DIGITAL ASSISTANT TO AID INDIVIDUALS WITH PRINT DISABILITIES TO INTERPRET PRINTED MATERIALS

Project Id: 2022-024

Project Proposal Report

Sanduni Madara P.G.

B.Sc. (Hons) Degree in Information Technology

Department of Computer Science and Software Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

March 2022

DIGITAL ASSISTANT TO AID INDIVIDUALS WITH PRINT DISABILITIES TO INTERPRET PRINTED MATERIALS

Project Id: 2022-024

Project Proposal Report

B.Sc. (Hons) Degree in Information Technology

Department of Computer Science and Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

March 2022

A declaration, copyright statement, and the statement of the supervisor.

We declare that this is our work, and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or institute of higher learning, and to the best of our knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Name	Student ID	Signature
Sanduni Madara P.G	IT19392172	

The above candidates are researching the undergraduate Dissertation under my supervision.

Name of supervisor: Dr. Anuradha Jayakodi

Name of co-supervisor: Ms. Shashika Lokuliyana

Signature of supervisor: Date:

Signature of co-supervisor: Date:

Abstract

According to current statistics in the World Health Organization (WHO), approximately more than 2.2 billion people worldwide are visually impaired or blind [1]. Not only the blind or visually impaired, but also the ones with low literacy, may have difficulty reading paper documents. The print disability prevents a person from obtaining information from printed material in the traditional manner and necessitates the use of alternative methods of accessing the information.Vision impairments, blindness, physical dexterity issues, learning disabilities, brain injuries, cognitive impairments, and literacy difficulties can all lead to print disability [1].

Hence, they have the same need as everyone else to have access to all types of information for the same reasons: leisure, education, employment, and so on, it is critical to remove impediments to their needs. In a word, they also have the feeling of being an active part of the society we live in. There have been technological advancements in scanning and reading printed materials using a variety of software and apps. Existing applications are incapable of reading equations, images, and tables as accurately as sighted people [2].

Therefore, an innovative, effective procedure is needed to be designed to fulfill their rights. So, this study especially considers presenting an improved smart assistant which provides audio assistance to navigate through a smart assistant which functions auto focused image capturing, reading mathematical equations and table data of printed materials, classifying selected text or paragraph, images, graphs and reading aloud generated digitized text.

Keywords— vision impairment, print disability, printed material, smart assistant, print disabled individuals

Table of Contents

LIST OF FIGURES

Page

LIST OF TABLES

1. Introduction

1.1.Background & Literature survey

Vision, the most powerful of our senses, is essential in every aspect of our lives. Without that, it is difficult for us to interact with the rest of the world on a daily basis. Those who live long enough will be affected by at least one eye condition. When considering the world population,at least 2.2 billion people worldwide have a vision impairment or a blindness. From these, at least one billion cases, or about half of all these cases of vision impairment could have been avoided if they had been treated earlier [1]. To overcome this problem, Braille is the most widely used tactile communication technique for visually impaired and sighted individuals. However, due to the high expense of braille, it is not accessible to every single person [3]. So, in some circumstances, their only option for interpreting printed materials is to depend on a third party, which will be examined more in this study.

Globally, the concept of a smart assistant or digital assistant is gaining traction [4]. One of the most important and difficult tasks in developing such applications is creating a user interface that is suitable for visually impaired users, both in terms of providing input and interpreting output feedback. Today, mobile devices have become the standards for the implementation of assistive technologies to help people with physical and cognitive disabilities. They are increasing the capacity and sensor capabilities, as well as standard possibilities for touch based input and auditory-tactile output [4].

According to the referred information, a digital assistant is mandatory for print disabled individuals to interact with the rest of the world on a daily basis. As shown in Figure 1.1 which denotes the data retrieved by the survey implemented, it is validated that a digital assistant is mandatory for the individuals with print disabilities to interpret an image. It depicts that they are in a level of lack of effective resources to spend their daily lives. Therefore, it is mandatory to have a proper digital assistant to fulfill their needs and rights.

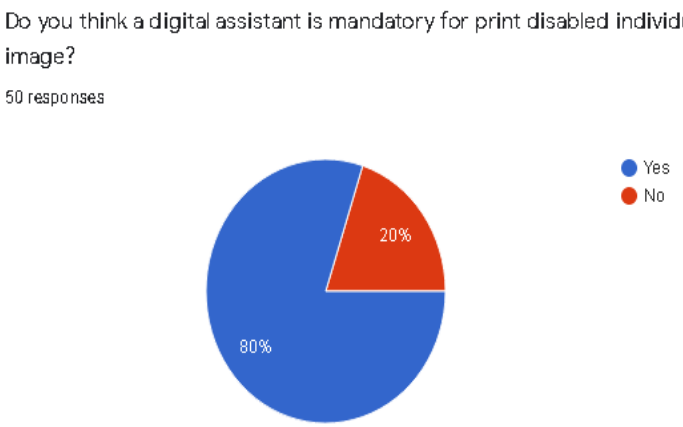


Figure 1.1 – Survey results on the importance of digital assistant toimage interpretation

