MAUU5900

Master's Thesis

in

Universal Design of ICT

May 2021

Improving image accessibility through touchup information

Nikhil Amatya

Department of Computer Science
Faculty of Technology, Art and Design

OSLOMET

Preface

This thesis will report on improving image accessibility through touchup information on mobile devices. The thesis topic "Improving image accessibility through touchup information" is proposed by Oslomet university to find the gap and address the problem faced by the visually impaired due to the inaccessible images. While working on the thesis, I get more familiar with the concept of Universal Design of ICT in terms of image accessibility and learn how to conduct the thesis research.

First and foremost, I consider myself fortunate to have been given the opportunity to conduct a thesis. This research will not be possible without the help of the professors' guidance, supervisor, numerous institutions, and individual respondents. I am grateful to my supervisor Raju Shrestha for his guidance, encouragement, and helpful advice. I also like to thanks my wife, Sampada Shrestha, who supported me every day.

Thirdly, I will like to thanks all the respondents who participate in my research, especially during the difficult situation of the Covid-19 pandemic. The situation was very challenging, but every participant showed great interest in prototype testing. I will also like to show my gratitude to my senior Suraj Gupta, who help me to contact and convenience the respondents to participate in the research.

Finally, I am very grateful to my friend and family who have helped me directly and indirectly conduct this thesis successfully.

Summary

Image is pictorial scripts or visual artifacts that are two-dimensional that resembles a subject that is usually a physical object. People with visibility disabilities find it challenging to access the image as it requires an equivalent image description to make it accessible. Due to its small size, it is even more challenging to make the image accessible on mobile devices. But there are different methods to make the image accessible for all. The developer can give alternative text for images on the web and content description on mobile to the images made accessible. But the problem arises when we try to access the image from the galley or the recently clicked image. We cannot find any method to add the content description or alternative text on the recent images on the mobile phone. So we can see a clear gap for image accessibility in mobile devices. During the literature study, many works have been done for the object detection on the image, image segmentation, and we can find many attempts to fill the gap but still visually impaired person are still facing image accessibility problem in mobile devices.

This research aims to fill the existing gap by design and develop an android prototype application to provide an accessible interface to access images for visually impaired people. The developed prototype mainly focuses on the interface part rather than other elements like image processing, object detection, or user notifications. The developed user interface will detect the finger movement in the application and will locate all the coordinates the finger has passed, and if the finger passes through the object, the prototype will notify the user about the object.

The research was conducted with the help of 24 respondents. The results from this research show that the developed interface which uses gesture (touch and slide) was easy to use and make the image accessible to the participants in the research. The touchup information makes it easy to understand the image, and combining two or more feedback methods helps visually impaired people to identify objects in the image.

Table of Contents

List of Table

List of Figures

List of Acronyms

ICT	Information and communications technology
UD	Universal Design
WHO	World Health Organization
WCAG	Web Content Accessibility Guidelines
XML	Extensible Markup Language
CNN	Convolution Neural Networks
IDE	Integrated Development Environment
IOS	iPhone Operating System

Chapter 1: Introduction

According to the World Health Organization (WHO), at least 2.2 billion people have vision impairment ("World Health Organization," 8 October 2019). In our modern society, people are surrounded by visual media tools and depended on these tools. Digital visual information is everywhere, and we must use it daily. However, smartphone applications are not capable of addressing all the accessibility problems as the study found that visually impaired users are finding difficulties in using the interface of the application as they want to feel what they are doing (Sultan, Siddiq, Rashid, & Farooque, 2015).

The person who has difficulty seeing or loses their vision is known as visual impairment. Such a person must face problems in day to day life like reading, walking, etc. We live in a highly competitive world where we must face many challenges and obstacles. Likewise, modern technologies will help to make our life and work easier and so for visual impairment. One of the best ways to convey messages or information can be through an image. Image accessibility can play a significant role and adds value to the development of technology, especially for people with low vision impairment. It is difficult for low vision people to get information about a thing in the right place and at the right time to make their day-to-day life more comfortable. It can be an effective way to provide the message and information through smartphones.

Research shows that images help make any kind of content more satisfying and quite easy to understand and learn what users want to get messages or information primarily for the people who have learning disabilities. The image plays as a cue card for a person having a vision problem (Gustafsdottir, 2020). In addition, it is essential to use proper alternative attributes. Otherwise, it will be difficult for low vision users to access or use the application to get the right information(Gustafsdottir, 2020).

Similarly, it is also necessary to analyze how people with low vision will interact with images. First, developers must consider how the visually impaired person screen reader users while using images, and text alternatives must be included for screen readers because text alternatives play a vital role for users when images do not load due to slow internet or any other technical problem. The study also shows that low vision impairment usually magnifies the images, and it is more beneficial for color blinded users ("Images | Usability & Web Accessibility," n.d.).

After image accessibility, object detection must be studied and scrutinized to make the application more effective. Recent research has shown that by the use of object detection, there has been progressing and success in many directions and applications (Verschae & Ruiz-del-Solar, 2015). Researchers mentioned that one of the main aims of object detection is to find and figure out all object instances of provided object classes, which can be either one or more than one despite location, pose, partial occlusions, and many more. Research shows that, in most of the computer vision system, as object detection allows to gain more information about the detected object and the scene so, it is the first task that is performed by the computer vision system. However, researchers have also mentioned that, even though object detection could be helpful for new applications and has been an integrated part of assistant driving technologies, especially for users of smartphones, developers and designers are not being used in many areas. Developers and designers must consider the object detection system for the betterment and development of the country and technology as well (Verschae & Ruiz-del-Solar, 2015).

Due to the rapid development of technology, the smartphone has been part of our life. By the year 2021, it is estimated that 40% of the world population will own a smartphone ("Smartphones - Statistics & Facts," 2020). Android dominates the share market for mobile operating systems taking nearly 72% share market worldwide ("Mobile Operating System Market Share Worldwide," 2020). Using a smartphone can be challenging for people with vision impairment.

These days many people are attracted to using a smartphone. Using applications of touchscreen smartphones can be challenging for the visually impaired. Still, on the other side, it will help them to get information quickly if it is designed with proper techniques, tools, methods, and many more. Such apps will help the visually impaired to save their life and effort. Hence, this research will help to investigate how developers can improve image accessibility in a mobile device with the use of touch up information for the visually impaired.

1.1. Statement of the Problem

In the present context, the smartphone's application has provided us facility to access several fields like transportation, healthcare, accounting, business, restaurant, games, find out the location, and many more. One of the challenging tasks for the developer is to make sure that mobile applications are accessible by the visually impaired. While designing a mobile app,

visually impaired users must not be neglected. According to "Article 9 - Accessibility | United Nations Enable" 2019), they had implemented the right regarding access to information intended for all users with appropriate technologies and accessible format for disable users so, that is why accessibility has been one of the main concern for developers these days.

If a minor technical problem or accessibility issues arise in the app, it will make it difficult to understand by visually impaired users. The problem such as the application cannot recognize the image, confusion in volume key with forwarding back key, the difficulty in zoom to access text, and many more. It means if the application is not accessible, we may lose users with visual impairments.

Accessing graphical information like images, maps and figures is a significant challenge for the visually impaired, and available applications or software do not convey meaningful information about graphical material (Giudice, Palani, Brenner, & Kramer, 2012). There are various tools and software available for accessing text-based content, but these solutions are relatively limited when handling the graphical material.

In the web application, the developer can give the alternative text for the images and make images in the web-readable for the visually impaired person. Similarly, the content description can be used in the mobile application in order to screen reader application to read the images. But the problem arises when the visually impaired person tries to access the image from the galley or the recently clicked image. There are no methods available to add the content description or alternative text of the recent pictures in real-time.

Yan and Ramachandran (2019) have studied the present condition of image accessibility in mobile applications. Study shows that accessibility issues in android mobile apps have a direct and indirect impact on people with disabilities problem. Researchers have also mentioned that such kinds of issues in the application do not harm the popularity of use because it is identified through the rating system in the application store. Also, Researchers suggested that for the betterment of a visually impaired person, developers must have proper accessibility training. Training must be related to the advance image view, button, and image button. Likewise, developers must focus on several issues like what are the missing elements in the application, why there is a lack of element focus, insufficient space between components, and many more. If developers focus on such issues and provide advance and effective mobile applications to the visually impaired, then the number of users will increase as it will help them to make their day to day lifestyle easier.

Dobosz (2017) has mentioned that developing and designing a mobile application for visual impairment has become a severe issue. Improvement in the field of mobile app technology might not go hand in hand with facilities for visually disabled people because any time enters new obstacles or limitations may emerge. Furthermore, improving image accessibility would be of interest to the business field and service for successful outcomes such as the healthcare department, and android application developers, and the government. However, there is much research carried out in improving image accessibility, but only a few kinds of research have been carried out, connecting image accessibility with touchup information. The researchers have not sufficiently examined and study regarding problem facing by visually impaired users in mobile applications.

From the foregoing, it is evident that there are conflicting views on providing images accessible in a mobile device for persons who are visually impaired. As mobile is a user-friendly device and easy to carry wherever we travel, so for the betterment of nationally and internationally, it is the developer's responsibility to design effective, easy to use, and understand features in the app for such users. Also, much research and study have been carried out by researchers regarding web accessibility for visual impairment users but few successful studies regarding image accessibility in the mobile app for such users. It shows that there is a requirement for further research in this regard.

1.2. Objectives

This thesis will provide insight to improve image accessibility in mobile devices through touch up information. The outcome of how disabled users interact with the mobile screen touch and get information from the image. So, the key objective of the study is to investigate the effect on image accessibility on mobile devices by touch up information. For achieving image accessibility for the visually impaired person, this thesis will design and develop a prototype of an android application that will provide an accessible interface to access the image on mobile devices.

1.3. Research Questions

Taking into consideration the background of the study, statement of the problem, and objectives, the following research questions try to give answerable inquiry to the problem faced by the visually impaired person while accessing the image in the mobile devices. To meet the prerequisites of this issue, the following two research questions will be answered:

R1: How does the developed touchup interface application help visually impaired users access the image on mobile devices?

R2: How easy is it for a Visual Impaired people to understand the image through the developed touchup interface?

1.4. Organization of the Chapters

Chapter 1: Introduction

This chapter discusses the overall concept of the topic and the problem for which the study is conducted, and the objective of the research.

Chapter 2: Literature Study

This chapter is about the theoretical background of the study which is divided into four parts image segmentation, object detection, user interface, and user notification. This thesis is mainly focus on the user interface, so the preference is given to this topic.

Chapter 3: Research Methodology

The third chapter research methodology describes the methodology adopted for this study. This chapter also describes the different methods used for prototype design as well as how the data is collected for the data analysis which will help to conclude the thesis.

Chapter 4: Data presentation and analysis

This chapter contains the interpretation and presentation of the data gathered and finding during the study.

Chapter 5: Discussion and Conclusion

The final chapter consist of discussion and conclusion are drawn with necessary recommendation and suggestion.

Chapter 2: Literature Study

The report tries to model the concept of improving image accessibility through touch up information on a mobile device. This thesis is conceptualized using four independent, yet related processes. The initial stage of the application is image segmentation when the image is open. The image is segmented to simplify and make the image easier to analyze. There are different methods and algorithms practicable for image processing. This literature study tries to find the best fit and practical image processing technique available currently. The second and most crucial process is object detection in this thesis. There have been several approaches for object detection, and many researchers are still working on it to find the best solution for it. The third part describes the interface part. How can the users interact with the mobile application? The users can interact with the mobile application in different ways, likes touching a different part of the screen, tapping, drawing gestures, and many more. The literature review tries to find the most suitable and best method to get the information from the image on the screen of the mobile device. The fourth and last one explains how the user can receive information from the application. Since the application of the thesis is going to be developed on the Android mobile operating system so, the literature study is mainly done in the specific so; the literature study is primarily done within the mobile operating system.

2.1.   Image Segmentation

Image Segmentation is a technique of dividing the image given to the system into multiple segments. The main goal of image segmentation is to make the image more simplified and easier to analyze. It is one of the crucial parts of image analysis (Raut, Raghuwanshi, Dharaskar, & Raut, 2009). There are several segmentation techniques available, and still, more research is being done to find the best methods for the segmentation.

