

**HEARIAM: A WEARABLE DEVICE THAT CAN RECOGNIZE TEXT TO ASSIST
BLIND PERSON IN READING**

A Thesis Proposal
Presented to the Faculty of the
Information and Communications Technology Program
STI College Davao

In Partial Fulfilment
of the Requirements for the Degree
Bachelor of Science in Computer Science

Daniel F. Maquiling
Kimber Chan L. Granaderos

March 23, 2019

ENDORSEMENT FORM FOR PROPOSAL DEFENSE

TITLE OF RESEARCH: Heariam: e-Learning System with Artificial
Intelligence To Assist Blind Children

NAME OF PROPONENTS: **Daniel F. Maquiling**
Kimber Chan L. Granaderos

**In Partial Fulfilment of the Requirements
for the degree Bachelor of Science in Computer Science
has been examined and is recommended for Proposal Defense.**

ENDORSED BY:

Lexfrevail P. Nervida
Thesis Adviser

APPROVED FOR PROPOSAL DEFENSE:

Lexfrevail P. Nervida
Thesis Coordinator

NOTED BY:

Leonardo E. Risma Jr.
Program Head

March 23, 2019

APPROVAL SHEET

This thesis proposal titled: HEARIAM: A WEARABLE DEVICE THAT CAN RECOGNIZE TEXT TO ASSIST BLIND PERSON IN READING prepared and submitted by Damiel F. Maquiling **and** Kimber Chan L. Granaderos, **in partial fulfilment of the requirements for the degree of Bachelor of Science in Computer Science, has been examined and is recommended for acceptance and approval.**

Lexfrevail P. Nervida

Thesis Adviser

**Accepted and approved by the Thesis Review Panel
in partial fulfilment of the requirements for the degree of
Bachelor of Science in Computer Science**

Leonardo Jr. E. Risma

Panel Member

Engr. Roje Escudro

Panel Member

Engr. Elvi Lito E. Ubas

Lead Panelist

Noted:

Lexfrevail P. Nervida

Thesis Coordinator

Leonardo Jr. E. Risma

Program Head

March 23, 2019

ACKNOWLEDGEMENTS

Ms. Lexfrevail P. Nervida, our thesis coordinator and thesis adviser, we would like to express our deepest gratitude for all the things you have done for us in guiding and helping in developing our project, as we continue to make this project possible we did made a lot of trouble in your part as our adviser and coordinator but still you believe and trusted us to finish our project in times of difficulties and situations we faced, you stayed with us even to the times we are down and partly giving up and continue to uplift our determination and confidence in developing our project;

To the Review Panel, we would like to extend our gratitude towards opening our minds with different ideas. You've guided us regarding our development of this paper, you've given us point of views that we need in order to complete our works, you've given us ideas and encouragement to finish our paper correctly and in time.

To our Beloved Parents, Mrs. Maquiling, Mrs. Granaderos, we would like to thank you all for supporting us physically, emotionally, and also financially, all of you mold us to become who we are today and we are always grateful for that, we are thankful for understanding our behavior and attitude towards things while we continue developing our project.

To all our friends, classmates, schoolmates, best friends and especially to our batch mates/block mates, we thank you for scaring us, making us laugh and helping us ease in times we are in a difficult state of doing our project, the times when all of us are stressed and in the verge of giving up we did inspire and helped each other; and we never lose hope in achieving our goals and dreams;

Most of all Our heavenly Father, for being in our hearts and mind as we face different circumstances and situations in developing our project, without your guidance and wisdom we weren't able to find the right path in making this project possible.

The researchers,

Daniel F. Maquiling

Kimber Chan L. Granaderos

ABSTRACT

Title of research: : HEARIAM: A WEARABLE DEVICE THAT CAN RECOGNIZE
TEXT TO ASSIST BLIND PERSON IN READING

Researchers: Daniel F. Maquiling
Kimber Chan L. Granaderos

Degree: Bachelor of Science in Computer Science

Date of Completion: March 2019

Key words: Machine Learning, Computer Vision, Image Processing

According to an article from India today, statistics show that Visual Impairment affects about 285 million of the world's population and about 90 percent of the world's visually impaired people live in developing countries. Globally, uncorrected refractive errors are the main cause of visual impairment while cataracts and age-related macular degeneration as the leading cause for blindness (who, 2017). This study highlights the challenges that blind and visually impaired people experience and the research and studies that aims to improve their quality of life.

the purpose of this study is to highlight the challenges that the blind and visually impaired people experience not just in the academic level but also in the labor market and the resolution that the proponents propose. it is highly known that most skills and talents that blind people have are ignored because they cannot see, and this is also the reason why they are not fully recognized in the society, the proponents aim to destroy the stereotype by giving them a chance to produce an outcome equal to those with sight, which is thru the device -- heariam.

it is hoped that this study will improve the quality of life for the blind and visually impaired in a more independent approach and that may this study contribute to the future reference of such aim.

TABLE OF CONTENTS

Title Page.....	1
APPROVAL SHEET	3
ACKNOWLEDGEMENTS	4
ABSTRACT	5
TABLE OF CONTENTS	6
INTRODUCTION.....	8
Background of the problem	10
Overview of the current state of the technology	11
Objectives of the study	12
Scope and Limitations of the study	12
LITERATURE REVIEW	14
Review of related literature, studies or systems	14
Guidelines for Developing e-Learning System for Visually Impaired	14
Watson Text-To-Speech - IBM	15
E-Learning for the Visually Impaired: A Holistic Perspective	15
Artificial Intelligence in Education of Children with Learning Disabilities	17
Synthesis	18
METHODOLOGY	20
RESULTS AND DISCUSSIONS	26
CONCLUSION AND RECOMMENDATIONS	28
Hardware/Software.....	29
Calendar of Activities.....	30
Budgetary Estimate	37
APPENDIX	38
Figure 7.0 Automatically run the code after booting up the device	40

Human Resources.....	41
REFERENCES.....	46

INTRODUCTION

In today's world, as the population increases, people who are visually impaired also increases. Most common causes are: cataracts, as the leading one followed by glaucoma and age-related macular degeneration. Apart from the common causes of blindness, genetic factors play a role in many kinds of eye diseases which can be inherited by infants. With this, inborn blindness increases.

People who are visually impaired are having a hard time adjusting to the society especially in the working industry. In the labor market, most companies prefer sighted people over visually challenged ones given that most corporate jobs require sight to create, review and access files. Even though companies prefer sighted employee this did not stop visually challenged people to join the labor market. Visually challenged employees tend to do two times the effort of sighted employees but sometimes, it cannot be denied that sighted ones are more efficient and quicker in performing corporate jobs which leads to workplace prejudice, according to an article from euractiv, "Workplace prejudice keeps blind people out of employment", wherein they said that the blind and heavily impaired sighted people have the highest unemployment rate among disabled groups, despite their great desire to be part of the labor market because of workplace prejudice which often leads to discrimination, or worse, termination (Jacobsen, 2013). However, prejudice for the blind and visually challenged does not root from the workplace only because it can also be observed in school and in society itself. In fact, according to an article from livestrong, "challenges that visually challenged children face", most blind and visually disabled children perform low in the academe because apart from the access and environmental challenges, some children that has such disability also have trouble remembering what they read. Children with vision problems can find it difficult to write as well or might have problems with math and other subjects. Schoolwork can be a challenge for a child who cannot keep a clear focus, deals with double vision or blurred print on pages (Keefer, 2019).