However, as mentioned earlier Braille is the way print disabled individuals are reading and writing, it is a tactile writing system. In that case those individuals cannot get more effective or accurate information on what they want to read. Especially when related to images in printed documents. As shown in the survey, Figure 1.2 depicts it too. Some had commented on some reasons saying that braille is not always helpful for vision impaired people to identify an image. They have said that when it comes with images it does not accurately give the sufficient information to the user. According to the survey's retrieved commentary information, it is clear that even the blind require alternative access or accessible formats, and braille is not the best solution for them to gain information from printed materials. This alternative access can be a human or an assistive tool. But, people can't be around every time for you to get help in interpreting a printed document. And when it comes to a legal or a confidential document, in that case trusting a person and involving them into those is a risk. So, it depicts that, smart assistant is very important for a print disabled person to interpret printed materials.

Accordance with the collected information by the survey, the pie chart in Figure 1.3 indicates that most respondents suggest that Using a assistive tool is easier for a print-disabled person to interpret an image in a printed documents, interact with the rest of the world on a daily basis while the lowest percentage of response is for using braille method.

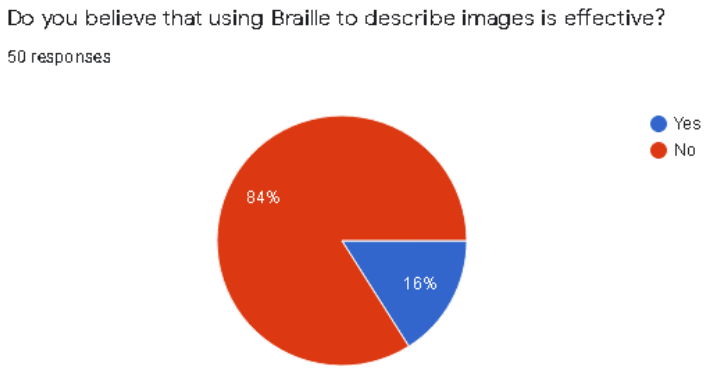


Figure 1.2 – Survey results on the effectiveness of braille for image interpretation

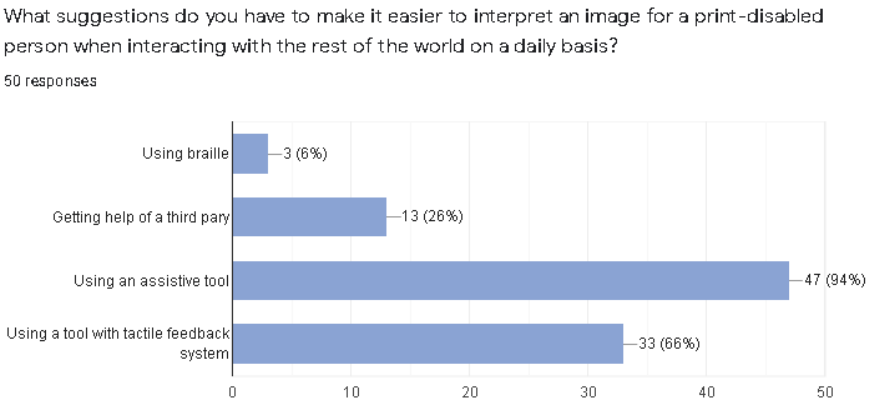


Figure 1.3 – Survey results on the suggestions for vision impaired people to interpret images

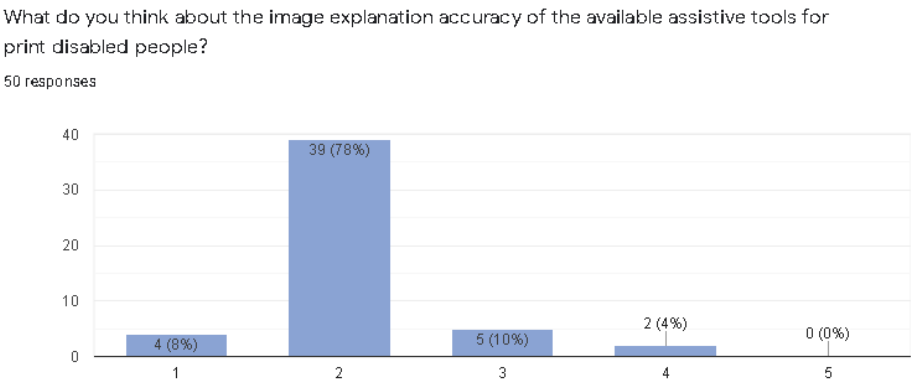


Figure 1.4 – Survey results on image explanation accuracy of the available tools

Furthermore, technological advancements have allowed for the scanning and reading of printed materials using a variety of software and apps. Existing applications are incapable of doing all reading equations, images, and tables as accurately as sighted people [2]. Focusing with the existing digital assistants, as shown in the Figure 1.4, mandatory responses are for low accuracy. Many say that the available tools are not giving an accurate interpretation of images in printed materials. The literature surveys of paper [5] highly depicts how existing similar systems function, as well as their strengths and weaknesses.