Taşkın, Yiğit, Taşkin, and Taşkın (2017) in their work, develop and tested an Android application for real-time image segmentation using different filters, calculation, and marking from the OpenCV library. They describe the application process in the flowchart below

```
┌─────────────────────────┐
│ INPUT FRAME IN          │
│ RGBA COLOR SPACE        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ RGB to HSV              │
│ Conversion              │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Threshold Filter        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Weighted Sum            │
│ Calculation             │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Gaussian Blur Filter    │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Contour Marking         │
└─────────────────────────┘
```
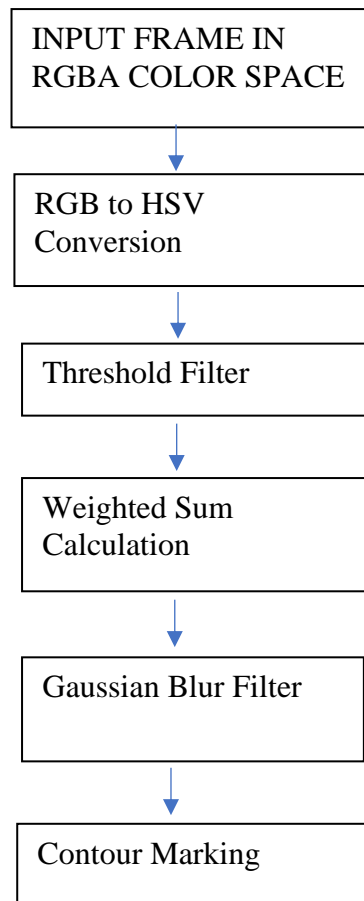
Figure 1: Flow Chart of image segmentation based on color (Taşkın et al. (2017).

In the first step, they converted each color value to 255 and obtained a conversion value between 0 to 1. Then they converted RGB to HSV using the OpenCV function cvtColor. After converting RGB to HSV, they select the red color for filtering. For thresholding, they used the inRange OpenCV function. The function addWeighted is used to calculate the weighted sum of two threshold output matrices such that the performance of color filtering has increased. They reduced the noise of the weighted sum value using a Gaussian blur. They used preexisting functions for blurring in this process. After applying the gaussian blur, contours, which is a curve joining all the consecutive points, are marked on the same color or intensity. After testing the application, they found that mobile devices are capable of coloring filtering. Dasila, Trivedi, Soni, Senthil, and Narendran (2017) used the OpenCV framework for image segmentation in their real-time environment perception system for the visually impaired. The OpenCV library helps segment the image by using the stereo matching mechanism in which the standard 2D model is converted into 3D and save to the point cloud structure. In the proposed system, image segmentation is performed in an interval of time to enhance the

perception of the environment. The segmentation process in the system is autonomous in which clearing the edges, refining, and segregating the image are performed to achieve the best and efficient result. The segmentation of the image is presented in the number of steps were at first, the device captures the image, and the system takes it as an input frame. Edge detection was done by applying Canny's Edge detection algorithm in the input frame. After the edge detection, the CV-RETR-TREE contour retrieval method has to perform to find out contours in the input frame. If there are any contour size less than a threshold, the contour is erased, and the obtained contours are overlaid in the original image. According to the contours, the image is segmented, and the mean center position is stored with a unique tag. Since the proposed system uses deeply nested contours, it cannot detect small objects like pins, pens if kept in a messy area.

Parikh, Shah, and Vahora (2018), in their deep learning-based object detection model, used morphological operations to segment images. The researcher adopted the Fuzzy Canny edge detection technique and morphological operations (dilation and erosion). To obtain the fuzzy-canny output, the fuzzy rule-based inference system is applied at first to get a fuzzified image, and then the canny algorithm is used to fuzzified the image. Using the morphological operation, they find the contours and draw a rectangle around the contours and finally crop the object from the original image.

2.2.   Object Detection

Object detection is a computer vision technique that locates the objects in the images or videos. When we see the pictures or videos, we can identify and find the object within a matter of moment. Similarly, object detection aims to replicate the intelligence of humans using the computer. For a visually impaired person, it is impossible to detect the objects in the images so, object detection using the system is essential for them to tackle the accessibility issue in accessing the image.

An object in the image is difficult to detect as there are many variations between images belonging to the same object category. The factors that make object detection difficult are viewpoint and scale, illumination, partial occlusions, and multiple instances (Taşkın et al., 2017).

Li and Zhang (n.d.) stated that CNN (convolutional neural network) could be the best solution to obtain high accuracy for object detection. The CNN based model was introduced by

Krozhevsky in 2012 then since it has dominated the leaderboard. Although the CNN model has higher accuracy, it has disadvantages like higher consummation cost, significant memory demands, and lower efficiency. The CNN based model uses tens or hundreds of layers for object detection, which increases computation cost, and the model has a lot of parameters which demand high memory to process.

Czúni, Kiss, Á, and Gál (2014) show the possibilities of lightweight object recognition on their demo application with the modified descriptor and different object segmentation. The researchers build the demo application in a client-server architecture in which they tested several image retrieval methods, and image segmentation approaches. For selecting the most appropriate lightweight method, they built the test databases with many test images. From their testing, they were able to show that the client-server application was capable of fast and robust object recognition and found that CEDD (Color and Edge Directivity Descriptor) was the most robust lightweight descriptor.

Elrefaei, Almasawa, and Al-gohany (2017) were able to detect objects in an image based on their color, shape, or local features. For color-based detection, they process the picture, which was in the phone gallery using the HSV color domain. They also use different existing techniques and algorithms for the detection of an object on the image. While doing so, they were able to detect eleven different colors, two-dimensional shapes like circles, rectangles, triangles, and squares, and they also correctly match local features.

Rajwani, Purswani, Kalinani, Ramchandani, and Dokare (2018) introduced a system in Android smartphones to help blind people in obstacle detection and navigation. The application detects the objects in the surroundings and alerts the users of the possible danger in the pathway. The system uses a camera to identify the object nearby and give feedback to the user in audio form, which can be easily understood by the visually impaired person. The researchers select the four-layered architecture for the system. The android application was developed using all preexisting tools readily available. The researcher uses the OpenCV library for image processing, and Google Cloud Vision API was used to compare and detect objects. The limitations of the system are that the application must always be open and have enough battery. The researchers also find that the user must always carry the android device, so it will be more convenient for the wearable device as the hand will be free.

Yashas, Maalik, Ashfaq, De Silva, and Ragulan (2017) build an IT solution for visually impaired people. The app takes the current location of the user by using the GPS of the mobile device

and gives the direction to the user for traveling from place to place. For doing so, the application uses a mobile phone camera with voice instruction. For object detection, the system uses Haar-Cascade, BLOB, and SURF algorithm. The authors have combined several methods to increase the application's accuracy and efficiency such that they claim that the application is more accurate than the white cane, which is used by blind people to navigate through the roads. Though the system is efficient and precise, the system does not work on dim and night lights. To overcome these limitations, the researcher hopes to future development using LED such that the build guider application can work in the nighttime.

Tosun and Karaarslan (2018) developed a prototype application aiming to make life easier for visually impaired people. The cheap and accessible virtual third eye is a mobile application with a designed chest apparatus. The researchers use image processing and machine learning technologies while building the virtual third eye. The app can also work on the background process when it is started. The application will also give audio feedback when the screen is a lock, and there is not any login process or account creation process such that it is as simple as possible. The user can hear the audio within the milliseconds after the application recognizes the object such that there will be no delay, and the user will notify directly. Presently, the application has two profiles, indoor and outdoor, and the profile needed to be selected manually by the users. The researchers also plan to make the selection process automatic by using the sensor in the future and add the feature of notifying the user of the distance and direction of the object.

Awad et al. (2018) were able to provide a useful set of features like light detection, color detection, object recognition, and banknote recognition in a single mobile application for visually impaired people. The different functions in one app make this application cheap, less complicated, and increases the practicality of the mobile application. For object detection and recognition, the object with their associated labels database is kept locally on the mobile phone such that the system response faster, and there will be no need for the internet connection for running the application smoothly. The researcher used the CNNdroid library, which is an open-source library for running trained Convolution Neural Networks (CNN) on Android. When the user captures the image of the object using the app, the object is compared with a predefined training set, which returns the user a list of the top five results of the object.

Shishir, Fahim, Habib, and Farah (2019) enable robust recognition of their research by segmenting the object from the background using the TensorFlow machine learning API such that the application runs without any remote server. The recognition problem is formulated such that the user knows the object through text to speech.
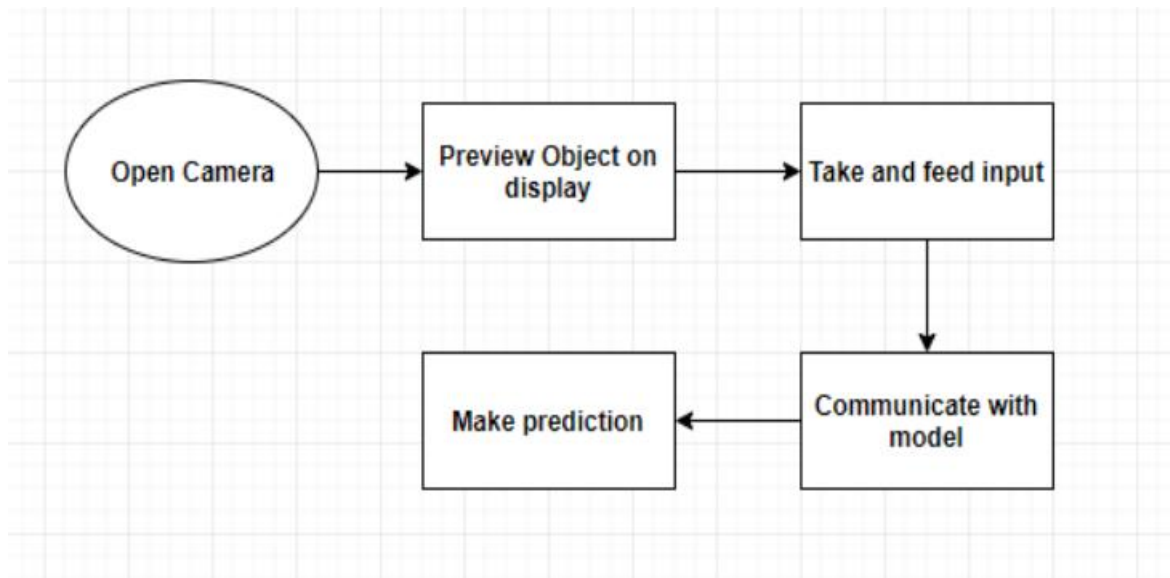

Figure 2: Flow chart showing an object detection algorithm (Shishir et al., 2019).

For object detection, at first, the google camera API is used that can detect camera hardware information and open the camera. View class is used to take the frame and later by using the TensorFlow machine learning framework for object detection. For detection, the input is feed to the algorithm, then communicates with the trained model, and finally makes the prediction.

TapTapSee is an application for mobile devices for recognizing and captioning the object capture by the camera or the image saved in the phone. The image from the phone is sent to CloudSight image recognition API via the internet and compared with the existing data. The result is shown in the form of text describing the scene in the image. A simple double-tap on the screen will capture the image, which will automatically analyze and identify the object. The application uses the mobile phone accessibility tool, VoiceOver for IOS, and TalkBack for Android to speak the object aloud ("TapTapSee - Blind and Visually Impaired Assistive Technology - powered by CloudSight.ai Image Recognition API," 2019). The application is effortless to use, but if there is no internet connection available, then the app is useless. The object recognition is done via the internet such that stable internet is always required to run the application smoothly.

Benhamida, Várkonyi-Kóczy, and Kozlovszky (2020) proposed a mobile-based application for real-time traffic statistics for autonomous driving cars system. The researchers used TensorFlow together with the transfer learning technique, which makes it easier to train the dataset on the pre-trained model. They used Single Shot MultiBox Detector (SSD) MobileNet V2 based model in their mobile-based system, which can detect multiple objects per image. The model returns the object coordinates of the rectangle that surrounds the detected object with the name of the object. The model which is trained is then converted to the TensorFlow Lite quantized model using TensorFlow Lite converter. The converted model is then compatible with mobile devices, which is four-time faster detection compared to other models on mobile devices.

2.3. User Interface

A user interface in the Android smartphone is a graphical and touch-sensitive display that allows users to interact with the devices. The mobile devices have a small screen, and users have to give input to the mobile devices using the touch in which they have to touch a particular part of the screen. This interacting method leads to accessibility issues for the visually impaired person. As they cannot see the screen, it will be problematic for these people to interact with the device. So, the user interface needs to be designed to allow users with diverse abilities to use it.

