



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE INFORMÁTICA Y ELECTRÓNICA
CARRERA ELECTRÓNICA Y AUTOMATIZACIÓN

**“DESARROLLO DE UN PROTOTIPO DE DISPOSITIVO PARA
RECONOCIMIENTO DE TEXTO TIPO IMPRENTA Y SU
CONVERSIÓN EN AUDIO ORIENTADO A PERSONAS CON
DISCAPACIDAD VISUAL MEDIANTE TÉCNICAS DE VISIÓN
ARTIFICIAL”**

Trabajo de titulación

Tipo: Dispositivo Tecnológico

Presentado para optar al grado académico de:

INGENIERO EN ELECTRÓNICA Y AUTOMATIZACIÓN

AUTOR:

DANILO ORLANDO LUNA SORIA

Riobamba – Ecuador

2022



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE INFORMÁTICA Y ELECTRÓNICA
CARRERA ELECTRÓNICA Y AUTOMATIZACIÓN

**“DESARROLLO DE UN PROTOTIPO DE DISPOSITIVO PARA
RECONOCIMIENTO DE TEXTO TIPO IMPRENTA Y SU
CONVERSIÓN EN AUDIO ORIENTADO A PERSONAS CON
DISCAPACIDAD VISUAL MEDIANTE TÉCNICAS DE VISIÓN
ARTIFICIAL”**

Trabajo de titulación

Tipo: Dispositivo Tecnológico

Presentado para optar al grado académico de:

INGENIERO EN ELECTRÓNICA Y AUTOMATIZACIÓN

AUTOR: DANILO ORLANDO LUNA SORIA

DIRECTOR: MSc. PABLO EDUARDO LOZADA YANEZ

Riobamba – Ecuador

2022

© 2022, Danilo Orlando Luna Soria

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento, siempre y cuando se reconozca el Derecho de Autor.

Yo, Danilo Orlando Luna Soria, declaro que el presente trabajo de titulación es de mi autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor asumo responsabilidad legal y académica de los contenidos de este trabajo de titulación. El patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 20 de enero de 2022.



Danilo Orlando Luna Soria

060391626-3

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE INFORMÁTICA Y ELECTRÓNICA
CARRERA ELECTRÓNICA Y AUTOMATIZACIÓN

El Tribunal del Trabajo de Integración Curricular certifica que: El Trabajo de Integración Curricular; tipo: Dispositivo Tecnológico, **DESARROLLO DE UN PROTOTIPO DE DISPOSITIVO PARA RECONOCIMIENTO DE TEXTO TIPO IMPRENTA Y SU CONVERSIÓN EN AUDIO ORIENTADO A PERSONAS CON DISCAPACIDAD VISUAL MEDIANTE TÉCNICAS DE VISIÓN ARTIFICIAL**, realizado por el señor: **DANILO ORLANDO LUNA SORIA**, ha sido minuciosamente revisado por los Miembros del Tribunal del Trabajo de Integración Curricular, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal autoriza su presentación.

	FIRMA	FECHA
Ing. Diego Ramiro Ñacato Estrella, MSc. PRESIDENTE DEL TRIBUNAL	 Firmado electrónicamente por: DIEGO RAMIRO NACATO ESTRELLA	2022-01-20
Ing. Pablo Eduardo Lozada Yáñez, MSc. DIRECTOR DEL TRABAJO DE TITULACIÓN	Firmado digitalmente por PABLO EDUARDO LOZADA YANEZ	2022-01-20
Ing. Jorge Luis Hernández Ambato, Ph.D. MIEMBRO DEL TRIBUNAL	 Firmado electrónicamente por: JORGE LUIS HERNANDEZ AMBATO	2022-01-20

DEDICATORIA

Dedico el presente trabajo de titulación a mis padres Bolívar y Cecilia que a lo largo de mi vida me han brindado su apoyo incondicional, por guiarme por el camino del bien y ayudarme con cariño y amor a cumplir mis metas. A mis hermanos que día a día me motivan a seguir adelante y en un futuro apoyarlos de la misma manera en su formación académica.

Danilo

AGRADECIMIENTO

Agradezco a Dios por brindarme salud y permitirme compartir este logro con mis padres y hermanos que son los seres más importantes de mi vida. A mis amigos, de manera especial a Cristhian y Armando quienes me brindaron su apoyo desinteresado durante el último peldaño de la carrera.

A los docentes de la facultad que ayudaron en mi formación integral dentro de la electrónica, de manera especial al Ing. Pablo Lozada MSc. y al Ing. Jorge Hernández PhD. quienes con su entereza y conocimiento contribuyeron en la culminación exitosa del presente trabajo de titulación.

Danilo

TABLA DE CONTENIDO

ÍNDICE DE TABLAS	ix
ÍNDICE DE FIGURAS	xi
INDICE DE GRAFICOS	xiii
ÍNDICE DE ANEXOS	xiv
ÍNDICE DE ABREVIATURAS	xv
RESUMEN	xvi
SUMMARY	xvii
INTRODUCCIÓN	1

CAPÍTULO I

1. MARCO TEÓRICO REFERENCIAL	5
1.1. Discapacidad visual	5
1.2. Personas no videntes	5
1.2.1. Problemas de las personas no videntes en el entorno	6
1.2.1.1. <i>Movilidad en el entorno</i>	6
1.2.1.2. <i>Orientación</i>	6
1.2.1.3. <i>Comunicación</i>	7
1.2.1.4. <i>Reconocimiento de objetos</i>	7
1.2.2. Dispositivos orientados al reconocimiento de texto y sintetizador de voz	8
1.3. Inteligencia Artificial	9
1.3.1. Visión Artificial	9
1.3.1.1. <i>Funcionamiento de un sistema de visión artificial</i>	10
1.3.1.2. <i>Elementos de un sistema de visión artificial</i>	10
1.3.1.3. <i>Reconocimiento Óptico de Caracteres (OCR)</i>	13
1.3.1.4. <i>Métodos de reconocimiento óptico de caracteres</i>	15
1.3.1.5. <i>Sintetizadores de voz</i>	16
1.4. Hardware	17
1.4.1. Tarjetas de desarrollo	17
1.4.1.1. <i>Raspberry Pi</i>	17
1.4.1.2. <i>Arduino</i>	19
1.4.1.3. <i>STM32</i>	20
1.4.2. Sensores	21
1.4.3. Actuadores	22
1.4.4. Transductores	22

1.4.5.	<i>Cámaras para adquisición de imágenes</i>	23
1.4.5.1.	<i>Cámara web</i>	23
1.4.5.2.	<i>Cámara IP</i>	24
1.4.5.3.	<i>Cámara Raspberry Pi</i>	25
1.5.	Software	26
1.5.1.	<i>Sistema Operativo</i>	26
1.5.2.	<i>Lenguaje de Programación</i>	27
1.5.3.	<i>OpenCV</i>	28
1.5.4.	<i>Software para Simulación</i>	28
1.5.4.1.	<i>Proteus</i>	29
1.5.4.2.	<i>SolidWorks</i>	29
1.6.	Proyectos similares enfocados a personas con discapacidad visual	30

CAPÍTULO II

2.	MARCO METODOLÓGICO	31
2.1.	Selección de componentes	32
2.1.1.	<i>Raspberry Pi 3B+</i>	32
2.1.2.	<i>Raspberry Pi camera V2</i>	33
2.1.3.	<i>Altavoces USB</i>	34
2.1.4.	<i>Panel led</i>	35
2.1.5.	<i>Módulo relé</i>	35
2.1.6.	<i>Pantalla LCD</i>	36
2.1.7.	<i>Componentes electrónicos</i>	36
2.1.8.	<i>Arduino nano</i>	37
2.1.9.	<i>Fuentes de alimentación</i>	38
2.2.	Diseño electrónico del sistema	38
2.2.1.	<i>Diagrama esquemático</i>	39
2.2.2.	<i>Diseño PCB</i>	39
2.2.3.	<i>PCB – 3D</i>	40
2.2.4.	<i>Circuito electrónico en PCB</i>	41
2.3.	Diseño de la estructura del prototipo	41
2.4.	Integración de todos los componentes	43
2.5.	Programación del sistema de visión artificial	45
2.5.1.	<i>Instalación del sistema operativo</i>	46
2.5.2.	<i>Instalación de librerías necesarias</i>	47
2.5.3.	<i>Etapas de diseño del algoritmo</i>	47

2.5.3.1.	<i>Adquisición de imágenes</i>	47
2.5.3.2.	<i>Procesamiento de la imagen</i>	48
2.5.3.3.	<i>Proceso de reconocimiento óptico de caracteres</i>	49
2.5.3.4.	<i>Conversión del texto detectado en audio</i>	49
2.6.	Programación de la tarjeta de desarrollo	50
2.7.	Algoritmo final del sistema	52
2.8.	Resultado final del prototipo	52

CAPÍTULO III

3.	PRUEBAS Y RESULTADOS	54
3.1.	Prueba de iluminación para detección de texto	54
3.2.	Pruebas de detección de texto según la distancia	54
3.3.	Pruebas de detección de texto con fuentes y tamaños diferentes	56
3.3.1.	<i>Pruebas de texto con fuente Calibri</i>	56
3.3.2.	<i>Pruebas de texto con fuente Times New Roman</i>	59
3.3.3.	<i>Pruebas de texto con fuente Arial</i>	61
3.4.	Pruebas con diferente cantidad de palabras	63
3.5.	Pruebas con diferente formato de texto	65
3.6.	Análisis de resultados	66

CAPITULO IV

4.	EVALUACIÓN ECONÓMICA	67
4.1.	Relación beneficio/costo	69

CONCLUSIONES	70
---------------------------	----

RECOMENDACIONES	71
------------------------------	----

BIBLIOGRAFIA

ANEXOS

ÍNDICE DE TABLAS

Tabla 1-1: Sintetizadores de voz	17
Tabla 2-1: Versiones de Raspberry Pi	18
Tabla 3-1: Características de las versiones de Arduino	19
Tabla 4-1: Características de las versiones de STM32	21
Tabla 5-1: Clasificación de los sensores	21
Tabla 6-1: Clasificación de los actuadores	22
Tabla 7-1: Clasificación de los transductores	23
Tabla 8-1: Comparativa de cámaras web	24
Tabla 9-1: Comparativa de cámaras IP	24
Tabla 10-1: Versiones de Raspberry Pi camera.....	25
Tabla 11-1: Características de los Sistemas Operativos más utilizados.....	26
Tabla 12-1: Clasificación de lenguajes de programación	27
Tabla 1-2: Características técnicas de Raspberry Pi modelo 3B+	33
Tabla 2-2: Parámetros técnicos de cámara Raspberry Pi V2	34
Tabla 3-2: Parámetros técnicos de los Altavoces	34
Tabla 4-2: Parámetros técnicos del panel led	35
Tabla 5-2: Parámetros técnicos del relé.....	35
Tabla 6-2: Parámetros técnicos de LCD.....	36
Tabla 7-2: Parámetros técnicos de Arduino nano	37
Tabla 8-2: Datos técnicos de fuente de alimentación	38
Tabla 9-2: Datos técnicos de fuente de alimentación para iluminación	38
Tabla 1-3: Prueba de iluminación	54
Tabla 2-3: Pruebas a diferente distancia desde la cámara hacia la hoja.....	55
Tabla 3-3: Pruebas con fuente Calibri – tamaño 9 pts.	56
Tabla 4-3: Pruebas con fuente Calibri – tamaño 11 pts.	57
Tabla 5-3: Pruebas con fuente Calibri – tamaño 12 pts.	58
Tabla 6-3: Prueba con fuente Times New Roman – tamaño 9 pts.	59
Tabla 7-3: Prueba con fuente Times New Roman – tamaño 11 pts.	59
Tabla 8-3: Prueba con fuente Times New Roman – tamaño 12 pts.	60
Tabla 9-3: Prueba con fuente Arial – tamaño 9 pts.	61
Tabla 10-3: Prueba con fuente Arial – tamaño 11 pts.	62
Tabla 11-3: Prueba con fuente Arial – tamaño 12 pts.	62
Tabla 12-3: Pruebas con diferente cantidad de palabras	64
Tabla 13-3: Pruebas con texto en Negrilla	65

Tabla 14-3: Pruebas con texto en Cursiva	65
Tabla 1-4: Costos de materiales y herramientas	67
Tabla 2-4: Costos de equipos	68
Tabla 3-4: Costos de implementación.	68
Tabla 4-4: Costo total de construcción del prototipo.....	68
Tabla 5-4: Relación beneficio/costo.....	69

ÍNDICE DE FIGURAS

Figura 1-1: Dispositivo Procer 2.....	8
Figura 2-1: Técnicas de iluminación.....	10
Figura 3-1: Sensor de cámara.....	11
Figura 4-1: Proceso de digitalización de imagen.....	11
Figura 5-1: Eliminación de ruido en una imagen.....	12
Figura 6-1: Imagen segmentada.....	12
Figura 7-1: Logo Tesseract + Python by Google.....	16
Figura 8-1: Computador de placa reducida Raspberry Pi 3B.....	18
Figura 9-1: Arduino mega 2560.....	19
Figura 10-1: Tarjeta de desarrollo STM32f411re.....	20
Figura 11-1: Cámara web.....	23
Figura 12-1: Leguaje de programación.....	27
Figura 13-1: Logo OpenCV.....	28
Figura 14-1: Logo Proteus.....	29
Figura 15-1: Logo SolidWorks.....	29
Figura 1-2: Proceso de elaboración del prototipo.....	31
Figura 2-2: Componentes de Raspberry Pi 3B+.....	32
Figura 3-2: Cámara Raspberry Pi V2.....	33
Figura 4-2: Componentes de Arduino nano.....	37
Figura 5-2: Diagrama electrónico diseñado en Proteus.....	39
Figura 6-2: Diseño de circuito impreso en Proteus.....	40
Figura 7-2: Diseño de circuito impreso en 3D en Proteus.....	40
Figura 8-2: Montaje de componentes en PCB.....	41
Figura 9-2: Diseño de la estructura del prototipo en SolidWorks.....	42
Figura 10-2: Partes de acrílico para la estructura cortadas en laser.....	42
Figura 11-2: Montaje de la PCB junto a Raspberry Pi.....	42
Figura 12-2: Bandeja porta papel – Rieles para colocar bandeja.....	43
Figura 13-2: Base de acrílico para sujetar la cámara.....	43
Figura 14-2: Cámara colocada.....	44
Figura 15-2: Sensor final de carrera.....	44
Figura 16-2: Montaje de elementos del prototipo.....	45
Figura 17-2: Prototipo ensamblado.....	45
Figura 18-2: Pantalla principal del sistema operativo.....	46
Figura 19-2: Captura de la imagen sin filtro.....	47

Figura 20-2: Imagen binarizada.....	48
Figura 21-2: Prototipo terminado.....	52

INDICE DE GRAFICOS

Gráfico 1-2: Diagrama de flujo del procesamiento de la imagen	48
Gráfico 2-2: Diagrama de flujo del reconocimiento de caracteres.....	49
Gráfico 3-2: Diagrama para la conversión de texto en audio.....	50
Gráfico 4-2: Diagrama de flujo de la tarjeta de desarrollo	51
Gráfico 5-2: Diagrama de flujo final del sistema	53
Gráfico 1-3: Eficiencia del dispositivo a diferente distancia	56
Gráfico 2-3: Eficiencia del prototipo con fuente Calibri y diferente tamaño.	58
Gráfico 3-3: Eficiencia con fuente Times New Roman y diferente tamaño.	61
Gráfico 4-3: Eficiencia del dispositivo con fuente Arial con diferente tamaño.	63
Gráfico 5-3: Tiempo de procesamiento de acuerdo a la cantidad de palabras.	64

ÍNDICE DE ANEXOS

Anexo A: Prototipo finalizado y operando.

Anexo B: Persona no vidente iniciando el sistema.

Anexo C: Persona no vidente insertando la bandeja con documento de texto.

ÍNDICE DE ABREVIATURAS

A	Amperios
CA	Corriente Alterna
CAD	Diseño Asistido por Computadora
KB	Kilo Byte
LCD	Pantalla de cristal líquido
mA	Mili amperios
mm	Milímetros
MHz	Megahercio
OCR	Reconocimiento óptico de caracteres
PCB	Placa de Circuito Impreso
OMS	Organización Mundial de la Salud
OpenCV	Open-Source Computer Vision
RGB	Rojo, Verde, Azul
SO	Sistema Operativo
TTL	Transistor Transistor Logic
TTS	Text To Speech - (Texto a Voz)
USB	Universal Serial Bus
VAC	Voltaje de Corriente Alterna
VDC	Voltaje de Corriente Directa
W	Vatios
2D	Dos dimensiones
3D	Tres dimensiones

RESUMEN

En el presente trabajo de titulación se desarrolló un prototipo para el reconocimiento de texto tipo imprenta con sintetizador de voz orientado a personas con discapacidad visual empleando técnicas de visión artificial. El prototipo fue construido utilizando una computadora de placa reducida denominada Raspberry Pi 3B+, que es de fácil acceso, con un software libre de Linux llamado Raspbian. Se instalará en Raspbian el software Python que permitirá realizar el programa que consiste en reconocer caracteres partiendo de la captura de una imagen de texto a través del algoritmo Tesseract de OCR. El algoritmo extraerá de la imagen el texto para después por medio de GTTS (Google Text to Speech) convertirlo en audio en forma de voz para que el usuario pueda acceder a la información. En cuanto al hardware, se diseñó una estructura capaz de contener los elementos electrónicos que forman parte del dispositivo como: Raspberry pi 3B+, cámara Raspberry Pi, tarjeta de desarrollo Arduino nano, altavoces, sensor final de carrera, switch, pulsador, LCD y un panel led que se encargara de la iluminación del sistema. Se realizó pruebas de funcionamiento con los tres tipos de fuentes más utilizados, Calibri, Times New Roman, y Arial, conjuntamente se determinó la iluminación correcta y el rango de distancia en que debe estar la cámara con respecto a la hoja de texto para obtener una buena eficiencia. Es así como con una iluminación superior a los 180 lúmenes, a una distancia de 25 ± 5 cm de la cámara, Calibri con un tamaño de 12 pts., en formato normal con un texto de 150 ± 50 palabras, tiene una eficiencia superior al 95%. Se recomienda realizar un aprendizaje offline automático denominado entrenador, haciendo que el algoritmo pueda aprender por sí solo.

Palabras claves: <VISIÓN ARTIFICIAL> <RASPEBERRY PI> <DISCAPACIDAD VISUAL> <ARDUINO> <PYTHON (SOFTWARE)> <ALGORITMO DE PROGRAMACIÓN>.



Firmado electrónicamente por:
HOLGER GERMAN RAMOS UVIDIA

1664-DBRA-UPT-2021

2021-08-27

SUMMARY

The present graduate research project was to develop a prototype for recognizing type text printing press with a voice synthesizer aimed at people with visual disabilities using artificial vision. The prototype was built using a single-board computer called Raspberry Pi 3B+, easily accessible, with a free Linux software called Raspbian. The Python software will be installed on Raspbian to develop the program that consists of recognizing characters starting from the capture of a text image through the algorithm OCR Tesseract. The algorithm will extract the text from the image later through GTTS (Google Text to Speech) convert it to audio in a speech to access the information. In respect of hardware, a structure capable of containing the electronic elements was designed that are part of the device such as Raspberry pi 3B+, Raspberry Pi camera, Arduino nano development board, speakers, limit switch sensor, switch, push-button, LCD, and a led panel to be in charge of lighting the system. Functional tests were carried out with all three types of fonts most used, Calibri, Times New Roman, and Arial, together with the correct lighting and the range of distance the camera should be from the sheet of text for good efficiency. Thus, with illumination greater than 180 lumens, at a distance of 25 ± 5 cm from the camera, Calibri font with a size of 12 pts., in normal format with a text of 150 ± 50 words, has an efficiency of more than 95%. It is recommended to carry out offline learning called trainer, making the algorithm can learn by itself.

Keywords: <ARTIFICIAL VISION> <RASPEBERRY PI> <VISUAL DISABILITY>
<ARDUINO> <PYTHON (SOFTWARE)> <PROGRAMMING ALGORITHM>.



Firmado electrónicamente por:

LENIN

IVAN LARA

INTRODUCCIÓN

ANTECEDENTES

Las personas con algún tipo de discapacidad, han tenido que esforzarse para acceder a la información a través de las nuevas tecnologías, aunque también es cierto que sin estas discapacidades no hubieran desarrollado otros sentidos como el oído, el tacto, incluso el ingenio necesario para superar las nuevas barreras que estas tecnologías imponen hoy en día (Zambrano Sonroza, y otros, 2017).

Según la (ORGANIZACION MUNDIAL DE LA SALUD, 2018), cada cinco segundos se produce un caso de ceguera entre personas adultas y cada minuto en la población infantil. El 90% de estas personas viven en las zonas más pobres del planeta o en vías de desarrollo. Si no se llevan a cabo las intervenciones oportunas, y de mantenerse las actuales estimaciones sobre el crecimiento de la población mundial y de la expectativa de vida, las cifras del número de personas con discapacidad visual total que se estima entre 40 y 45 millones a nivel mundial en 2020 podrían incrementarse considerablemente.