Friends of Davao School for the blind is located at KM 10, Bago Aplaya, Talomo District Davao City. Mrs. Lolita D. Inocentes, Ph.D, who is the current school principal states that the school is the only one in Mindanao. The school aims to provide knowledge and information about the tech industry, specifically on its usage that will benefit and aid the blind students in their learning, especially reading. For the 38 years that friends of davao school for blinds stood,

continuous update to the emerging technology has been done in order to give what might benefit the students best. The school is government supported but is currently looking for sponsorship through their Facebook page and promoting it to share awareness to the netizens about the students who is striving to learn despite their blindness. In Friends of Davao School for the blind, most of them are genetically blind. Even if it is the case, it won't stop them from exploring and learning things such as reading and playing musical instruments. They also have their computer class that is currently handled by Mr. Mark Gomez, a blind computer instructor and then shared to us that they are still using the manual process of teaching the students how to read using the Braille system and screen reader application.

With the given problem, the proponents aim to develop a device that can help blind people take advantages of the present technology. The team will develop a device that can recognize text to assist a blind person in reading. The project will have the DIY VR Box (Virtual Reality Box) as the shell of the camera and microprocessor. With just the use of their hearing, the user will know the text in front of them as if they are seeing it. The device will look like a Virtual Reality glasses that can capture text images and convert it into sound. The sound will be considered as the translator for the blind person. The device just has one simple task to do to help a blind person. It is to recognize text. As it is obviously known that a completely blind person cannot read and with the help of the device, the blind can freely read text

Background of the problem

Friends of Davao School for the Blind is a school intended to help blind people learn things such as navigation, mobility and reading. The school still uses braille system for reading and adding the screen reader application for their computer class. Braille is a system of touch reading and writing for blind persons in which raised dots represent the letters of the alphabet. It also contains equivalents for punctuation marks and provides symbols to show letter groupings. For centuries, Braille has been useful in making blind people capable of reading and writing but the manufacturing processes of Braille will at some point fall short. The proponents believe this because according to an article from madehow, “How is Braille manufactured?”, unless a manuscript has been written originally in Braille, it needs to be translated. This is generally done by a typist using a special six-key typing machine may re-type the manuscript in Braille which makes manufacturing Braille laborious (Advameg, 2019). However, the Davao School for the Blind found another way to aid students to read books that are not originally written in Braille that is, using a screen reader application. Screen reader as defined, is a software application that enables people with severe visual impairments to use a computer. Screen readers work closely with the computer’s Operating System (OS) to provide information about icons, menus, dialogue boxes, files and folders. The device provides access to the entire OS that it works with, including many common applications. In Addition to this, all screen readers use a wide variety of keyboard commands to carry out different tasks. Tasks include reading part or whole of a document, navigating web pages, opening and closing files, editing and listening to music. A visually impaired computer user will use a combination of screen reader commands and operating system commands to accomplish the many tasks a computer is capable of performing. All current operating systems have their own keyboard shortcuts, which are available to everyone not just screen reader users (Watson, 2005)

Overview of the current state of the technology

Friends of Davao School for the Blind is a residential school and recognized by the Department of Education which caters preparatory class up to grade 10. The system called “Heariam” is to help the students make use of the current technology as they are still using the manual process. The academe uses the braille system which is a way of reading used by the blind people that can be felt by finger. The school’s secretary who is also the braille reading instructor teaches the students on how to use the braille during their class. The students will touch the dotted alphabet and recognize the text. The system takes a lot of time to use and recognize the text since the students’ needs to feel the dotted alphabet one by one. Since the academe have computer classes, their computer instructor assists the students to access the computer, they use the screen reader application with braille system. A Screen reader application is an application that uses a wide variety of keyboard commands to carry out different tasks. Tasks include reading part or whole of a document, navigating web pages, opening and closing files, editing and listening to music. A visually impaired computer user will use a combination of screen reader commands and operating system commands to accomplish the many tasks a computer is capable of performing. All current operating systems have their own keyboard shortcuts, which are available to everyone not just screen reader users. Despite the modern approach of the academe to improve the learning environment of the blind, the proponents believe that is still have rooms for improvement for the following reasons: (1) Manufacturing papers that comes in Braille, will at some point fall short, (2) Using a screen reader application is only accessible with assistance, (3) Apart from full dependence when accessing the screen reader, it also uses wide variety of commands which is inefficient for the blind and (4) bounded to be in the school premises only. Thus, the proponents aim to develop a device that follows the universal design suggested by W3C Web Content Accessibility Guidelines (Lazar et al., 2004) that is focused exclusively on Accessibility, Flexibility and Ease of use, and Visibility of Information.

Objectives of the study

The study aims to develop a device that follows the universal design suggested by W3C Web Content Accessibility Guidelines (Lazar et al., 2004) that is focused exclusively on Accessibility, Flexibility and Ease of use, and Visibility of Information that can also recognize text to assist blind person in reading independently. The proposed study is not only beneficial to the school but specifically to the students because the device is aimed to help the blind read texts in 10 frames interval instantly which is more efficient than using screen reader. Also, the device is based from a theoretical model VIVID (Vision Impaired using Virtual IT Discovery) that aims to maximize accessibility for majority of the multitude of disabilities which claims to help blind students achieve academic results that rivals of those that are sighted. This objective can be achieved by providing a module for pre-processing, for identifying or recognizing text and converting identified text into sound using text to speech synthesizer. Specifically, the proponents of the study want to decipher the following:

1. Develop a device that can convert text to speech through live camera/computer vision and will identify text in every FPS (Frame Per Second).
2. Develop a device that will make use of an Optical Character Recognition approach to identify text.
3. Develop a device that would help the blind and visually impaired children to be more capable in the academic field.

Scope and Limitations of the study

Scope

The project will focus on developing a device that can recognize or identify text to help the user make use of the advance technology and have an accurate and easier way of reading text. The proposed project wearable device that can recognize text to assist blind person in reading has module for pre-processing, recognize text and converting text to sound:

Pre-processing

- The device's code will run automatically after the raspberry pi boot up.
- The device provides a live camera or computer vision that will identify or track texts in every FPS (Frame Per Second).
- The device will start again in pre-processing stage when text-recognition and convert text image to sound is done.

Text Recognition

- The device can recognize both upper- and lower-case letters, numbers.
- The device will use EAST Deep Learning Algorithm to predict red and green pixels
- The device will use thresholding method for recognizing texts.

Convert Text-Image to Sound

- After Identifying the text, it will be converted into sound using text to speech library from python.

Limitations

Developing a project has a lot of advantages and disadvantages as well as having limitations in its capability.

- The device will only recognize fonts that are trained.
- The device will only recognize up to 18 font size above.
- The device can recognize text only up to half a meter.
- The device recommends black fonts and white background.
- The device should be used up to 8hrs a day.
- The device would need extra battery to power up.

LITERATURE REVIEW

Review of related literature, studies or systems

Friends of Davao School for the Blind always aim to help blind people learn things like a normal student. Aside from the school, there are people and organizations that created a system that will help blind people become efficient in their day-to-day life with the use of advanced technology. There are also researchers who want to further their knowledge about the studies to help blind people and the ones who found solutions. Some studies are still being tested and will also become a tool that can help blind people recognize objects or text and detect obstacles. To further understand this project, the proponents found different materials that can be used for the project and is related to the E-learning system with AI that can assist blind person in reading that will become a big help in producing an output.