Dreyer, Grønli, and Bygstad (n.d.) proposed a unique solution to make the touch-base phone more applicable for visually impaired users. The researcher designed a UI that makes the application more applicable by operating through minimum gestures, not relying on physical keys, and haptic and verbal feedback. They conducted research based on Design research methodology. They developed a prototype on Android where interaction with UI is possible through two sensors, i.e., sound and feedback. By testing the prototype, they were able to limit the need for gestures to execute commands by giving information when the user taps the screen.

The designers and developers are still facing challenges while creating an accessible touchscreen interface. This is mainly due to a lack of understanding of how blind people prefer to use the touch screen (Kane, Wobbrock, & Ladner, 2011). To find out how blind people want to use the touch screen, the researcher conducted two user studies, which compare how blind and sighted people use gestures on the touchscreen. First, participants were asked to invent

gestures to perform everyday computing tasks, and secondly, participants were to perform a set of reference gestures. After analyzing the data, the researchers were able to give a design guideline for an accessible touch screen as follows:

- Avoid using print writing such as letters, numbers, or punctuation in gestures. If the print writing gesture is used, then the user should be trained, or there should be an option to select alternative gestures.

- The physical edges and corners of the smartphone can be very useful for the visually impaired person as these places can be located easily. The critical function of the application can be placed in these landmarks to increase the accessibility of the application.

- Finding a specific target on the screen can be a challenging task for blind people. We can reduce the demand for location areas by increasing the size or using the approximating targeting method in which the user can touch near the target and explore the adjacent area with a finger to locate the target.

- Blind people's speed of drawing the gesture may be varied than sighted people. Subsequently, using the gesture speed to recognizes the gesture may result in increased mistakes for blind users.

- Imitating the traditional spatial layout like the QWERTY keyboard or telephone keypad makes the interface easier for the blind user screen (Kane, Wobbrock, et al., 2011).

Shin, Lim, Lee, and Kyung (2013) proposed a new approach for information and interaction in mobile touch devices using the concept of 'background display' and 'foreground display.' Background display is indirect representation elements such as color, size, and location, whereas the foreground display is a direct representation element such as text. The researcher used a calculator program for the experiment in which they have size, shape, and information about the location of buttons as the background display. Altogether, 6 participants were blindfolded to examine the effect of the background display. The users were able to input rapidly, and all participants preferred a background display.

Buzzi, Buzzi, Leporini, and Trujillo (2015) explored the gesture preferences of visually impaired people on smartphones. The researcher developed the whole system that will collect the data of gestures used by respondents. The simple wireless capture gesture system, based on the client-server architecture, was able to take data from 3 devices at the same time.
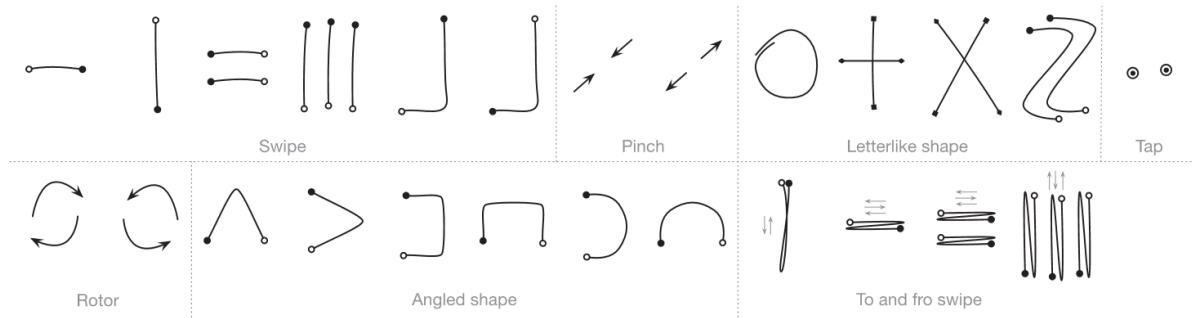
Figure 3: Different gestures used in the research (Buzzi et al., 2015).

The participants find to and fro, and rotor gestures most challenging group to perform and swipe mainly with straight or right-angled strokes, and taps group were most uncomplicated to the participants. The visually impaired peoples generally prefer simple gestures with a single finger and a single stroke.

Shishir et al. (2019) present an android application for recognizing objects using real-time objects and text detection by scanning. This application is different from others in the sense that nearly all other application requires to take a picture of the object and save it in the memory to detect the object. The application developed by the researcher does not need to take photos of the object; it just requires scanning the object and recognize them. This visual impairment user will not have to delete the images from memory. In the future, the researcher's purpose of adding voice command such that I will be more useful for fully blind people.

2.4. User Notification

The user notification is a method of providing the information back to the user when the user interacts with the device. When the user gives some input to the application, the application returns the information by processing the input. The app can notify the user in many forms, like displaying something in devices, voice output, vibrating the phone, or combing the different types of informing systems.

Kane, Bigham, and Wobbrock (2011) believe that the touch screens are inaccessible to blind people as the users need to locate the object on the screen visually. For the solution to this problem, the researcher developed an application called Slide Rule. The researcher tries to remove the accessibility problem by using the audio output and gesture input in their Slide Rule app. The blind and visually impaired encounter several issues due to a lack of accessible audio or tactile feedback while performing certain tasks on mobile phones. The research

shows that blind users were able to perform tasks faster and significantly improves the accessibility for blind users.

Giudice et al. (2012) designed and assessed an economical and intuitive methodology for accessing graphical materials by using the Vibro-audio interface. The system provides synchronous triggering vibration patterns and audio information whenever the users touch some elements on the touchscreen. After conducting the research, the researcher found out that the Vibro-audio interface can be the practical multimodal solution for giving visual information.

Awada, Issa, Tekli, and Chbeir (2013) evaluated the touch screen vibration accessibility for blind users. The goal of the research was to prove that the vibration touch screen can effectively perceive the contour-based image. To determine the statement, the researcher performed three experiments in which the respondents recognize basic shapes, simple geometric objects, and spatial relations between the simple geometric object. The result of the research shows that the accessibility method of the vibrating touch screen is feasible for accessing contour-based images.

The main problem encountered by blind users while using screen touch mobile is due to a lack of hardware key such that they cannot locate the specific point in the touch screen. An alternative method of providing the information is vocal, which may also be inaccessible in case of a noisy environment, meeting, during classes, and. So, the researchers purpose a vibro-tactile channel to improve interaction with the screen touch devices' speeches (Buzzi, Buzzi, Leporini, & Paratore, 2013). The investigator combined the haptic and audio feedback and developed the flexible add-on for enriching the UIs to improve accessibility. The customizable accessibility cue can be inserted into an XML file to render Android as a reactive colored ball, as shown below.

```
<org.cnr.iit.accessible.CustomAccessibleButton
  android:id="@+id/ballBtn01"
  android:layout_width="fill_parent"
  android:layout_height="20dp"
  android:layout_gravity="top"
  android:layout_margin="5dp"
  android:paddingTop="10dp"
  cnr:customMsg="custom spoken message"
  cnr.vibPattern="0,100,100,100"
  cnr:customSound="mySound" />
```


Figure 4: XML attributes and graphically rendered vibrating and speaking cues on a UI(Buzzi et al., 2013).

The researcher conducted a pilot study with two blind persons and found enthusiastic and positive feedback for the proposed solution. They suggested future work of using different short sounds to announce useful information like shifting focus from one area to another.

Chapter 3: Research Methodology

This chapter will present a clear idea about which method and techniques are implemented for conducting this research. The research methodology is an efficient way to tackle the research problem. The research methodology is directly connected with the objective of the study, which will discuss how the outcome of the research can be obtained (Sileyew, 2019). The research methodology chapter starts with the research design, data collection method, data analysis technique, and pilot testing. This chapter presents all the necessary steps to be followed throughout the research work, which will help to accomplish an objective of the thesis. Finally, in the last part, the attention is given to ethical consideration, which plays an essential role in the research.

3.1. Research Design

The primary objective of this study is to investigate the effect on image accessibility on mobile devices by touch up information. Furthermore, the result of the research will provide information regarding how the developed prototype will help visually impaired users access the image on mobile devices. There are different research methods available for conducting research, and finding a suitable way is essential for finding the conclusion of the research.

Qualitative research is investigative research, which is generally used to understand view and perception. This method of analysis helps to discover new thoughts and personal opinions. In this qualitative research method, different kind of unstructured or semi-structured practices for data collection like individual interviews, group discussion, diary and other are used for data collection whereas for the quantitative data collection is done in more structured approach such as surveys, online polls, telephone interviews, and systematic observation. The quantitative research method uses numerical data or data that can be converted into useable statistics to conclude the research.

Qualitative research is a real-life inquiry that focuses on the "why" and "how" rather than the "what" of social phenomena and depends on the direct experience of human beings. Qualitative research collects data from the participant's point of view and gathers non-numerical data. Quantitative analysis is more like a structured survey, whereas qualitative research is more than a structured study as it uses dynamic processes such as interviews and group discussions. Mainly, the qualitative research method is designed for the target audiences' behaviors connected to topics or issues. It utilizes an in-depth analysis of a small

group of individuals for building theories such that the result of the research is not predictive but descriptive. So, for this thesis, the qualitative research method is chosen as the qualitative research method offers a deeper understating of the topic and can be done with smaller sample size.

3.1.1. Prototype design

For the development of the prototype, the java programming language is used, and Android Studio is used as the integrated development environment (IDE). Currently, Android Studio is the best IDE, which provides the best-unified environment for developing android applications. As discussed in the earlier chapter, for object detection and image segmentation, there are many different methods available, and still more research are being done on the topic. For this prototype development, TensorFlow, which is an open-source platform for machine learning, is chosen. TensorFlow bundles together machine learning and deep learning models and algorithms which can be trained. TensorFlow Lite is a set of tools that helps to run the TensorFlow model on mobile devices and enable machine learning inference on the mobile device, which has limited power and processing speed. TensorFlow Lite is a lightweight version specially developed and designed for mobile devices, which makes machine learning on devices easier and on-device, which means no data is sent to and fro from a server. Using TensorFlow Lite in the prototype have the following advantages

- Latency: there is no round-trip of data to the server so, there will be no delay in object detection

- Privacy: the data remains inside the devices, so there will be no issue on the privacy

- Connectivity: Internet connection is not required as all the process is done within the device

- Power consumption: the prototype will not need an internet connection, which will save a lot of power. ("Object detection | TensorFlow Lite," n.d.)

While developing the prototype, already build a pre-trained model from TensorFlow is used, which detects 90 classes of objects. The developed prototype uses the same model SSD MobileNet model as in the research of Benhamida et al. (2020). The pre-trained models can be customized as per the need and re-train again with new objects.

Accessibility is crucial while developing a prototype, which should be considered early in the design phase. WCAG 2.0, which is updated to WCAG 2.1 in 2018, covers mobile accessibility.

The most commonly occurring accessibility issues on the mobile application are missing labels, inadequate touch targets, and poor contrast between foreground and background ("Starting Android Accessibility," 2020).

```xml
<ImageView
    android:id="@+id/add_gallery"
    android:layout_width="match_parent"
    android:layout_height="100dp"
    android:layout_centerHorizontal="true"
    android:layout_centerVertical="true"
    android:contentDescription="@string/add_photo_gallery"
    android:src="@drawable/ic_add_photo" />
```

Figure 5: Using label contentDescription on the ImageView

One of the issues of accessibility is the missing label, which can eliminate by given special care while developing the application. The prototype developed uses an ImageView in which the label contentDescription is given such that when screen reader application will read the contentDescription while using the application.

Add Image

Please press anywhere to open gallery

The second accessibility issue is inadequate touch targets, which can be a significant issue for people having a vision problem. Finding a small target like buttons or the icons on the interface can be very difficult if they are not placed in the right place. The developed prototype user can press anywhere on the screen to open the gallery where they can select the image from the gallery. So, there will be no problem for the user to open the gallery in the prototype.

For the poor contrast between foreground and background, "Make apps more accessible" n.d.) suggest that if the text is smaller than 18pt or text is bold and shorter than 14, then the color contrast ratio should be at 4.5:1, and for other text, the contrast ratio should be 3.0:1. The developer can also use the Accessibility Scanner app developed by Google to check any accessibility issues due to contrast. While checking the prototype with the Accessibility Scanner app, the application does not show any problem with the contrast.