En Ecuador el total de personas con discapacidad visual registradas en el registro nacional de discapacidad realizado por él (CONADIS, 2020), asciende a 55843 presentando grados de discapacidad entre el 30% y 100%, en el aspecto local existen 1746 casos registrados en la provincia de Chimborazo.

Actualmente existen aplicaciones móviles para personas no videntes tanto para plataformas Android, así como para IOS, que ayudan a que estas personas puedan desenvolverse, lamentablemente son de paga siendo inaccesibles a las personas de bajos recursos. La más conocida por personas no videntes se llama “Voice Dream Scanner”, que tiene un precio de 7 dólares americanos, (VOICE DREAM, 2020), es una empresa que se dedica a diseñar aplicaciones enfocados a un entorno social. Las personas con problemas visuales ya no podrán decir que abandonan la lectura por haber perdido la visión, así como tampoco si se agudiza su enfermedad visual hasta el punto de no poder leer las letras de un libro convencional. La lectura por voz sintética se puede hacer con voces de alta calidad con acceso a un total de 54 lenguas diferentes; entre ellas, catalán, castellano, inglés o francés. Además, se puede elegir entre diferentes voces femeninas o masculinas, según el gusto del lector.

FORMULACIÓN DEL PROBLEMA

¿Cómo realizar un dispositivo capaz de reconocer texto tipo imprenta aplicando técnicas de visión artificial y cuya información sea transferida a un usuario no vidente mediante audio?

SISTEMATIZACIÓN DEL PROBLEMA

¿Cuáles son el estado del arte relacionado a los prototipos de dispositivos enfocados al reconocimiento de texto para personas con discapacidad visual mediante visión artificial y las técnicas empleadas para tal efecto?

¿Cuáles son los requerimientos que deben ser considerados para el desarrollo del prototipo de dispositivo reconocimiento de texto para personas con discapacidad visual mediante visión artificial y la transferencia del contenido mediante audio?

¿Cómo se puede desarrollar el prototipo de dispositivo que permita reconocer texto tipo imprenta mediante técnicas de visión artificial y convertirlo a voz para usuarios no videntes considerando los componentes de hardware y software disponible?

¿Qué tipo de pruebas son necesarias para la validación del prototipo desarrollado considerando el contexto de operación, sujetos de pruebas y ambientes controlados de pruebas, así como evaluación de los resultados?

JUSTIFICACIÓN TEÓRICA

El desarrollo de este proyecto se enmarca en la realidad que viven las personas con discapacidad visual, y más al nivel de nuestro país, donde los recursos económicos no permiten que estas personas puedan acceder con facilidad a sistemas que le permitan desenvolverse adecuadamente en un ambiente laboral, personal, entre otros.

En la actualidad existen muchas apps (aplicaciones móviles) tanto para sistemas Android como para IOS, que ayudan a las personas con discapacidades a desenvolverse en la sociedad. Las más conocidas son: Linguoo que es una app que lee a viva voz páginas web. TTS (Text To Speech) una aplicación para sistemas Android que lee artículos, dispone de varios idiomas y de diferentes voces, Voice Dream Scanner que escanea un texto a través de la cámara del celular y lo transforma en voz para que la persona no vidente pueda escucharlo. La mayoría de las aplicaciones son de paga, frenando la adquisición de muchas aplicaciones móviles a personas no videntes de bajos recursos (INFOTECARIOS, 2017).

Para las computadoras personales existen también programas que transforman el texto en voz que también reproducen las diferentes teclas o comandos del pc según interactúe la persona. JAWS es un programa de paga que se ha convertido en un asistente personal para las personas no videntes, JAWS (Job Acces With Speech) es un software lector de pantalla para ciegos o personas

con visión reducida creado por Freedom Scientific de San Petersburg en Florida – Estados Unidos (DOCE, 2015).

Todas las apps o programas mencionados anteriormente utilizan smartphones o computadoras, pero no existe en el mercado un dispositivo fijo de oficina que ayude a una persona no vidente a convertir el texto de una hoja en voz, por lo que necesitan de una persona auxiliar que lo lea. En Ecuador (CHINCHERO IZA, 2019) y (PILATASIG ESCOBAR, 2017), crearon como proyecto final un prototipo que realice la conversión de texto a voz, por medio de visión artificial. Aunque los dos proyectos son totalmente diferentes la idea principal era ayudar a las personas con discapacidad visual.

El impacto social es muy importante en el desarrollo y validación de esta tesis, puesto que es una gran ayuda para aquellas personas que se vean afectadas por estas dificultades y al mismo tiempo se podría pensar en que se abren las puertas para implementar más proyectos enfocados al área social que son muy necesarios para el bienestar y mejoramiento de la calidad de vida del ser humano.

Por tal motivo se plantea el diseño y construcción de un prototipo de reconocimiento de texto basado en técnicas de visión artificial con el fin de proporcionar una herramienta útil sobre todo en el ámbito laboral a una persona con discapacidad visual con un menor costo para que pueda adquirido.

JUSTIFICACIÓN APLICATIVA

La presente propuesta de trabajo de titulación tiene como finalidad el desarrollo de un prototipo que permita el reconocimiento de texto enfocado a personas con discapacidad brindándoles así una herramienta tecnológica para facilitar su desenvolvimiento en el trabajo específicamente. Además, para que el usuario no vidente interactúe con el prototipo se emitirá señales de voz indicando las acciones que está realizando ese instante.

El prototipo contara con una cámara la cual obtendrá la información de las hojas de texto colocadas en el prototipo y este a su vez enviara las imágenes a un microprocesador u otro dispositivo el cual realizara el procesamiento de dichas imágenes mediante técnica o técnicas de visión artificial; como resultado proporcionara el texto, el cual será emitido al usuario mediante voz ocupando el mismo microprocesador como se muestra en el siguiente esquema.

Dentro de las técnicas de reconocimiento de texto se pondrán a prueba los más importantes y los más usados dentro del campo de visión artificial. Para que el prototipo sea funcional se realizara un análisis en base a los requerimientos de cada técnica de visión artificial basados en el hardware y software adecuados.

Para la verificación de la eficiencia del prototipo se realizará las pruebas pertinentes en personas con discapacidad visual por lo tanto este proyecto tendrá un gran impacto social en beneficio a

estas personas, mejorando su estilo de vida y brindándoles más seguridad al realizar sus actividades en especial en el ámbito laboral.

OBJETIVOS

OBJETIVO GENERAL

Desarrollar un prototipo de dispositivo para reconocimiento de texto tipo imprenta y su conversión en audio orientado a personas con discapacidad visual mediante técnicas de visión artificial.

OBJETIVOS ESPECÍFICOS

- Investigar el estado del arte relacionado a los prototipos de dispositivos enfocados al reconocimiento de texto para personas con discapacidad visual mediante visión artificial y las técnicas empleadas para tal efecto.
- Determinar los requerimientos que deben ser considerados para el desarrollo del prototipo de dispositivo reconocimiento de texto para personas con discapacidad visual mediante visión artificial y la transferencia del contenido mediante audio.
- Implementar el prototipo de dispositivo que permita realizar el reconocimiento de texto tipo imprenta convirtiendo la información a voz a partir de un diseño que considere el hardware y software disponible.
- Validar el funcionamiento del prototipo desarrollado considerando el contexto de operación, sujetos de pruebas y ambientes controlados de pruebas, así como evaluación de los resultados obtenidos.

CAPÍTULO I

1. MARCO TEÓRICO REFERENCIAL

1.1. Discapacidad visual

En el mundo existen personas con diferentes tipos de discapacidades ya sean físicas, sensoriales, intelectuales, psíquicas o viscerales. Según (Organización Nacional de Ciegos Españoles, 2011) actualmente las discapacidades han adquirido un rol importante con un impacto positivo, donde la discapacidad no es un símbolo de enfermedad, teniendo beneficios sociales como tecnológicos. En torno a la discapacidad visual se han mantenido durante siglos mitos y creencias que han estigmatizado a las personas que la padecían. Algunos de ellos han perdurado hasta nuestros días y sigue siendo una condición muy temida por la sociedad. Desde la prehistoria, las personas ciegas han estado discriminadas y vivían abocadas a la mendicidad (Organización Nacional de Ciegos Españoles, 2011).

No obstante, las personas ciegas habían discurrido desde muy antes sistemas propios para realizar algunas actividades. Para comunicarse, utilizaron procedimientos aislados y sofisticados, pero difíciles de propagar. Las posibilidades de desplazarse eran también muy restringidas, pero hay constancia de que utilizaron cuerdas, animales, palos y barras para realizar pequeños itinerarios, recurriendo a los lazarillos para desplazamientos más largos (Organización Nacional de Ciegos Españoles, 2011).

La relación entre discapacidad visual y dependencia ha sido una constante en todas las sociedades. Esta concepción y la baja prevalencia son las causas de que, hasta épocas muy recientes, no se hayan potenciado las posibilidades de las personas que la padecen. La situación actual es sustancialmente distinta, pero queda mucho por hacer (Organización Nacional de Ciegos Españoles, 2011).

1.2. Personas no videntes

Son personas que sufren de falta de visión o ceguera, considerada una discapacidad física que se debe a la pérdida total o parcial de la visión. Según la (ORGANIZACION MUNDIAL DE LA SALUD, 2018), la discapacidad visual tiene cuatro grados, leve, moderada, grave y ceguera. Hay casos donde la carencia de visión se da desde el nacimiento y otros debido a que por alguna enfermedad o accidente pierden la vista. A nivel mundial, se calcula que aproximadamente 1300 millones de personas viven con alguna forma de deficiencia de la visión de lejos o de cerca (Organización Mundial de la Salud, 2018).

Con respecto a la visión de lejos, 1885 millones de personas en el mundo presentan algún grado de pérdida de visión, de los cuales 217 millones sufren de problemas de visión grave y los 36

millones restantes son invidentes. Por otro lado, 826 millones de personas padecen una deficiencia de la visión de cerca. El crecimiento y envejecimiento de la población aumentarán el riesgo de que más personas se vean afectadas por una deficiencia visual (Organización Mundial de la Salud, 2018).



Figura 1-1. Personas no videntes

Fuente: (Organización Mundial de la Salud, 2018)

1.2.1. Problemas de las personas no videntes en el entorno

Las personas no videntes al no poder contar con el sentido de la vista se le presentan varios inconvenientes al desenvolverse en el entorno.

1.2.1.1. Movilidad en el entorno

La movilidad se entiende al conjunto de desplazamientos y acciones que ejecutan los seres vivos en su entorno, las personas utilizan sus sentidos y sus partes motrices para desplazarse de un lugar a otro de forma segura y eficiente, al desplazarnos ocupamos todos nuestros sentidos, y uno de los principales es la vista. Las personas con discapacidad visual sufren muchas dificultades, limitando la hora de hacer sus actividades cotidianas. Lamentablemente el entorno y la sociedad está enfocada a personas que no poseen discapacidad, llevando a los no videntes a adaptarse obligadamente al entorno ya sea con el apoyo de dispositivos o personas (Terven Salinas, y otros, 2013, p. 32).

1.2.1.2. Orientación

La orientación es la capacidad de reconocer la posición de las personas con respecto al entorno, es decir para que puedan dirigirse de un lugar a otro necesitan la orientación, pues esto lo consiguen con la ayuda de señales, gráficos o textos que se encuentran en diarios, revistas, anuncios, artículos, que les beneficia poder llegar al objetivo.

(Terven Salinas, y otros, 2013, p. 35) afirma que las personas con discapacidad visual no pueden orientarse por lo que tienen que depender de una persona que le brinde información constantemente para poder llegar al objetivo correcto.

1.2.1.3. Comunicación

Tener información del entorno, de la posición de las cosas, del texto de un papel o de algún hecho que se esté suscitando es fundamental para conocer y darse cuenta del mundo. Esta información permite analizar situaciones y encontrar conclusiones para poder actuar. La mayoría de diarios, revistas, libros, entre otros, se encuentran disponibles día a día en puntos de venta donde normalmente cualquier persona puede tener acceso, pero no existen documentos orientados a personas con discapacidad visual provocando un problema social.

Este problema se debe a que el braille o el sistema de escritura para personas no videntes genera costos de producción demasiados altos, ya que ocupa un papel más grueso a diferencia del papel normal de periódico. Actualmente el braille ha decaído ya que la tecnología avanzada por medio de las computadoras o los dispositivos móviles, creando efectos positivos y negativos en el ambiente de las personas invidentes. En un artículo, mencionado por (Terven Salinas, y otros, 2013, p. 36), indica que el 10 % de niños con discapacidad visual aprenden y utilizan el sistema braille, el restante sufren de deficiencia de lectura braille.

Las personas no videntes que no poseen de ordenadores o dispositivos móviles, obligadamente les toca apoyarse en otra persona que posea todas las facultades y que se dedique a leerles textualmente el texto (Terven Salinas, y otros, 2013, p. 36).

Así mismo, en la comunicación, al tener una conversación, normalmente las personas utilizan expresiones con sus extremidades donde realizan mímicas o movimientos dando énfasis al mensaje que transmiten, para una persona no vidente esto se convierte en una limitación a la hora de comunicarse (Terven Salinas, y otros, 2013, p. 36).

1.2.1.4. Reconocimiento de objetos

Las personas con discapacidad visual tienen problemas al reconocer objetos, no obstante gracias a su discapacidad el resto de sus sentidos se ha visto potenciados con una mejor capacidad de respuesta ante el entorno al que se desenvuelven diariamente. Todos los objetos poseen características diferentes, ya sea por su forma, su textura, su olor y sabor, sin embargo, existen objetos que difieren en una o dos características convirtiéndose en una gran dificultad para la persona no vidente a la hora de reconocer el objeto.

1.2.2. Dispositivos orientados al reconocimiento de texto y sintetizador de voz

Muchas universidades a nivel mundial, se han dedicado a desarrollar dispositivos tecnológicos para ayudar a las personas no videntes, proyectos innovadores de orientación, de movilidad, para reconocer objetos, entre otros.

Aproximadamente en 1992 aparece lo que ya podemos considerar un equipo compacto (Personal Reading Edge), de la firma Xerox, que en un solo bloque reúne ordenador, escáner fijo y síntesis de voz. El tamaño de este equipo tiene el tamaño de una impresora normal y es mucho más versátil debido, entre otras características, a que el programa que controla todo el equipo va alojado en una tarjeta PCMCIA (tarjeta de memoria para computadores personales). Esto permite almacenar diversas versiones del software sin necesidad de tener que cambiar la máquina o llevarla a reprogramar. (The Henry Ford, 2019).

Por estas mismas fechas se fabrica el primer equipo portátil basado casi exclusivamente en tecnologías estándares, se trata del Galileo, de la firma australiana Robotron, que incorpora una placa de PC (computadora personal) con un sistema operativo Windows 3.11 y un programa de OCR (reconocimiento óptico de caracteres), además de la interfaz y la síntesis de voz de Elan Informatique (Bermúdez Cabra, y otros, 2012, p.17).

El CIDAT (Centro de Investigación, Desarrollo y Aplicación Tiflotécnica) de la ONCE una fundación española, contribuye en el año 2000, con el proyecto Lectoscan, al mundo de las máquinas automáticas de lectura. Es este un equipo que también consta de elementos estándares en su interior, aunque la carcasa no parezca la de un PC. Podríamos decir que es el primer equipo mixto entre las máquinas y los programas de ordenador, dado que las entradas y salidas físicas del PC del interior de Lectoscan están al alcance del usuario, de forma que este puede conectar un monitor, un teclado y un ratón, y usar Lectoscan como un PC con un microprocesador PENTIUM III a 1 GHz (Bermúdez Cabra, y otros, 2012, p. 17).



Figura 1-1. Dispositivo Procer 2

Fuente: (Bermúdez Cabra, y otros, 2012, p. 17)

Cabe mencionar que al pasar el tiempo la tecnología ha evolucionado a pasos gigantescos, donde ahora se puede encontrar teléfonos celulares de alta, media y baja gama, capaces de instalar aplicaciones que utilizan componentes internos del dispositivo como la cámara, flash, sonido entre otros. Los desarrolladores de aplicaciones han creado herramientas enfocándose en personas con discapacidad más aun a personas no videntes utilizando asistentes de voz. La (Asociación D.O.C.E, 2015) comenta que la aplicación móvil más usada es KNFB Reader ya que escanea palabras impresas en papel así también como en otras superficies comunes en el medio, para después reproducir el sonido de la palabra al usuario, lo negativo es que la aplicación es pagada y muy costosa.

1.3. Inteligencia Artificial

En el libro de (Rouhiainen, 2018, p. 5) se dice que la inteligencia artificial es la capacidad de entrenar a una máquina, dispositivo, programas etc. de modo que pueda tomar decisiones basadas en experiencias o entrenamiento, similar a un ser humano. Sin embargo, a diferencia de las personas, los dispositivos basados en inteligencia artificial no necesitan descansar y pueden analizar más información. Asimismo, la proporción de errores es significativamente menor en las máquinas que realizan las mismas tareas que sus contrapartes humanas.

1.3.1. Visión Artificial

Una de las aplicaciones de la inteligencia artificial es la visión, haciendo referencia a la vista humana, uno de los sentidos más importantes que poseen las personas y muchos organismos vivos. Los ojos del ser humano captan información del entorno para posteriormente entender e interactuar entre sí (Nogué, y otros, 2012, p. 10).

Diariamente las personas utilizan representaciones gráficas para transmitir un mensaje o información, así, por ejemplo, los colores de un semáforo que controlan el flujo vehicular, o una señal informativa que se compone de un dibujo o un texto que regula las normas de comportamiento humano. Estas señales son interpretadas por el cerebro a través del ojo humano (Nogué, y otros, 2012, p. 10).

La visión artificial imita la vista humana de forma digital o computacional, a través de una cámara, por la que procesa imágenes o videos y realiza acciones de acuerdo a las necesidades del usuario (Nogué, y otros, 2012, p. 10).

Estas imágenes captadas por la cámara y son analizadas a profundidad por un microcontrolador, microprocesador, dispositivo móvil o computadora que a través de aplicaciones descifran la información que contiene la imagen, gracias a procesos digitales como el reconocimiento de objetos, la interpretación de colores, conteo, entre otros. (Solano Aguilar, 2018, p. 25).

1.3.1.1. Funcionamiento de un sistema de visión artificial

El sistema de visión artificial cumple una serie de procesos, inicialmente captando imágenes por medio de una cámara para luego ser procesada con técnicas avanzadas para al final obtener dicha imagen con características diferentes y así cumplir el requerimiento del usuario (Nogué, y otros, 2012, p. 12).

1.3.1.2. Elementos de un sistema de visión artificial

El sistema de visión artificial está compuesto de varios elementos como es la iluminación, el sensor de imagen, tarjeta de adquisición de imagen, algoritmos de imágenes, computadora o módulo de proceso y el sistema de respuesta en tiempo real.

Iluminación: Son las condiciones de luz en que se encuentra el objetivo convirtiéndose en un aspecto importante, ya que el entorno debe tener condiciones de luz uniforme para que la información no se disperse (Chincheró Iza, 2019). Existen varios tipos de iluminación, como: led, fibra óptica, fluorescente o laser, así como las diferentes técnicas como: iluminación frontal o directa, de ángulo o lateral, de ángulo bajo o por campo oscuro como se muestra en la figura 2-1 (Solano Aguilar, 2018, p. 27).

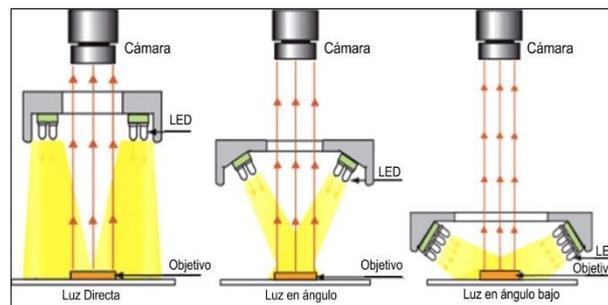


Figura 2-1. Técnicas de iluminación

Fuente: (Solano Aguilar, 2018)

Sensor de imagen: Elementos perceptibles a la luz que dependiendo de la intensidad de la luz generan una señal eléctrica la cual se la puede digitalizar, tomando la información, así como las características del objeto (Chincheró Iza, 2019, p. 14). La cámara se le considera un sensor de imagen como se muestra en la figura 3-1, ya que es un traductor capaz de transformar la imagen y digitalizarla para que sea procesada. Existen varios tipos de cámaras dependiendo de características que requiera el usuario, lineales, matriciales, infrarrojas, térmicas, 3D, entre otros.