According to the findings of the literature surveys [6] and the information gathered from the survey, implementing a new smart assistant is an important thing for visually impaired people. So, this study aims to present an improved smart assistant with a haptic feedback system which provides audio assistance to navigate through a smart assistant to coexist in the society without feeling dependent on others too much. It functions auto focused image capturing, reading mathematical equations and table data of printed materials, classifying selected text or paragraph, images, graphs and reading aloud generated digitized text.

1.2.Research Gap

Comparing recent research on image interpretation, this component of the system is mainly focused on filling the image interpretation gap in the overall system. Studying the literature points, various tools have been built to interpret images in printed documents, and several assistive tools have also been implemented for print disabled people. However, the majority of them are missing some critical factors that should be improved for use by print-disabled people.

Image explanation can be done many ways, but when it comes to blindness, images should be described in such a way that a person who has been blind since birth can understand. The majority of existing tools, even those that are great and are available on the Play Store, do not provide an explanation for this factor [8]. Some tools doesnt even describe the basic colors in it [7], [9]. Since "Image captioning algorithm based on multi-branched CNN and Bi-LSTM" paper [7] is a great paper which uses an attention mechanism to get the key features of the image to describe it, is not done for print disabled people. So, it also lack the sufficient explanation for a blind person to understand an image in a printed document.

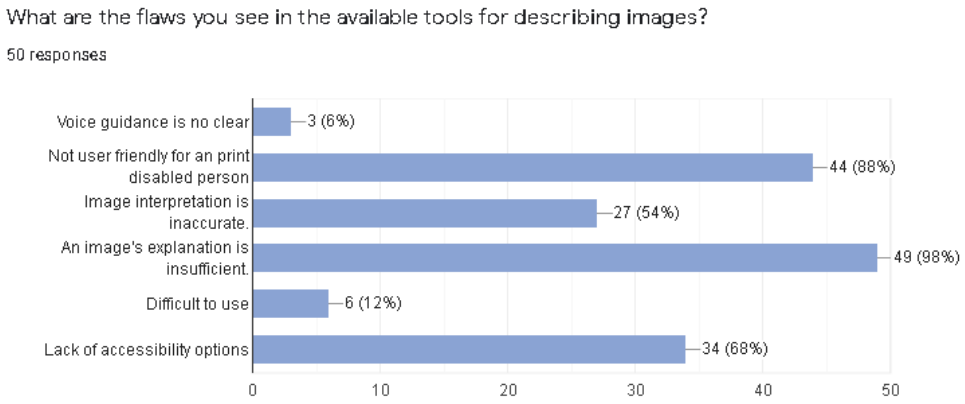


Figure 1.5 – Survey results on flaws in available tools

Addition to this, the majority of respondents in the survey stated that the image explanations of the available tools are insufficient. Figure 1.5 exemplifies this point. In consideration of Figure1.6, many people suggest that feature descriptions, color descriptions, and sufficient explanations be improved in existing tools for vision impaired people to accurately interpret images.

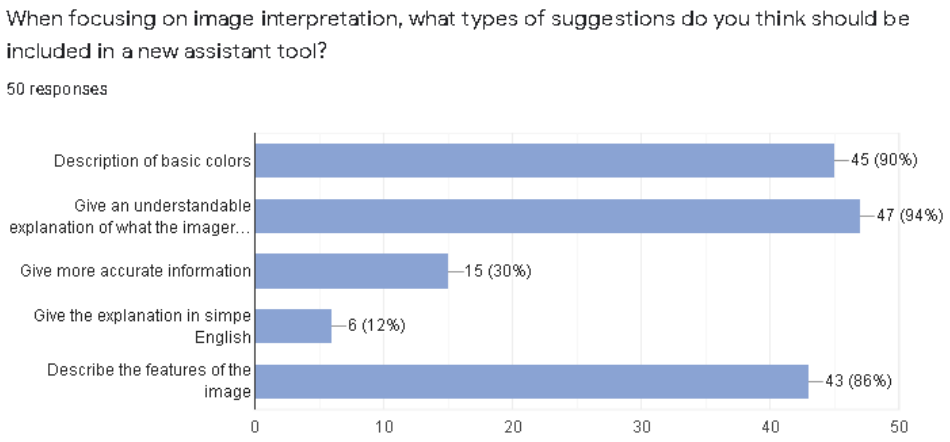


Figure 1.6 – Survey results on the suggestions for a new assistant tool

Table 1.1 – Research gap compared to existing systems

	Research A	Research B	Research C	This solution
Describing colors within the image	✗	✓	✗	✓
Describe the surrounding content other than the main subject	✗	✓	✓	✓
Introduce an attention mechanism to select key features	✓	✗	✗	✓
Describe the images in a way even for people who are blind since the birth can feel	✗	✗	✗	✓
Optimized for mobile/cloud use	✗	✓	✓	✓