3.2. Data Collection Method

In the process of conducting this research, data were collected from the respondents from the Blind institution of Nepal located in the Kathmandu valley and from the residence of Norway living in Oslo. The name of the blind institution is not given due to privacy reasons. The primary objective of the blind institution in Nepal is to provide vocational training and technical and skill development for disabled people. The meeting with the respondent who lives in Oslo was arranged with the help of the former student of Oslomet University who has participate respondent in his previous research. Some of the respondents were blindfolded for those who had normal vision. In this research, the data collection method used was a semi-structured interview by conducting a zoom meeting and direct questioning. The questionnaires were distributed and administered to the respondents. The researcher personally handed over the Questionnaire and application task to the respondent. The questionnaire was created using the Likert scale (5=Strongly Agree, 4= Agree, 3= Neutral, 2= Disagree, 1= Strongly Disagree). It took about a three-week time period to collect the data. A total number of 24 completed questionnaires were collected and used for further analysis.

All questions were evaluated on a Likert scale, which includes the respondent's eyesight profile. The personal eyesight profile in section A contains three items; (1) type of visual impairment, (2) loss the vision by birth, or other (3) use of the mobile device.

Likewise, user interface questions were evaluated on a Likert scale from one to five. The mobile user interface is a graphical and touch-sensitive display in the mobile devices where the user can interact with the mobile and give some input to the device. The small screen size on mobile devices makes the user interface design more challenging. The questionnaire about the user interface tries to find whether the user finds the gesture touch and slide easy to use and if the content of the image is accessible to the user or not. The prototype gives the user voice instruction while using it. Whether the inbuild instruction is clear or not. The user is notified when the user touches the object in the image through the voice and the haptic feedback. The questionnaire also tries to find the answer to whether combining two notification sounds and haptic feedback make it easier for users or not.

Due to the covid-19 pandemic situation, it was difficult to conduct face-to-face interviews. Special care and many corona rules were needed to be followed for the interview. Data collection was conducted in two phases, first in-person who were from Oslo and in the second phase through online mode where respondents were from both Nepal and Oslo. Questionnaires were paper-based and written in English and interviewed in English. For a few respondents, Norwegian translation using google translator was provided on demand to make a better understanding of respondents. Most of the respondents were Nepalese, and some of them needed Nepali translation as the researcher is a native Nepali speaker, so these respondents were given explanations in the Nepalese language. The questions were asked to the respondents, and filled out answers on questionnaires by conducting a zoom meeting as well as a direct interview. Approximately 40 people were contacted for participation. But 16 people refused to participate due to fear of corona and other health issues. Most of the respondents gave a positive response and participated in the research.

In the first phase of the data collection, the interview was conducted by video conferencing through the zoom application. The supervisor of the blind institute played a major role in supporting video conferencing and help to install the mobile application on their devices and was there all the time when required. In the second phase of the data collection, the interview was conducted face to face with the respondents. The participant who has normal vision was

blindfolded while the user was testing the prototype. Following are the detail of how the research was conducted to collect the data, which will conclude the outcome of this thesis.

User testing formality

- Welcome
    - Welcome the respondent to the research
    - Clarify the research and explain what we are going to do
- Consent form
    - Introduce the consent form, which is sent via e-mail to the supervisor and respondent (for video conferencing via zoom)
    - Explain the content of the consent form to the respondent
- Instructions
    - Explain how to use the prototype application in detail
    - Those who had a problem with vision were interviewed, and the person who was able to read the question were handout the questionnaire
    - Both negative and positive feedback was taken from all the respondent

During the experiments, no respondents were in extreme need of language translation support. Therefore, the researcher handled all experiments and tasks by himself with the help of the supervisor in case of video conferencing via zoom. However, there was some technical disturbance while conduction the interview through video conferencing as there was some connection lost due to poor internet speed in Nepal. So, it was a bit time consuming than face to face interview. Due to the present scenario of the pandemic, a number of contacted respondent refused to participate on the research as the government asked people to stay at home as much as possible and limit the contact with other as much as possible such to limit the infection.

3.3. Data Analysis Techniques

The study uses both descriptive and inferential analysis. This study was designed to understand the opinion of the visually impaired respondents who use mobile in day-to-day life. The total number of observations was twenty-four, where seventeen observations were from Nepal, and seven were from Oslo.

Likewise, the questionnaire consists of three parts. Part A consists of information regarding the respondent's personal eyesight situation. Part B consists of questions for the individual

respondents to ascertain user interface, and part C seeks answers for the overall general use of an application of an individual respondent. Questionnaires relating to the study were sent to some individuals through zoom. The questionnaires begin and end with a courteous note thanking them for their time and participation.

Data is the collection of facts, observation, or other information relatedtoh the issue or problem which is used especially in reference to experimental observations. The data that describes the information and is often open-ended, which allows participants to express themselves, is qualitative data. This form of data is also called categorical data and can be classified based on labels, attributes, and other identifiers rather than a number. The qualitative data can mainly be divided into two types nominal and ordinal data.

Nomial data which is also known as "named" data, is classified as information that is used to name or mark variables but has no numerical value. The nominal data does not usually have an inherent ordering. For example, frequency of mobile use is a nominal variable having the number of categories, but there is no clear way to order them. Ordinal data is a type of categorical data that has been given a numerical value. Ordinal data has variables that are specified in a specific order. Ordinal variables are normally numbered in order to signify the list's order. The numbers, on the other hand, are assigned as labels that are not mathematically calculated. An example of the ordinal data is data called using the Likert scale. Both the nominal and ordinal data are non-parametric variables which are only differentiates in terms that ordinal data is put in a kind of order by its location distinguishes them.

The questionnaire of this research consist of Likert scale which gives the ordinal data from the respondents as well as it also consists of some question which gives the profile description of the respondent's answer is nominal data. For analyzing the ordinal data, generally, mode, median, quartiles are used. The nominal data is analyzed by grouping the input variable and calculating the percentage or mode of the distribution. For the non-parametric variable, different tests can be found to test the hypothesis. Wilcoxon signed-rank test, Friedman 2-way ANOVA, Wilcoxon rank-sum test, and Kruskal-Wallis 1-way test are the ordinal data tests. Since there are four types of visual impairment involved during the data collection Kruskal-Wallis test is used to test the hypothesis, which is a nonparametric test for ordinal data. Microsoft Excel is used for the calculation and get the result.

3.4.   Pilot Test

The pilot test is performed in order to determine if there are any imperfections, restrictions, or other weaknesses in the interview design and make important changes before doing the research on the full scale. A pilot test is done to improve the quality and efficiency of the research, and it is conducted on a smaller scale. A total of 5 respondents were chosen for the pilot test. The questionnaire was distributed to two students studying at Oslomet University, and their spouse and another pilot test respondent have Amblyopia, also known as lazy eye disorder of sight, in order to find any flaws or limitations on the research. Four participants who have normal vision were blindfolded, and one participant who has a lazy eye was only blindfolded on one eye. After the pilot tests, the respondents were asked for their feedback. Each respondent reported that all provided information and answers are clear and easy to understand. In addition, they liked the concept of getting the information and voice reply through a touch-up application. The total experiment time taken by pilot participants was from 20 to 30 minutes. The participants gave feedback that some of the questions on the questionnaire were difficult to understand as they may be too technical and difficult to answer. So, there were make some adjustments to the question such that it is understandable for the general public.

3.5.   Ethical Consideration

While conducting research, the researcher and participants are personally involved in a different stage, which arises the ethical challenges. During the research of this thesis, a qualitative research method was used, which will be utilized to explain, clarify, and elaborate on the people's experiences using the system. The research should be done in a way such that it not only do good but also should avoid doing any harm. While including participants in the research, ethical principles must be considered, which include informed consent, minimizing the harm, anonymity and confidentiality, voluntary participation, and only assess relevant components ("Step 6: Issues of research ethics for your dissertation | Lærd Dissertation," 2012). All the participants in the research must fully be informed and made aware of the thesis such that they can decide whether they take part in the thesis or not. Participants can leave their participation without any explanations or feedback; also, no pressure must be applied for those who want to discontinue their participation. The participants in the thesis should not be harmful, both physically and psychologically. The participants' identity should

not be revealed or accessed on documentation of the thesis. This thesis must follow all the ethical principles and must not cross any ethical boundary. No data are collected on the research that will link the individual person.

Chapter 4: Data Presentation and Analysis

This chapter deals with the data presentation and analysis of the data collected from the users. Overall, 24 participants take part in the prototype testing. The main objective of this research is expected to be fulfilled with the outcomes derived from the analysis of the data.

4.2. Eyesight profile of Respondent

The first questionnaire gives the answers of the types of visual impairment on the respondents. The respondents can select one of the four options. The first option is totally blind, which means that they can't see any light or form and refers to complete loss of sight. Secondly, low vision is a severe visual impairment that cannot be improved using glasses or contacts, and the better-seeing eye visual acuity is 20/70 or worse. Thirdly, partial vision means the individual who is partially sighted either in one or both eyes, and normal vision means the individual who can see objects clear near and far away. As mention earlier, participants with normal vision were blindfolded for data collection.
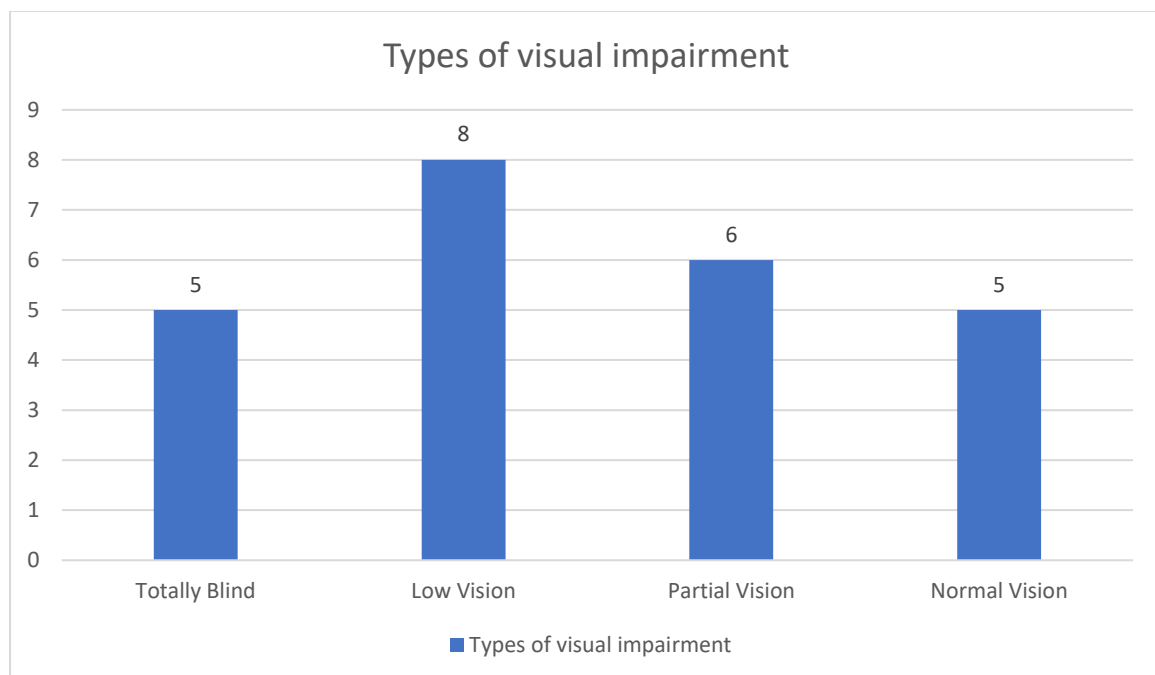

Figure 7: The bar graph showing the number of participants with a different type of vision

The graph above shows the number of participates divided by the types of visual impairment. A total of 24 participants take part in the prototype testing in which five respondents were totally blind, eight have low vision, five have partial vision, and four respondents have normal

vision who uses blindfolds during prototype testing. The number of the low vision participants was the most with eight, whereas the other three types of vision impairment have similar numbers of participants.