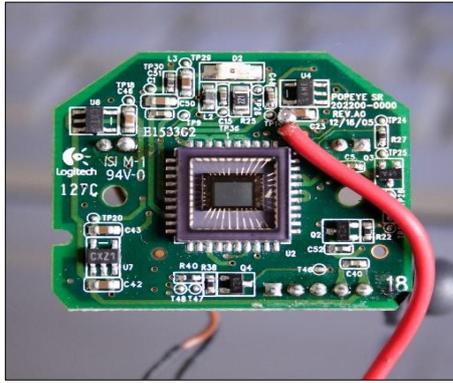


Figura 3-1. Sensor de cámara.

Fuente: (Chinchero Iza, 2019)

Tarjeta de adquisición de la imagen: Se encarga de guardar la imagen en una tarjeta de memoria, convirtiéndola en una imagen digital para después procesarla y analizarla, (Chinchero Iza, 2019, p. 16). En la figura 4-1 se observa la conversión de una imagen a una digital, por el cual pasa por un filtro Red, Green, Blue (RGB), y por un sensor CCD (dispositivo de carga acoplada) o CMOS (semiconductor complementario de óxido metálico) para tener como resultado una imagen digital (Solano Aguilar, 2018, p. 28).

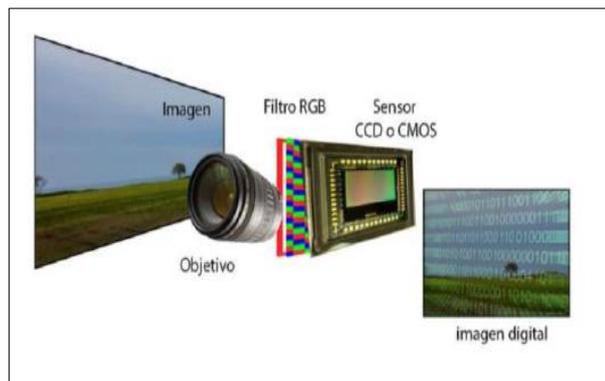


Figura 4-1. Proceso de digitalización de imagen.

Fuente: (Solano Aguilar, 2018)

Algoritmo de análisis de imágenes: Para extraer la información que es el objetivo principal de la imagen, es necesario transformarle o modificarle usando técnicas de procesamiento, así como filtros que eliminen el ruido de la imagen (Chinchero Iza, 2019, p. 18).

Los algoritmos de análisis de imágenes sacan la información más importante de la escena dependiendo de los requerimientos del usuario, por lo que sigue varias etapas. La primera etapa es la de pre-procesado o de filtro que mejora la imagen o elimina el ruido, realzando las características del objetivo, así también, mejorando el brillo y el contraste. Hay usuarios que aplican un filtro de escala de grises o también denominada binarización que es convertir a la

imagen en blanco y negro para obtener un mejor reconocimiento de objetos (Solano Aguilar, 2018, p. 29).



Figura 5-1. Eliminación de ruido en una imagen

Fuente: (Chinchero Iza, 2019)

La segmentación es el próximo paso, y es aquella que se encarga de seleccionar ciertas áreas de la imagen, especialmente el área en donde se encuentran los objetos importantes a reconocer o analizar. Esta selección llevara toda la información al siguiente proceso para que sea examinada y extraída. En la figura 6-1 se muestra una imagen aplicada el proceso de segmentación (Solano Aguilar, 2018, p. 29).

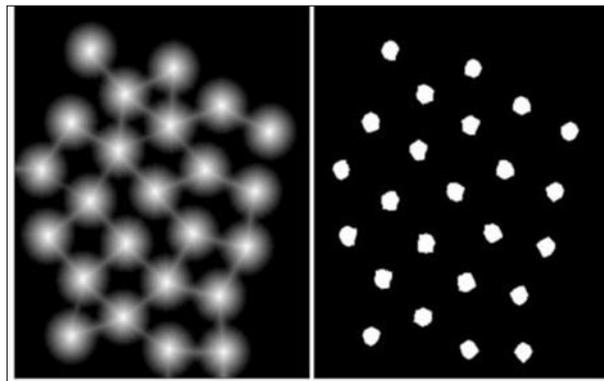


Figura 6-1. Imagen segmentada.

Fuente: (Solano Aguilar, 2018)

La siguiente etapa es la descripción que busca los rasgos del objetivo segmentado o seleccionado, diferenciando todas las características que este posee, para dejar claro el objetivo. La última etapa es el reconocimiento y se encarga de clasificar los elementos encontrados gracias a algoritmos de inteligencia artificial u otros que etiquetarán o agruparán a los elementos para después ordenarlos en grupos esperando la acción para la que fueron seleccionados. (Solano Aguilar, 2018, p. 30).

Monitor en el proceso: Es el sistema que recibe, procesa y analiza la información por medio de algoritmos de inteligencia artificial de las imágenes en tiempo real y reconoce el objetivo de acuerdo los requerimientos del usuario (Chinchero Iza, 2019, p. 18).

Sistema de respuesta en tiempo real: Un sistema en tiempo real es aquel que actúa en el momento, es decir, el tiempo que demora el procesamiento de la imagen desde que se toma la foto hasta el reconocimiento del objeto al instante (Chincheró Iza, 2019, p. 18).

1.3.1.3. Reconocimiento Óptico de Caracteres (OCR)

El reconocimiento óptico de caracteres que en inglés significa (Optical Character Recognition) es una tecnología de reconocimiento de letras o símbolos que comprende un conjunto de técnicas estadísticas, así como, identifica la forma que posee cada carácter para después compararla en una base de datos automáticamente los diferentes caracteres en una base existente de caracteres alfanuméricos (Sánchez Fernández, y otros, 2008, p. 11).

El OCR no reconocen exactamente las letras de un determinado alfabeto, sino que reconoce las formas de los símbolos al compararles con otras., teniendo en cuenta que la precisión que se obtiene en la práctica al intentar distinguir entre un conjunto de símbolos no es del 100%, concluyendo que, cuanto más numeroso es el conjunto de símbolos entre los que se decide, mayor es la probabilidad de que se produzca un fallo en la elección. (Sánchez Fernández, y otros, 2008, p. 11). En todo sistema de reconocimiento óptico de caracteres (OCR) se distinguen al menos estas 4 etapas:

- Adecuación de la imagen (pre-procesado).
- Selección de la zona de interés (segmentación).
- Representación digital de la imagen (extracción de características).
- Distinción del carácter contenido en la imagen (reconocimiento).

Para cada una de las cuatro etapas es posible aplicar multitud de técnicas ya existentes o desarrollar alguna específica en función de las condiciones en las que se presentan los datos de entrada, que en el caso de OCR se puede traducir por las imágenes de entrada (Sánchez Fernández, y otros, 2008, p. 11).

Preprocesamiento: Las técnicas de OCR son útiles para digitalizar textos de algún libro, revista, hoja de texto es decir caracteres impresos o formularios rellenos manualmente (caracteres manuscritos). Tanto en un caso como en el otro el desglose de los caracteres individuales es mucho más sencillo que en el caso de texto manuscrito continuo, para el que es necesario la aplicación de técnicas de pre-procesado y segmentaciones más complejas que en el caso de OCR. (Sánchez Fernández, y otros, 2008, p. 12).

En esta fase de adecuación de la imagen, el objetivo es eliminar cualquier tipo de ruido o imperfección que no pertenezca al carácter, así como normalizar el tamaño del mismo. Además,

para el caso de OCR, la normalización de la imagen también puede implicar un binarizado de la misma, es decir separar la letra del fondo de la imagen (Sánchez Fernández, y otros, 2008, p. 12). Para el binarizado de la imagen, se utilizan diversos algoritmos:

- Etiquetado: para la división de la imagen en regiones de componentes conectadas.
- Erosión o expansión: para la eliminación de pequeños grupos de píxeles.
- Umbralizado de histograma: para eliminar los objetos más brillantes o más oscuros de la imagen.

Segmentación: La fragmentación o segmentación de la imagen constituye una de las mayores dificultades del reconocimiento, y se hace necesaria para poder reconocer cada uno de los caracteres de la imagen binaria. Consiste en descomponer un texto en diferentes entidades lógicas que deben ser lo suficientemente invariables, para ser independientes del escritor, y lo suficientemente significativas para su reconocimiento (Sánchez Fernández, y otros, 2008, p. 12).

Para reconocer caracteres es necesario, en primer lugar, su localización dentro del texto del documento, teniendo en cuenta, en esta operación el orden en el que se disponen en el mismo y los espacios en blanco y finales de línea, para que pueda recomponerse el texto tal y como se encontraba en el documento original (Sánchez Fernández, y otros, 2008, p. 12). Existen tres magnitudes que determinan el orden de los caracteres dentro de un texto:

- Los renglones de los que consta.
- Las palabras de un renglón.
- Las letras de una palabra.

La segmentación del texto manuscrito es un caso más complejo que el tratado en OCR, donde los caracteres, bien se encuentran claramente separados en la imagen original o bien es posible separarlos de manera relativamente fácil, ya que su escritura es regular y presenta características aprovechables para este fin (Sánchez Fernández, y otros, 2008, p. 12).

Extracción de características: Teniendo seleccionados los caracteres o segmentada la imagen, ésta queda normalizada, es decir, estará la información que será “reconocida”. En este proceso se realiza una codificación de la manera más óptima de las características más esenciales del texto partiendo de una imagen del texto binarizada, es decir, a niveles de grises. Partiendo del reconocimiento de formas, la matriz bidimensional es un vector dimensional donde guarda todos los componentes que posee la matriz. La dimensión de los vectores normalmente es grande por lo que genera un gran costo computacional al realizar el procesamiento del mismo, e intentar

reconocer o clasificar caracteres por lo que provoca que el resultado del proceso no sea satisfactorio (Solano Aguilar, 2018, p. 32).

Debido a esto, es que existen gran variedad de técnicas denominadas “técnicas de selección y extracción de características”, mediante las cuales se puede obtener una representación del objeto a reconocer de la manera más correcta y eficiente. La fase más difícil es la extracción de las características en los sistemas de OCR, ya que es muy difícil escoger un conjunto de características que sean óptimas para el reconocimiento (Sánchez Fernández, y otros, 2008, p. 14). Para que una característica se pueda considerar buena debe poseer:

- *Discriminación:* Deben ser características que diferencien suficientemente una clase de otra.
- *Igualdad:* Deben tener igual valor para mismas clases.
- *Independencia:* Las características deben estar no correlacionadas (independientes) unas de otras.
- *Pequeño espacio para características:* El número de características debe ser pequeño para la rapidez y facilidad de clasificación.

Reconocimiento: Una vez se tienen las características más importantes de la imagen a analizar hay que determinar el carácter correspondiente por medio de técnicas de OCR.

1.3.1.4. Métodos de reconocimiento óptico de caracteres

KNN: Para OCR, existe un método muy conveniente, no paramétrico y supervisado, que proporciona resultados muy adecuados para la aplicación que se está tratando, el algoritmo K-NN es un método muy usado ya que es fácil de usar y posee un gran número de propiedades estadísticas capaces de enfrentarse a distintos tipos de problemas de reconocimiento de caracteres. (Sánchez Fernández, y otros, 2008, p. 16).

El algoritmo funciona teniendo un conjunto de caracteres de muestra y un nuevo objeto de clase desconocida por lo que se busca entre el conjunto de prototipos los “k” más parecidos a nuevo objeto. A éste se le asigna la clase más numerosa entre los “k” objetos seleccionados (Sánchez Fernández, y otros, 2008, p. 16).

Tesseract OCR: Tesseract es un método de OCR en línea y de código abierto que puede reconocer archivos de imagen en múltiples formatos para después convertirlos en texto. Fue desarrollado por HP y luego adoptado por Google hasta el día de hoy. Tesseract en el 2006 fue una de las técnicas de OCR más precisas siendo compatible con muchos idiomas del mundo considerándose un software multiplataforma y open source, teniendo un tiempo de ejecución aceptable, pero

presentando problemas si hay un error en la segmentación. El formato único que procesa es un TIFF y se compila y ejecuta en Linux, Windows y Mac OS X (Gutierrez, y otros, 2011, p. 22).



Figura 7-1. Logo Tesseract + Python by Google.

Fuente: (Gutierrez, y otros, 2011)

GOOCR: Desarrollado por Joerg Schulenburg, que se basa en un conjunto de reglas. Es un software portable a diferentes sistemas operativos, el tamaño de las fuentes que soporta esta técnica es entre 20 y 60 pixeles y acepta muchos formatos de imágenes, sin embargo, presenta inconvenientes con letras cursivas, texto escrito a mano y adicionalmente es muy sensible a imágenes que contengan ruido y grandes ángulos de inclinación (Gutierrez, y otros, 2011, p. 22).

OCRAD: Creado por Antonio Díaz, es un método de extracción de características geométricas. Es muy rápido, pero muy sensible al ruido y es difícil de adaptar nuevos símbolos (Gutierrez, y otros, 2011, p. 23).

1.3.1.5. Sintetizadores de voz

La función de un sintetizador de voz es convertir automáticamente un texto a una voz sintetizada, una voz artificial reproducido por un ordenador o un dispositivo. Para la conversión de texto a voz o más conocido como “Text To Speech” primero se analiza el texto de entrada para determinar la estructura de la frase y la composición fonética de cada palabra para luego transformar esta representación lingüística abstracta en VOZ (Guzmán Arreola, 2004, p. 17). A continuación, se detallan los diferentes tipos de sintetizadores de voz más utilizados.

Tabla 1-1: Sintetizadores de voz

	PICO TTS	eSpeak TTS	SVOX Classic TTS	gTTS
Disponibilidad	Limitada	Android	Android	Librería de Python
Precio	Gratuito	Gratuito	2.99 cada voz	Gratuito
Español	Si	Si	Si	Si
Fabricante	Google/SVOX	Eyes-Free Project	SVOX	Google

Fuente: (Hernández, 2011)

Realizado por: Luna, Danilo 2022

1.4. Hardware

Según (Manosalvas Salazar, 2017, p. 15), “el hardware constituye todo dispositivo físico que compone el mismo, que en conjunto con el software realizan tareas específicas”. El hardware posee características de tamaño, consumo energético, funcionalidad, capacidad de procesamiento, entre otros. Un sistema computacional tradicional posee de un hardware que se refiere a la parte palpable del sistema.

1.4.1. Tarjetas de desarrollo

(Concepción, 2019) define a una tarjeta de desarrollo como una placa o circuito electrónico que está compuesto de un microcontrolador que ejecuta una serie de instrucciones de un programa diseñado para lograr un propósito, acompañado de elementos electrónicos que permiten la alimentación de voltaje al microcontrolador, así como su correcto funcionamiento para interactuar con las entradas y salidas del microcontrolador y así poder conectar los diferentes sensores y actuadores. A continuación, se detallan las más utilizadas.

1.4.1.1. Raspberry Pi

Raspberry Pi es un minicomputador basado en arquitectura de un microprocesador ARM comúnmente utilizado en teléfonos inteligentes. Creada por la fundación educacional sin fines de lucro Raspberry, ubicada en Caldecote, Cambridgeshire, Inglaterra. La Raspberry Pi es un “single board computer” (computador de una sola placa) que están basados en SoC (sistema de un solo chip).

El minicomputador Raspberry Pi ha sido diseñado con un objetivo específico, servir de plataforma para aquellos interesados en la programación o el control de periféricos de bajo nivel. Sin embargo, su apariencia cruda y su aparente “poco rendimiento” en comparación con cualquier PC o laptop actual no debe ser motivo de desprecio, teniendo en cuenta que muchos robots exploradores y sondas espaciales son controlados por minicomputadores de mucho menores prestaciones que la Raspberry Pi (Salcedo Tovar, 2015, p. 14).



Figura 8-1. Computador de placa reducida Raspberry Pi 3B

Fuente: (Salcedo Tovar, 2015)

En la figura 9-1 se muestra una Raspberry Pi siendo este un dispositivo accesible en el mercado y tiene un precio asequible, dependiendo de las versiones y características que esta posea. En la siguiente tabla 2-1 se detalla las características de las versiones disponibles, como dato adicional tenemos que la primera versión fue lanzada al mercado el 29 de febrero de 2012.

Tabla 2-1: Versiones de Raspberry Pi

	Raspberry Pi 4B	Raspberry Pi 3B+	Raspberry Pi 3B	Raspberry Pi Zero
SoC	Broadcom BCM2711	Broadcom BCM2837B0	Broadcom BCM2837	Broadcom BCM2835
CPU	Quad Cortex A72 1.5Ghz	Quad Cortex A53 1.4Ghz	Quad Cortex A53 1.2Ghz	ARM11 1Ghz
Arquitectura	ARMv6-A	ARMv6-A	ARMv6-A	ARMv6
GPU	VideoCore IV	VideoCore IV	400MHz VideoCore IV	250MHz VideoCore IV
RAM	1-2 o 4 GB SDRAM	1GB SDRAM	1GB SDRAM	512MB SDRAM
Almacenamiento	micro SD	micro SD	micro SD	micro SD
Ethernet	1000 dual band	10/100/300	10/100	no
Wireless	802.11n/Bluetooth 4.0	802.11n/Bluetooth 4.0	802.11n/Bluetooth 4.0	no

Video Output	2 micro HDMI	HDMI	HDMI	HDMI
Audio Output	HDMI/Auricular	HDMI/Auricular	HDMI/Auricular	HDMI
GPIO	40 pines	40 pines	40 pines	40 pines
Precio	\$ 135,00	\$ 90,00	\$ 80,00	\$ 25,00

Fuente: (Ochoa Duarte, 2015)

Realizado por: Luna, Danilo 2022

1.4.1.2. Arduino

Arduino es la tarjeta de desarrollo más conocida a nivel mundial, ya que cuenta con hardware y software libre, es decir los diagramas de diseño y conexión entre componentes que forman parte de este, así como, el software en donde se puede programar el mismo, pueden ser utilizados y adquiridos por cualquier persona. La mayoría de microcontroladores que ocupan este tipo de tarjetas son AVR's - Atmel, en la mayoría de los casos el ATmega 328 y ATmega 2560. Hay muchos modelos de Arduino, cada uno con diferentes características dependiendo del fin en que se utilice. A continuación, en la tabla 3-1 se presentan las características de las tarjetas Arduino más usadas y en la figura 9-1 un ejemplo de dichas tarjetas (Concepción, 2019).

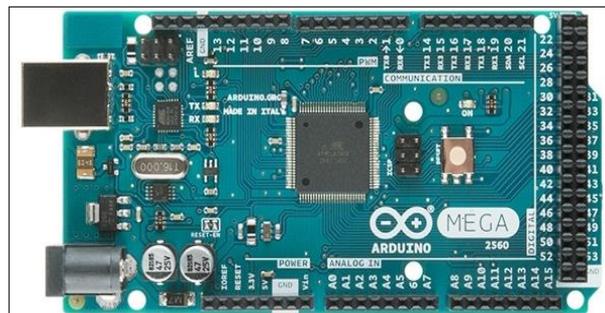


Figura 9-1. Arduino mega 2560

Fuente: (Concepción, 2019)

Tabla 3-1: Características de las versiones de Arduino

	Arduino Mega 2560	Arduino Uno	Arduino Nano	Arduino micro
Microcontrolador	ATmega2560	ATmega328P	ATmega328P	ATmega32U4
Voltaje de funcionamiento	5v	5v	5v	5v
Voltaje de entrada	7-12v	7-12v	7-12v	7-12v
Pines I/O digitales	54	14	14	20
Pines I/O digitales PWM	15	6	6	7

Pines de entrada analógica	16	6	8	12
Memoria flash	256 KB	32 KB	32 KB	32 KB
SRAM	8 KB	2 KB	2 KB	2 KB
EEPROM	4 KB	1 KB	1 KB	1 KB
Velocidad del reloj	16 MHz	16 MHz	16 MHz	16 MHz
Consumo	93 mA	46 mA	19 mA	16 mA
Precio	\$25,00	\$17,00	\$12,00	\$10,00

Fuente: (Concepción, 2019)

Realizado por: Luna, Danilo 2022

La programación de las tarjetas de Arduino se las realiza a través de un programa libre llamado IDE o un interfaz de desarrollo integrado, que permite desarrollar programas para después, cargarlos a la tarjeta Arduino. Los programas desarrollados utilizan lenguaje C, un lenguaje de bajo nivel y de fácil uso ya que cuenta con tutoriales y librerías existentes en línea (Concepción, 2019).

1.4.1.3. STM32

Los microcontroladores de la familia STM32 cuentan con una gran cantidad de versiones, en la figura 11-1 se muestra uno de los modelos más comunes, dicho esto es posible encontrar un modelo que se adapte bien a las características del proyecto a desarrollar. Están basados en procesadores ARM-Córtex-M de 32 bits, permitiendo una multitud de funcionalidades en aplicaciones embebidas de tiempo real combinado con un bajo consumo además de varias características que se detallan en la tabla 4-1 (Arranz Ortega, 2019, p. 28).