Furthermore, Table 1.1 briefly compares the above mentioned issues in the existing systems with the suggested solution. By reviewing the outcomes, it is clear that this solution is implemented with far more innovative functionalities than other currently available investigations.

1.3. Research Problem

According to the survey conducted , it concludes that vision imapaired individuals have many difficulties when interpreting an image in a printed document. As a result of the survey, Figure 1.7 depicts that many people think identifying an image is very important for print disabled people. Figure 1.8 says that vision impaired people are most affected to interpret an image in the educational sector and health services. Paper [10] also depicts that because everytime they cannot ask help from a third party to interpret an image. Also few say that when working with a legal or a confidential document they also face many problems. Because trusting a person in such a case is not secure. Without any help by a third party, they cannot interpret an image in a printed document[2].

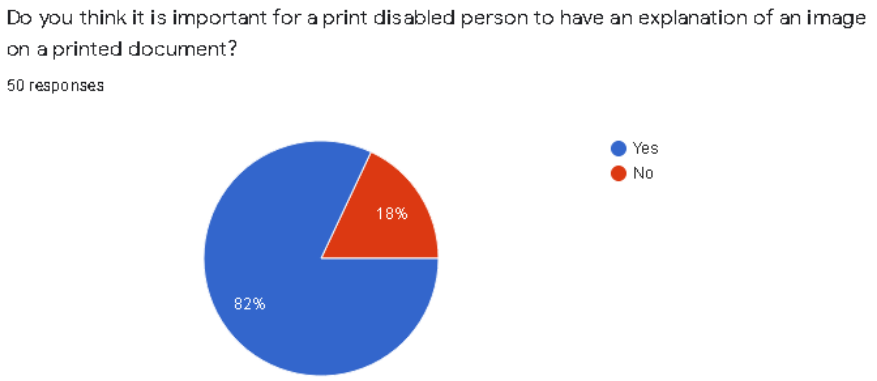


Figure 1.7 – Survey results on the importance of image explanation

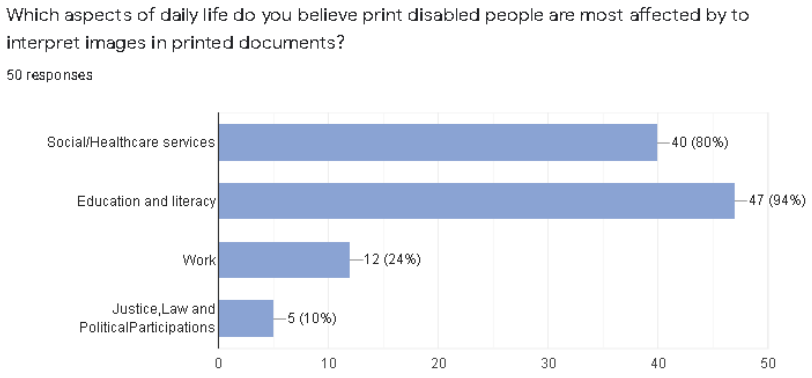


Figure 1.8 – Survey results on how print disabled people affected by interpreting images

Most of the print disabled individuals use braille to make their da today work easy[11]. According to the survey commentary section and the Figure 1.2, it says that braille cannot give a descriptive explanation to an image in a printed document and using braille is not practical when interpreting images. Concurring to the results of the surveys, Figure 1.6 and literature reviews the accessibility issues with the existing tools, lack of understanding, insufficicent explanation of the images and incapability to distinguish the needed information of an image in a printed document without depending on others are some other problems that are collected through the survey.

2. Objectives

2.1. Main Objective

The main objective of the research is to explore and identify the feasible system with a proper image interpretation model to facilitate reading a printed document. This study aims to identify the key features of an image and considering the other objects around the main features to give a proper explanation. The main target is to give the descriptive explanation of the image in simple English to understand it for a blind person who is blind since the birth.

2.2. Specific Objectives

In addition to the main objectives, there are some specific objectives related to the implementation

- To identify a proper image identification model to extract image features

First we have to identify the existing image interpretation models and introduce a new or improved image recognition algorithm to identify the images in a document with a proper interpretation model. And also we should identify the key features of the image to give an accurate explanation to the image. Captured images will be identified as the imputs and the new identifying model can be used to extract the features of the captured image.