The second questionnaire tries to find if the participants have visual impairment from birth or not. This question is asked to the respondents to know their familiarity with the object.
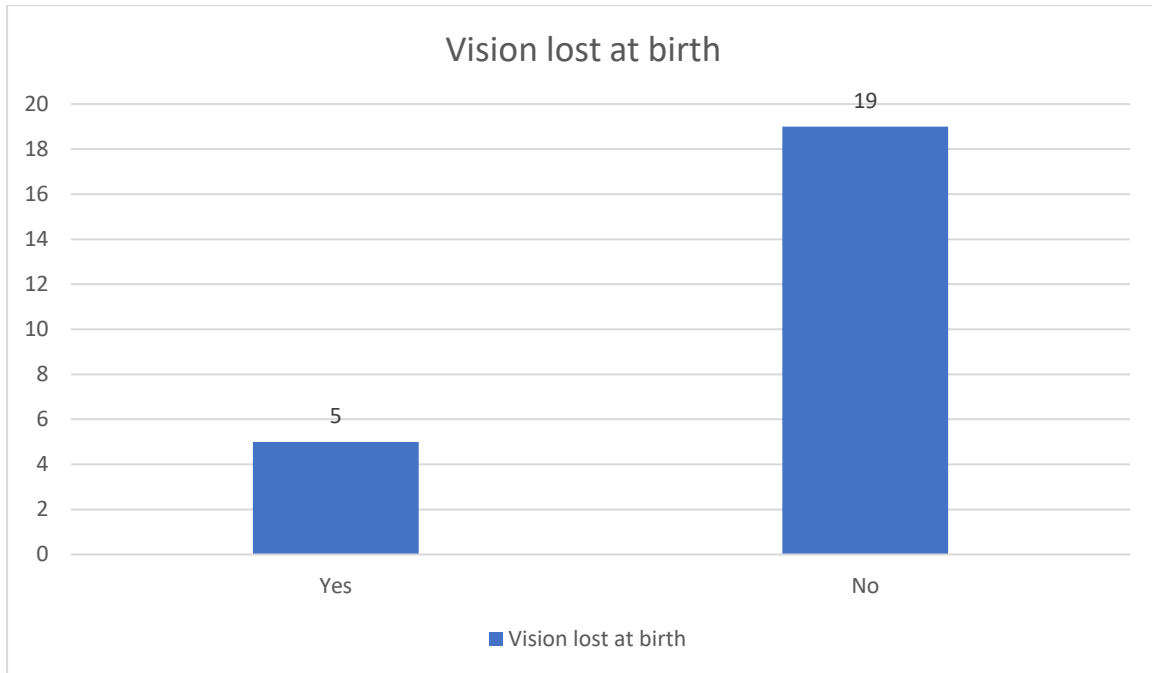

Figure 8: The bar graph showing the number of participant vision lost at birth

The graph shows the number of 5 of the respondent has lost their vision at birth, and other 19 participants have lost vision not at birth, and they have lost their vision later as a result of injury, disease or a medical condition.

The third question on the questionnaire finds the answer of how frequently the participants use the mobile phone in daily life. The answerer can select the answer from the four options, namely frequently, occasionally, rarely, and never.
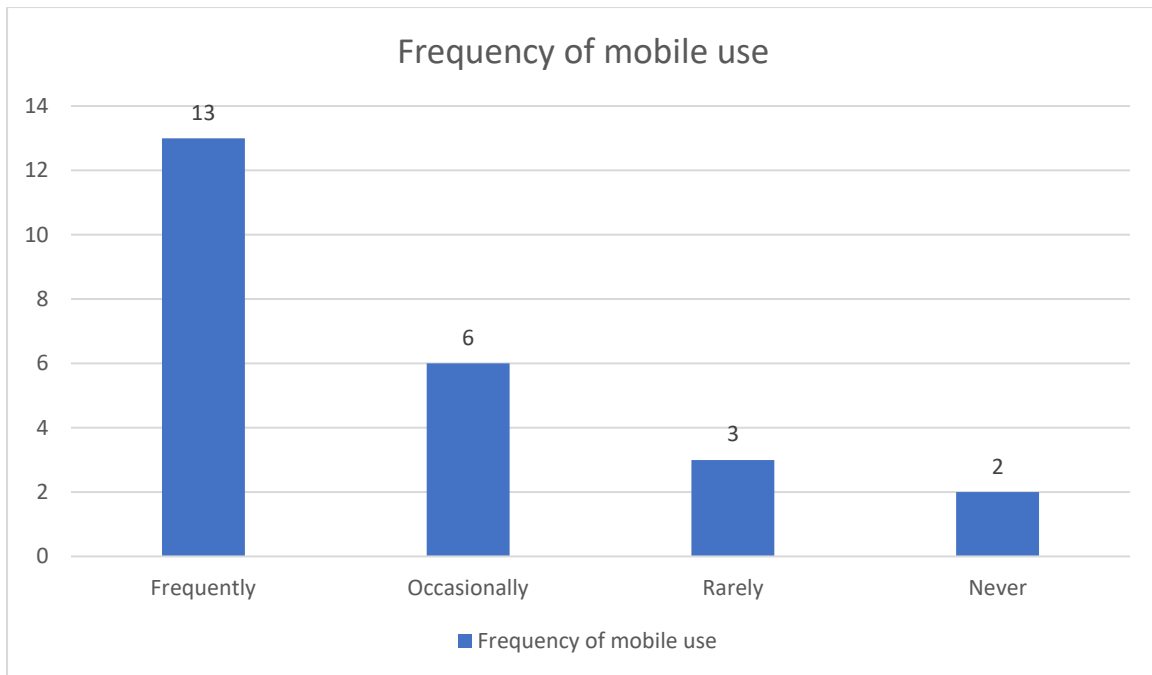
Figure 9: The bar graph showing the frequency of mobile use by the participants

The graph illustrates that most of the respondents on the prototype testing use the mobile phone frequently or occasionally. Thirteen of the participants uses the mobile phone often and uses it as a part of their life. Six of the respondents uses mobile phone time to time. Only two of the respondents don't have any experience with mobile uses, and three of the participants rarely use mobile phones.
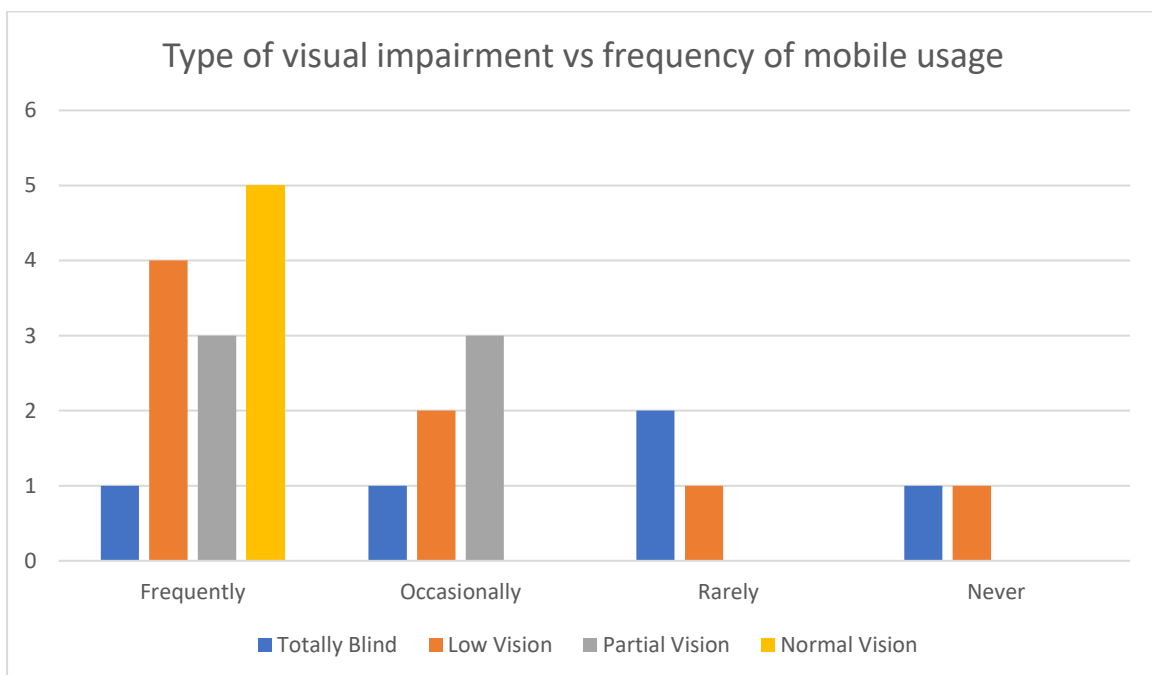

Figure 10: The bar chart showing the type of visual impairment vs. frequency of mobile usage

The graph above describes the type of visual impairment vs. frequency of mobile usage. Out of thirteen frequent mobile users, 1 is totally blind, 4 are low vision, 3 are partial vision, and 5 are normal vision. Out of 6 participants who use the mobile phone occasionally, one is totally blind, 2 are low vision, and 3 are partial vision. 3 participants uses the mobile phone rarely, 2 are totally blind, and one has low vision. Two participants who have not used mobile devices before one is blind and another is low vision.

4.2. Analysis of Data

The first Likert scale question Q1 in the questionnaire was to find if the use of the gesture in the prototype touch and slide was easy for the participants to get the information of the object in the image.
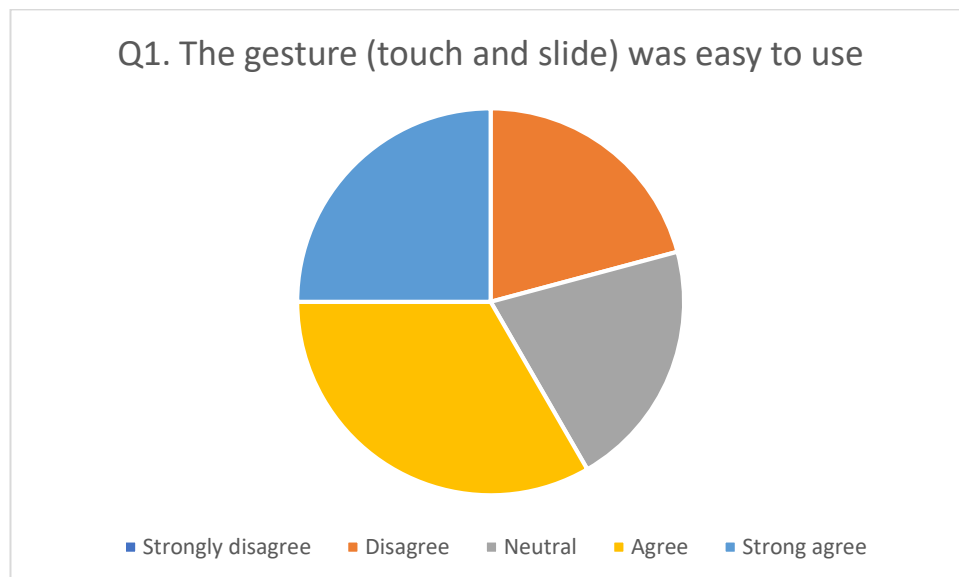

Figure 11: The pie chart showing the Likert scale result of the question, the gesture (touch and slide) was easy to use

The above pie chart shows the result of the use of touch and slide gestures on the prototype, and nearly more than half of the participants either agree or strongly agree with the statement. An equal number of the participants either disagree or remain neutral with the statement. Touch and slide gesture is used in the prototype to get the information from the touchscreen of the mobile device. Most of the users find it easy to slide the finger around the screen and get information about the image.
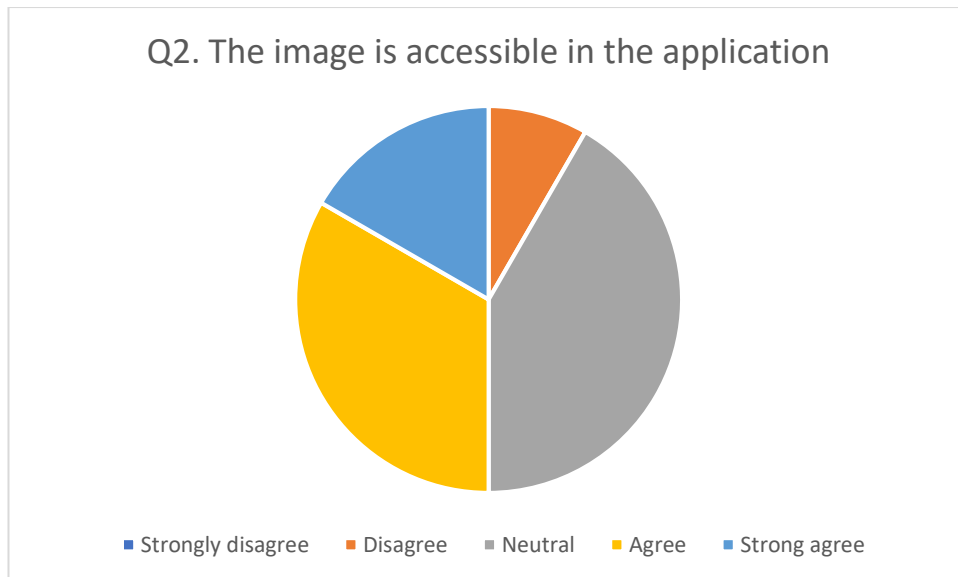
Figure 12: The pie chart showing the Likert scale result of the question, image is accessible in the application