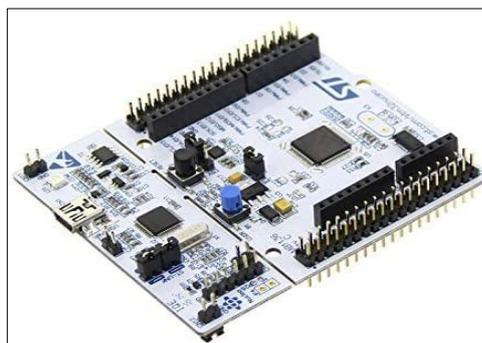


Figura 10-1. Tarjeta de desarrollo STM32f411re

Fuente: (Arranz Ortega, 2019)

Tabla 4-1: Características de las versiones de STM32

	STM32f407	STM32f411re	STM32f103c8t6
Fabricante	STMicroelectronics	STMicroelectronics	STMicroelectronics
CPU	ARM® Cortex®-M4 32-Bits	ARM® Cortex®-M4 32-Bits	ARM® Cortex®-M3 32-Bits
Voltaje	1.8 V – 3.6 V	1.7 V – 3.6 V	2 V – 3.6 V
Velocidad de Reloj	168 Mhz	84 Mhz	72 Mhz
Memoria de Programa	1MB (1MB x 8)	512KB (512K x 8)	64KB (64K x 8)
Memoria Flash	Si	Si	Si
RAM	192K x 8	96K x 8	20K x 8
Temperatura de operación	-40°C ~ 85°C (TA)	-40°C ~ 85°C (TA)	-40°C ~ 85°C (TA)
Número de E/S	140	50	37
Precio	\$75,00	\$50,00	\$11,00

Fuente: (Arranz Ortega, 2019)

Realizado por: Luna, Danilo 2022

1.4.2. Sensores

Un sensor es un dispositivo electrónico, mecánico o electromecánico que genera una señal de voltaje, resistencia o corriente, a través de un fenómeno natural físico como la temperatura, la distancia, la presión, etc., aunque ciertamente no existe una definición clara, se denomina a un sensor como transductor, ya que es capaz de convertir una señal de una forma física en una señal correspondiente pero de otra forma física distinta, por lo tanto, un sensor de acuerdo a los parámetros que será medido, entrega una señal transducible que está en función de la variable de medida, en la tabla 5-1 encontramos una breve clasificación de los sensores (Vinueza Escobar, 2017, p. 17).

Tabla 5-1-: Clasificación de los sensores

Según el tipo de funcionamiento	Según el tipo de señal eléctrica que generan	Según el rango de valores que proporcionan	Según el nivel de integración	Según el tipo de variable física medida
<ul style="list-style-type: none"> • Activos • Pasivos 	<ul style="list-style-type: none"> • Digitales • Analógicos • Temporales 	<ul style="list-style-type: none"> • De contacto (On-Off) • De medida 	<ul style="list-style-type: none"> • Discretos • Integrados • Inteligentes 	<ul style="list-style-type: none"> • Mecánicos • Eléctricos • Magnéticos • Térmicos • Acústicos

				<ul style="list-style-type: none"> • Ultrasónicos • Químicos • Ópticos • Láser
--	--	--	--	--

Fuente: (Luna Danilo, 2022)

Realizado por: Luna, Danilo 2022

1.4.3. Actuadores

Los actuadores también llamados accionadores constituyen la interfaz entre las señales de control del dispositivo y el proceso. Podemos distinguir cuatro tipos de actuadores según la fuente de energía con la que trabajan: electrónicos, eléctricos, neumáticos (aire) e hidráulicos (aceite) (Brunete, y otros, 2020, p. 25). En la tabla 6-1 se observa una breve clasificación de los actuadores más comunes en el medio.

Tabla 6-1: Clasificación de los actuadores

Neumáticos	Hidráulicos	Eléctricos
Fuente de Energía Aire de Presión	Fuente de Energía Fluido (aceite mineral)	Fuente de Energía Electricidad
Cilindros neumáticos Motores neumáticos Válvulas neumáticas	Cilindros Hidráulicos Motores Hidráulicos Válvulas Hidráulicas	Motores de corriente continua Motores de corriente alterna Motores paso a paso Relés Contactores Lámparas y luminarias

Fuente: (Brunete, y otros, 2020)

Realizado por: Luna, Danilo 2022

1.4.4. Transductores

Los transductores son sistemas o dispositivos que realizan la conversión de una magnitud física cualquiera a una magnitud eléctrica, generalmente intensidad, voltaje o impedancia, de forma que la última constituya una réplica, tan perfecta como sea posible de aquella. En la tabla 7-1 se detallan los principales transductores (Granda Miguel, y otros, 2010, p. 14).

Tabla 7-1: Clasificación de los transductores

TIPO	CARACTERISTICA	EJEMPLOS
Electroacústicos	Electricidad en sonido	Altavoz, Parlantes
Electromagnéticos	Electricidad en energía magnética	Electroimán, Buzzer
Electromecánico	Electricidad en energía mecánica	Motor Eléctrico
Fotoeléctrico	Electricidad en luz	Foco, LCD
Piezoeléctrico	Electricidad en resistencia	Transforma electricidad en sonido

Fuente: (Granda Miguel, y otros, 2010)

Realizado por: Luna, Danilo 2022

1.4.5. Cámaras para adquisición de imágenes

Una cámara posee un sensor electrónico para digitalizar las imágenes y guardarlas en una memoria. El usuario tiene la posibilidad de almacenar las imágenes en la misma cámara o de guardarlas en diferentes equipos digitales, ya sea una computadora (ordenador), una memoria USB, etc. Por lo general las cámaras digitales permiten registrar no solo fotos, sino también videos. La calidad de los documentos generados suele medirse de acuerdo a la cantidad de píxeles, aunque hay otros factores que deben tenerse en cuenta como son la sensibilidad y el tamaño del sensor (Pérez Porto, y otros, 2016, p. 22).

1.4.5.1. Cámara web

La noción de cámara web, alude a la cámara fotográfica de tipo digital que se conecta a una computadora (ordenador) para registrar y transmitir imágenes mediante Internet, en la figura 12-1. Dicha transmisión puede realizarse de manera privada es decir entre dos o más computadoras o de manera pública difundiendo las fotos en un sitio web o una red social (Merino, y otros, 2016, p. 27).



Figura 11-1. Cámara web

Fuente: (Merino, y otros, 2016)

Por tanto, se presenta una comparativa de varias cámaras web en la tabla 8-1.

Tabla 8-1: Comparativa de cámaras web

Modelo	Características	Angulo de visión	Precio
Logitech c270	Posee una resolución máxima de 720p/30fps, es compatible con Windows 7 o posterior, MacOS 10.6 o posterior y Chrome OS.	55°	\$60,00
Spedal Full HD	Incorpora un lente con una resolución de 1920x1080 pixeles compatible con Windows 7, 8, 10 o superior y configuración Plug & Play.	100°	\$60,00
Aukey webcam Full HD	Sensor de imagen CMOS de 1/2,9" con un ángulo de visión de 65°, video a 1080p, micrófono estéreo, configuración Plug & Play y es compatible con Windows, MacOS y Android.	Rota 360°	\$65,00
Papalook PA452	Posee una resolución máxima de 1080p/30fps, es compatible con Windows 7 o posterior, MacOS 10.6 o posterior y Chrome OS.	65°	\$70,00
Logitech c920s	Posee una resolución máxima de 1080p/30fps con enfoque automático, es compatible con Windows 7 o posterior, MacOS 10.6 o posterior y Chrome OS.	75°	\$115,00

Fuente: (Merino, y otros, 2016)

Realizado por: Luna, Danilo 2022

1.4.5.2. Cámara IP

Es un dispositivo que capta y transmite una señal de audio y video digital a través de una red IP estándar a otros dispositivos de red, por lo que utiliza una dirección IP, un servidor web y un protocolo de streaming de video, haciendo que los usuarios visualicen, almacenen y gestionen video de forma local o remota (Barreno, 2013, p. 19). En la tabla 9-1 se muestra varios modelos disponibles del mercado.

Tabla 9-1: Comparativa de cámaras IP

Modelo	Características	Angulo de visión	Precio
Sannce 1080p	Giro de 350° horizontal y 110° vertical. Resolución máxima de 1080p. Audio bidireccional. Grabación en tarjeta microSD de hasta 64gb. Detección de movimiento y notificación a través de su app.	60°	\$60,00

YI Dome Camera 1080p	Giro rotacional de 345° en horizontal y 115° en vertical. Notificación de movimiento a través de su app. 2 vías de audio. Visión nocturna.	60°	\$75,00
Fredi cámara IP WiFi 1080p	Dispone de Wifi y puerto Ethernet. Certificación IP65. Visión nocturna hasta 15m..	120°	\$55,00
Ring Spotlight Cam	Resolución máxima de 1080p/30fps. Compatible con Windows, MacOS y Chrome OS.	60°	\$35,00
Reolink 5MP E1 Pro	Resolución máxima de 1080p/30fps. Enfoque automático. Compatible con Windows, MacOS y Chrome OS.	78°	\$115,00

Fuente: (Barreno, 2013)

Realizado por: Luna, Danilo 2022

1.4.5.3. Cámara Raspberry Pi

Los computadores de placa reducida, han creado dispositivos externos que complementan o amplifican acciones a la hora de su funcionamiento. Es así que, crean periféricos con sus propias marcas, que facilitan su conexión y su trabajo haciéndolo una herramienta más compacta a la hora de crear proyectos o procesos. Raspberry ha creado una cámara fácil de conectar a su placa reducida para realizar proyectos en los que se requiera un sistema de visión artificial, suprimiendo el uso de una webcam tradicional (Solano Aguilar, 2018, p. 40).

Existen varias versiones de estas cámaras, dependiendo de los requerimientos que el usuario y se conectan por medio un bus de datos a un sócalo de la placa reducida ofreciendo comodidad al momento de usarlo (Solano Aguilar, 2018, p. 40).

En la siguiente tabla que corresponde a 10-1 se muestran características de las cámaras para Raspberry Pi disponibles en el mercado.

Tabla 10-1. Versiones de Raspberry Pi camera

	Cámara Pi alta calidad	Cámara Pi V2	Cámara Pi No IR V2	Cámara Pi
Sensor	Sony IMX477	Sony IMX219	Sony IMX219 sin infrarrojo	OmniVision OV5647
Enfoque	Ajustable	Si	Si	Si
Resolución	12,3 megapíxeles	8 megapíxeles	8 megapíxeles	5 megapíxeles
Cable plano	Si	Si	Si	Si
Precio	\$80,00	\$52,00	\$52,00	\$11,00

Fuente: (Solano Aguilar, 2018)

Realizado por: Luna, Danilo 2022

1.5. Software

(Manosalvas Salazar, 2017) en su trabajo dice que los computadores y algunos dispositivos de uso personal poseen un software que ejecuta tareas repetitivas bajo restricciones como son:

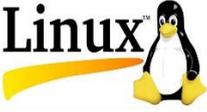
- Cantidades pequeñas de memoria, generalmente en el orden de los KB.
- Procesamiento limitado.
- Limita el consumo de energía en todo momento.

1.5.1. Sistema Operativo

Un Sistema Operativo (SO) es un software que proporciona un acceso sencillo y seguro al soporte físico del ordenador (hardware), ocultando al usuario detalles de la implementación particular y creando la ilusión de existencia de recursos ilimitados (o abundantes). Otra definición, es el de un programa que actúa como intermediario entre el usuario de la computadora y el hardware de la computadora (Castillo, 2015, p. 12).

Los sistemas operativos más utilizados para computadoras personales son, Windows, Linux y Mac OS, así como para dispositivos móviles se usa IOS y Android los cuales presentan características que se detallan en las tablas 11-1.

Tabla 11-1: Características de los Sistemas Operativos más utilizados

	Logo	Descripción	Licencia
Windows		Es el software más utilizado a nivel mundial, desarrollado por Microsoft, Compatible con procesadores Intel y Ryzen. Actualmente existe la V. 10.	Pagada
Linux		Un software OpenSource, posee Código abierto, multiplataforma, multiusuario y multitarea. Existen varias versiones de Linux, una de ellas es Raspbian.	Libre
Mac OS		Sistema operativo ocupado para dispositivos Apple, no es compatible con otro hardware.	Pagada
Android		Sistema operativo para dispositivos móviles, Es similar a Linux ya que es OpenSource	Libre

Fuente: (Castillo, 2015)

Realizado por: Luna, Danilo 2022

1.5.2. Lenguaje de Programación

Se denomina lenguaje de programación a las reglas gramaticales definidas que permiten escribir un conjunto de órdenes o instrucciones consecutivos para crear o controlar de manera física o lógica los sistemas informáticos de una máquina (Sierra, 2009, p. 26).



Figura 12-1. Leguaje de programación

Fuente: (Sierra, 2009)

Los lenguajes de programación se clasifican en dos niveles: el de bajo nivel o lenguaje de máquina que está estrechamente vinculado el hardware con el software, teniendo un control directo con el equipo y el lenguaje de alto nivel que las instrucciones se encuentran en un idioma similar al del programador facilitándolo a la hora de escribir un algoritmo para luego ser transformados a lenguaje máquina (Sierra, 2009, p. 26). En la tabla 12-1 se muestran ejemplos de los programas más conocidos:

Tabla 12-1: Clasificación de lenguajes de programación

Alto Nivel	Bajo Nivel
Las instrucciones se encuentran en un idioma similar al del programador facilitándolo a la hora de escribir un algoritmo.	Estrechamente vinculado el hardware con el software. Lenguaje nativo de las máquinas.
<ul style="list-style-type: none">• Java• PHP• Python	<ul style="list-style-type: none">• C+• Fortran• Cobol• Lisp

Fuente: (Sierra, 2009)

Realizado por: Luna, Danilo 2022

1.5.3. *OpenCV*

OpenCV que sus siglas son “Open Source Computer Vision” que significa visión artificial abierta, que es una librería de código abierto creado en un principio por Intel destinada a realizar proyectos que ayuden a detectar movimiento, reconocer objetos, y realizar reconstrucciones en 3D, algunas de las aplicaciones realizadas con visión artificial, por lo que es una plataforma muy popular. Además de los algoritmos de visión artificial se resalta el aprendizaje computacional que se enfoca en entrenar a la máquina con el objetivo de mejorar el funcionamiento por su experiencia. (Solano Aguilar, 2018, p. 43).

Se le considera una herramienta de fácil uso con muy buena eficiencia incluso realizando procesos en tiempo real. Sus algoritmos son desarrollados en lenguaje C haciéndolo compatibles sistemas operativos, como son Windows, Linux o MAC OS X, incluso Android para dispositivos móviles (Solano Aguilar, 2018, p. 43).



Figura 13-1. Logo OpenCV

Fuente: (Solano Aguilar, 2018)

1.5.4. *Software para Simulación*

El uso de programas de CAD (diseño asistido por computadora) está muy extendido en todas las ingenierías, especialmente en las técnicas. Para cualquier materia existe la posibilidad de optar por varios paquetes de software de características similares. Por ejemplo, en el campo de la electrónica, excluyendo el software especializado y limitándonos a un tipo de software más genérico, disponemos entre otros de: OrCAD, Protel, Electronics Workbench Multisim, Tango, Circuit Maker, Pads, Microsim, Proteus, etc. Razones tales como: efectividad, precio, experiencia previa, facilidad de uso, etc., hacen que los profesionales se decidan por una u otra opción (Barrón Ruiz, 2004, p. 28).

1.5.4.1. Proteus

Proteus VSM es un sistema de diseño electrónico basado en la simulación analógica, digital, mixta de circuitos, que brindan la posibilidad de interacción con muchos de los elementos que integran el circuito. Incluye componentes animados para la visualización en si comportamiento en tiempo real, además de un completo sistema de generación y análisis de señales. También cuenta con un módulo para el diseño de circuitos impresos (Rossano, 2013, p. 18).

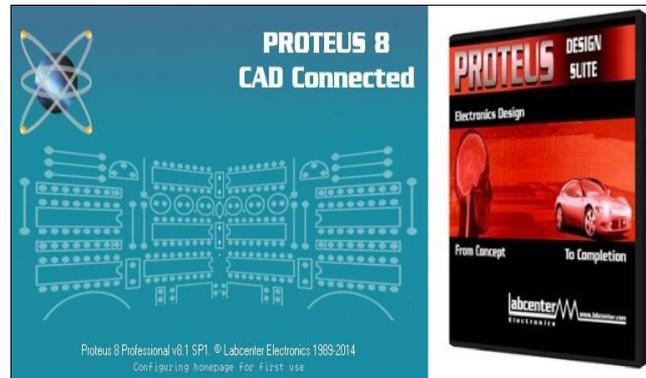


Figura 14-1. Logo Proteus

Fuente: (Rossano, 2013)

1.5.4.2. SolidWorks

(Solano Aguilar, 2018, p. 43) dice que Solid Works es un software de diseño de piezas u objetos, que permite realizar simulaciones en 2 dimensiones y 3 dimensiones, basado en un diseño asistido por computadora, permitiendo modelar cualquier tipo de objeto de una forma más sencilla y rápida, con la ventaja de poder visualizarla digitalmente teniendo una perspectiva del resultado final de la pieza u el objeto.



Figura 15-1. Logo SolidWorks

Fuente: (Solano Aguilar, 2018)

1.6. Proyectos similares enfocados a personas con discapacidad visual

En la ciudad de Quito – Ecuador, (Chincheró Iza, 2019) desarrolló un dispositivo capaz de convertir el texto de una imagen en audio, por medio de algoritmos de visión artificial, dispositivo que ayuda a personas no videntes a tener conocimiento del texto que hay en su entorno, es en este trabajo donde expone conceptos teóricos sobre la visión artificial, en donde engloba el procesamiento de imágenes, reconocimiento óptico de caracteres, diseñado el programa en Python.

Un grupo de graduados de la Universidad de Córdoba en Argentina en el 2018 desarrollaron un dispositivo portátil de lectura auditiva capaz de convertir el texto impreso en audio. A este dispositivo lo llamaron Procer Reader el cual adquiere la imagen utilizando un escáner. La voz se produce por medio de un altavoz incorporado o puede hacer uso de auriculares, el dispositivo es capaz de leer libros, diarios, revistas y archivos digitales docx, pdf y txt (Vélez Caicedo, 2019).

En México se desarrolló un dispositivo de lectura y reconocimiento visual mediante algoritmos e inteligencia artificial por parte de la empresa israelí OrCam, empresa que apoya y que está enfocada a personas con discapacidad visual. El dispositivo funciona de forma autónoma y sin necesidad de contar con una conexión en línea con un ordenador. Tiene reconocimiento facial, lectura de libros, periódicos, carteles, permite identificar billetes y carteles que se encuentran en la calle (Vélez Caicedo, 2019).

CAPÍTULO II

2. MARCO METODOLÓGICO

En este capítulo se detalla el proceso para el desarrollo del prototipo del dispositivo para el reconocimiento de texto tipo imprenta y su conversión en audio, basando en el método científico analítico, experimental y cuantitativo, partiendo de la experimentación y análisis de un gran número de casos se siguió una serie de pasos importantes y necesarios: La selección de componentes, el diseño electrónico del sistema, el diseño de la estructura del prototipo, la integración de los componentes, programación del sistema de visión artificial, y la programación para el control del prototipo. En la figura 1-2 se detalla brevemente las etapas que se realizó:

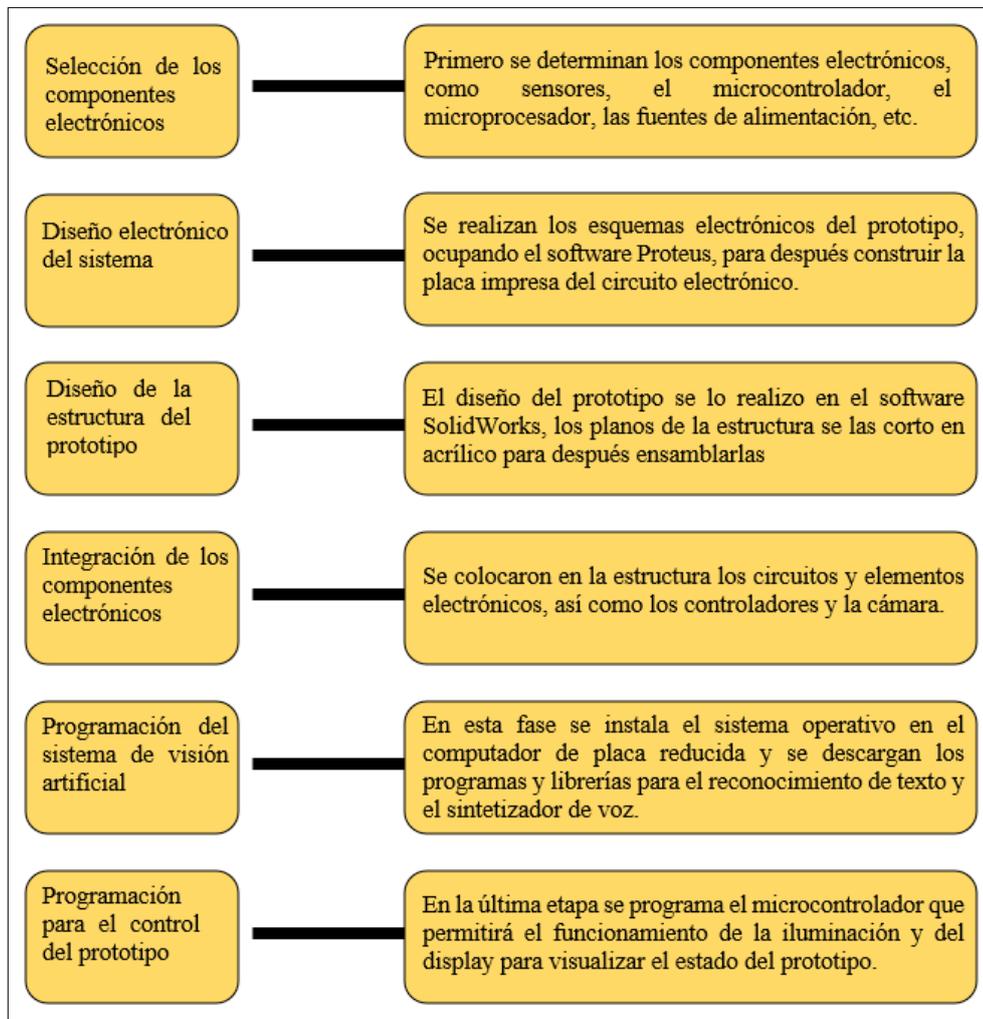


Figura 1-2. Proceso de elaboración del prototipo

Realizado por: Luna, Danilo, 2022

2.1. Selección de componentes

En la primera etapa se procedió a elegir los componentes electrónicos que serán instalados en el prototipo, así como el computador de placa reducida que se encargará de realizar el reconocimiento de texto, la cámara y la tarjeta de desarrollo que hará la parte del control de la luz y un indicador de estado mediante una LCD.