- To identify a proper language model to generate a descriptive explanation

To convert the extracted image features into a natural labguage, we need to identify a proper language model to generate a descriptive explanation. Here, the extracted features can be output through an audio assistant in simple English.

- To develop a gesture-controlled process with a haptic feedback system in the mobile application

The aim of implementing a gesture controlled process is to make the user more comfortable and used for the interfaces to navigate through them. Moreover, the gesture-controlled interaction process will be added to the system by including the tactile feedback system, which will help the disabled user to navigate through the application by feeling the vibration. It will make the user to confirm whether he selects the correct button for the navigation.

3. Methodology

The overall research, which is under computer vision is to implement a digital assistant ti aid indiviuals with print disabilities to interpret printed documents. The individual contribution to this project is to explore and identify models to extract the image features and convert them in to a natural language. The captured image of the printed document is sent to a model to extract the features of the image. Then it identifies the main features with the respective sub objecs to describe the image efficiently. In this component, Convolutional Neural Network (CNN) will be used to extract features of the image. It is a neural network used in image recognition and extraction. And it automatically detects the important features without any human supervision. Then the extracted image is converted in to simple English through an audio assistant even to understand it for a blind person who is blind since birth. So, as the language model Long Short-Term Memory (LSTM), which is a recurrent neural network capable of converting data into natural language will be used. Trough the capturing process the user will be guided by an audio assistant and the gesture controlled process will also be guided via audio. As the datasets Flickr8k, Flickr30k and MSCOCO, which are the most commonly used datasets for image captioning tasks could be used.

3.1. System Architecture

3.1.1. Software Solution

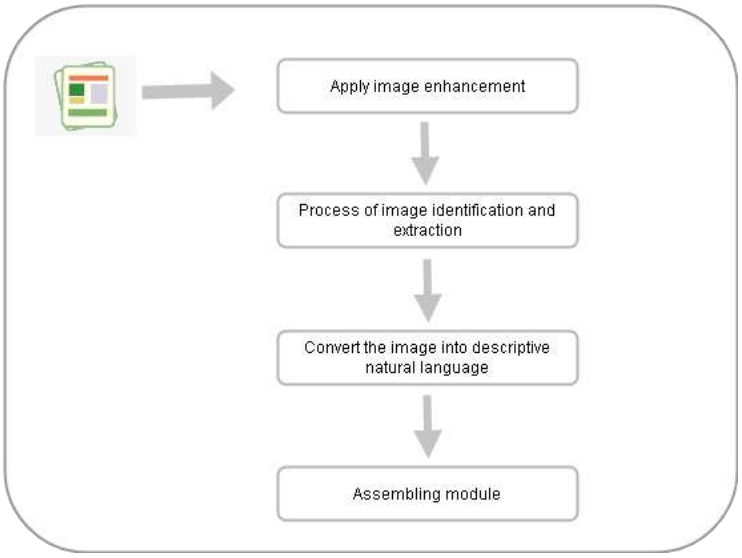


Figure 3.1 – System overview diagram

3.2. The flow of the project

3.2.1. Requirement gathering and analysis

The most important aspect of the research study is gathering and analyzing requirements. This phase captures all conceivable requirements of the system to be built, and before we begin implementing the function, we analyze the requirements that must be met for a perfect solution. The approaches described below are used to collect requirements.

- Read research papers and articles related to image interpretation
- Identifying the existing tools
- Distributed a survey to gather information

Going through the research papers is the best way to gather requirements to get a clear idea about the relevant component. While reading the papers we can identify the existing tools and by studying on them we can identify the gap that to be fulfilled by the component. And data and information can be collected through the survey to develop a proper tool that can be competitive to the existing tools.

3.2.2. Feasibility study

Schedule Feasibility:

The suggested system should be completed within the time frame specified. To ensure a quality product, each phase should be time-bound. The Gantt chart will show the time restrictions for each task.

Technical Feasibility:

Research team members should have some understanding of mobile application development technologies and training a model using machine learning methods. To complete the suggested application, all members of the research team should have the knowledge in computer programming languages for the implementations.

Economy feasibility:

There should be cost constraints for the product's resources. All members should be within the price range. The approach should be less expensive and more comprehensive.

3.2.3. Implementation

The system implementation is done by focusing on the gathered information to interpret images on printed documents. We have to identify a proper dataset and use it to train the model to have the expected outcome of the tool.

3.2.4. Testing

The product will be tested using different kinds of testing methods such as unit testing, integration testing and user acceptance testing. Because appropriate testing ensures that flaws and issues are found early in the application's life cycle. If there are any issues during the testing phase, these should be addressed before the product is released.

