

Mobile app with reading speech-translated OCR images for visually impaired people**Aplicación móvil con lectura de imágenes OCR traducidas a voz para personas con discapacidad visual**

VAZQUEZ-GUZMAN, Francisco†*, OLGUIN-GIL, Liliana Elena, VAZQUEZ-ZAYAS, Eduardo and NICANOR-PIMENTEL, Brawhim Jesseth

Tecnológico Nacional de México/Instituto Tecnológico de Tehuacán, Mexico.

ID 1st Author: *Francisco, Vázquez-Guzmán* / **ORC ID:** 0000-0002-3886-4774, **Researcher ID Thomson:** 4006862, **CVU CONACYT ID:** 1094851

ID 1st Co-author: *Liliana Elena, Olguín-Gil* / **ORC ID:** 0000-0003-4649-1434, **Researcher ID Thomson:** 4006628, **CVU CONACYT ID:** 410583

ID 2nd Co-author: *Eduardo, Vázquez-Zayas* / **ORC ID:** 0000-0002-6534-5582, **Researcher ID Thomson:** 4028090, **CVU CONACYT ID:** 1094961

ID 3rd Co-author: *Brawhim Jesseth, Nicanor-Pimentel* / **ORC ID:** 0000-0002-5512-9149

DOI: 10.35429/JTI.2020.21.7.26.31

Received July 25, 2020; Accepted December 30, 2020

Abstract

This research allows to have an overview of the different technologies that can be used to benefit people with visual disabilities. In the association "Sentir con los ojos del corazón" located in Tehuacán, Puebla, México, people with visual disabilities are served who do not have the technological tools available to understand their environment, such as restaurant menus, signs on doors, reading a book and any setting that contains a text, making life difficult in a world where most texts are oriented towards visual people. There are few applications for people with visual disabilities that allow them to improve their lives in the different areas in which they operate. Therefore, it is proposed to design a mobile application that interacts with a virtual assistant to translate the images into text to speech through optical character recognition (OCR), allowing them to function in different educational, work, social environments, among others. This project allows the Inclusion of people with visual disabilities to improve the quality of life using applications for mobile devices and to be self-sufficient in their daily life, later managing to translate in different languages, with different intensities and tone of voice, using different platforms.

Resumen

Esta investigación permite tener un panorama de las diferentes tecnologías que se pueden utilizar para beneficiar a las personas con discapacidad visual. En la asociación "Sentir con los ojos del corazón" ubicada en Tehuacán, Puebla, México, se atienden a personas con discapacidad visual que no tienen al alcance las herramientas tecnológicas para comprender su entorno, como menús de restaurantes, letreros en puertas, leer un libro y cualquier escenario que contenga un texto, haciendo la vida difícil en un mundo donde la mayoría de los textos están orientados a personas visuales. Existen pocas aplicaciones para personas con discapacidad visual que permitan mejorar su vida en los diferentes ámbitos en que se desenvuelven. Por lo cual se propone diseñar una aplicación móvil que interactúe con un asistente virtual para traducir las imágenes en texto a voz mediante un reconocimiento óptico de caracteres (OCR), permitiendo desenvolverse en diferentes entornos educativos, laborales, sociales, entre otros. Este proyecto permite la Inclusión de las personas con discapacidad visual para mejorar la calidad de vida utilizando aplicaciones para dispositivos móviles y ser autosuficientes en su vida diaria, logrando más adelante traducir en diferentes idiomas, con diferentes intensidades y tono de voz, utilizando diferentes plataformas.

Visual impairment, OCR, App, Inclusion**Discapacidad visual, OCR, App, Inclusión**

Citation: VAZQUEZ-GUZMAN, Francisco, OLGUIN-GIL, Liliana Elena, VAZQUEZ-ZAYAS, Eduardo and NICANOR-PIMENTEL, Brawhim Jesseth. Mobile app with reading speech-translated OCR images for visually impaired people. Journal of Technology and Innovation. 2020. 7-21:26-31.

* Correspondence to Author (E-Mail: francisco.vg@tehuacan.tecnm.mx)

† Researcher contributing as first author.