Guidelines for Developing e-Learning System for Visually Impaired

Guidelines for developing e-Learning System for Visually Impaired is an analytical paper that defines the guidelines for adapting an LMS to blind and visually impaired users in nine criteria with the aid of existing Web 2.0 technology. The problems identified in this analytical paper is based on a research carried out in the course of working with various associations for the blind and visually impaired in Zagreb University in Croatia. The goal of this analytical paper is to create guidelines that will improve the LMS for blind and visually impaired by meeting the universal design, which is focused exclusively on Flexibility in use, Ease of use, Visible Information. Additionally, the proponents of this analytical paper that the goals of modeling framework of assistive technologies shall be met such as, application in any assistive technology system, the possibility of classifying assistive technology systems, defining the basic structure of assistive technology systems in device specifications, developing new assistive technology systems with the aim of meeting the needs of end users and supporting the process of providing assistive technology to the user, aimed at the user accepting the solution provided. The guidelines were defined using the analysis of the LMS currently in Croatia namely, The

Moodle-based Merlin, e-Student System in used at the Faculty of Transport and Traffic Sciences and AAI@EduHr user login system. The evaluation was carried out using the validation tool on (<http://validator.w3.org>) and the AChecker tool for evaluating e-accessibility which later found 48 known problems and 300 potential problems for Merlin, 0 known problems and 30 potential problems for AAI@EduHr and 11 known problems and 63 potential errors for E-student. In conclusion, this analytical paper found a number of flaws with the LMS currently being used in Croatia with regard to accessibility to the blind and visually impaired and considered the 9 guidelines, namely, Criterion 1 – HTML row and column design, Criterion 2 – Keyboard shortcuts, Criterion 3 – Direction of navigating a user interface, Criterion 4 – Information architecture, Criterion 5 – Interface optimization, Criterion 6 – HTML tags, Criterion 7 – Auditory feedback, Criterion 8 – Confirming data entry and Criterion 9 – Accessibility evaluation to be vital and highly recommended for implementation for the development of e-Learning systems.

Watson Text-To-Speech - IBM

Watson Text-To-Speech Synthesizer is an IBM that provides APIs that uses the speech-synthesis to synthesize text into speech using a different kind of languages, dialect and voices. It supports both female and male voices. It also supports a WebSocket interface that provides both plain text and SSML (Speech synthesis markup language) input, including the SSML elements and word timings. Speech Synthesis Markup Language is an XML based markup language that provide text comments for speech-synthesis applications. One example that uses this API is the Watson-Chatbot that performs the functionality of Google Assistant that also uses a Google API text to speech synthesizer. The proponents will be using this since it is free and more user friendly than Google text to speech synthesizer.

E-Learning for the Visually Impaired: A Holistic Perspective

The E-Learning for the Visually Impaired: A Holistic Perspective is an analytical research that investigated the common problems that blind and visually impaired students faces using interview and observation method in a 2-year duration. The proponents interviewed the visually challenged students and the sighted instructors to give emphasis on the severity of the problem and suggested resolution, which lead for the proponents to come to an understanding that most LMS is designed for the sighted and lacks accessibility for the students with disabilities. To

address the problem, the proponents believe that there should be changes in its design and accessibility, in an effort to supplement the guidelines and methods to improve the current situation, the proponents have used former research and studies such as Daniels and Elliott (2003) set of generic guidelines and processes for testing the accessibility of e-learning websites based upon an earlier version of W3C Web Content Accessibility Guidelines in addition to the earlier work on usability and mobility by other researchers, Lazar (Lazar et al., 2004) presented the Web Accessibility Integration Model based upon the influence of stakeholders perspective, Web accessibility and societal foundation, WAI (2006) which presented an accessible e-learning model that is contextual in nature and centers on the aspect of stakeholder involvement which comprises accessibility drivers in the form of guidelines, standards and legal requirements and Kelly, Phipps and Swift (2005) introduced the most holistic model for accessibility and e-learning, this model is circular in format to illustrate that learning is a holistic, cultural, political activity which also includes and social aspects for learners. All of which previous studies and research contributed to the incorporation of VIVID (Vision Impaired using Virtual IT Discovery). VIVID is a theoretical model that maximizes accessibility for majority of the multitude of disabilities which students may have, VIVID utilizes a holistic approach that overcomes many of the shortcomings of previously developed models and standards while building on the significant contributions previously made. The cyclic development process specified within the VIVID model is designed to be interactive and reflective course development thus, allowing modification based on experience which is expected results are for visually impaired students to achieve the same learning outcomes as the sighted peers. A key factor in VIVID is it's capable of a flexible delivery timeline is necessary due to differing nature of each individual's students' disabilities such as recording the reading speed, cognition, physical access methods, and completion of tasks. The use of this model will facilitate employment and social inclusion through educational programs that include the peer support and networking opportunities currently enjoyed by sighted students in mainstream education.

Artificial Intelligence in Education of Children with Learning Disabilities

Artificial Intelligence in Education of Children with Learning Disabilities is a research paper mainly focused on using technology and A.I for creating an improved education field for Children, helping students with learning disabilities achieve outcomes as with those who do not have learning disabilities in a general case that traditional learning is not effective. In an effort to create an improved learning environment, the proponents tried to identify the weaknesses and strengths of the subject matter and incorporate it to the methods they will use to diagnose and develop educational plan for students with learning disability. However, the current AI – ANN (Artificial Neural Network) and SVM (Support Vector Machine) although gained lots of attention in AI world but the results that they give when diagnosing a students' weakness in learning is not highly accurate, that is why they never appeared to be used in diagnosing Learning Disability. With this, a new approach was explored to diagnose children's learning disability and that is another expert system was designed to help diagnose children with learning disability using a learned knowledge and heuristics to evaluate child's process using gaming strategies, once deficiencies has been identified by the system, the child can start using the system in relation to improving the deficiency sighted. Another is the proponents believed that one factor in learning is to the motivation to learn so they developing a system that uses cognitive approach which recognizes the subject's strengths and weaknesses as well as adapt to the subject's interest to keep one's motivation. In addition to systems and devices, the proponents believe that using a figure (be it as a humanoid or a virtual assistant) that closely relates to human appearance help the subject more focus on the task. In conclusion, the research found that with the rising number of kids with disabilities, the community must adapt from its mainstream education to help create an environment that fosters and gathers results and provide equal opportunities to all. Furthermore, this research highlighted the work done and potential development AI and technology can bring to the current education that the world has.

Synthesis

The proposed project “Heariam: Computer Vision And Text To Speech Artificial Intelligence” is new. According to Lolita D. Inocentes, the principal of Friends of Davao School for the Blind, it is the only school in Mindanao that provides education for blind people. The teacher for their computer class, Mark Homez, stated that they are willing to use the proposed project as long as it is ready to use.

With the above-mentioned studies, it is understood that creating an accessible device for students with disabilities that also converts text-to-speech needs to follow a universal and holistic design and use multiple open source libraries to build the project. Guidelines for developing an accessible LMS for the blind and visually impaired students is a study that tackles the common problems that students with disabilities encounter and its resolution to achieve an outcome that is equal to those with sighted students. VIVID, is a theoretical model that incorporated most previous studies and research to improve the educational experience of blind students and maximize the results by tracking all activities in a flexible delivery time and find out the strengths and weaknesses of the student which will help the instructor further assist the needs of the student. Text-to-speech synthesizer is a study or software that supports different kinds of open source library for text and speech synthesis. The goal of the project is to create a software that will improve the educational experience of the student and help them create outcomes that equals sighted students.

The device works by using opencv for the computer vision that captures the texts in the image, represented as red pixels. Then, the proponents used EAST deep learning algorithm to map and track. Lastly, after mapping and tracking all red pixels, the proponents used tesseract as the OCR (Optical Character Recognition engine) to recognize all the red pixels as text and convert it to speech with the use of python speech library. The device will restart the process once the camera sights or identifies a new text.

There are a lot of studies that aim to improve the educational and overall experience of the blind and visually impaired students, but the key point is that they all aim to help blind students achieve results equal to those sighted ones and have their abilities acknowledged by society, which inspired the proponents to derive Heariam, a device aimed not only to assist but to also create a whole new independent world for the blind and visually challenged. Therefore,

understanding the very objective of this study and learning different algorithms that will contribute to creating an accessible and improved system for the blind and visually challenged will determine the project's effectiveness, efficiency and accuracy.

METHODOLOGY

The methodology that will be used is Prototyping because it is one of the SDLC model that is useful when developing and enhancing a pre-existing project. Prototyping is used when the developers do not know the requirements at first. In this process, a prototype is first developed, tested and revised repeatedly as per clients request until a final revision is accepted. The accepted revision will then be basis in developing the final product.