2.1.1. Raspberry Pi 3B+

Para el prototipo se escogió trabajar con una computadora de placa reducida específicamente con el modelo Raspberry Pi 3B+ capaz de contar con puertos accesibles para conectar una pantalla HDMI, un teclado, un mouse, incluso cuenta con Wifi y Bluetooth. Dado a su tamaño se hace una herramienta fácil de transportar y cuenta con accesorios externos que hacen que sea una placa poderosa para el desarrollo de proyectos, siendo una tarjeta de fácil acceso con un precio asequible. En la figura 2-2. se muestra la Raspberry Pi 3B+.

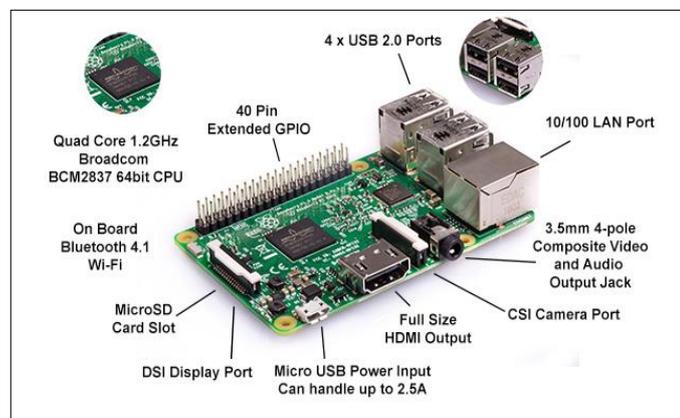


Figura 2-2. Componentes de Raspberry Pi 3B+

Fuente: (Luna, Danilo 2022)

En la tabla 1-2 se detallan sus características, entre ellas se puede destacar su procesador que si bien es cierto no es el más potente del mercado, pero cumple con las expectativas en la mayoría de las tareas a ejecutar, una de ellas es el procesamiento de imágenes al realizar visión artificial, que por tener un software libre se pueden descargar programas y librerías que maximizan su funcionalidad. El precio también es un punto positivo ya que no supera los 100 dólares americanos siendo de fácil acceso a la hora de adquirirlo.

Tabla 1-2: Características técnicas de Raspberry Pi modelo 3B+

Características	Descripción
Procesador	Chipset Broadcom BCM2837 64Bit Quad Core a 1.2 GHz
GPU	Videocore IV
RAM	1GB SDRAM a 400Mhz
Almacenamiento	Micro SD
USB 2.0	4 puertos USB
Voltaje de alimentación	2.5 A – 5V
GPIO	40 pines
Puerto Ethernet	10/100 Ethernet (RJ-45)
Wi-Fi	Integrado 802.11n
Bluetooth	Integrado Bluetooth 4.1
Precio	\$80,00

Fuente: <https://electronicacaribe.com/wp-content/uploads/2019/09/pi3.jpg>.

Realizado por: Luna, Danilo 2022

2.1.2. *Raspberry Pi camera V2*

La cámara se encarga de capturar la imagen para el reconocimiento de texto, esta cámara por ser fabricado por la misma empresa que la computadora de placa reducida es 100% compatible con la Raspberry Pi 3B+. Se muestra en la figura 3-2, y sus principales características se describen en la tabla 2-2.

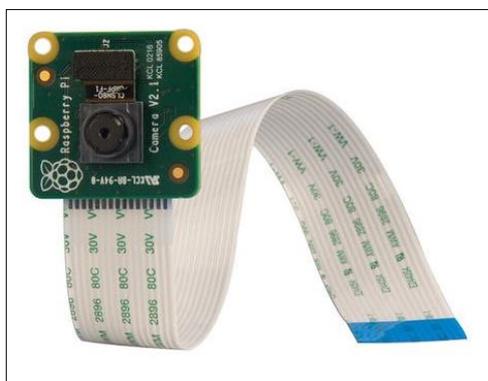


Figura 3-2. Cámara Raspberry Pi V2

Fuente: (Merino, y otros, 2016)

Tabla 2-2: Parámetros técnicos de cámara Raspberry Pi V2

Características	Descripción
Marca	Adafruit
Modelo	Versión 2.1
Dimensiones	25 x 23 x 9 mm
Peso	3.4 g
Resolución	8 megapíxeles
Sensor	Sony IMX219
Resolución del sensor	3280 x 2464 píxeles
Video	1080p30, 720p60 y resoluciones 640 x 480p90

Fuente: (UNIT ELECTRONIC, 2016)

Realizado por: Luna, Danilo 2022

2.1.3. *Altavoces USB*

Los altavoces escogidos son capaces de transmitir la información al usuario, son unos altavoces de computadora, gracias a que la Raspberry Pi cuenta con puertos USB, y un conector de audio de 3.5 mm que sirven para conectar unos parlantes USB. Las características de los altavoces escogidos se visualizan en la siguiente tabla 3-2.

Tabla 3-2: Parámetros técnicos de los Altavoces

Características	Descripción
Marca	leerFei
Voltaje de entrada	5 VDC
Potencia de salida	2W*2
Conector de audio	3.5 mm
Impedancia	4 ohm

Fuente: (Luna Danilo, 2022)

Realizado por: Luna, Danilo 2022

2.1.4. Panel led

Para el sistema de iluminación se utilizó un panel led a 12 voltios con una potencia de 6 Watos, capaz de expandir luz blanca a toda el área de trabajo. En la siguiente tabla 4-2 se muestran las características del panel.

Tabla 4-2: Parámetros técnicos del panel led

Características	Descripción
Voltaje de Operación	12 VDC
Potencia	6 W
Color de Luz	Blanca
Lúmenes máx.	450 y 550 Lm

Fuente: (Luna Danilo, 2022)

Realizado por: Luna, Danilo 2022

2.1.5. Módulo relé

Para poder accionar el panel led se utilizó un módulo relé, capaz de accionar el panel al instante de realizar el reconocimiento de texto, así como desactivar la luz al finalizar la acción. Gracias a que el módulo cuenta con un sistema de aislamiento electrónico que protege a la tarjeta de desarrollo, éste se lo alimenta directamente con 5 voltios a través de la misma tarjeta. A continuación, en la tabla 5-2 se muestran las principales características.

Tabla 5-2: Parámetros técnicos del relé

Características	Descripción
Voltaje de Operación	5 V DC
Señal de Control	TTL (3.3V o 5V)
Modelo Relay	SRD-05VDC-SL-C
Capacidad máx.	10A/250VAC, 10A/30VDC.
Corriente máx.	10A/(NO), 5A(NC)
Tiempo de acción	10 ms / 5ms

Fuente: (Electronics, 2022)

Realizado por: Luna, Danilo 2022

2.1.6. Pantalla LCD

Para poder visualizar el estado en que se encuentra el prototipo, se eligió un LCD (Liquid Crystal Display) o pantalla de cristal líquido de 16x2, es decir, 16 columnas con 2 filas, sus especificaciones se muestran en la tabla 6-2

Tabla 6-2: Parámetros técnicos de LCD

Características	Descripción
Fabricante	VISHAY
Serie	LCD-016M002B
Voltaje de entrada	5 VDC (4.7 VDC – 5.3 VDC)
Corriente de entrada	1.2 – 3mA
Voltaje Led	4.2 – 4.6 VCD
Resistencia Led	25 Ω
Dimensiones	80 x 36 x 13.5 mm

Fuente: (Luna Danilo, 2022)

Realizado por: Luna, Danilo 2022

2.1.7. Componentes electrónicos

Para la detección de la hoja de papel, para el encendido y apagado del sistema, para inicializar el reconocimiento y para encender los indicadores led, se utilizaron los siguientes componentes electrónicos:

- Resistencias de 220 Ohmios
- Diodos Led
- Final de carrera
- Borneras
- Pulsador
- Switch
- Cable USB Arduino
- Cable para conexiones

Para escoger la tarjeta de desarrollo a utilizar es necesario conocer cuantas señales de salida y entrada se tendrán. En este prototipo se utilizaron 5 salidas y 2 entradas.

2.1.8. Arduino nano

La tarjeta de desarrollo que se encargó del control de iluminación, así como de los diferentes sensores, es el modelo Arduino Nano mostrado en la figura 4-2, por tener un tamaño considerablemente pequeño, económico, completo y compatible con circuitos electrónicos, las características de Arduino nano se muestran en la tabla 7-2.

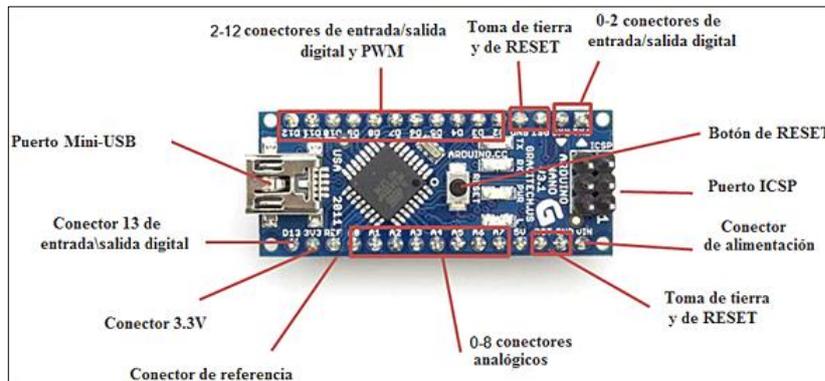


Figura 4-2. Componentes de Arduino nano.

Fuente: (Luna Danilo, 2022)

Tabla 7-2: Parámetros técnicos de Arduino nano

Características	Descripción
Microcontrolador	Atmel ATmega328
Dimensiones	45 x 18 mm
Voltaje de Operación	5V
Voltaje de entrada	7-12V
Entradas/ Salidas Digitales	14
Entradas Analógicas	8
Canales PWM	6
Memoria Flash	32 KB
Consumo	19 mA
Velocidad de Reloj	16 MHz

Fuente: (Luna Danilo, 2022)

Realizado por: Luna, Danilo 2022

Se escogió Arduino Nano debido a que cuenta con las entradas y salidas necesarias en este prototipo, además de ser pequeño y de fácil implementación en el circuito electrónico.

2.1.9. Fuentes de alimentación

Se necesitó de dos fuentes de alimentación, una que proporciona directamente voltaje hacia la Raspberry Pi 3B+ y otra capaz de encender el panel Led. La tarjeta de desarrollo se alimenta directamente a través de la Raspberry, ya que posee puertos USB.

Las características de la fuente de alimentación para la Raspberry Pi 3B+ se muestran en la siguiente tabla 8-2 y la fuente de alimentación capaz de encender el panel led se describe en la tabla 9-2.

Tabla 8-2: Datos técnicos de fuente de alimentación para Raspberry Pi

Características	Descripción
Voltaje de Entrada	100-240 VAC
Voltaje de Salida	5 VDC
Corriente	3 A
Frecuencia	50/60 Hz

Fuente: (Luna Danilo, 2022)

Realizado por: Luna, Danilo 2022

Tabla 9-2: Datos técnicos de fuente de alimentación para iluminación

Características	Descripción
Voltaje de Entrada	110 VAC
Voltaje de Salida	12 VDC
Corriente	2 A
Frecuencia	60 Hz

Fuente: (Luna Danilo, 2022)

Realizado por: Luna, Danilo 2022

2.2. Diseño electrónico del sistema

En esta etapa se construyó el circuito electrónico que formará parte del dispositivo de reconocimiento de texto. Por lo que se hizo el uso del software Proteus, el cual permite diseñar el circuito esquemático, para después realizarlo en PCB (circuito impreso en placa).

2.2.1. Diagrama esquemático

La placa electrónica estará formada por la tarjeta de desarrollo Arduino, las fuentes de alimentación, y el módulo relé, los demás componentes se les ubicara de forma independiente, pero estarán todos conectados entre sí.

El esquema electrónico se muestra en la figura 5-2, donde se visualizan todos los componentes antes mencionados

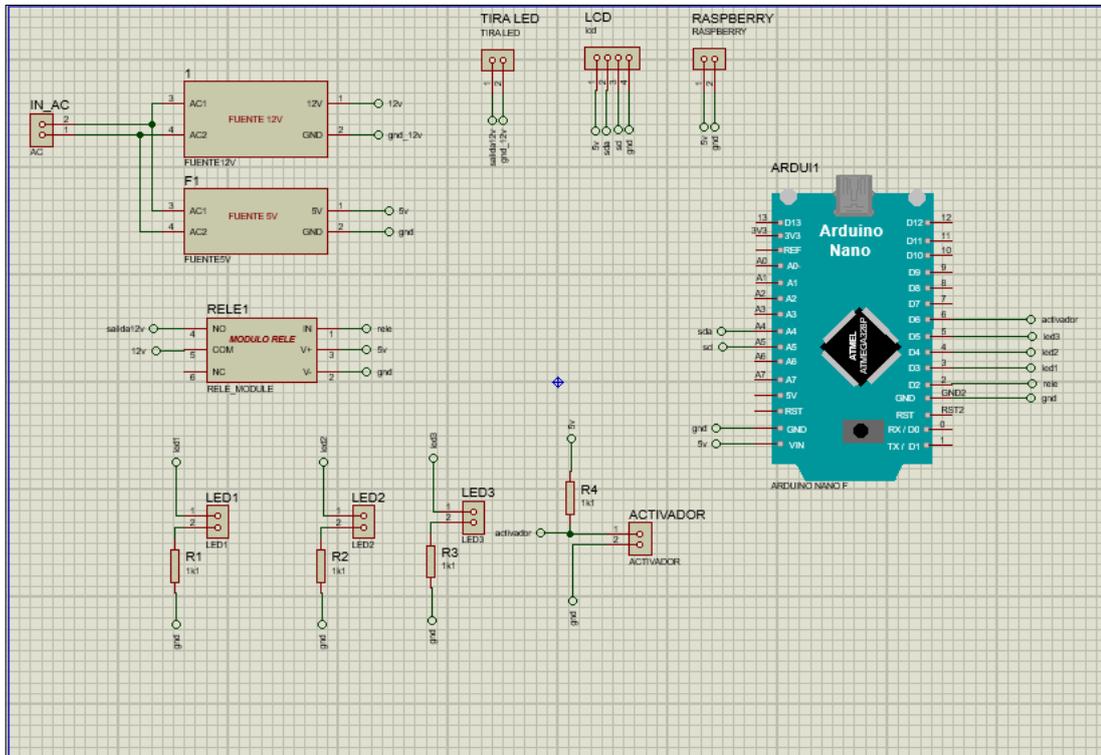


Figura 5-2. Diagrama electrónico diseñado en Proteus

Realizado por: Luna, Danilo 2022

2.2.2. Diseño PCB

Una vez realizado el esquemático del circuito electrónico, se procedió con el diseño del circuito impreso, para ello se ubicó los componentes y se elaboró las pistas (líneas de conexión) entre elementos como se muestra en la figura 6-2.

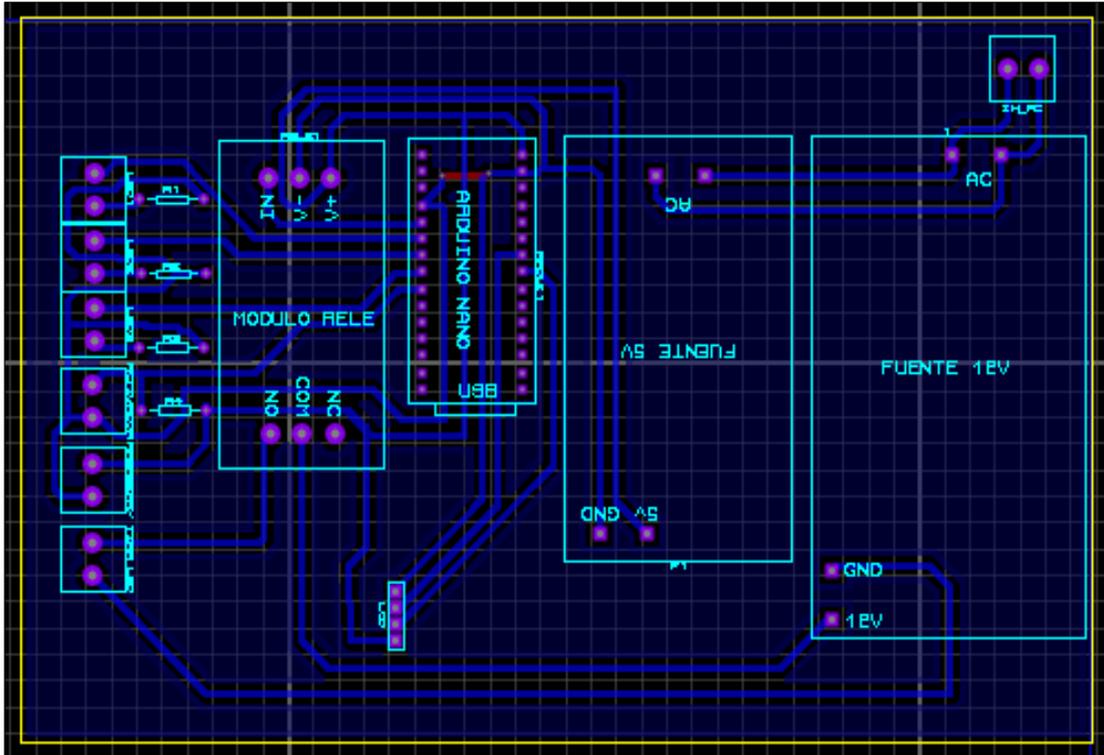


Figura 6-2. Diseño de circuito impreso en Proteus

Realizado por: Luna, Danilo 2022

2.2.3. PCB – 3D

Proteus posee una herramienta que permite realizar una visualización en 3D del diseño PCB como se muestra en la figura 7-2, esto es de mucha ayuda debido a que brinda una mejor perspectiva de la placa del circuito electrónico.

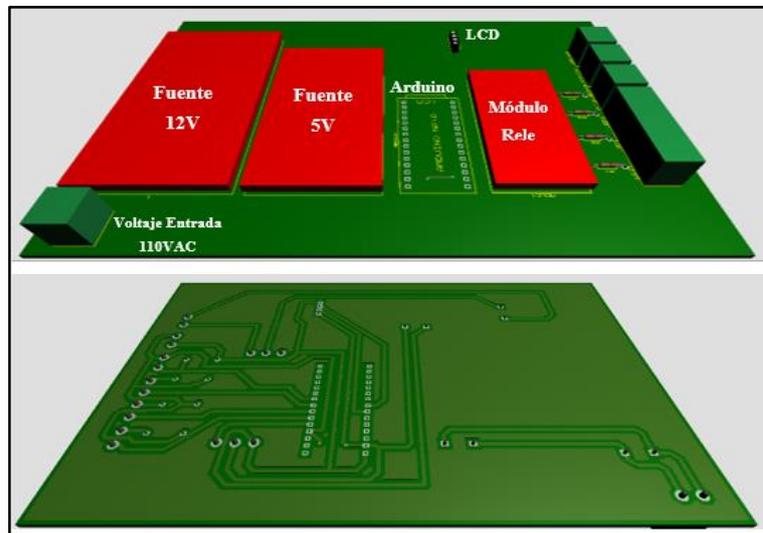


Figura 7-2. Diseño de circuito impreso en 3D en Proteus

Realizado por: Luna, Danilo 2022

2.2.4. Circuito electrónico en PCB

Una vez diseñada la PCB en Proteus, se elabora la misma de forma física, y se montan los elementos previamente seleccionados. En la figura 8-2 se muestra el resultado del circuito electrónico en placa elaborado.