The second Likert statement was if the image was accessible in the application or not. When the user touches the object on the image by using gesture touch and slide, then the prototype gives feedback to the user with sound and haptic feedback. Nearly half of the respondents either agree or strongly agree with the statement, and most of the prototype users were neutral with the statement. No prototype user strongly disagrees with the statement that the image is not accessible in the application.
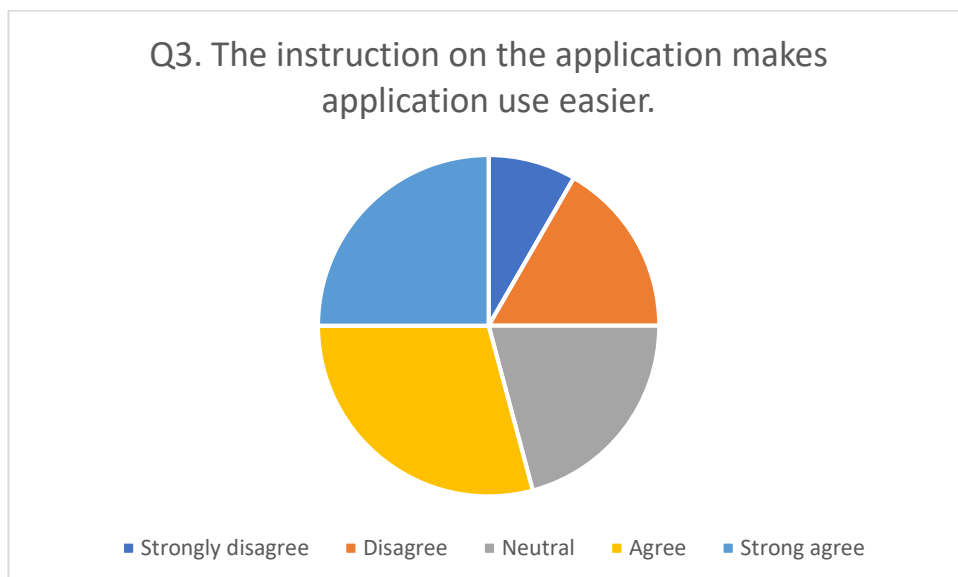

Figure 13: The pie chart showing the Likert scale result of the image is accessible in the application

The prototype gives users instruction on the speech such that they find it easier to use the application. The Likert scale question was whether the instruction on the application helps

the users while using the prototype or not. The pie chart shows that more than half of the participants find that the different speech instruction on the prototypes makes the application more easy to use.

Q4. It was easy to understand the image through touch up information

■ Strongly disagree ■ Disagree ■ Neutral ■ Agree ■ Strong agree
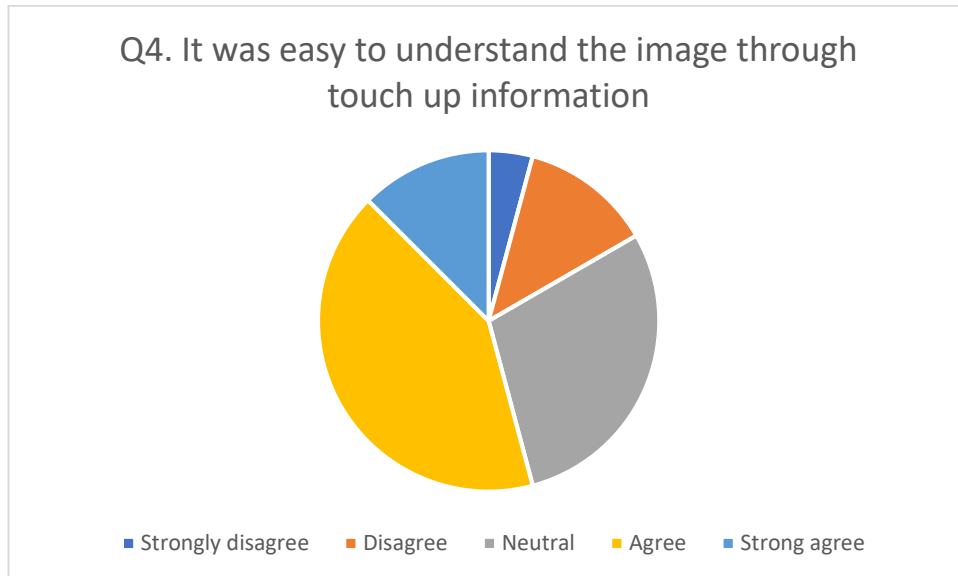
Figure 14: The pie chart showing the Likert scale result of the question, it was easy to understand the image through touch up information

The next statement was to know to whether the touch up information in the developed prototype makes the participants understand the image easier or not. The pie chart above illustrates the result of the prototype testing for the Likert question; it was easy to understand the image through touch up information. From the pie chart it is clear that majority of the participant agree or strongly agree with the statement. Nearly 5% of the respondents strongly disagree with the statement.
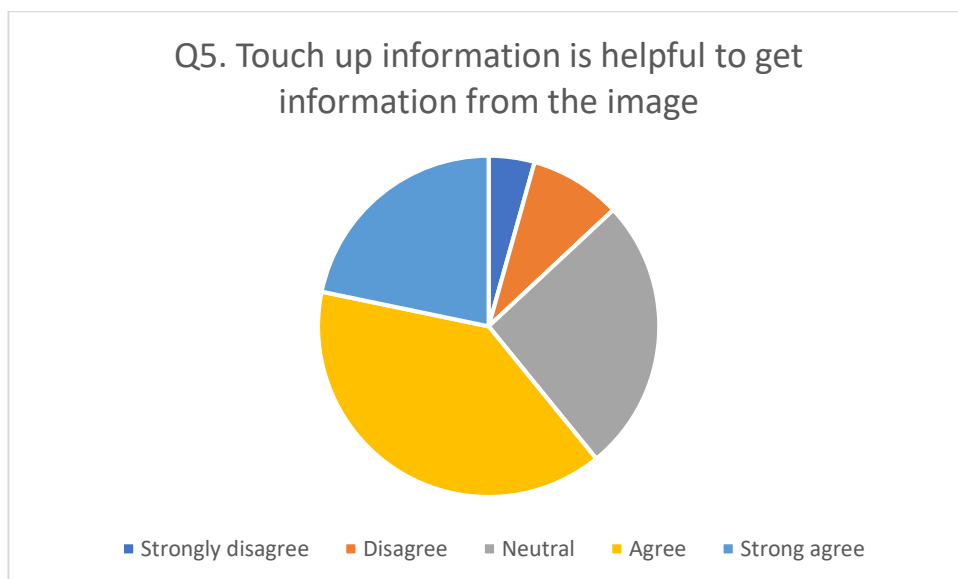
Q5. Touch up information is helpful to get information from the image

■ Strongly disagree ■ Disagree ■ Neutral ■ Agree ■ Strong agree

The fifth question tries to find the answer if the touch up information is helpful to get information from the image or not. When the respondents touch the image in the prototype, the prototype returns the object name in the speech and haptic feedback. The result shows that most of the respondents agree with the statement.
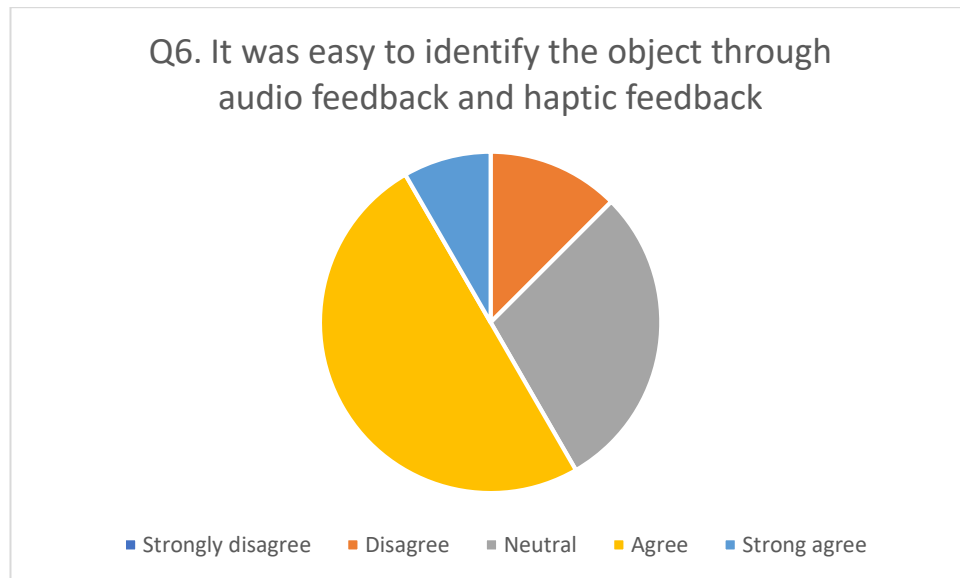
Figure 16: The pie chart showing the Likert scale result of the question, it was easy to identify the object through audio and haptic feedback

The sixth Likert scale question was to find if the combination of the audio and haptic feedback makes the user easy to identify the object. The pie chart shows reveal that half of the participants agree that combining two methods of user notification makes then easier to know the object in the image.
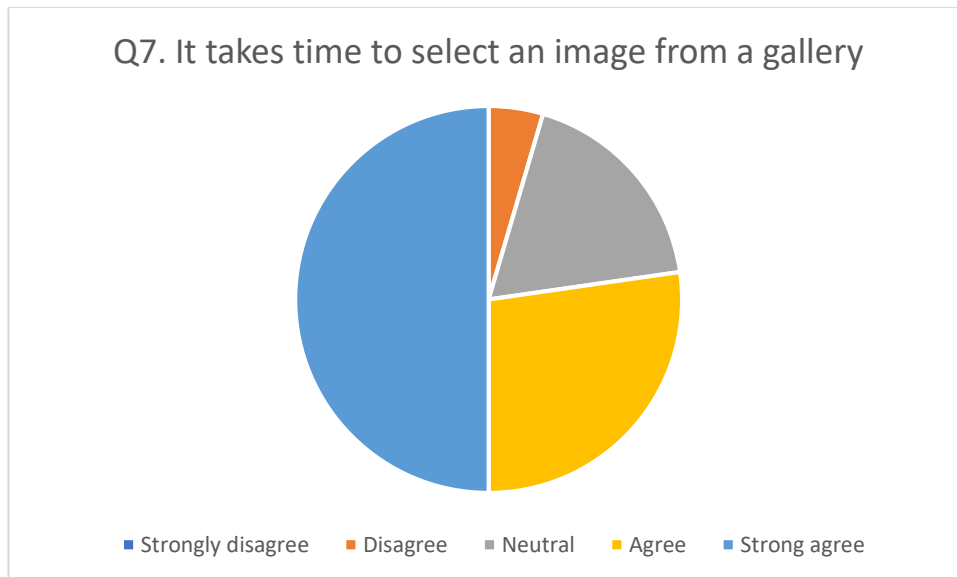
Figure 17*:* The pie chart showing the Likert scale result of the question, it takes time to select an image from a gallery