Introduction

López Delgado, A., Olmedo, E., Tadeu, P., & Fernández Batanero, J. M. (2019), p. 196) visual impairment as follows:

The World Health Organization in its descriptive note number 282 divides vision capacity into 4 levels: normal vision, moderate visual impairment, severe visual impairment, and blindness. Grouping the antepenultimate and penultimate into the descriptive of "low vision", which together with the term blindness represent cases of visual impairment.

On this subject, Cardona Mesa (2019) states that: "There are 285 million visually impaired people in the world of whom 39 million are completely blind and 246 million who have low vision".

And one thing we care about as a society is that: "In 2010, people who had some form of disability in Mexico exceeded 5 million, representing 5.1% of the population" (Esparza-Maldonado, A. L., Margain-Fuentes, L. Y., Alvarez-Rodríguez, F. J., & Benítez-Guerrero, 2018, p. 151).

According to the latest INEGI survey (INEGI, 2010), the state of Puebla is in fourth place nationally with a limitation in activity to see 63,575 people. Therefore, the ITTEH-CA-9 Academic Corps, in agreement with the association "Sentir con los ojos del corazón", is intended to benefit visually impaired people with this research.

The academic body has conducted three research related to visual and hearing impairment, which has allowed us to know the problems faced by these types of people.

In recent years the constant change of technology as a result of the different platforms and equipment that evolve rapidly, allow to develop research projects in different areas, thus implementing tools to be used in an important society improving the quality of life of people (Zambrano, D.M; Daza-Alava, Y. D; Pinargote-Zambrano, J.D; Lituma-Ramirez, 2019). There are several technological tools aimed at supporting the visually impaired. (Escandell Bermúdez, María Olga, Fortea Sevilla, María del Sol, Castro Sánchez, 2014, p. 490) state:

"In the case of visual impairment, ICTs made available to these people are quite diverse and extensive, ranging from visual adapters, text-to-sound converters, to the use of Braille-specific printers."

Today there are many conventional smartphones and tablets (Apple and Google) that allow low-vision users to adjust color, Contrast, size, brightness, visibility, readability, and blind users can turn on voice output (VoiceOver on Apple and Google Talk-Back on Android) that reads information aloud, from emails, messages to eBooks, plus voice-controlled digital assistants (SIRI and Alexa) are used to read text, open apps, search, online, send messages, and initiate a call. Also for writing activities, the application incorporates a dictation function that converts voice into text (Natalina Martiniello, Werner Eisenbarth, Christine Lehane, 2019, p. 2).

Due to the high use of mobile devices with high-resolution cameras and high processing capacity, the idea of applying OCR technology to mobile photos arises.

However, existing applications are geared towards visuals without hearing by the interface they present. When regular users require transcribing text on their smartphone that comes from contact cards, document identification, or lottery code, they will use one of the available text entry methods. Optical Character Recognition (OCR) technology can help reduce text transcription time and have been available on mobile devices for several years, in various commercial applications available, some applications recognize structured documents (e.g. business cards)(Bellino, A., 2019, p. 167982).

The goal of the project is to develop an easy and intuitive application with a simple interface that allows access through a screen reader, which uses the camera of the device to take a photograph of the text and translates to voice audibly.

Methodology

Ilya Bibik (2018) states: "Scrum's approach is to divide the project into smaller logical fragments (mini projects) and execute them in short iterations of ideally one to three weeks.

These iterations are called Sprints." Therefore, it was decided to use the Scrum methodology for this project.

In the release plan it was necessary to identify the user stories, which specify the software requirements and in which the customer describes the features that the software must possess.

User stories were based on two crucial aspects:

- Conversation: Between the client and the developer, to analyze and expand details, which is done verbally or documented if required.
- Confirmation: Performed by means of acceptance tests to verify and confirm the correct monitoring of the customer's requirements.

To track requirements and needs, a user story planning was structured with which each iteration was able to determine the necessary changes in the project.