In this process model, the initial project is partially implemented before or during the analysis phase to give the customers an opportunity to check or see the product early in the life cycle. The process will start by interviewing the customers about the requirements needed and then develop an incomplete paper model. This process will be used to build the initial prototype that supports the basic requirements that the customers wants or needs. After the customers figure out the basic problems, the prototype will be enhanced in order to meet other requirements or eliminate the problems. The process of finding problems and enhancing the prototype will continue until the user approves and be satisfied by the final product.

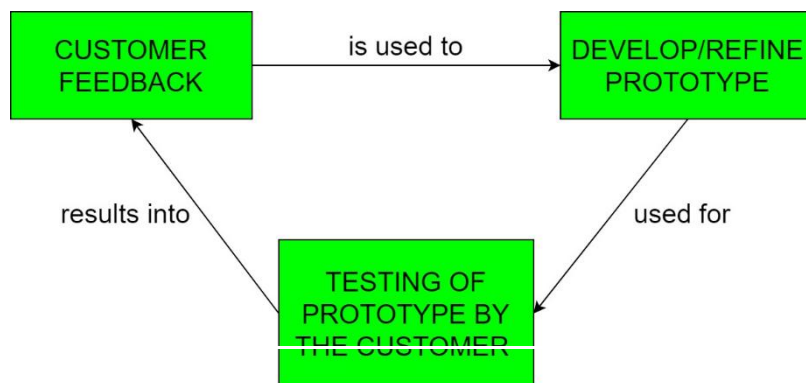


FIGURE 1.0: Prototyping Diagram

The proponents of HEARIAM: E LEARNING SYSTEM WITH ARTIFICIAL INTELLIGENCE FOR THE BLIND CHILDREN, in their prototyping methodology, their approach is to do evolutionary prototyping. In this kind of prototyping endeavors the major specialty of the technology of the proponents use which are the Speech Synthesizer, Firebase Cloud, and Oxford Dictionary API.

In stages of the development, there are three cycle types of development on the proponents' use. The first stage is Customer Feedback, the second stage is Develop or Refine Prototype, and the third stage is Testing of Prototype by the Customer. On the contrary of definition why proponents highlight the evolution prototyping in developing the system, it is to highlight the milestones of the development and to highlight the main function of system requirements of the user's ability to cope up in using the system.

Customer Feedback

Initialization of the proponents' project, the customer feedback relies on the proponents' goal or foresaw the initial problem of the disabled community especially the blind people at the community they approach to conduct an experiment. On the contrary of the blind people, the technology factor they use is too old and limits the grasp of its nature to its ability to interact with other people. Whereas our technology nowadays is continuously implacable which solves the problem in the past. And with this concept of idea to the blind people, the proponents initiate its path on its problem and its logic to concept its functional activities to the system.

Develop or Refine Prototype

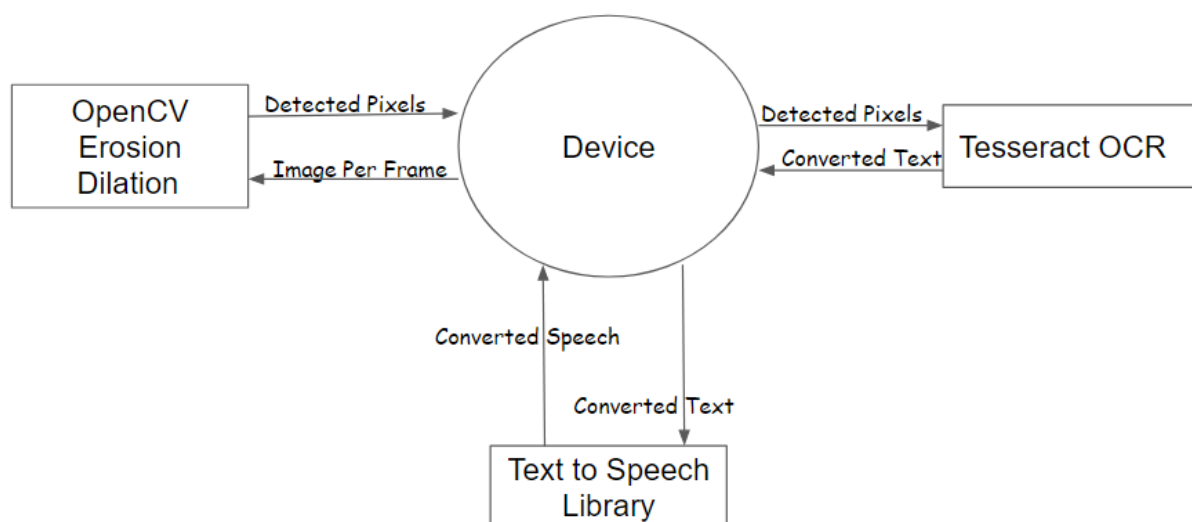
On the initial concept of the system stated at the customer feedback, the logic for its development rationalize the functionalities to its minimum or most required components for the ideal development. And with the concept of a use case of the development of the system, the proponent will be able to gather its peripheral concept to develop the ideals or concept of the system.

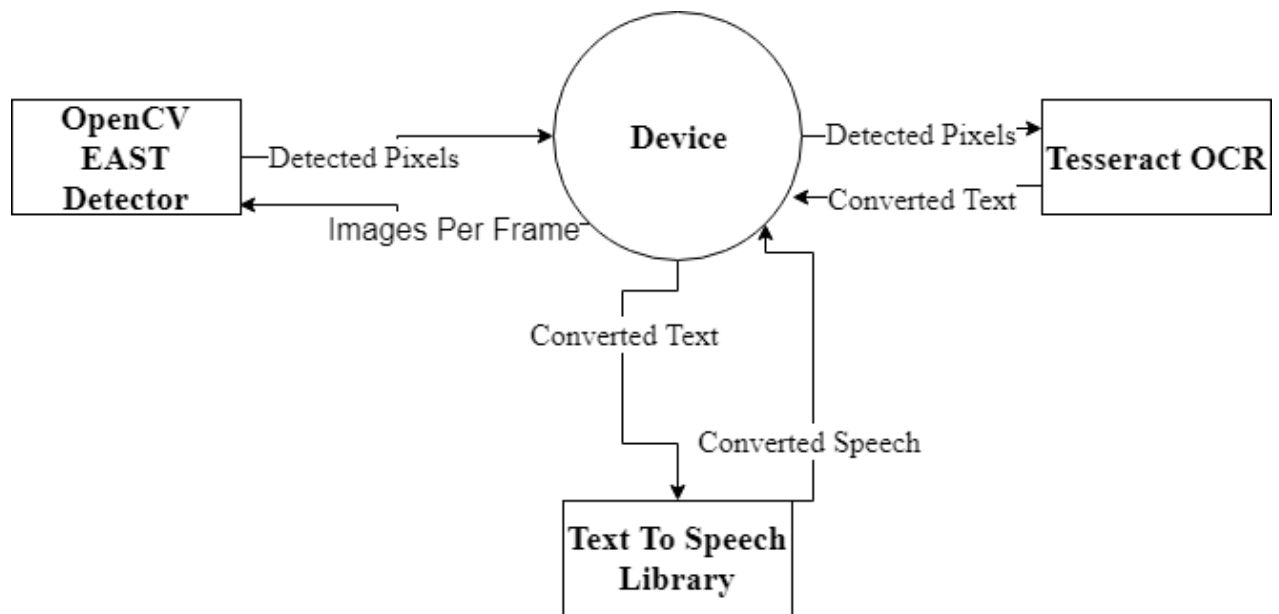
The minimized hardware concept on what the goal project of the proponents, the proponents used OpenCV which uses Computer Vision or enabling the camera from the device and with the help of EAST Deep Learning Algorithm it can classify all pixilated colors coming from the image captured by OpenCV. Then proponents use open source libraries to make the project possible in an efficient calendar of the development. The libraries are Tesseract and Python's Text to Speech library. In this open source libraries, the Python programming language was

chosen as a primary language for the development to support for all stated open source libraries. Tesseract is an OCR engine popularly used in many developers in Machine Learning it uses thresholding methods to get the text from an image. Python's Text to Speech Library is a library of millions of speech in different languages one of the popular used is English.