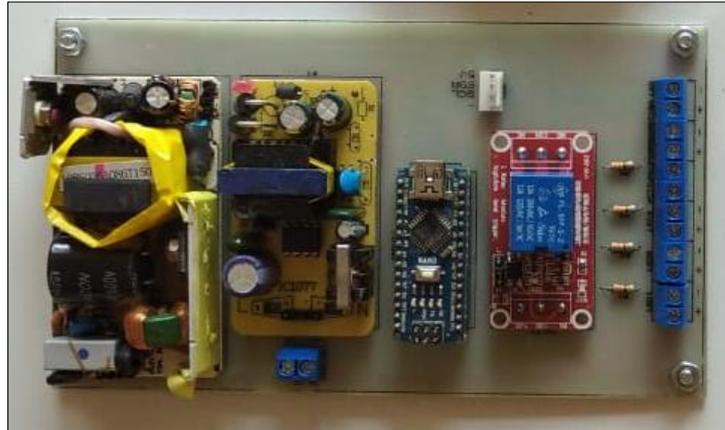


Figura 8-2. Montaje de componentes en PCB

Realizado por: Luna, Danilo 2022

2.3. Diseño de la estructura del prototipo

La estructura del prototipo brinda el acceso a la persona no vidente con el dispositivo de reconocimiento de texto, donde externamente se visualiza la pantalla LCD y los elementos de accionamiento mientras que internamente irán instalados los diferentes componentes electrónicos antes mencionados, colocándolos estéticamente para que no sean manipulados. La forma del prototipo tiene similitud con una impresora o un escáner, para que éste pueda ser ubicado en un escritorio de oficina siendo cómodo a la hora de utilizarlo.

El diseño se lo realizó en el programa SolidWorks 2021, en la figura 9-2 se puede observar el modelado de la estructura del dispositivo, las piezas diseñadas, por medio de una cortadora a laser, se les puede obtener de forma real en acrílico blanco para finalmente armarlas y pegarlas como se muestra en la figura 10-2.

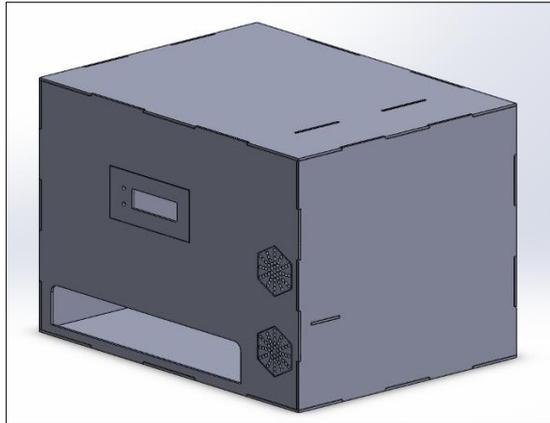


Figura 9-2. Diseño de la estructura del prototipo en SolidWorks

Realizado por: Luna, Danilo 2022

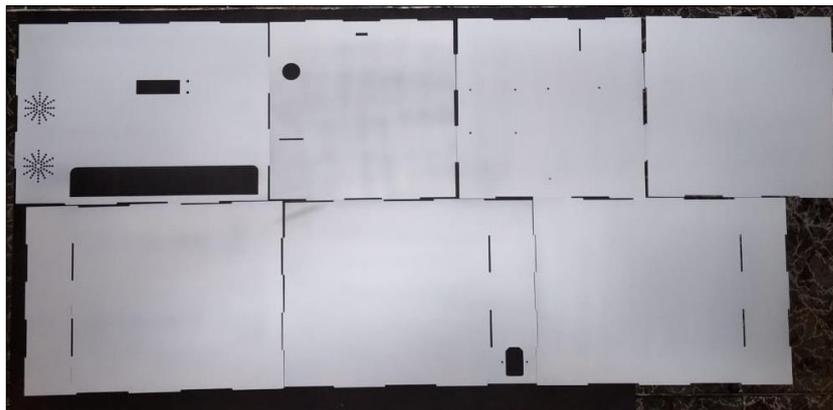


Figura 10-2. Partes de acrílico para la estructura cortadas en laser

Realizado por: Luna, Danilo 2022

La estructura presenta un apartado en el lado derecho donde se colocaron los elementos electrónicos, como se muestra en la figura 11-2, además en la parte inferior se instaló la bandeja para la hoja de papel.

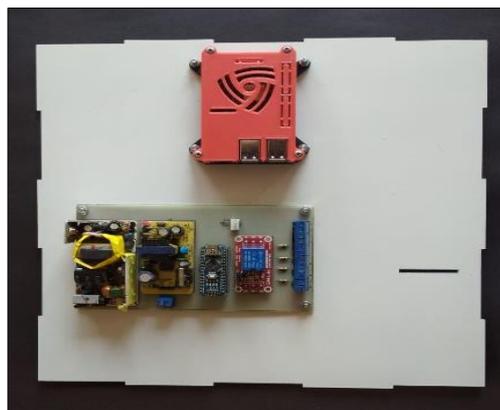


Figura 11-2. Montaje de la PCB junto a Raspberry Pi

Realizado por: Luna, Danilo 2022

La bandeja contiene unos rieles que permiten un correcto desplazamiento al momento de colocar una hoja de papel como se observa a continuación en la figura 12-2.

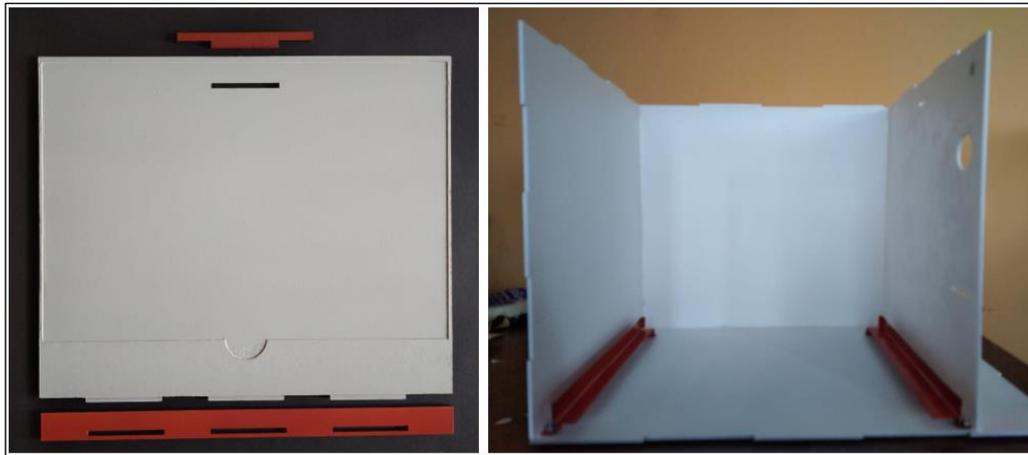


Figura 12-2. Bandeja porta papel – Rieles para colocar bandeja

Realizado por: Luna, Danilo 2022

La cámara además de estar colocada en la parte superior del prototipo cuenta con un soporte de acrílico donde mediante cuatro tornillos permite regular hasta la altura deseada teniendo una movilidad aproximadamente de 5cm como se muestra en la figura 13-2.

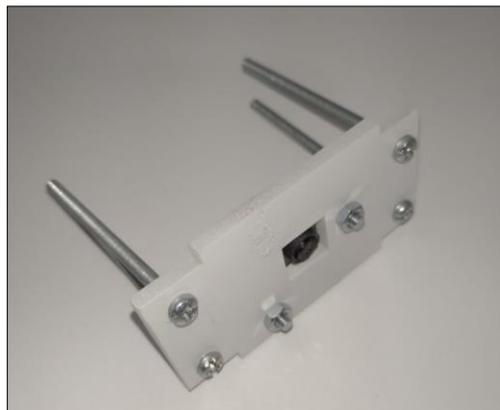


Figura 13-2. Base de acrílico para sujetar la cámara

Realizado por: Luna, Danilo 2022

2.4. Integración de todos los componentes

Una vez realizada la estructura, se procedió a instalar todos los elementos electrónicos, es así como se colocó la Raspberry Pi 3B+, el circuito electrónico, los altavoces, la cámara, el sensor que detecta la bandeja de entrada de la hoja, la pantalla LCD, el botón de inicio, y el panel led.

La cámara se colocó en la parte superior como se muestra en la figura 14-2, y ésta tiene la posibilidad de ser regulada para poder tener el mejor enfoque a la hora del reconocimiento de texto. De la misma manera junto a la cámara se instaló el panel led principal, encargado de la iluminación del sistema, este se enciende al momento de realizar el reconocimiento de texto.



Figura 14-2. Cámara colocada

Realizado por: Luna, Danilo 2022

El sensor encargado de la detección de la hoja se le ubicó en la parte final del riel de la bandeja, así se conoce el momento en que la hoja fue ingresada. En la figura 15-2 se muestra la ubicación del sensor.



Figura 15-2. Sensor final de carrera

Realizado por: Luna, Danilo 2022

La Raspberry Pi 3B+ y el circuito electrónico donde se encuentra la tarjeta de desarrollo junto a las fuentes de alimentación se instalaron en el apartado derecho del dispositivo, como se muestra en la figura 16-2, aquí se harán las conexiones respectivas entre los componentes. En este mismo apartado se colocará los altavoces.



Figura 16-2. Montaje de elementos del prototipo

Realizado por: Luna, Danilo 2022

En la cara frontal del prototipo se instaló la pantalla LCD y dos diodos led que sirven como indicadores del encendido del dispositivo y el inicio el reconocimiento de texto. Luego de situar los componentes electrónicos en conjunto con la bandeja de entrada y colocar la cara frontal, el resultado del dispositivo se muestra en la figura 7-2.

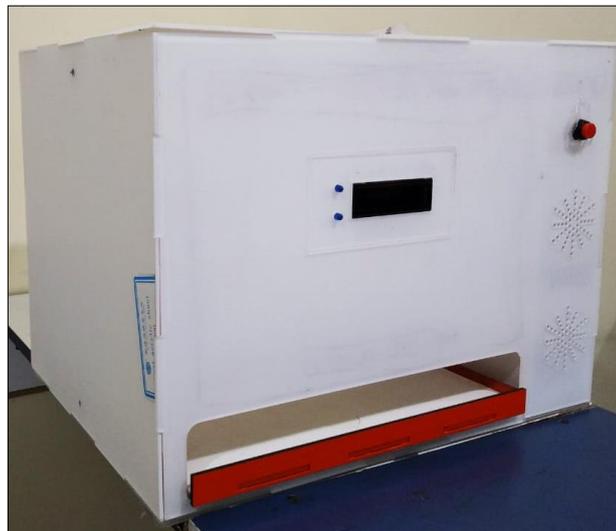


Figura 17-2. Prototipo ensamblado

Realizado por: Luna, Danilo 2022

2.5. Programación del sistema de visión artificial

Una vez colocados todos los elementos que conforman el dispositivo, se procedió a programarlo, aquí se desarrollaron las sentencias para el reconocimiento de texto. La Raspberry Pi 3B+ es la encargada de la visión artificial y para poder programarla se necesitó seguir una serie de pasos detallados a continuación:

- Instalación del sistema operativo Raspbian
- Instalación de librerías para visión artificial
- Programación en Python



Figura 18-2. Pantalla principal del sistema operativo

Realizado por: Luna, Danilo 2022

2.5.1. *Instalación del sistema operativo*

Raspberry al ser una mini computadora requiere de un sistema operativo. Se eligió instalar Raspbian, ya que tiene una interfaz más accesible y amigable a la hora de instalar los diferentes programas, así como las librerías necesarias. Raspbian es gratuito y para poder utilizarlo se instala en una memoria micro SD la cual es insertada en la Raspberry. Los pasos para instalar Raspbian se detallan a continuación:

- Instalar Noobs, un software que instala Raspbian en la Micro SD
- Escoger el sistema operativo Raspbian e instalar en la tarjeta Micro SD
- Colocar la Micro SD en la Raspberry
- Encender la Raspberry

La ventaja de este sistema operativo es que ya cuenta con algunos programas y complementos necesarios para realizar las primeras configuraciones además viene preinstalado una versión de Python siendo éste el programa encargado de compilar y ejecutar los scripts compatibles con el mismo.

2.5.2. *Instalación de librerías necesarias*

Python maneja complementos externos que deben ser descargados de acuerdo a lo requerido. Raspberry tiene que contar con las siguientes librerías:

- Numpy: Librería para la informática científica, matrices, números, etc.
- Cv2: OpenCV: Librería encargada del reconocimiento de imágenes.
- Pytesseract: Librería que proporciona el reconocimiento de caracteres OCR.
- GTTS: Librería que convierte el texto a voz.
- Os: Librería que permite acceder a funciones del sistema operativo.
- Serial, time: Librería encargada de la comunicación serial.
- Picamara: Librería encargada de las funcionalidades de la cámara.

Estas librerías se las puede descargar a través del terminal del sistema operativo, y son fáciles de acceder debido a que son gratuitas.

2.5.3. *Etapas de diseño del algoritmo*

Para explicar mejor la programación del sistema de reconocimiento realizado en la Raspberry Pi 3B+, se lo ha dividido por etapas de desarrollo que son mostradas a continuación:

2.5.3.1. *Adquisición de imágenes*

En esta etapa se capturó la imagen a través de la cámara Raspberry Pi V2 misma que se guarda en la memoria del mini computador. La figura 19-2 muestra un ejemplo de foto captada por la cámara del prototipo.

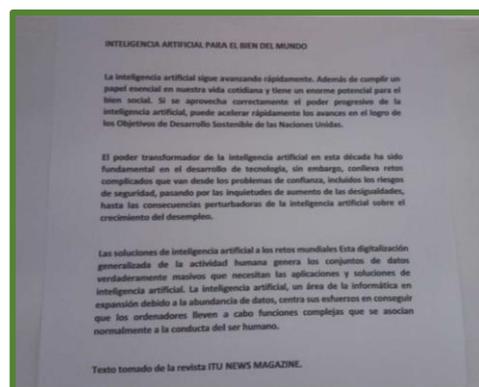


Figura 19-2. Captura de la imagen sin filtro

Realizado por: Luna, Danilo 2022

2.5.3.2. Procesamiento de la imagen

La imagen captada por la cámara es procesada, por tal motivo pasa por un filtro llamado binarización que consiste en cambiar la imagen a escala de grises para después limpiar el proceso (eliminación de ruido), resaltando el color blanco y negro, se puede observar en la figura 20-2 para finalmente realizar la segmentación, eligiendo las regiones de interés, en este caso los caracteres. El gráfico 1-2 corresponde al diagrama de flujo para procesar el texto de la imagen.

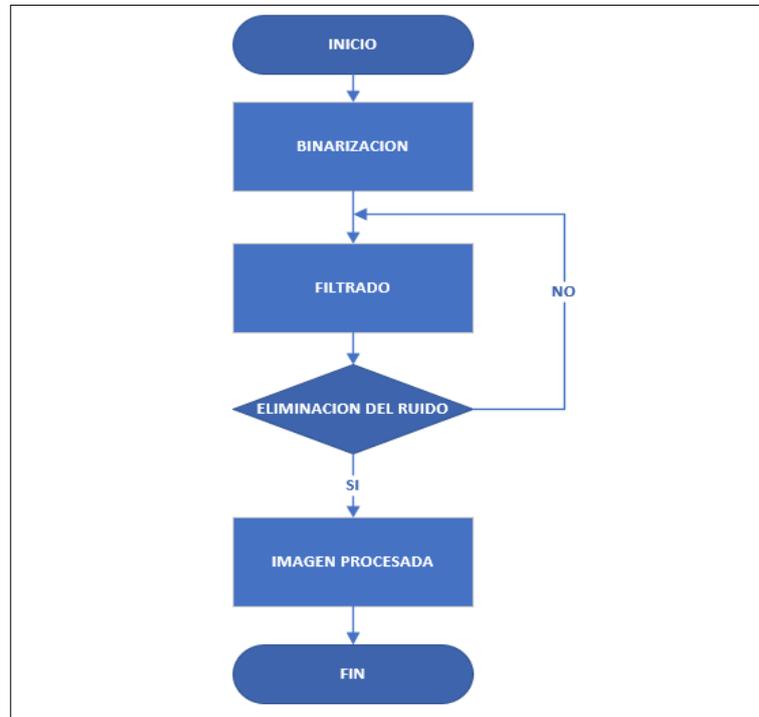


Gráfico 1-2. Diagrama de flujo del procesamiento de la imagen

Realizado por: Luna, Danilo 2022

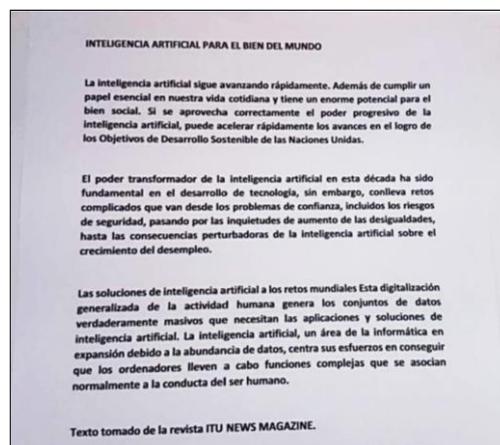


Figura 20-2. Imagen binarizada.

Realizado por: Luna, Danilo 2022

2.5.3.3. Proceso de reconocimiento óptico de caracteres

Preparada la imagen se realiza el reconocimiento de caracteres, implementando una comparación de la forma de los caracteres de la imagen con la forma de los caracteres ya existentes en la librería Tesseract de Python, parte importante del reconocimiento OCR. A la librería Pytesseract se le debe configurar el idioma para que extraiga el texto debidamente. El tiempo del procesamiento dependerá de la complejidad de la imagen ya que deberá descifrar y buscar las letras que conforman el texto. En el gráfico 2-2 se muestra el proceso del reconocimiento óptico de caracteres.

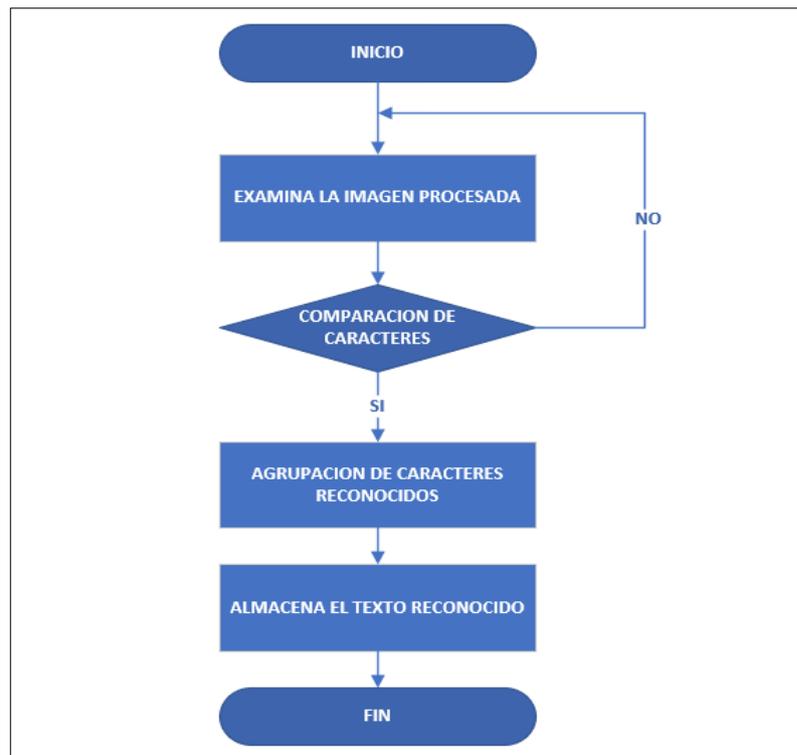


Gráfico 2-2. Diagrama de flujo del reconocimiento de caracteres

Realizado por: Luna, Danilo 2022

2.5.3.4. Conversión del texto detectado en audio

Para realizar la conversión de texto a audio se utilizó una librería llamada GTTS de Python, que usa la API de Google (Google Text-To-Speech). Con esta librería se logró tomar textos cortos y extensos, para luego ser almacenado en un archivo de audio en formato MP3. Pero lo más importante de la plataforma es que brinda la posibilidad de personalizar la conversión y determinar ciertos parámetros para el idioma. En el gráfico 3-2 se describe el proceso de la conversión de texto a voz.

