efficient method that is able to be adapted to both conversion processes. In the end, a model that uses HMM and ANN techniques to convert STT conversion is suggested.

Speech to Text conversion receives input from a microphone as speech, and later transformed into text which is displayed on the desktop. Speech processing is the research of signals and the different methods utilized for processing them.In this process, a variety of applications like speech coding, speech synthesizer, speech recognition and technology for recognition of speakers that use speech processing are employed.Among the mentioned above, speech recognition may be the most crucial one.

## Machine Learning for Handwriting Recognition

|  |  |
| --- | --- |
| **Paper Title** | Machine Learning for Handwriting Recognition |
| **Date of publication** | 2020 |
| **Place of publication** | International Journal of Computer (IJC) |
| **Volume number and year** | Vol.12 No.05 (2021), 2037-2042 |
| **Authors** | Preetha S, Afrid I M, Karthik Hebbar P, Nishchay S K |
| **Link To Paper** | [Link](https://core.ac.uk/download/pdf/327266589.pdf) |

Summary:

This paper aims to understand training computers to read handwritten letters and numbers. It also offers the distinction between Online character recognition and Offline character recognition. While Offline character recognition involves parsing image of document to series of texts and words, Online character recognition is a bit complicated process as it is a dynamic process. It involves of recognizing character data at the time of writing itself. It needs specialized writing pad and an electric pen. On the basis of movement of pen, the written character is recognized

## Smartphone-Based Obstacle Detection for the Visually Impaired

|  |  |
| --- | --- |
| **Paper Title** | Smartphone-Based Obstacle Detection for the Visually Impaired |
| **Date of publication** | 2015 |
| **Place of publication** | International Journal of Computer (IJC) |
| **Volume number and year** | vol 9279 |
| **Authors** | Caldini, A., Fanfani, M., Colombo, C. |
| **Link To Paper** | https://doi.org/10.1007/978-3-319-23231-7\_43 |

Summary:

This paper involves around exploiting the hardware and software of smartphones. It uses computer vision to extract images 3D representation of the scene and detect possible obstacles. images are captured by the smartphone camera and processed with a modified Structure from Motion algorithm that takes as input also information from the built-in gyroscope. Thus, the software is able to estimate the obstacles’ location and label them to avoid the user colliding with these obstacles.

## Clothing matching for visually impaired persons

|  |  |
| --- | --- |
| **Paper Title** | Assistive Clothing Pattern Recognition for Visually Impaired People |
| **Date of publication** | 2014 |
| **Place of publication** | IEEE |
| **Volume number and year** | vol. 44, no. 2, pp. 234-243 |
| **Authors** | X. Yang, S. Yuan and Y. Tian |
| **Link To Paper** | doi: 10.1109/THMS.2014.2302814 |

Summary:

As we all know matching clothes is one of the important steps when deciding what to wear but since visually impaired people face difficulties when it comes to color, this paper focuses on recognizing clothing patterns in four categories (plaid, striped, patternless, and irregular) and identifies 11 clothing colors. A camera mounted upon a pair of sunglasses is used to capture clothing images. The clothing patterns and colors are described to blind users verbally. This system can be controlled by speech input through microphone.

## Real-Time Walk Light Detection with a Mobile Phone

|  |  |
| --- | --- |
| **Paper Title** | Robust traffic lights detection on mobile devices for pedestrians with visual impairment |
| **Date of publication** | 2016 |
| **Place of publication** | ScienceDirect |
| **Volume number and year** | vol. 148, no. 2, pp. 123-135 |
| **Authors** | Sergio Mascetti, Dragan Ahmetovic, Andrea Gerino, Cristian Bernareggi, Mario Busso, Alessandro Rizzi, |
| **Link To Paper** | https://doi.org/10.1016/j.cviu.2015.11.017 |

Summary:

This paper contributes in helping pedestrians facing visual problems to cross streets when traffic light is safe to cross. The software uses [image processing](https://www.sciencedirect.com/topics/engineering/image-processing) and recognition, it proposes a robust setup for image capture that makes it possible to acquire clearly visible traffic light images regardless of daylight variability due to time and weather. The technique is also practical in supporting road crossing.

# 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/)