There is a study conducted in Zagreb University in Croatia. In their study, it specifies that there is a need to adjust the current LMS for the blind students. As stated, it should be easily accessible for the blind and to follow a universal design, guidelines was based from the former studies and research provided by [www.3.org](http://www.w3.org) (Web Content Accessibility Guidelines – WCAG 2.0). In relation to this study, Heariam: E-learning system with artificial intelligence, is provides easier access and flexible educational environment for both the students and the instructors.

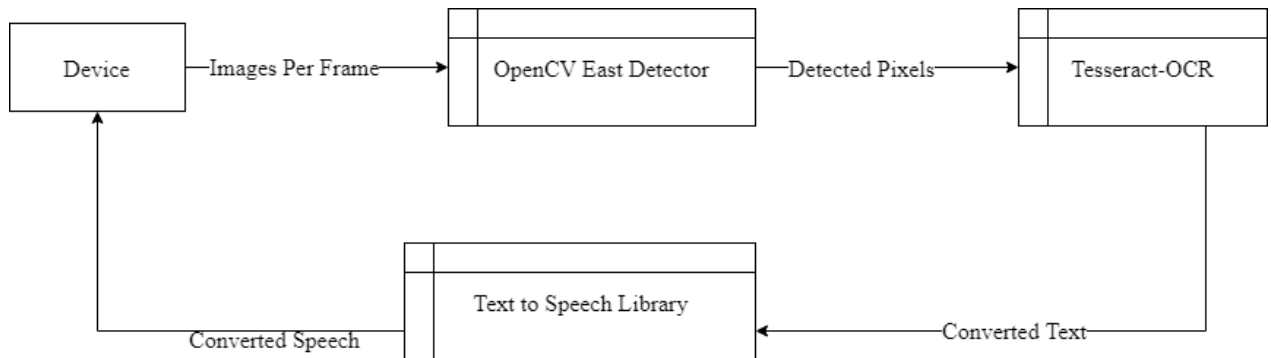
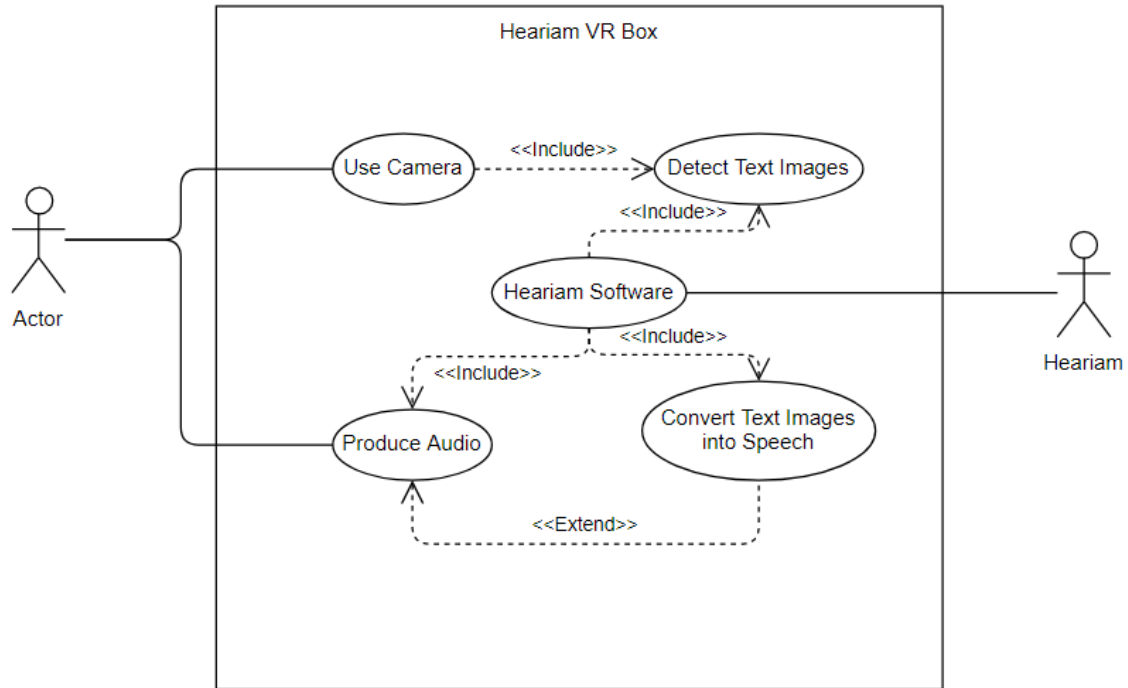
Figure 2.0 Context Diagram





This is a broad overview with the flow or the process of the device. It demonstrates the interaction between the process and external process entities.

Figure 2.1 Data Flow Diagram/Use Case



This shows the relationship between entity sets that are process every frame to recognize text and convert it to speech.

Figure 3.0 Setting Up OpenCV and EAST

This is the code for enabling OpenCV and calling EAST algorithm. Wherein the device will detect all the text in range.

Figure 3.1 Detected Text

This is what it looks like when it detects text from every image in every Frame Per Second.

Figure 4.0 Getting the ROI for each frame

This function will get the x and y axis for each frame to determine all red pixels and use it as a classifier.

Figure 4.1 Algorithm for EAST Detector

This function shows that every image in every frame the x and y axis will acquire.

Figure 5.0 Using the Tesseract-OCR engine

This is where the proponents call Tesseract-OCR engine to perform thresholding method to process and convert all red pixels to text from every image in Frame Per Second.

Figure 6.0 Using the Pyttsx3 Text to Speech Library

This is where we call the Text to Speech open source library from python and turn the resulted text from tesseract and then convert it in to speech.

Testing of Prototype by the Customer

In the study conducted by International Journal of Cyber Society and Education, they found that to address the common problems blind students encounter, a holistic approach must be used.

In their study, the holistic approach incorporated previous research and studies including the WCAG 2.0 guidelines and included other factors such as influences from stakeholders, accessibility drivers, standards and legal requirements to create a theoretical model that overcomes the shortcomings of previous models and studies. The proponents believe that this model will illustrate that learning is a holistic, cultural, political activity which also includes social aspects for learners, this theoretical model was later named VIVID (Vision Impaired using Virtual IT Discovery), a model capable of a flexible delivery timeline necessary due to differing nature of each individual's students' disabilities, such capability as recording the reading speed, cognition, physical access methods, and completion of tasks.

With the result of the former studies and VIVID model, the proponents aim to create a better environment for the blind and hopes that this will facilitate employment and social inclusion through educational programs that include the peer support and networking opportunities currently enjoyed by sighted students in mainstream education.

RESULTS AND DISCUSSIONS

In this chapter shows how the system works including the use of the device. As seen on Chapter 3, what each code does and represent. This also includes different test cases that show the behavior of the system.

The different tables show the different test cases of the device. These are the behavior of the system on how the system works.