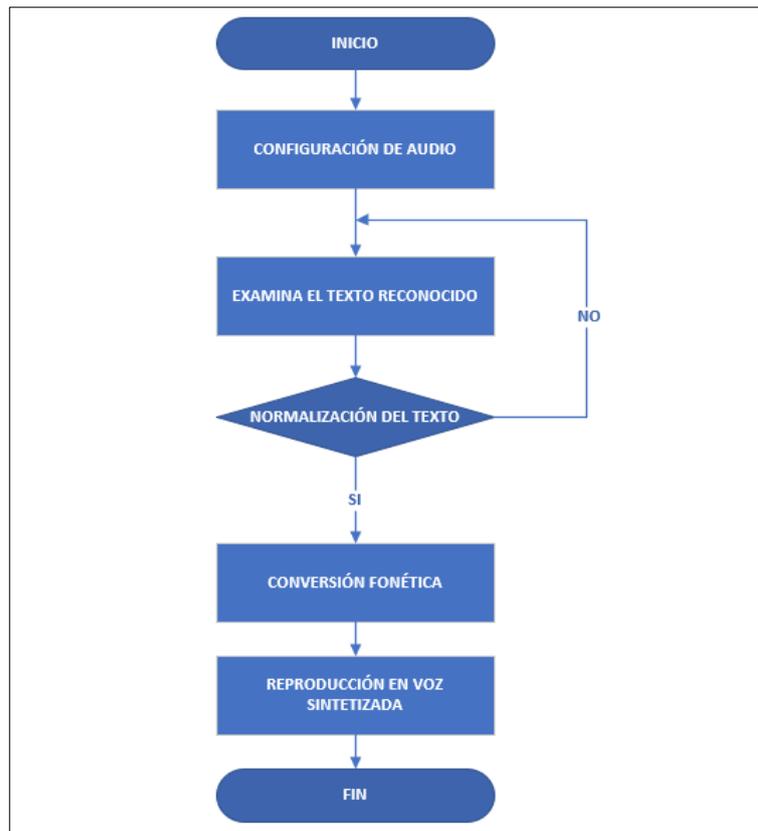


Gráfico 3-2. Diagrama para la conversión de texto en audio

Realizado por: Luna, Danilo 2022

2.6. Programación de la tarjeta de desarrollo

La tarjeta de desarrollo Arduino es un complemento del dispositivo, éste se encarga de controlar la iluminación del sistema así también los indicadores visuales del prototipo. Para comunicarse con la Raspberry Pi simplemente se lo realizó por medio de comunicación serial ya que la tarjeta Arduino está conectado directamente al puerto serial de Raspberry Pi 3B+.

Arduino cuenta con un software libre llamado IDE de Arduino que es fácil de programar ya que usa lenguaje C++. A continuación, se detalla el proceso de la tarjeta de desarrollo por medio de un diagrama de flujo como se muestra en el gráfico 4-2.

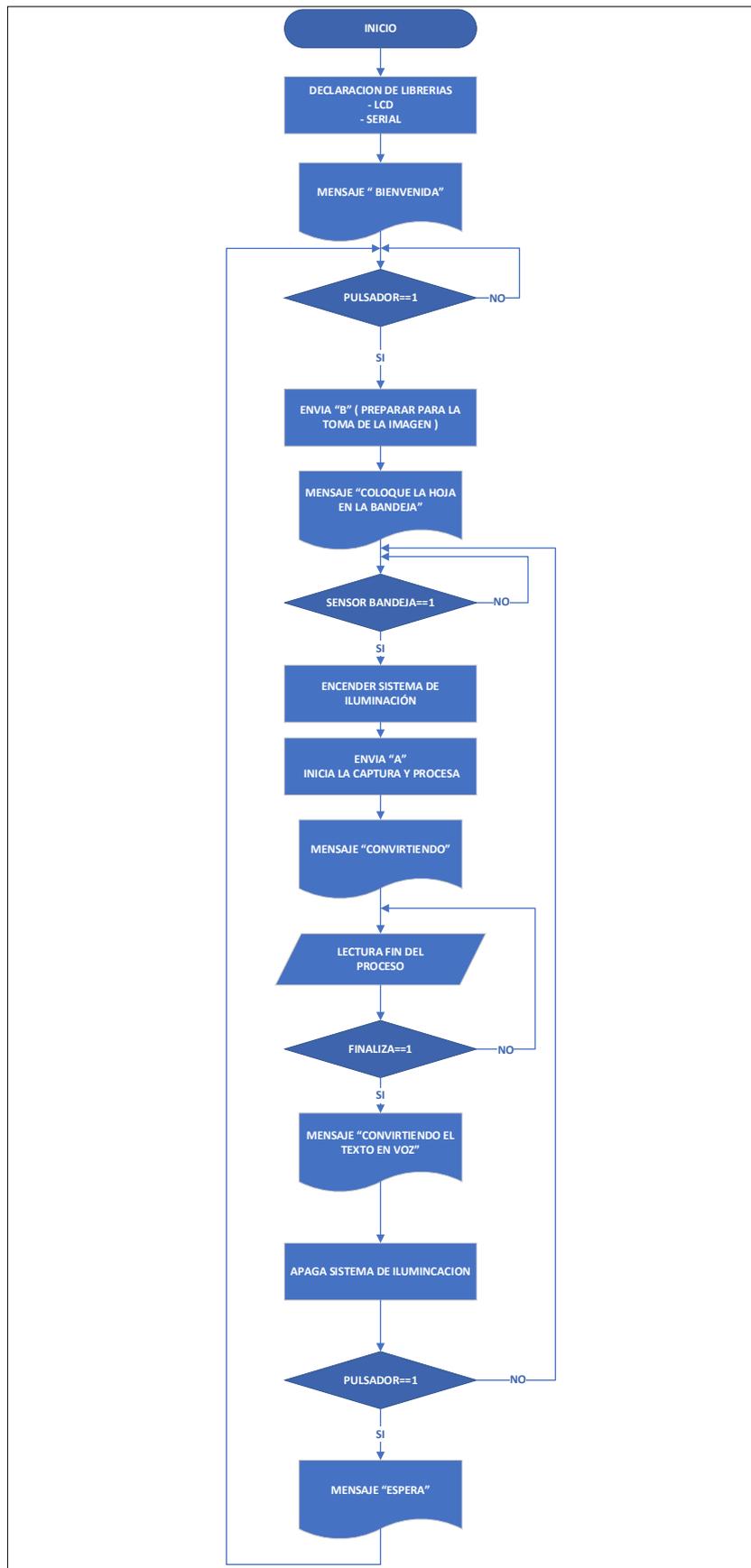


Gráfico 4-2. Diagrama de flujo de la tarjeta de desarrollo

Realizado por: Luna, Danilo 2022

2.7. Algoritmo final del sistema

A continuación, se muestra en el gráfico 5-2 de la siguiente página, el proceso final del sistema de reconocimiento de texto junto con el sintetizador de voz.

2.8. Resultado final del prototipo

El prototipo del dispositivo de reconocimiento de texto y sintetizador de voz, después de ser implementado y programado se muestra en la figura 21-2.



Figura 21-2. Prototipo terminado

Realizado por: Luna, Danilo 2022

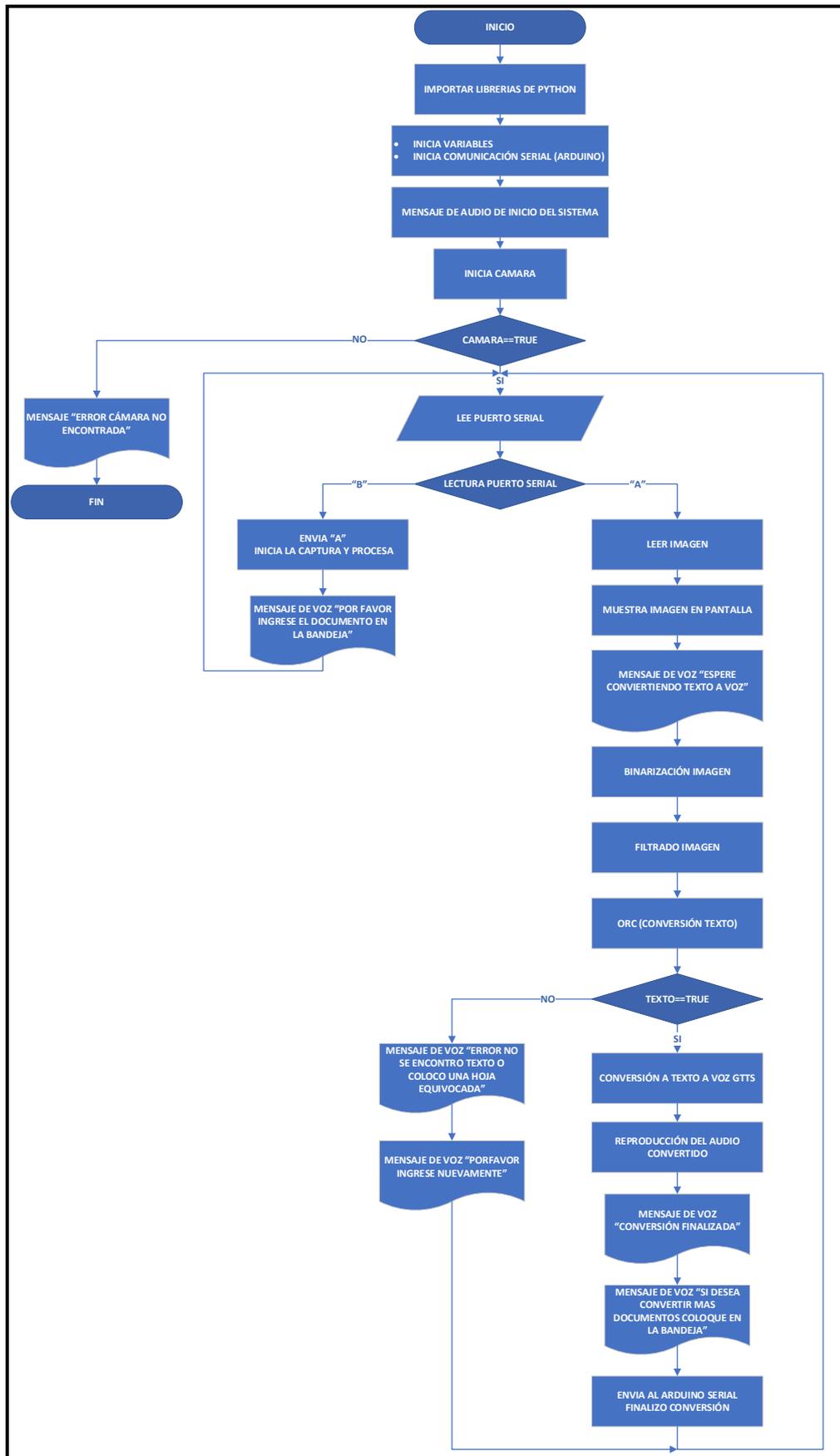


Gráfico 5-2. Diagrama de flujo final del sistema

Realizado por: Luna, Danilo 2022

CAPÍTULO III

3. PRUEBAS Y RESULTADOS

En el presente capítulo se muestran los resultados de las pruebas del prototipo terminado, para ello los parámetros considerados son: distancia entre la hoja con texto y la cámara, iluminación del entorno de la hoja, tipo de fuente, cantidad y formato de texto. Todo ello tomando en cuenta la eficiencia del dispositivo en un ambiente controlado.

3.1. Prueba de iluminación para detección de texto

La primera prueba que se desarrolló es la de iluminación ya que este aspecto es fundamental dentro del entorno para la detección de texto, logrando de esta manera reducir las imperfecciones en la captura de la imagen. En la tabla 1-3 se muestra los resultados obtenidos en el luxómetro al variar la luminosidad de la luz led de 0 a 540 lúmenes, evidenciando que a partir de 180 lúmenes la cámara detecta de forma clara el texto tomando en cuenta que la luz led se encuentra frontal a la hoja.

Tabla 1-3: Prueba de iluminación

Watts	Lúmenes	Detecta	Detecta con fallas	No detecta
0	0			✓
1	90		✓	
2	180	✓		
3	270	✓		
4	360	✓		
5	450	✓		
6	540	✓		

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

3.2. Pruebas de detección de texto según la distancia

Estas pruebas se realizaron en base a tres distancias: 20, 25 y 30 cm como se muestra en la tabla 2-3, para cada una de las distancias se tomaron cinco muestras con fuente Calibri 12 pts., donde se obtuvo una eficiencia del dispositivo de 92,79% a una distancia de 20 cm y una eficiencia

máxima del 94,82% a una distancia de 25 cm, cabe recalcar que a 30 cm de distancia muestra una eficiencia del 93,60%, de esta manera podemos demostrar que el prototipo posee un buen funcionamiento en este intervalo de distancia. Además, en el grafico 1-3 se puede evidenciar la eficiencia del prototipo en base a la distancia.

Tabla 2-3: Pruebas a diferente distancia desde la cámara hacia la hoja

n	Distancia (cm)	Palabras del texto	Palabras no detectadas	Palabras incorrectas	Palabras correctas	Eficiencia (%) ($\frac{P.correctas}{P. del texto}$)
1	20	197	1	14	182	92,39
2	20	197	0	14	183	92,89
3	20	197	0	14	183	92,89
4	20	197	1	13	183	92,89
5	20	197	1	13	183	92,89
PROMEDIO			1	14	183	92,79
1	25	197	1	10	186	94,42
2	25	197	0	10	187	94,92
3	25	197	0	10	187	94,92
4	25	197	0	10	187	94,92
5	25	197	0	10	187	94,92
PROMEDIO			0	10	187	94,82
1	30	197	1	12	184	93,40
2	30	197	0	12	185	93,91
3	30	197	1	12	184	93,40
4	30	197	0	12	185	93,91
5	30	197	1	12	184	93,40
PROMEDIO			1	12	184	93,60

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

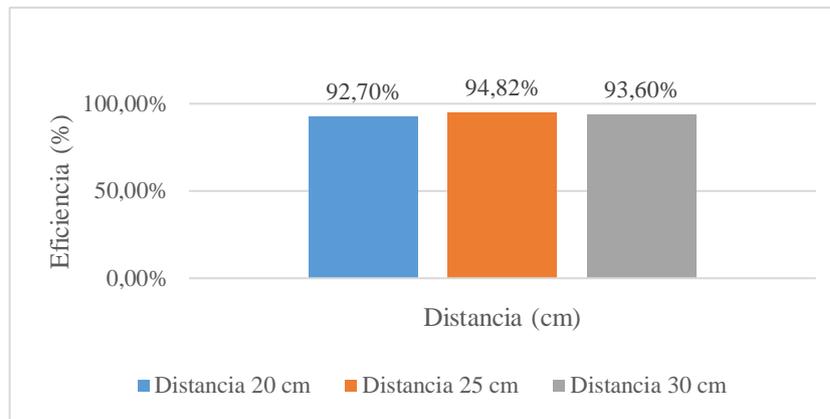


Gráfico 1-3. Eficiencia del dispositivo a diferente distancia

Realizado por: Luna, Danilo 2022

3.3. Pruebas de detección de texto con fuentes y tamaños diferentes

Estas pruebas se realizaron concretamente con tres tipos de fuente: “Calibri”, “Times New Roman” y “Arial”, debido a que son más utilizadas y conocidas en nuestro medio. Además, tomando en cuenta que es un prototipo y por ciertas limitaciones de hardware no resulta recomendable aumentar más tipos de fuentes en las pruebas ya que provocaría un procesamiento mucho más tardado. Las tres fuentes escogidas fueron probadas con diferentes tamaños de letra, para ello se consideró 9, 11 y 12 pts., siendo los más comúnmente utilizados en la elaboración de documentos de texto. A continuación, se detalla las pruebas realizadas.

3.3.1. Pruebas de texto con fuente Calibri

De acuerdo a lo antes mencionado se realizaron pruebas con tres tamaños de texto con esta fuente, cuyos resultados reflejan que el tamaño más adecuado debe ser superior a 9 pts., debido a que la eficiencia del prototipo con estos parámetros supera el 90% de eficiencia como se muestra en las tablas 3-3, 4-3 y 5-3 respectivamente. Cabe mencionar que se utilizó 15 hojas con texto de 115 palabras para determinar estos porcentajes.

Tabla 3-3: Pruebas con fuente Calibri – tamaño 9 pts.

n	Palabras del texto	No. Palabras detectado	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $(\frac{E.Abs*100}{P.texto})$	Eficiencia (%) $(\frac{P.correctas*100}{P.del\ texto})$
1	115	9	43	63	104,51	52	45,22	54,78
2	115	12	42	61	105,77	54	46,96	53,04
3	115	10	45	60	107,25	55	47,83	52,17
4	115	13	47	55	119,01	60	52,17	47,83

5	115	10	44	61	106,23	54	46,96	53,04
6	115	11	48	56	108,14	59	51,30	48,70
7	115	9	45	61	106,05	54	46,96	53,04
8	115	14	49	52	124,49	63	54,78	45,22
9	115	13	46	56	107,53	59	51,30	48,70
10	115	16	49	50	125,23	65	56,52	43,48
11	115	12	43	60	107,42	55	47,83	52,17
12	115	10	44	61	106,41	54	46,96	53,04
13	115	11	46	58	110,58	57	49,57	50,43
14	115	13	46	56	108,39	59	51,30	48,70
15	115	12	41	62	105,16	53	46,09	53,91
PROMEDIO		12	45	58	110,14	56,87	49,45	50,43

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 4-3: Pruebas con fuente Calibri – tamaño 11 pts.

n	Palabras del texto	No. Palabras detectado	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $\left(\frac{E.Abs \cdot 100}{P.texto}\right)$	Eficiencia (%) $\left(\frac{P.correctas \cdot 100}{P.del\ texto}\right)$
1	115	2	8	105	105,49	10	8,70	91,30
2	115	2	8	105	103,54	10	8,70	91,30
3	115	3	9	103	104,77	12	10,43	89,57
4	115	2	8	105	104,77	10	8,70	91,30
5	115	3	8	104	106,23	11	9,57	90,43
6	115	3	8	104	108,14	11	9,57	90,43
7	115	2	8	105	106,05	10	8,70	91,30
8	115	2	8	105	124,49	10	8,70	91,30
9	115	2	8	105	107,53	10	8,70	91,30
10	115	2	8	105	125,23	10	8,70	91,30
11	115	2	8	105	107,42	10	8,70	91,30
12	115	2	9	104	106,41	11	9,57	90,43
13	115	2	8	105	110,58	10	8,70	91,30
14	115	2	8	105	108,39	10	8,70	91,30
15	115	2	8	105	105,16	10	8,70	91,30
PROMEDIO		2	8	105	108,95	10,33	8,99	91,30

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 5-3: Pruebas con fuente Calibri – tamaño 12 pts.

n	Palabras del texto	No. Palabras detectado	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $\left(\frac{E.Abs*100}{P.texto}\right)$	Eficiencia (%) $\left(\frac{P.correctas*100}{P. del texto}\right)$
1	115	2	5	108	104,78	7	6,09	93,91
2	115	0	5	110	105,31	5	4,35	95,65
3	115	0	5	110	105,25	5	4,35	95,65
4	115	2	5	108	106,21	7	6,09	93,91
5	115	0	5	110	105,23	5	4,35	95,65
6	115	0	5	110	105,35	5	4,35	95,65
7	115	0	5	110	104,05	5	4,35	95,65
8	115	2	5	108	104,39	7	6,09	93,91
9	115	0	5	110	104,53	5	4,35	95,65
10	115	0	5	110	105,35	5	4,35	95,65
11	115	0	5	110	104,52	5	4,35	95,65
12	115	0	5	110	106,01	5	4,35	95,65
13	115	0	5	110	105,18	5	4,35	95,65
14	115	0	5	110	105,29	5	4,35	95,65
15	115	0	5	110	105,16	5	4,35	95,65
PROMEDIO		0	5	110	105,11	5,4	4,70	95,65

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

En el gráfico 2-3 se puede observar el porcentaje de eficiencia del dispositivo después de realizar las pruebas con el texto en tamaño 9, 11 y 12 pts., siendo el de mayor efectividad con un 95,65% el tamaño 12 pts.

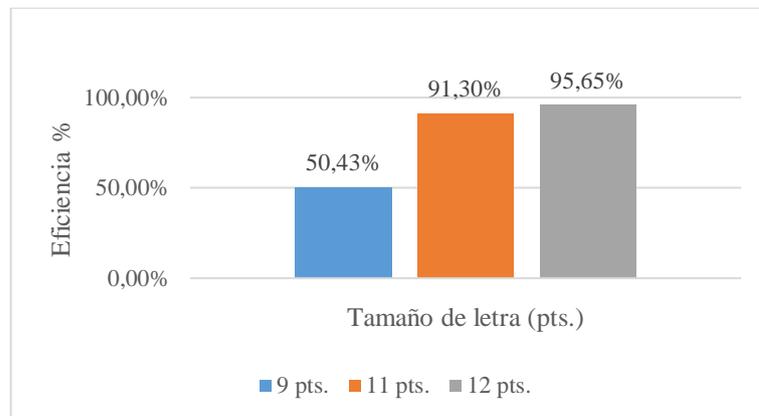


Gráfico 2-3. Eficiencia del prototipo con fuente Calibri y diferente tamaño

Realizado por: Luna, Danilo 2022

3.3.2. Pruebas de texto con fuente Times New Roman

En este apartado se utilizaron tres tamaños para obtener los datos de las tablas 6-3, 7-3 y 8-3 donde se puede observar los resultados obtenidos con esta fuente, las cuales ubican en su punto más alto de eficiencia al texto con tamaño 12 pts., claramente muy superior que con tamaño 9 pts., tomando en cuenta que se utilizaron 15 hojas con 115 palabras de texto. Cabe mencionar que el desempeño del dispositivo con tamaño 11 pts. muestra una eficiencia aceptable.