3.3. Project requirements

3.3.1. Functional requirement

- Extract data from the image
- Identify the objects of the image
- Describing the colors of the image
- Describe the main features of the image
- Generate meaningful captions
- Further description of the image using surrounded objects near main subject

3.3.2. Non-functional requirement

- Accuracy
- Availability
- Well optimized for cloud/mobile use

3.3.3. User requirements

- User should have a mobile phone to use the application
- User should have an English knowledge to understand the guidelines
- User should be able to hear
- User should have a simple knowledge to use a mobile applation
- User should be able to touch and feel the screen to navigate through the app

4. Budget and justification

Table 4.1 – Budget justification

Item	Cost(RS)
App publishing cost on google play	5000.00
Backend hosting cost	10000.00
Paper publishing cost	5000.00
Total	20000.00

REFERENCE LIST

[1] "Blindness and vision impairment," *Who.int*. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment. [Accessed: 11-Feb-2022].

[2] N. D. U. Gamage, K. W. C. Jayadewa, and J. A. D. C. A. Jayakody, "Document reader for vision impaired elementary school children to identify printed images," in *2019 International Conference on Advancements in Computing (ICAC)*, 2019, pp. 279–284.

[3] K. Smelyakov, A. Chupryna, D. Yeremenko, A. Sakhon, and V. Polezhai, "Braille character recognition based on neural networks," in *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, 2018, pp. 509–513.

[4] Á. Csapó, G. Wersényi, H. Nagy, and T. Stockman, "A survey of assistive technologies and applications for blind users on mobile platforms: a review and foundation for research," *J. Multimodal User Interfaces*, vol. 9, no. 4, pp. 275–286, 2015.

[5] P. Sankalpani, I. Wijesinghe, I. Jeewani, R. Anooj, M. D. J. T. H. Mahadikaara, and J. A. D. C. Anuradha Jayakody, "'smart assistant': A solution to facilitate vision impaired individuals," in *2018 National Information Technology Conference (NITC)*, 2018, pp. 1–6.

[6] P. Nawandar and V. V. Gohokar, "Design and development of multisensory smart assistive technology for blind persons," in *2018 International Conference on Research in Intelligent and Computing in Engineering (RICE)*, 2018, pp. 1–4.

[7] S. He, Y. Lu, and S. Chen, "Image captioning algorithm based on multi-branch CNN and bi-LSTM," *IEICE Trans. Inf. Syst.*, vol. E104.D, no. 7, pp. 941–947, 2021.

[8] B. Makav and V. Kilic, "A new image captioning approach for visually impaired people," in *2019 11th International Conference on Electrical and Electronics Engineering (ELECO)*, 2019.

[9] P. Shah, V. Bakrola, and S. Pati, "Image captioning using deep neural architectures," in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2017, pp. 1–4.

[10] A. Jayakody, S. Lokuliyana, A. A. T. Sampath, G. T. S. Silva, S. A. L. Rajanthika, and H. M. T. B. Seneviratne, “Mobile application for vision impaired people to facilitate to learn the English language,” *Int. J. Comput. Appl.*, vol. 138, no. 12, pp. 12–17, 2016.

[11] J. John, “Recognition of Documents in Braille,” *arXiv [cs.CV]*, 2017.