Once the needs and requirements were contemplated through the user stories, the modules that the system would have were defined. The system modules were proposed according to the priority and order in which they would be developed for implementation, being as follows:

1. System functions

a. Home

- i. Speaker's Welcome
- ii. Tutorial

b. Data capture

- i. Approach to the input text
- ii. Processing using OCR
- iii. Translation and preparation of the output

c. Exit

- i. Text storage
- ii. Playing text in speech

With the evaluation of the first user stories and the definition of the system modules, the first launch plan was made, estimating an average time for the elaboration of each module based on a week of 5 days and a day of approximately 3 hours, as shown in Table 1.

Form	#	User history	Estimated time Weeks
System functions	1	Home – Welcome	1
	2	Home – Tutorial	1
	3	Capture - Digitization	1
	4	Captura – OCR	2
	5	Capture - Translation and preparation	1
	6	Output - Text storage	1
	7	Output - Voice playback	1
Total estimated time			8

Table 1 Launch plan

An iteration plan was made in which the communication of the actors (client and developer) was essential to follow up on the development of the modules and reallocate the risk and priority of the modules.

Depending on the user stories and their assessment, a three-iteration plan was developed to be revalued and if necessary restructured; at the end of each one a working meeting was held with the client to carry out the corresponding operational and acceptance tests.

Form	#	User history	Iterations		
			1	2	3
System functions	1	Home – Welcome	X		
	2	Home – Tutorial	X		
	3	Capture – Digitization		X	
	4	Captura – OCR		X	
	5	Capture - Translation and preparation		X	
	6	Output - Text storage			X
	7	Output - Voice playback			X

Table 2 Iteration Plan

Depending on the proposed planning, iterations were constantly being reviewed through the participation of the product owner, the master scrum and the development team.

Results

It is important to implement the necessary efforts to promote technological development in order to support visually impaired people to be included in society in a natural way.

This article proposes the application of the OCR feature on images in which text is captured using a mobile application to audibly reproduce such information, to help visually impaired people.

For application development and speech, Google's OCR and Text-To-Speech libraries were implemented.

The project on Android is divided into the following points:

- Splash design. The application logo was designed, and the app input animation was programmed where it presents the project name while loading the application.
- Central design of the application. Added the app icon in the main view, as well as a button with an ImageView where camera preparation and text capture will be programmed. This photo will be viewed in the ImageView.
- Integration of Google's OCR library. After you have taken the photo and assigned it in the ImageView, the Google OCR library will be added to the Gradle repository; being a standalone library to Android, is compiled and made use of it and the TextBlock and Frame objects that are from the OCR library are installed. The Frame is instantiated which will take the contents of the bitmap so that with an array of type BlockText can be adapted and obtained the lines of text recognized by the device. Finally, the contents of the array are mapped in the TextView to show the user the detected text on the screen. At the end, the function that receives the obtained text is executed, and with the Tex-To-Speech class, an object of that class is instantiated so that it can initialize the event where it receives the Text, and the language setting to which it will read the detected text is assigned.

For the mobile app to work properly, the following requirements are required:

- Software specifications
 - Mobile device with at least Android 7 operating system
 - Space available for the application of at least 20 MB for installation
- Hardware Specifications:
 - Camera and speakers, on mobile device
 - Voice recognition sensor

The test was performed on a cell phone running Android 10 operating system. When you run it, the Splash screen is displayed, which presents the name of the application and its logo (Figure 1).



Figure 1 Home screen

The data entry is then captured. The support of a person acting as an assistant must be supported to correctly capture the text to be translated by voice. Figure 2 shows the functions of the application. The Scan Text button opens your device's camera to capture the photo of the text you want to interpret audibly.



Figure 2 Data capture screen

Once the photo is captured, the application will interpret the text of the image and convert it to voice to play it automatically (Figure 3).



Figure 3 Text-to-speech

The accessibility of the mobile application for reading text images was satisfactory, compared to other programs, due to its easy use on mobile devices with visually impaired people.

Acknowledgment

A scientific article is a work that is not only the result of the efforts of the members of the Academic Corps of Information Technologies of the Tecnológico Nacional de México / Instituto Tecnológico de Tehuacán, but needs the help of many people, both professionally and personally. With these lines we want to show our thanks to all of them.