Experimental

Table 1.0 Booting up

No	Activities	Expected Result	Actual Result	Interface
1	Turning on device	The code from the device should automatically run using nano.	The code will run automatically	Figure 7.0

Table 2.0 Pre-processing

No	Activities	Expected Result	Actual Result	Interface
1	OpenCV will turn the camera on	The device's Camera should be on	Camera will be turn on	Figure 3.0
2	EAST algorithm will detect all the red pixels.	All red pixels in each image in frame per second should be detected	Red pixels should all be detected	Figure 3.1

Table 3.0 Converting Text

No	Activities	Expected Result	Actual Result	Interface
1	Tesseract -OCR engine will convert all text red pixels to text	Converted Text will be shown	Converted Text output	Figure 5.0

Table 4.0 Converting Speech

No	Activities	Expected Result	Actual Result	Interface
1	Text results will turn in to speech	Device will speak the text captured by the camera	Text	Figure 6.0

CONCLUSION AND RECOMMENDATIONS

The proponents made a device that has addressed the major objectives of the study. The achieved result is more or less 80% of the expected functionality, heariam is capable of tracking, identifying and reading texts however, it is yet to be as quick, as instant and as accurate than the project claims to; the proponents believe that heariam has to go through more tests, implementation and development for it to reach the expected result.

The proponents had achieved the necessary platforms that operates heariam, which was discussed in different chapters of the paper. Opencv is capable of successfully capturing the contents of the image. EAST deep learning algorithm is capable of successfully mapping all the text shown as red pixels. Tesseract is capable of successfully recognize all the red pixels and convert it to speech using python speech library. Currently, the device performs to its expected functions but the proponents believe that the current data processing of the device should be further developed in the near future to make it more ideal and convenient to the user.

Hardware/Software

Hardware

Camera Module V2 Raspberry Pi – Capture image text for image processing.

Raspberry Pi 3 Model B- Main board or the microprocessor.

Headphone- Hardware output device that plug into a computer line out to privately listen to audio. It is used to listen for audio feedbacks.

Software

Raspbian OS version Raspbian- is a free operating system based on Debian optimized for the Raspberry Pi hardware. An operating system is the set of basic programs and utilities that make your Raspberry Pi run.

Python3 version 3.7.2 - is a high-level programming language designed to be easy to read and simple to implement. It is open source, which means it is free to use, even for commercial

applications. It is considered a scripting language, like Ruby or Perl and is often used for creating Web applications and dynamic Web content.

OpenCV version 3.2- (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in commercial products.

Tesseract-OCR- Tesseract is an optical character recognition engine for various operating systems. It is free software, released under the Apache License, Version 2.0, and development has been sponsored by Google since 2006. In 2006, Tesseract was considered one of the most accurate open-source OCR engines then available.

Calendar of Activities

This part is where the proponents listed all the activities before, during and after the process of developing and planning of the project.

Thesis Project Defence

Title Defense (1st week of January 2019)

The proponents presented the title to the panelists. The title presented is “A Wearable Device That Can Recognize Object and Text to Assist A Blind Person”. This is the original title that was proposed but due to its feasibility and accuracy issues, the title was changed to “Heariam: A Wearable Device That Can Recognize Text to Assist Blind Person in Reading”.

Mock Defense (3rd week of February 2019)

The proponents were being tested if the documentation and presentation are all checked and completed.

Revision (4th week of February 2019)

The proponents have some changes in the documentation after the mock defense. Instead of having object and obstacle detection, the project was changed to text recognition for blind people. There are few revisions in the documentation but still manages to complete it.

Final Defense for thesis 0 (3rd week of March 2019)

The proponents are getting ready for the final day to present the proposed project.

Revision Thesis1 (3rd week of July 2019)

The proponents have some changes in the documentation after the final defense.

Documentation

Chapter 1 (2nd week of January 2019)

The proposed title was already accepted and the proponents started doing the chapter 1 of documentation which contains the background of the problem, overview of the current state of technology, objectives of the study, and scope and limitations of the study. For the background of the problem, the proponents gathered information of the main problem that the client encountered. Overview of the current state of the technology is where the proponents researched things about the client to look for any contact information for visiting the client to know the current technology they are using. As the proponents looked for the problems encountered by the client, the objectives were formed to help the client find solutions to their problem. After knowing the problem and objectives, the proponents decipher all the things needed in developing the project by listing down all of the scope and limitations.

Chapter 2 (3rd week of January 2019)

The proponents already have the things needed in planning the development. To further understand about the project, the proponents researched for things related to the project in order to broaden knowledge about what the study is. Some materials are quite similar with the study but is still lacking some proof and research.

Chapter 3 (3rd week of January 2019)

Along with Chapter 2 the proponents also did the Chapter 3 Documentation

Checking/Changes Documentation 1 (2nd week of February 2019)

Before the mock defense the proponents double check with the adviser the documentation.

Checking/Changes Documentation 2 (2nd week of March 2019)

After the revision the proponents reviewed and did some changes for the documentation.

Environment Activity

Consulting Machine Learning Community (2nd week of June 2019)

The proponents consult various Machine Learning Community for the purpose of how to build the device properly

Canvas Equipment's (3rd week of June 2019)

The proponents went to CreateLabz to canvas the equipment's used to develop the device.

Setting up/Getting started in Coding Development (3rd week of August 2019)

The proponents initially started to develop the code for the device.

Research + Discovery

Clients Current State of Technology (1st week of February 2019)

Along with the Initial Visit of the client the proponents also discover how they operate their computer class.

East Deep Learning Algorithm (1st week of August 2019)

The proponents research about the East Deep Learning Algorithm and how to use it.

Python (2nd week of July 2019)

The proponents started learning python language for machine learning.

OpenCV (2nd week of July 2019)

The proponents learned how to use OpenCV.

Tesseract OCR (2nd week of July 2019)

The proponents research how does Optical Character Recognition works and how tesseract operates it.

Image Processing (2nd week of July 2019)

The proponents research Image processing and how it helps our device.

Raspberry pi Concept and Parts (3rd week of July 2019)

The proponents learned how to operate Raspberry pi and what parts should be used.

Code Development

Implementing Computer Vision with OpenCV (3rd week of July 2019)

The proponents started coding of the device's vision.

Implementing East Deep Learning Algorithm (3rd week of July 2019)

The proponents implemented the algorithm for the device.

Implementing Tesseract (1st week of September 2019)

The proponents implemented the OCR engine for the device.

Checking Prototype (1st week of September 2019)

After all the implementation for the code the proponents noticed that there are bugs needed to fix or make the device recognize text more accurately and fast.

Improving FPS (2nd week of September 2019)

The proponents Improve the accuracy of the algorithm and it was a success.

Improving OCR (2nd week of September 2019)

The proponents Improve the accuracy for OCR engine and it was a success.

Design

Planning Size and look

After The development the proponents Initially design the DIY VR BOX

Making of the Device's DIY box (2nd week of September)