5. Appendices

5.1. Gantt chart

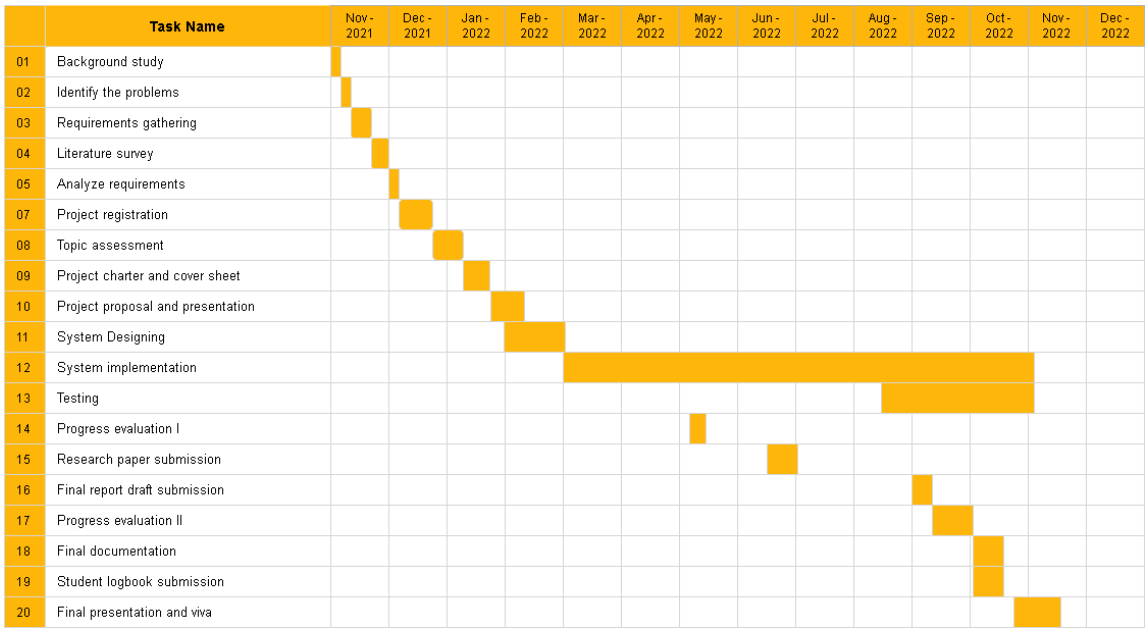


Figure 5.1 – Gantt chart

5.2. Work Breakdown Structure

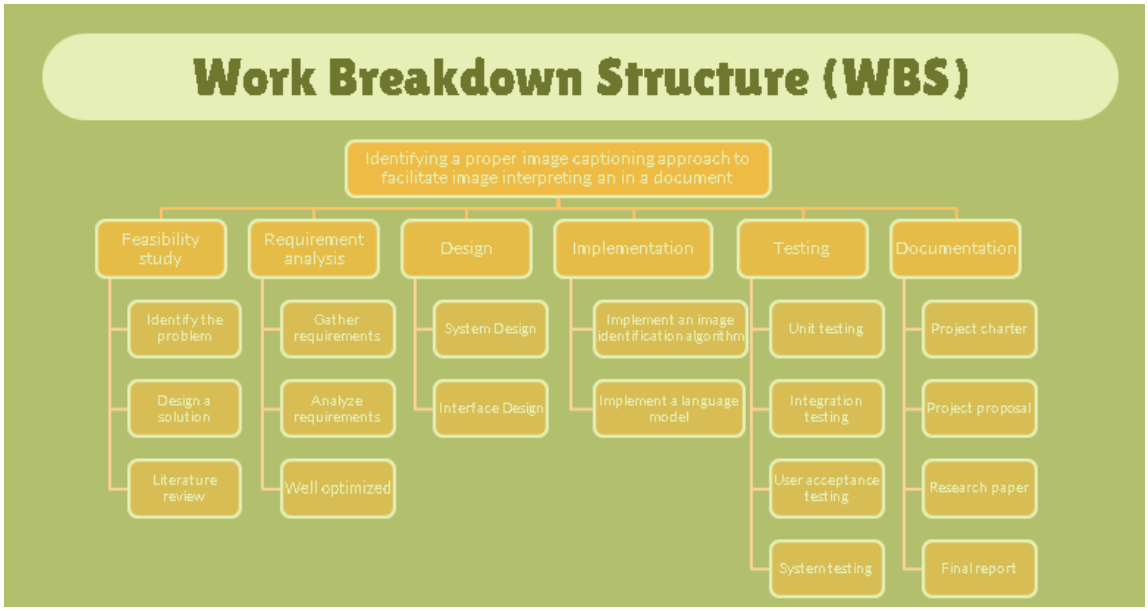
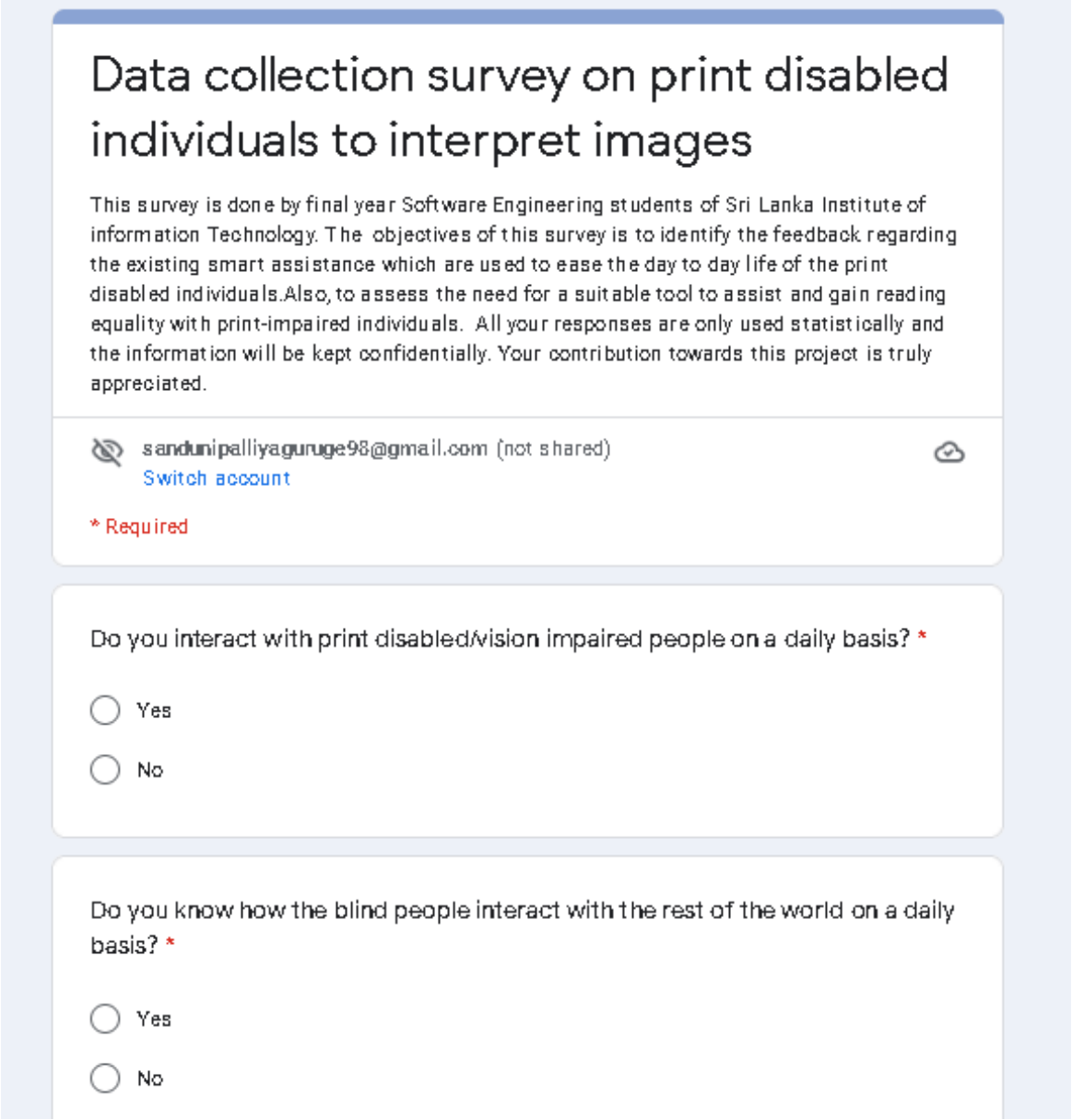