To Tecnológico Nacional de Mexico, the central body of our national system, because thanks to its economic support programs we have managed to grow as teachers and discover that we can contribute to the development of knowledge and technology in research and inclusion.

To Master Yeyetzin Sandoval González, director of the Tecnológico Nacional de México / Instituto Tecnológico de Tehuacán, for her unconditional support in all our academic projects.

To Master Lucrecia Guadalupe Valenzuela Segura, who in addition to being our reference in the subject of inclusion and the "machinist" of this train, is an excellent person with us.

To our families, who usually do not understand our projects, our long sessions in front of the computer, or our delays at lunchtime, thank you for your patience.

Conclusions

This article has talked about visually impaired supports. In this same tenor, in correspondence with the needs, professionals are also required to take advantage of technology to develop typhotechnological tools, depending on the type or level of visual impairment.

It was shown that to develop mobile applications and projects in general, it is necessary to employ appropriate methodologies to ensure a reduced delivery time and quality, based on constant reviews with the customer, coordinators and the team.

A portion of the application's execution was shown, detailing that the end result is the audible reproduction of text that might be in a magazine, sign, ad, notice, or real-life object, which a visually impaired person could not locate and understand. A simple yet understandable, voice-based and thoughtful interface was designed for the visually impaired.

Efforts to do this type of aid should not be spared, they are valuable and necessary. The inclusion of visually impaired persons could be achieved gradually if research and technological developments are increased. If this article provides a small grain of sand, then we will be satisfied.

References

A. Cardona, R. V. (2019). Mobility assistance devices in visually impaired people: a bibliographic review. *Polytechnic Journal*, 15(28), 107–116. *Polytechnic Journal*, 15(28), 107–116.

<https://doi.org/10.33571/rpolitec.v15n28a10>

Bellino, A., & Herskovic, V. (2019). CameraKeyboard: A novel interaction technique for text entry through smartphone cameras. *IEEE Access*, 7, 167982–167996. <https://doi.org/10.1109/ACCESS.2019.2954101>

Bibik, I. (2018). *How to Kill the Scrum Monster: Quick Start to Agile Scrum Methodology and the Scrum Master Role*. In Apress Media LLC. <https://doi.org/10.1007/978-1-4842-3691-8>

Escandell Bermúdez, María Olga, Fortea Sevilla, María del Sol, Castro Sánchez, J. J. (2014). DIGITAL GAP IN VISUALLY IMPAIRED PEOPLE. *International Journal of Developmental and Educational Psychology*, 1(1), 489–497. DOI: 10.17060 / ijodaep.2014.n1.v1.396

Esparza-Maldonado, A. L., Margain-Fuentes, L. Y., Alvarez-Rodríguez, F. J., & Benítez-Guerrero, E. I. (2018). Development and evaluation of an interactive system for the visually impaired. *Technological*, 21(41), 149–157.

http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0123-77992018000100010&lng=en&tlng

INEGI, I.N. de E. and G. (2010). No Title. *Health and Social Continuity (Disability)*. <https://www.inegi.org.mx/app/areasgeograficas/?ag=21#tabMCcollapse-Indicadores>

López Delgado, A., Olmedo, E., Tadeu, P., & Fernández Batanero, J. M. (2019). Propuesta de las condiciones de las Aplicaciones móviles, para la construcción de un Entorno de Accesibilidad Personal para usuarios con discapacidad visual en las Smart Cities. *Aula Abierta*, 48(2), 193. <https://doi.org/10.17811/rifie.48.2.2019.193-202>

Natalina Martiniello, Werner Eisenbarth, Christine Lehane, A. J. W. W. (2019). Exploring the use of smartphones and tablets among people with visual impairments: Are mainstream devices replacing the use of traditional visual aids? *Assistive Technology*, 1–13. <https://doi.org/10.1080/10400435.2019.1682084>

Zambrano, D. M., Daza Álava, Y. D., Pinargote Zambrano, J. D., & Lituma Ramirez, E. D. (2019). Prototipo para orientación de personas con discapacidad Visual mediante una aplicación para móvil. *Revista Científica*, 2(35), 247–257. <https://doi.org/10.14483/23448350.14523>.