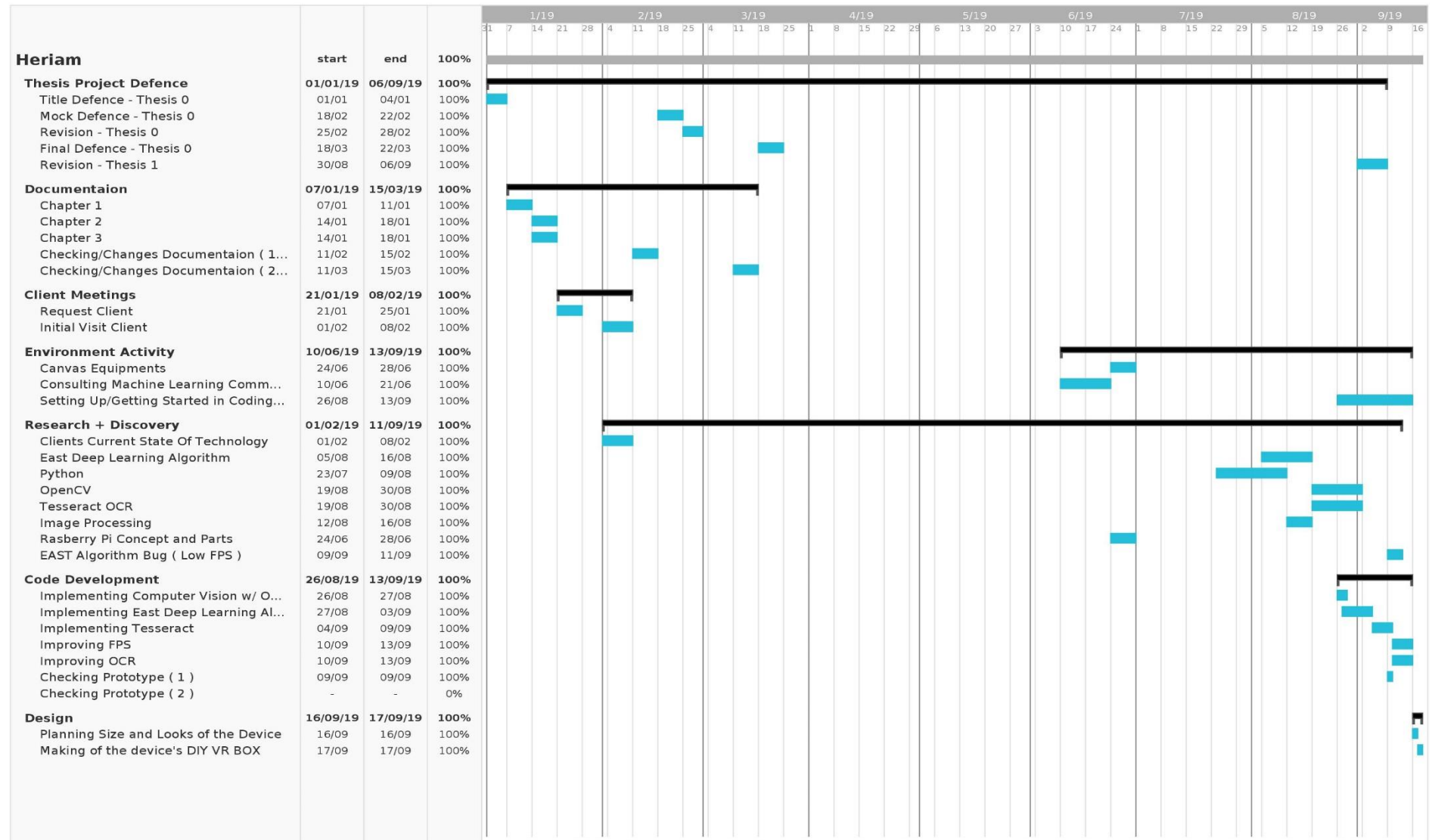
The proponents initially make the DIY box for the vessel of the device.

Changes in grouping and budgeting (2nd week of July 2019)

The proponents lost a group member resulting in changing the proposed title for that reason the proponents don't have enough budget to provide the expenses required to develop the device and decided that instead a device, the proponents will develop an AI system.

Developing and Documenting (3rd week of July (2019)

The proponents started developing the new proposed title and revise the documentation from device to AI system.



Legend: Blue - Completed

Budgetary Estimate

Quantity	Specifics	Approximate Cost
1	Rasberry Pi Model b+	3999.00
1	Rasbery Pi Camera V2	2799.00
1	HDMI converter	799.00

APPENDIX

Figure 3.0 Setting Up OpenCV and EAST

```
ap = argparse.ArgumentParser()
ap.add_argument("-e", "--east", type=str, required=True,
    help="path to input EAST text detector")
ap.add_argument("-v", "--video", type=str,
    help="path to optional input video file")
ap.add_argument("-c", "--min-confidence", type=float, default=0.5,
    help="minimum probability required to inspect a region")
ap.add_argument("-w", "--width", type=int, default=320,
    help="resized image width (should be multiple of 32)")
ap.add_argument("-h", "--height", type=int, default=320,
    help="resized image height (should be multiple of 32)")
args = vars(ap.parse_args())
```

Figure 3.1 Detected Text

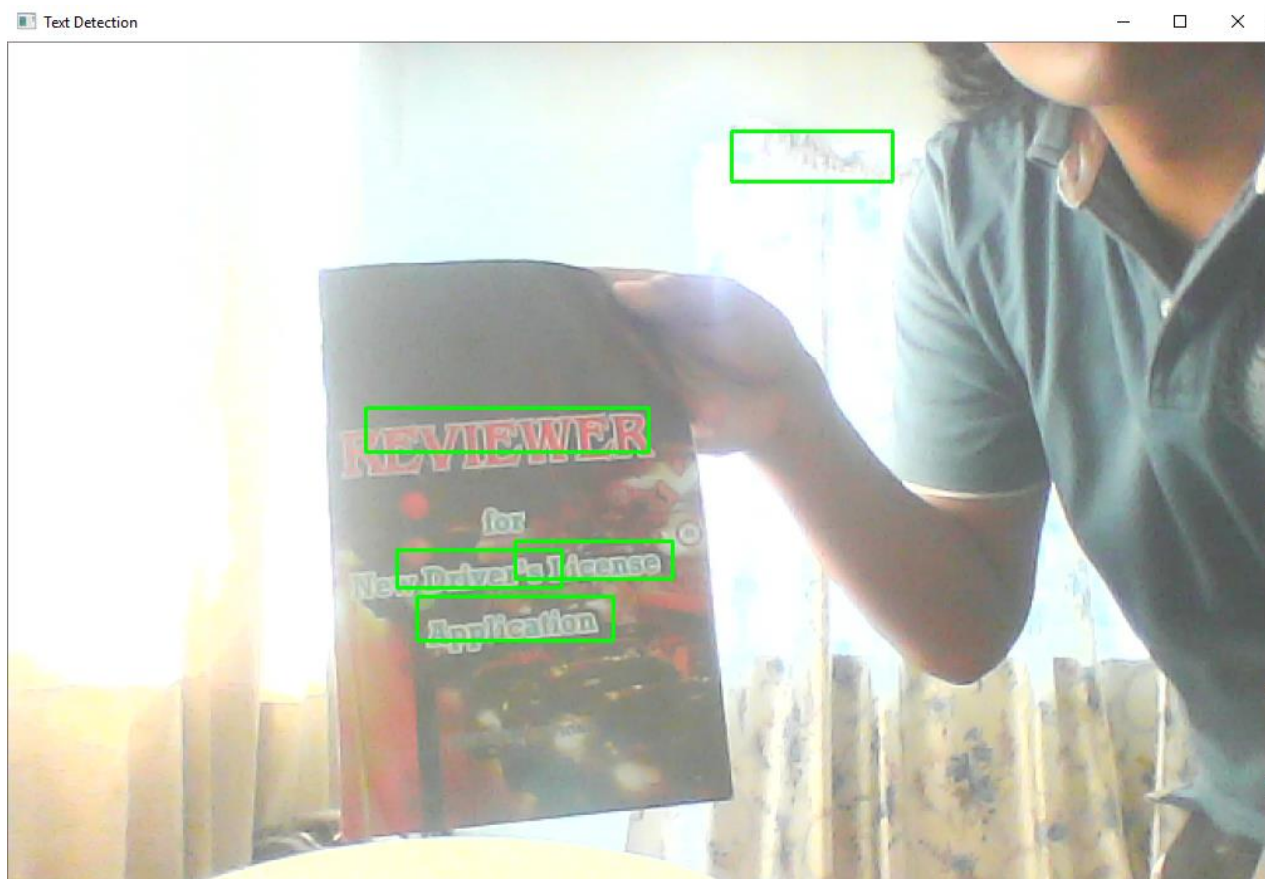


Figure 4.0 Getting the ROI for each frame

```
roi = orig[startY:endY, startX:endX]
config = ("-l eng --oem 1 --psm 7")
text = pytesseract.image_to_string(roi, config=config)
```