Figure 5.2 – Work Breakdown Structure (WBS)

5.3. Online Survey



Data collection survey on print disabled individuals to interpret images

This survey is done by final year Software Engineering students of Sri Lanka Institute of information Technology. The objectives of this survey is to identify the feedback regarding the existing smart assistance which are used to ease the day to day life of the print disabled individuals.Also, to assess the need for a suitable tool to assist and gain reading equality with print-impaired individuals. All your responses are only used statistically and the information will be kept confidentially. Your contribution towards this project is truly appreciated.

sandunipalliyaguruge98@gmail.com (not shared)
Switch account

* Required

Do you interact with print disabled/vision impaired people on a daily basis? *

- Yes
- No

Do you know how the blind people interact with the rest of the world on a daily basis? *

- Yes
- No

Which aspects of daily life do you believe print disabled people are most affected by to interpret images in printed documents? *

- [ ] Social/Healthcare services
- [ ] Education and literacy
- [ ] Work
- [ ] Justice, Law and Political Participations
- [ ] Other:

What suggestions do you have to make it easier to interpret an image for a print-disabled person when interacting with the rest of the world on a daily basis? *

- [ ] Using braille
- [ ] Getting help of a third pary
- [ ] Using an assistive tool
- [ ] Using a tool with tactile feedback system
- [ ] Other: _____

What do you think about the image explanation accuracy of the available assistive tools for print disabled people? *

	1	2	3	4	5	
Low	○	○	○	○	○	High

Do you think a digital assistant is mandatory for print disabled individuals to interpret an image? *

- ◯ Yes
- ◯ No

Have you ever seen a blind person using such a tool to interpret images in printed materials? *

- ◯ Yes
- ◯ No

Do you think it is important for a print disabled person to have an explanation of an image on a printed document? *

- ◯ Yes
- ◯ No

Do you believe that using Braille to describe images is effective? *

- ◯ Yes
- ◯ No

If not. what are your thoughts/suggestions for it?

What are the flaws you see in the available tools for describing images? *

- [ ] Voice guidance is no clear
- [ ] Not user friendly for an print disabled person
- [ ] Image interpretation is inaccurate.
- [ ] An image's explanation is insufficient.
- [ ] Difficult to use
- [ ] Lack of accessibility options
- [ ] Other:

When focusing on image interpretation. what types of suggestions do you think should be included in a new assistant tool? *

- [ ] Description of basic colors
- [ ] Give an understandable explanation of what the imagery represents
- [ ] Give more accurate information
- [ ] Give the explanation in simpe English
- [ ] Describe the features of the image
- [ ] Other:

Thank you for your time and consideration.

Submit Clear form