Most of the respondents find that choosing the image from the gallery, which is the inbuild in the mobile devices is time-consuming. Nearly half of the respondents strongly agree that it takes time to select an image from the gallery. The respondents find it challenging to select the images from the gallery of the mobile devices. About a quarter of the respondents agree with the statement
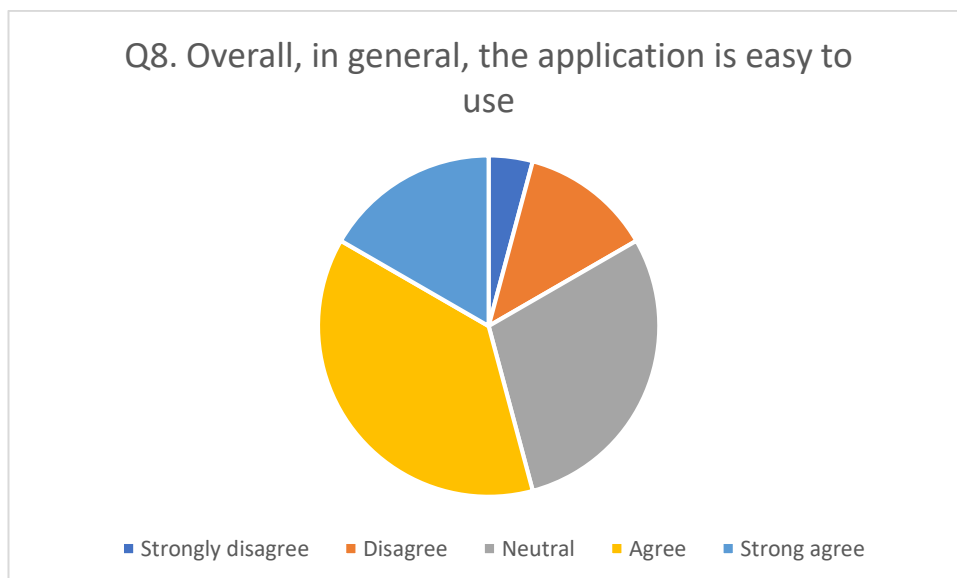
Figure 18:The pie chart showing the Likert scale result of the question; overall, in general, the application is easy to use

The eighth question is the general question in which if the overall application is easy to use or not is asked. This question tries to observe is overall application is easy to use or not. The pie chart above exhibits that the majority of the participant either agree or strongly agree

with the statement, and only one-fifth of the participant either strongly disagree or disagree with the statement.
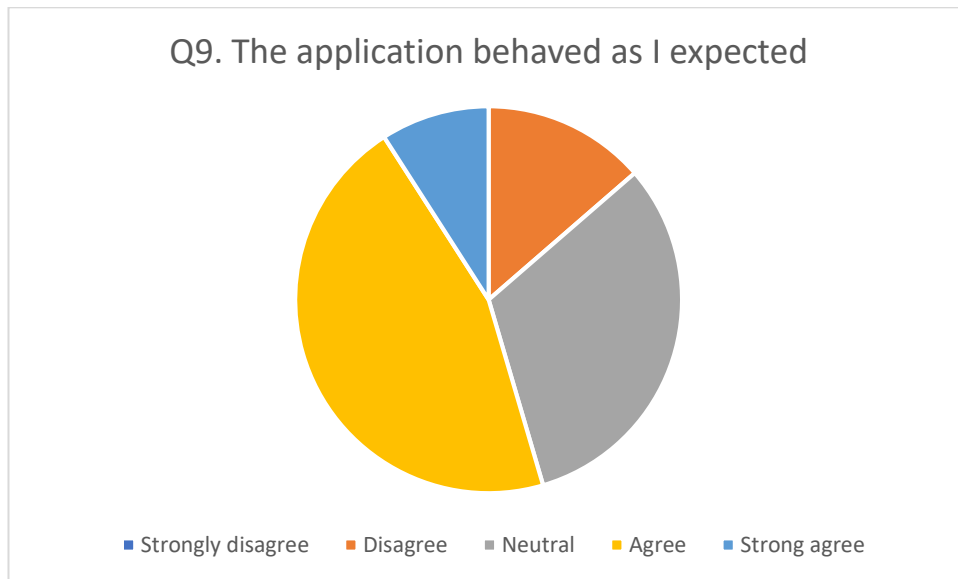
Figure 19*:* The pie chart showing the Likert scale result of the question, the application behaved as I expected

The ninth question attempts to find if the participant finds the application behaved as they expected. It is the general question in which attempts to know if the developed prototype was as the participant has expected. Nearly 9% of the participants strongly agree with the statement that the application work as they expected, and 45% of the respondents agree with the Likert question.
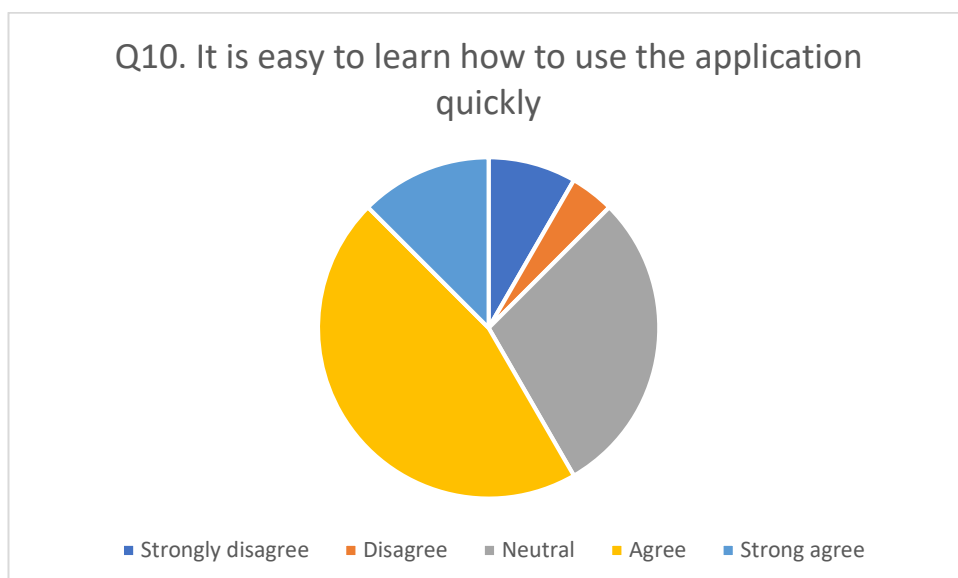
Figure 20*:* The pie chart showing the Likert scale result of the question, it is easy to learn how to use the application quickly

The last question of the above pie chart demonstrates that the prototype application was easy to learn or not. When the participant uses the prototype application, where they able to learn quickly to use the application. The pie chart above shows that more than half of the participants agree or strongly agree with the statement above. Just about 10% of the participant either strongly disagree or disagree with the statement as they find it difficult to use the application quickly.

4.2.1. Hypothesis testing

The Kruskal-Wallis H test is a rank-based nonparametric test which is also known as the "one-way ANOVA on ranks." This test is used to see whether two or more groups are statistically significant on continuous or ordinal data. To find the answer to the research question, we can evaluate the Likert scale data from the different questions. The hypothesis is created according to the Likert scale question.

H0: Different types of visually impaired respondents medians are equal for gesture (touch and slide) was easy to use

H1: Different types of visually impaired respondents medians are not equal for gesture (touch and slide) was easy to use

Kruskal-Wallis Test Q1

	Totally Blind	Low Vision	Partial Vision	Normal Vision	
rank sum	37.5	92.5	81	89	
count	5	8	6	5	24
r^2/n	281.25	1069.53125	1093.5	1584.2	4028.48125
H-stat					5.569625
df					3
alpha					0.05

Table 1: Test statistics table for Q1

We find the critical value in the table of critical values for the chi-square distribution for df=k-1=4-1=3 and=0.05 since we are comparing 4 groups, the critical value is 7.81. If H > 7.81, the decision rule is to reject H0.

Since we observed that the H value is 5.57, which is less than the critical rejection value of 7.81 so that we can conclude that the null hypothesis is not rejected.

So it is concluded that the null hypothesis different types of visually impaired respondents medians are equal for gesture (touch and slide) was easy to use.

Similarly, to determine if the image on the mobile phone is accessible or not, we can create the hypothesis as below.

H0: Different types of visually impaired respondents medians are equal for the image is accessible in the application

H1: Different types of visually impaired respondents medians are not equal for the image is accessible in the application

Kruskal-Wallis Test Q2

	Totally Blind	Low Vision	Partial Vision	Normal Vision	
rank sum	35	110.5	85	69.5	
count	5	8	6	5	24
r^2/n	245	1526.28125	1204.16667	966.05	3941.49792
H-stat					3.82995833
df					3
alpha					0.05

Table 2: Test statistics table for Q2

Since we observed that the H value is 3.83 which is less than the critical rejection value of 7.81 so that we can conclude that the null hypothesis is not rejected.

So it is concluded that the null hypothesis different type of visually impaired respondents medians are equal for gesture (touch and slide) was easy to use.

H0: Different types of visually impaired respondents medians are equal for the instruction on the application makes application use easier.

H1: Different types of visually impaired respondents medians are not equal for the instruction on the application makes application use easier.

Kruskal-Wallis Test Q3

	Totally Blind	Low Vision	Partial Vision	Normal Vision	
rank sum	25	104.5	59	87.5	
count	4	8	6	5	23
r^2/n	156.25	1365.03125	580.166667	1531.25	3632.69792
H-stat					6.97169384

df				3
alpha				0.05

Table 3:Test statistics table for Q3

Since we observed that the H value is 6.97, which is less than the critical rejection value of 7.81 so that we can conclude that the null hypothesis is not rejected.

So it is concluded that we have statistically significant evidence that shows that all different four groups have equal median for the instruction on the application makes application use easier.

H0: Different types of visually impaired respondents medians are equal for it was easy to understand the image through touch up information

H1: Different types of visually impaired respondents medians are not equal, for it was easy to understand the image through touch up information

Kruskal-Wallis Test Q4

	Totally Blind	Low Vision	Partial Vision	Normal Vision	
rank sum	49	95	46	86	
count	5	8	5	5	23
r^2/n	480.2	1128.125	423.2	1479.2	3510.725
H-stat					4.3201087
df					3

Table 4:Test statistics table for Q4

Since we observed that the H value is 4.32, which is less than the critical rejection value of 7.81 so that we can conclude that the null hypothesis is not rejected.

So it is concluded that we have statistically significant evidence that shows that all four different groups have equal median for it was easy to understand the image through touch up information.

H0: Different types of visually impaired respondents medians are equal for touch up information is helpful to get information from the image

H1: Different types of visually impaired respondents medians are not equal for touch up information is helpful to get information from the image

Kruskal-Wallis Test Q5

	Totally Blind	Low Vision	Partial Vision	Normal Vision	
rank sum	47	84.5	51	93.5	
count	5	7	6	5	23
r^2/n	441.8	1020.03571	433.5	1748.45	3643.78571
H-stat					7.21273292
df					3
alpha					0.05

Table 5: Test statistics table for Q5

Since we observed that the H value is 7.21, which is less than the critical rejection value of 7.81 so that we can conclude that the null hypothesis is not rejected.

So it is concluded that we have statistically significant evidence that shows that all four different group have equal median for touch up information is helpful to get information from the image.

H0: Different types of visually impaired respondents medians are equal, for it was easy to identify the object through audio feedback and haptic feedback

H1: Different types of visually impaired respondents medians are not equal, for it was easy to identify the object through audio feedback and haptic feedback

Kruskal-Wallis Test Q6

	Totally Blind	Low Vision	Partial Vision	Normal Vision	
rank sum	54.5	91.5	68	62	
count	5	7	6	5	23
r^2/n	594.05	1196.03571	770.666667	768.8	3329.55238
H-stat					0.3815735
df					3
alpha					0.05

Table 6: Test statistics table for Q6

Since we observed that the H value is 0.38, which is less than the critical rejection value of 7.81 so that we can conclude that the null hypothesis is not rejected.

So it is concluded that we have statistically significant evidence that shows that all four different groups have equal median for it was easy to identify the object through audio feedback and haptic feedback.

Chapter 5: Discussion and Conclusion

5.1. Discussion

The objectives of this study are how can developed touchup interface application help visually impaired users access the image on mobile devices and how easy it is for visually impaired people to understand the image through the developed touchup interface. The evaluation and analysis of data statically proved that the use of the gesture (touch and slide) was easy to use for all types of visually impaired people. Moving the finger around the touchscreen to get the object information about the image can be easy for the visually impaired person. The image on the mobile devices The prototype combines the two feedback method sound and haptic feedback which notify the user when the user touch or move the finger around the object in the image. Combining two or more feedback methods can be more effective for the special able persons. The audio instructions while using the application can also be helpful for the visually impaired person. The participants also find that it was easy to understand the image through touch up information where in the prototype, the user knows the object in the image by touching it. The application usage the inbuilt gallery to select the image in which most of the user face challenges to select the image as it was time consuming and difficult. Overall, in general, most of the participant find the application is to learn and use and behaved as they expected.