Figure 4.1 Algorithm for EAST Detector

```
def decode_predictions(scores, geometry):
    (numRows, numCols) = scores.shape[2:4]
    rects = []
    confidences = []

    for y in range(0, numRows):
        scoresData = scores[0, 0, y]
        xData0 = geometry[0, 0, y]
        xData1 = geometry[0, 1, y]
        xData2 = geometry[0, 2, y]
        xData3 = geometry[0, 3, y]
        anglesData = geometry[0, 4, y]

        for x in range(0, numCols):
            if scoresData[x] < args["min_confidence"]:
                continue

            (offsetX, offsetY) = (x * 4.0, y * 4.0)

            angle = anglesData[x]
            cos = np.cos(angle)
            sin = np.sin(angle)

            h = xData0[x] + xData2[x]
            w = xData1[x] + xData3[x]

            endX = int(offsetX + (cos * xData1[x]) + (sin * xData2[x]))
            endY = int(offsetY - (sin * xData1[x]) + (cos * xData2[x]))
            startX = int(endX - w)
            startY = int(endY - h)

            rects.append((startX, startY, endX, endY))
            confidences.append(scoresData[x])

    return (rects, confidences)
```

You, a few seconds ago • Uncommitted changes

Figure 5.0 Using the Tesseract-OCR engine

```
lower_white = np.array([0,0,255-sensitivity], dtype=np.uint8)
upper_white = np.array([255,sensitivity,255], dtype=np.uint8)

mask = cv2.inRange(hsv, lower_white, upper_white)

i=i + 1
result = ''
if i == 10:
    result = pytesseract.image_to_string(mask)
    print(result)
```

Figure 6.0 Using the Pyttsx3 Text to Speech Library

```
voices = engine.getProperty("voices")
engine.setProperty("voice", voices[1].id)

engine.say(result)
engine.runAndWait()
str = re.sub('[^0-9a-zA-Z -]+', '', result)
i=1 You, 10 days ago * initial commit
```

Figure 7.0 Automatically run the code after booting up the device

```
pi@dex:~$ sudo chmod 644 /lib/systemd/system/sample.service
pi@dex:~$ sudo systemctl daemon-reload
pi@dex:~$ sudo systemctl enable sample.service
Created symlink from /etc/systemd/system/multi-user.target.wants/sample.service to /lib/systemd/system/sample.service.
pi@dex:~$ sudo reboot
```


Human Resources

The following pages contains the curriculum vitae of the researchers and the Adviser's Acceptance Form.

Curriculum Vitae of
DAMIEL F. MAQUILING
 Blk 4 Denmark st. lot 22 Vista Verde Panacan
 damiel.maquiling@gmail.com
 09269873428

EDUCATIONAL BACKGROUND

Level	Inclusive Dates	Name of school/Institution
Tertiary	2014-present	STI College of Davao
High School	2009-2013	F. Bustamante National High School
Elementary	2003-2009	Juan Cruz Elementary School

PROFESSIONAL OR VOLUNTEER EXPERIENCE

Inclusive Dates	Nature of Experience/Job Title	Name and Address of Company or Organization
02/2019-04/2019	Intern Mobile Developer	Traxion Tech.
05/2016-06/2016	Intern Front End Developer	Jairosoft

AFFILIATIONS

Inclusive Dates	Name of Organization	Position
07/2018-04/2019	iRobot Club	Auditor
07/2018-10/2018	CS Department	Auditor
03/2019-Present	DevCon Davao	Program Lead

SKILLS

- Proficient in both oral and written communication
- Computer Literate
- Proficient in using Microsoft Word, PowerPoint, Excel and OpenOffice
- Basic Computer Troubleshooting
- Mobile Development
- Basic Knowledge of Java, HTML, CSS, JavaScript, PHP, SQL, Firebase programming languages and/or software applications

TRAININGS, SEMINARS OR WORKSHOP ATTENDED

Inclusive Dates	Title of Training, Seminar or Workshop
11/2018	U: Hackademia X STI Hosted by UnionBank
10/2018	First Aid: Basic Life Support Training
09/2018	Intro to Arduino Tech Seminar
09/2018	Intro to WordPress Seminar

09/2018 Image Processing Using Raspberry Pi Seminar

Curriculum Vitae of
KIMBER CHAN L. GRANADEROS
 San Vicente YCF, Brgy. San Antonio, Agdao, Davao City
 09218367667

EDUCATIONAL BACKGROUND

Level	Inclusive Dates	Name of school/ Institution
Tertiary	06/2014-Present	STI College of Davao
High School	06/2010-04/2014	Holy Cross of Agdao
Elementary	06/2004-03/2010	South San Juan Elementary School

PROFESSIONAL OR VOLUNTEER EXPERIENCE

Inclusive Dates	Nature of Experience/Job Title	Name and Address of Company or Organization
02/2019 - 04/2019	Internship	Kadena Hub

AFFILIATIONS

Inclusive Dates	Name of Organization	Position
03/2019-Present	DevCon Davao	DevCon Campus
06/2018-04/2019	iRobots	Secretary

SKILLS

Skills	Level of Competency	Date Acquired
Django Python	Intermediate	02/2019
Vue JS JavaScript	Intermediate	10/2018
C#	Intermediate	01/2015
Java	Intermediate	08/2014

TRAININGS, SEMINARS OR WORKSHOP ATTENDED

Inclusive Dates	Title of Training, Seminar or Workshop
01/2019	Spacelab: Hyperledger Composer
11/2018	Davao JS Geekup (Vue JS)
03/2018	Swiftkind Code Talk (Python)

ADVISER'S ACCEPTANCE FORM

NAME OF PROPONENTS:

Daniel F. Maquiling

Kimber Chan L. Granaderos

APPROVED RESEARCH TITLE:

HEARIAM: A WEARABLE DEVICE THAT CAN
RECOGNIZE TEXT TO ASSIST BLIND PERSON IN
READING

AREA OF STUDY:

Image Processing

CONFORME:

Lexfrevail P. Nervida

Thesis Adviser

APPROVED BY:

Lexfrevail P. Nervida

Thesis Coordinator

Date: 10 April, 2019

NOTED BY:

Leonardo Jr. E. Risma

Program Head

REFERENCES

(Braille Production, n.d)

Retrieved from:

<http://www.madehow.com/Volume-4/Braille-Publication.html>

Davydenko, E. (2012). Artificial Intelligence in education of children with learning disabilities.

Retrieved from:

<https://pdf.semanticscholar.org/cacb/b5ccdb7a5a566968af749b1c45abf4e1ea69.pdf>

Jacobsen, H. (2013). Workplace prejudice keeps blind people out of employment.

Retrieved from:

https://www.euractiv.com/section/health-consumers/news/workplace-prejudice-keeps-blind-people-out-of-employment/?fbclid=IwAR3RFxEAwCvy_QVHEOFOLHax4yLey4I4910asIHicsmxpeBIXquzcq-QzZU

Keefer, A. (2019). Challenges that visually challenged children face.

Retrieved from:

<https://www.livestrong.com/article/126096-characteristics-visually-impaired-child/>

Murray, I. (2013). E-learning for the vision impaired.

Retrieved from:

https://www.researchgate.net/publication/258022459_E-LEARNING_FOR_THE_VISION_IMPAIRED_A_HOLISTIC_PERSPECTIVE

Perisa, Perakovic, & Remenar. (n.d). Guidelines for Developing E-Learning system for Visually Impaired.

Retrieved from:

https://bib.irb.hr/datoteka/615063.S3_01_.pdf

Watson, L. (2005). What is a screen reader.

Retrieved from:

<https://www.nomensa.com/blog/2005/what-screen-reader>

Watson Text-To-Speech IBM (2015). Retrieved from <https://cloud.ibm.com/docs/services/text-to-speech?topic=text-to-speech-about>