5.2.   Conclusion

The main objective of this thesis is to develop an accessible interface for the visually impaired person to access the image and evaluate how the developed application helps the users understand the image through touchup information. At first, a literature study was carried out. The literature study was divided into four processes and was carried out to find the best method or principle to carry out each process successfully. Based on the literature review, the researchers have performed research on a variety of algorithms and techniques for image segmentation and object detection. There are many applications available for object detection, but only a few works offline. Making the application work offline can be one of the vital features of the application. The essential focus of the thesis is given to the user interface. So, most of the time of the development is separated for this part. During the literature survey, I could not find any research work done for interacting the mobile devices by moving

the finger on the touch screen. This method of interaction with mobile devices fill the gap of image accessibility in mobile devices.

As the research is based on a qualitative approach so, the instrument used to access the developed interface is based on a questionnaire. Data have been used from primary sources by conducting an interview among a sample of visually impaired users through video conference and by distributing self-administrative questionnaires to non-visually impaired persons. A questionnaire was planned to investigate the fact. The questionnaire contains the use of fixed questions that are presented to respondents in the same way, with no variation in question-wording.

For the outcome of the study, for both visually impaired and normal vision, people were made to use the prototype user interface. Afterward, the respondent's feedback on the significant area and problematic area on the use of image accessibility in a mobile app through touchup information was collected, which helped to draw the conclusion part of the research. Through the pilot test, it shows that they liked the concept of getting the information and voice message through a touch-up application. Also, all provided information and answers were clear and easy to understand. There was no difficulty in finding an image through instruction. To conclude, interaction method with mobile devices can fill the gap of image accessibility in mobile devices.

5.3. Limitations and Future work

Due to different circumstance there are several limitations with this study. The main limitation is due to ongoing COVID-19 pandamic which limit face-to-face contact with other people. Many respondents was contacted for the prototype testing but only few respondents approve for the testing. Many prototype tester was afraid to meet face-to-face. Most of the research was conducted virtually. So, the number of the participants was very low. The developed prototype currently uses the inbuilt gallery of the android mobile, which most of the prototype testing participants find time-consuming to select the image. So, an accessible in-app galley can be developed, which makes it easy for the user to select the image. Presently, the application works only on the static image but not on the dynamic image or real-time image. The application can be developed to work on real-time object detection. The application can use the inbuilt camera for real-time images. Right now, the prototype only gives the name of the object in the image but generating the whole sentences from the object

detected which describes the image can be the future development that can make the image more accessible and can give new dynamics to the prototype. Furthermore, more work can be done in part of the object detection as for the time being, TensorFlow lite pre-trained model, which only detects 90 classes of the object, is used in the prototype. Lastly, the prototype is only developed for an android operating mobile device which can be developed for cross-platform IOS device since the TensorFlow lite is used in the application which can be used for both android and IOS devices.

6. References

Article 9 - Accessibility | United Nations Enable. (2019). Retrieved from
https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/article-9-accessibility.html

Awad, M., El Haddad, J., Khneisser, E., Mahmoud, T., Yaacoub, E., & Malli, M. (2018). *Intelligent eye: A mobile application for assisting blind people.* Paper presented at the 2018 IEEE Middle East and North Africa Communications Conference (MENACOMM).

Awada, A., Issa, Y. B., Tekli, J., & Chbeir, R. (2013). *Evaluation of touch screen vibration accessibility for blind users.* Paper presented at the Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility.

Benhamida, A., Várkonyi-Kóczy, A. R., & Kozlovszky, M. (2020). *Traffic Signs Recognition in a mobile-based application using TensorFlow and Transfer Learning technics.* Paper presented at the 2020 IEEE 15th International Conference of System of Systems Engineering (SoSE).

Buzzi, M. C., Buzzi, M., Leporini, B., & Paratore, M. T. (2013). *Vibro-tactile enrichment improves blind user interaction with mobile touchscreens.* Paper presented at the IFIP Conference on Human-Computer Interaction.

Buzzi, M. C., Buzzi, M., Leporini, B., & Trujillo, A. (2015). *Exploring Visually Impaired People's Gesture Preferences for Smartphones*. Paper presented at the Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter, Rome, Italy.
https://doi.org/10.1145/2808435.2808448

Czúni, L., Kiss, P. J., Á, L., & Gál, M. (2014, 27-30 Oct. 2014). *Lightweight mobile object recognition.* Paper presented at the 2014 IEEE International Conference on Image Processing (ICIP).

Dasila, R. S., Trivedi, M., Soni, S., Senthil, M., & Narendran, M. (2017). *Real time environment perception for visually impaired.* Paper presented at the 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR).

Dobosz, K. (2017). Designing Mobile Applications For Visually Impaired People. In (pp. 103-126).

Dreyer, A., Grønli, T.-M., & Bygstad, B. MAKING TOUCH-BASED MOBILE PHONES ACCESSIBLE FOR THE VISUALLY IMPAIRED.

Elrefaei, L., Almasawa, M., & Al-gohany, N. (2017). Development of An Android Application for Object Detection Based on Color, Shape, or Local Features. *The International journal of Multimedia & Its Applications, 9*, 21-30. doi:10.5121/ijma.2017.9103

Giudice, N. A., Palani, H. P., Brenner, E., & Kramer, K. M. (2012). *Learning non-visual graphical information using a touch-based vibro-audio interface.* Paper presented at the Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility.

Gustafsdottir, G. (2020). Accessibility: Image Alt text best practices. *Accessibility: Image Alt text best practices.* Retrieved from https://support.siteimprove.com/hc/en-gb/articles/115000013031-Accessibility-Image-Alt-text-best-practices

Images | Usability & Web Accessibility. (n.d.). Retrieved from https://usability.yale.edu/web-accessibility/articles/images

Kane, S. K., Bigham, J. P., & Wobbrock, J. (2011). Fully accessible touch screens for the blind and visually impaired. *University of Washington*.

Kane, S. K., Wobbrock, J. O., & Ladner, R. E. (2011). *Usable gestures for blind people: understanding preference and performance.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Li, Z., & Zhang, R. Object Detection and Its Implementation on Android Devices.

Make apps more accessible. (n.d.). Retrieved from
https://developer.android.com/guide/topics/ui/accessibility/apps

Mobile Operating System Market Share Worldwide. (2020). *StatCounter Global Stats.* Retrieved from
https://gs.statcounter.com/os-market-share/mobile/worldwide

Object detection | TensorFlow Lite. (n.d.). Retrieved from
https://www.tensorflow.org/lite/models/object_detection/overview

Parikh, N., Shah, I., & Vahora, S. (2018). *Android smartphone based visual object recognition for visually impaired using deep learning.* Paper presented at the 2018 International Conference on Communication and Signal Processing (ICCSP).

Rajwani, R., Purswani, D., Kalinani, P., Ramchandani, D., & Dokare, I. (2018). Proposed System on Object Detection for Visually Impaired People. *International Journal of Information Technology (IJIT), 4*(1), 1-6.

Raut, S. A., Raghuwanshi, M., Dharaskar, R., & Raut, A. (2009, 22-24 Jan. 2009). *Image Segmentation – A State-Of-Art Survey for Prediction.* Paper presented at the 2009 International Conference on Advanced Computer Control.

Sanjari, M., Bahramnezhad, F., Fomani, F. K., Shoghi, M., & Cheraghi, M. A. (2014). Ethical challenges of researchers in qualitative studies: The necessity to develop a specific guideline. *Journal of medical ethics and history of medicine, 7*.

Shin, H., Lim, J., Lee, J., & Kyung, K. (2013, 11-14 Jan. 2013). *Background display for visually impaired people in mobile touch devices.* Paper presented at the 2013 IEEE International Conference on Consumer Electronics (ICCE).

Shishir, M. A. K., Fahim, S. R., Habib, F. M., & Farah, T. (2019). *Eye Assistant: Using mobile application to help the visually impaired.* Paper presented at the 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT).

Smartphones - Statistics & Facts. (2020). Retrieved from
https://www.statista.com/topics/840/smartphones/

Starting Android Accessibility. (2020). Retrieved from
https://developer.android.com/codelabs/starting-android-accessibility#0

Step 6: Issues of research ethics for your dissertation | Lærd Dissertation. (2012). Retrieved from
http://dissertation.laerd.com/process-stage6-step6.php

Sultan, N., Siddiq, K., Rashid, T., & Farooque, M. (2015). Evaluation of smart phone applications accessibility for blind users. *International Journal of Computer Applications, 127*(3), 9-16.

TapTapSee - Blind and Visually Impaired Assistive Technology - powered by CloudSight.ai Image Recognition API. (2019). Retrieved from https://taptapseeapp.com/

Taşkın, C., Yiğit, A., Taşkın, D., & Taşkın, A. (2017). *Image processing on Android devices with OpenCV.*

Tosun, S., & Karaarslan, E. (2018). *Real-Time Object Detection Application for Visually Impaired People: Third Eye.* Paper presented at the 2018 International Conference on Artificial Intelligence and Data Processing (IDAP).

Verschae, R., & Ruiz-del-Solar, J. (2015). Object Detection: Current and Future Directions. *Frontiers in Robotics and AI, 2*. doi:10.3389/frobt.2015.00029

World Health Organization. (8 October 2019). Retrieved from https://www.who.int/en/news-room/fact-sheets/detail/blindness-and-visual-impairment

Yan, S., & Ramachandran, P. (2019). The current status of accessibility in mobile apps. *ACM Transactions on Accessible Computing (TACCESS), 12*(1), 1-31.

Yashas, M., Maalik, K., Ashfaq, M., De Silva, K., & Ragulan, S. (2017). Blind guider: An IT solution for visually impaired people. *Int. J. Sci. Res. Pub., 7*(11), 700-703.

7. Appendices

7.1. Appendix 1

Consent Form

User testing of improving image accessibility through touchup information

Background and subject

You are invited to participate in a study of improving image accessibility in mobile devices through touch up information. This research will help people who are differently able to interact with the images on mobile devices and get the image's information. The research should take approximately 15-20 minutes to complete. The participants need to use the prototype application on the android phone and give the answer to the questionnaire after using the prototype.

Participation

Participation in the research will be entirely voluntary, and you can withdraw from taking part in the research or leave the research without giving any explanation at any time. The respondent is allowed to decline to provide an answer to any particular question for any reason.

Benefits

The participants will not receive any direct benefits from participating in the research, but the respondent's response to the research will help learn more about improving image accessibility in mobile devices.

Confidentiality

Confidentiality and the privacy of the participant will be protected during the research. All the information and the collected data will be kept strictly confidential. The research will not record any personal data that will link to the individual person. All the data collected in this research will be anonymous such that collected data will not link back to any person directly nor indirectly. The data collected, which will help to conclude this thesis, will be retained until the final submission of the thesis.

Persons to Contact

If you have any questions and want to talk about the research or if you feel any risk about taking part in this research or you might not have been treated fairly, then you can feel free to contact the researcher of this thesis, Nikhil Amatya, at +47 96820253 or through e-mail s329920@oslomet.no.

This consent form explains all details about the research, and you are being asked to participate. Please read this consent form carefully before taking part in the research. The respondents are allowed to ask a question at any time after joining the research.

7.2. Appendix 2

Questionnaire

Personal Information

What type of visual impairment do you have?

1. Totally blind   2. Low Vision   3. Partial Vision   4. Normal Vision

Have you lost the vision at birth?

1. Yes   2. No

Do you use (or have you used) mobile devices?

1. Frequently   2. Occasionally   3. Rarely   4. Never

5 Likert scale question

1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strong agree

	Question	1	2	3	4	5
	User interface					
1	The gesture (touch and slide) was easy to use					
2	The image is accessible in the application					
3	The instruction on the application makes application use easier.					
4	It was easy to understand the image through touch up information					
5	Touch up information is helpful to get information from the image					

6	It was easy to identify the object through audio feedback and haptic feedback					
7	It takes time to select an image from a gallery					
	General					
8	Overall, in general, the application is easy to use					
9	The application behaved as I expected					
10	It is easy to learn how to use the application quickly					