Tabla 6-3: Prueba con fuente Times New Roman – tamaño 9 pts.

n	Palabras del texto	No. Palabras detectado	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $\left(\frac{E.Abs*100}{P.texto}\right)$	Eficiencia (%) $\left(\frac{P.correctas*100}{P.del\ texto}\right)$
1	115	37	33	45	118,91	70	60,87	39,13
2	115	42	32	41	96,59	74	64,35	35,65
3	115	41	30	44	50,86	71	61,74	38,26
4	115	37	32	46	78,34	69	60,00	40,00
5	115	39	30	46	83,2	69	60,00	40,00
6	115	40	30	45	93,45	70	60,87	39,13
7	115	38	31	46	65,16	69	60,00	40,00
8	115	40	30	45	75,49	70	60,87	39,13
9	115	38	31	46	69,54	69	60,00	40,00
10	115	38	31	46	81,37	69	60,00	40,00
11	115	37	32	46	65,45	69	60,00	40,00
12	115	42	35	38	70,36	77	66,96	33,04
13	115	38	30	47	67,54	68	59,13	40,87
14	115	41	32	42	70,16	73	63,48	36,52
15	115	39	33	43	56,29	72	62,61	37,39
PROMEDIO		39	32	44	76,18	70,60	61,39	38,26

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 7-3: Prueba con fuente Times New Roman – tamaño 11 pts.

n	Palabras del texto	No. Palabras detectado	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $\left(\frac{E.Abs*100}{P.texto}\right)$	Eficiencia (%) $\left(\frac{P.correctas*100}{P.del\ texto}\right)$
1	115	16	9	90	318,1	25	21,74	78,26
2	115	4	11	100	128,43	15	13,04	86,96
3	115	4	13	98	130,99	17	14,78	85,22

4	115	8	15	92	124,3	23	20,00	80,00
5	115	6	16	93	133,12	22	19,13	80,87
6	115	4	11	100	123,34	15	13,04	86,96
7	115	4	11	100	127,02	15	13,04	86,96
8	115	4	11	100	131,5	15	13,04	86,96
9	115	4	13	98	129,58	17	14,78	85,22
10	115	5	14	96	128,35	19	16,52	83,48
11	115	4	13	98	124,42	17	14,78	85,22
12	115	5	16	94	126,11	21	18,26	81,74
13	115	5	15	95	131,17	20	17,39	82,61
14	115	4	12	99	129,36	16	13,91	86,09
15	115	4	11	100	125,86	15	13,04	86,96
PROMEDIO		5	13	97	140,78	18,13	15,77	84,35

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 8-3: Prueba con fuente Times New Roman – tamaño 12 pts.

n	Palabras del texto	No. Palabras detectado	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $\left(\frac{E.Abs \times 100}{P.texto}\right)$	Eficiencia (%) $\left(\frac{P.correctas \times 100}{P.del\ texto}\right)$
1	115	0	9	106	123,85	9	7,83	92,17
2	115	0	9	106	182,81	9	7,83	92,17
3	115	0	9	106	105,25	9	7,83	92,17
4	115	1	9	105	106,21	10	8,70	91,30
5	115	0	9	106	105,23	9	7,83	92,17
6	115	0	9	106	105,35	9	7,83	92,17
7	115	0	9	106	104,05	9	7,83	92,17
8	115	1	9	105	104,39	10	8,70	91,30
9	115	0	9	106	104,53	9	7,83	92,17
10	115	0	9	106	105,35	9	7,83	92,17
11	115	0	9	106	104,52	9	7,83	92,17
12	115	0	9	106	106,01	9	7,83	92,17
13	115	1	9	105	105,18	10	8,70	91,30
14	115	0	9	106	105,29	9	7,83	92,17
15	115	0	9	106	167,66	9	7,83	92,17
PROMEDIO		0	9	106	115,71	9,20	8,00	92,17

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

De acuerdo a los resultados reflejados en las tablas anteriores, en el gráfico 3-3 se muestra la eficiencia con esta fuente con los diferentes tamaños.

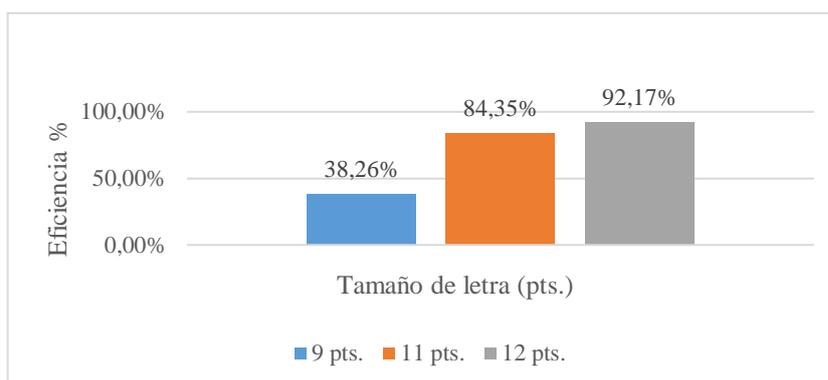


Gráfico 3-3. Eficiencia con fuente Times New Roman y diferente tamaño

Realizado por: Luna, Danilo 2022

3.3.3. Pruebas de texto con fuente Arial

Para estas pruebas se utilizaron tres tamaños diferentes como en el caso anterior, los resultados se encuentran en las tablas 9-3, 10-3 y 11-3. Además, la eficiencia del prototipo con cada tamaño se puede apreciar en el gráfico 4-3.

Tabla 9-3: Prueba con fuente Arial – tamaño 9 pts.

n	Palabras del texto	No. Palabras detectado	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $(\frac{E.Abs*100}{P.texto})$	Eficiencia (%) $(\frac{P.correctas*100}{P.del\ texto})$
1	115	12	24	79	205,47	36	31,30	68,70
2	115	13	25	77	105,77	38	33,04	66,96
3	115	11	24	80	107,16	35	30,43	69,57
4	115	12	24	79	119,01	36	31,30	68,70
5	115	12	25	78	106,21	37	32,17	67,83
6	115	12	25	78	108,14	37	32,17	67,83
7	115	13	24	78	106,05	37	32,17	67,83
8	115	12	24	79	121,22	36	31,30	68,70
9	115	12	24	79	107,53	36	31,30	68,70
10	115	12	24	79	125,23	36	31,30	68,70
11	115	12	24	79	118,12	36	31,30	68,70
12	115	12	24	79	107,17	36	31,30	68,70
13	115	12	24	79	109,34	36	31,30	68,70

14	115	12	24	79	123,19	36	31,30	68,70
15	115	12	24	79	115,18	36	31,30	68,70
PROMEDIO		12	24	79	118,99	36,27	31,54	68,70

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 10-3: Prueba con fuente Arial – tamaño 11 pts.

n	Palabras del texto	No. Palabras detectado	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $(\frac{E.Abs*100}{P.texto})$	Eficiencia (%) $(\frac{P.correctas*100}{P.del\ texto})$
1	115	0	11	104	132,65	11	9,57	90,43
2	115	0	9	106	105,77	9	7,83	92,17
3	115	0	9	106	107,25	9	7,83	92,17
4	115	0	9	106	119,01	9	7,83	92,17
5	115	0	10	105	106,23	10	8,70	91,30
6	115	0	7	108	108,14	7	6,09	93,91
7	115	0	7	108	106,05	7	6,09	93,91
8	115	0	7	108	124,49	7	6,09	93,91
9	115	1	7	107	107,53	8	6,96	93,04
10	115	0	10	105	125,23	10	8,70	91,30
11	115	0	10	105	107,42	10	8,70	91,30
12	115	0	7	108	106,41	7	6,09	93,91
13	115	0	7	108	110,58	7	6,09	93,91
14	115	0	7	108	108,39	7	6,09	93,91
15	115	0	7	108	135,9	7	6,09	93,91
PROMEDIO		0	8	107	114,07	8,33	7,25	93,04

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 11-3: Prueba con fuente Arial – tamaño 12 pts.

n	Palabras del texto	No. Palabras detectado	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $(\frac{E.Abs*100}{P.texto})$	Eficiencia (%) $(\frac{P.correctas*100}{P.del\ texto})$
1	115	0	6	109	121,08	6	5,22	94,78
2	115	1	5	109	123,47	6	5,22	94,78
3	115	0	6	109	212,56	6	5,22	94,78
4	115	0	6	109	124,78	6	5,22	94,78
5	115	1	5	109	125,49	6	5,22	94,78

6	115	0	6	109	116,89	6	5,22	94,78
7	115	0	6	109	121,83	6	5,22	94,78
8	115	0	6	109	124,49	6	5,22	94,78
9	115	1	6	108	124,12	7	6,09	93,91
10	115	0	6	109	125,23	6	5,22	94,78
11	115	0	6	109	121,09	6	5,22	94,78
12	115	0	6	109	123,69	6	5,22	94,78
13	115	0	6	109	128,46	6	5,22	94,78
14	115	0	6	109	128,93	6	5,22	94,78
15	115	1	6	108	135,91	7	6,09	93,91
PROMEDIO		0	6	109	130,53	6,13	5,33	94,78

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

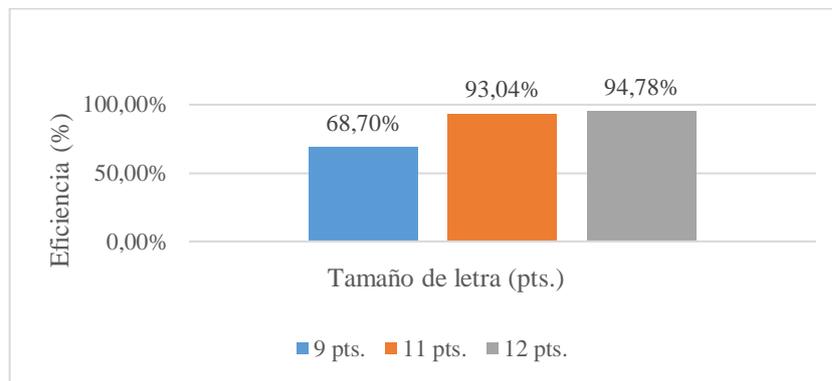


Gráfico 4-3. Eficiencia del dispositivo con fuente Arial con diferente tamaño

Realizado por: Luna, Danilo 2022

3.4. Pruebas con diferente cantidad de palabras

Tomando como referencia los resultados obtenidos en las pruebas de detección de texto, la fuente Calibri 12 pts. con una eficiencia del 95,65%, denota ser la mejor opción para realizar pruebas posteriores. Para ello se procedió a desarrollar más ensayos basándose en el tiempo de procesamiento y en el formato de texto. El tiempo de procesamiento se basa en la cantidad de palabras que hay en el texto. Se hicieron tres pruebas con diferentes cantidades de palabras como se puede observar en la tabla 12-3, donde se evidencia una eficiencia mayor al 95% en todas las pruebas, tomando en cuenta que el tiempo de procesamiento es proporcional a la cantidad de palabras del texto.

Tabla 12-3: Pruebas con diferente cantidad de palabras

n	Palabras del texto	Palabras correctas	Tiempo de procesamiento (s)	Eficiencia (%) $(\frac{P.correctas*100}{P.del\ texto})$
1	69	66	64,44	95,65
2	69	66	102,18	95,65
3	69	66	95,36	95,65
4	69	66	104,31	95,65
5	69	66	85,03	95,65
PROMEDIO		66	90,26	95,65
1	126	122	156,23	96,83
2	126	122	123,47	96,83
3	126	122	121,85	96,83
4	126	122	124,21	96,83
5	126	122	121,66	96,83
PROMEDIO		122	129,48	96,82
1	197	188	195,01	95,43
2	197	188	204,86	95,43
3	197	188	212,56	95,43
4	197	188	196,02	95,43
5	197	187	204,24	94,92
PROMEDIO		188	202,54	95,43

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

El tiempo de procesamiento de acuerdo al número de palabras correctas detectadas se muestra en el gráfico 5-3 en base a los resultados promedio de cada prueba.

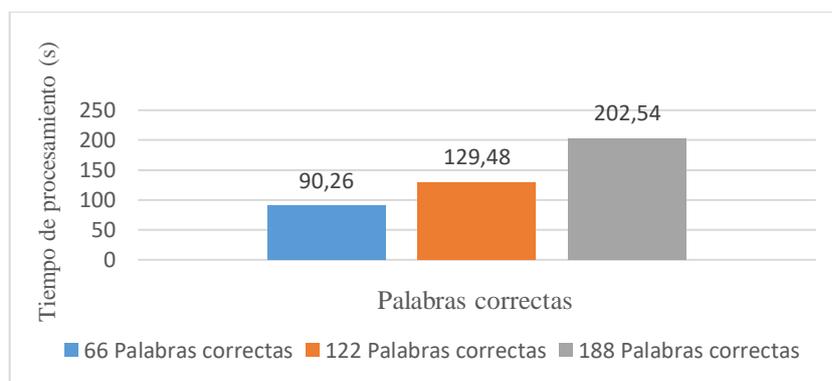


Gráfico 5-3. Tiempo de procesamiento de acuerdo a la cantidad de palabras

Realizado por: Luna, Danilo 2022

3.5. Pruebas con diferente formato de texto

En estas pruebas se emplearon los formatos de texto en negrilla y cursiva con fuente Calibri 12 pts., para determinar cuál de ellos presenta una mejor eficiencia.

En las tablas 13-3 y 14-3 se observan los resultados obtenidos al ejecutar el sistema con estos formatos, pero principalmente la eficiencia y el tiempo de procesamiento del dispositivo. Presentando una eficiencia superior al 90% en todos los casos y un tiempo de procesamiento de 150 segundos aproximadamente.

Tabla 13-3: Pruebas con texto en Negrilla

n	Palabras del texto	Palabras no detectadas	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $\left(\frac{E.Abs*100}{P.texto}\right)$	Eficiencia (%) $\left(\frac{P.correctas*100}{P.del\ texto}\right)$
1	197	0	10	187	154,06	10	5,08	94,92
2	197	0	10	187	167,33	10	5,08	94,92
3	197	0	10	187	148,36	10	5,08	94,92
4	197	0	10	187	142,13	10	5,08	94,92
5	197	1	10	186	158,24	11	5,58	94,42
PROMEDIO		0	10	187	154,02	10,20	5,18	94,92

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 14-3: Pruebas con texto en Cursiva

n	Palabras del texto	Palabras no detectadas	Palabras Incorrectas	Palabras Correctas	Tiempo de procesa. (s)	Error absoluto (P. text – P. corr)	Error relativo $\left(\frac{E.Abs*100}{P.texto}\right)$	Eficiencia (%) $\left(\frac{P.correctas*100}{P.del\ texto}\right)$
1	197	1	12	184	152,26	13	6,60	93,40
2	197	1	12	184	123,47	13	6,60	93,40
3	197	1	12	184	212,56	13	6,60	93,40
4	197	1	12	184	124,78	13	6,60	93,40
5	197	1	12	184	125,49	13	6,60	93,40
PROMEDIO		1	12	184	147,71	13	6,60	93,40

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

3.6. Análisis de resultados

Con las pruebas realizadas se tiene que la iluminación debe ser mayor o igual a 180 lúmenes para que los caracteres se reconozcan de manera correcta, la posición de la cámara debe estar a una distancia de 25 ± 5 cm de altura aproximadamente con respecto a la hoja. Para una eficiencia superior al 90% el tamaño de la letra debe ser mayor o igual 11 pts. en cualquier tipo de letra ya antes escogido, siendo Calibri, la fuente con mejor eficiencia superando el 95%. Al tener un tipo de formato ya sea negrita o cursiva la eficiencia baja 2%, es así que, teniendo un texto en letra Calibri con tamaño de 12 pts., en tipo negrita del 94% de eficiencia, y con tipo cursiva del 93.4%.

CAPITULO IV

4. EVALUACIÓN ECONÓMICA

En este capítulo se realizará el análisis de costos del prototipo, tomando en cuenta todos los aspectos que influyen en el costo total del mismo. En la tabla 1-4 se detalla los costos de los materiales y herramientas. Los valores correspondientes a costos de equipos y costos de implementación se observan en la tabla 2-4 y 3-4 respectivamente, finalmente en la tabla 4-4 se evidencia el costo total de construcción del prototipo.

Tabla 1-4: Costos de materiales y herramientas

Cantidad	Detalle	Valor unitario	Total
4	Resistencia	\$0,07	\$0,28
2	Diodo led	\$0,15	\$0,30
1	Modulo Relé	\$4,00	\$4,00
1	Fuente 5v - 3A	\$9,00	\$9,00
1	Fuente 12v - 2A	\$8,00	\$8,00
1	Display 16*2	\$8,00	\$8,00
1	Pulsador	\$1,25	\$1,25
1	Conector hembra para 110v	\$5,00	\$5,00
1	Sensor final de carrera	\$0,50	\$0,50
1	Cable de poder	\$2,50	\$2,50
1	Interruptor	\$0,75	\$0,75
7	Bornera	\$0,35	\$2,45
1	PCB 20*10 cm	\$2,55	\$2,55
1	Flex de 15 pines - 50 cm	\$6,50	\$6,50
1	Cable USB	\$2,00	\$2,00
1	Interruptor	\$3,00	\$3,00
1	Cable de timbre	\$3,00	\$3,00
1	Estaño	\$0,50	\$0,50
1	Cable UTP	\$1,50	\$1,50
1	Conector de red hembra	\$0,40	\$0,40
1	Conector de red macho	\$0,50	\$0,50
1	Cautín	\$0,35	\$2,45
1	Multímetro	\$18,00	\$18,00
1	Corte láser	\$40,00	\$40,00
1	Foco led 3watts	\$4,00	\$4,00
2	Riel metálico	\$5,00	\$10,00
30	Impresiones	\$0,05	\$1,50
1	Bandeja MDF de 3mm	\$5,00	\$5,00
12	Tornillo 1/2 * 1/2 pulgada	\$0,25	\$3,00
4	Tornillo 1/2 * 1/8 pulgada	\$0,15	\$0,60
4	Tornillo 1/2 * 2 pulgadas	\$0,35	\$1,40
6	Pegamento para acrílico	\$0,50	\$3,00
TOTAL			\$150,93

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 2-4: Costos de equipos

Cantidad	Detalle	Valor unitario	Total
1	Raspberry Pi 3b	\$80,00	\$80,00
1	Raspberry Pi camera v2	\$52,00	\$52,00
1	Micro SD 16 Gb	\$11,00	\$11,00
1	Arduino nano	\$12,00	\$12,00
1	Teclado	\$10,00	\$10,00
1	Altavoces	\$15,00	\$10,00
1	Router	\$20,00	\$20,00
TOTAL			\$195,00

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 3-4: Costos de implementación

Cantidad	Detalle	Tiempo (Horas)	Valor unitario	Total
1	Desarrollo de software	120	\$3,50	\$420,00
1	Implementación	24	\$2,50	\$60,00
1	Construcción	8	\$2,50	\$20,00
TOTAL				\$500,00

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Tabla 4-4: Costo total de construcción del prototipo

Descripción	Total
Costos de materiales y herramientas	\$150,93
Costos de equipos	\$195,00
Costos de implementación	\$500,00
TOTAL	\$845,93

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

4.1. Relación beneficio/costo

Tabla 5-4: Relación beneficio/costo.

Dispositivo	Costo
PROTOTIPO	\$845,93
PROCER 2	\$1100,00
Beneficio/Costo	\$254,07

Fuente: Luna Danilo, 2022

Realizado por: Luna, Danilo 2022

Como se observa en la tabla 5-4 el prototipo tiene un costo de \$845,93 tomando en cuenta un 95% de eficiencia y al realizar una comparación beneficio/costo con el dispositivo PROCER 2 fabricado en la Argentina.

Dicho dispositivo de lectura auditiva convierte texto impreso y/o digital en audio, ya que cuenta con un escáner portátil que captura la imagen y por medio de un mini computador que se encuentra previamente conectado realiza el procesamiento de la imagen y el texto detectado lo convierte en señales sonoras para el usuario. Actualmente tiene un precio de comercialización de 95000 pesos argentinos, que en dólares americanos representan alrededor de \$1100,00 incluidos costos de importación, representa un 24,18% de diferencia entre ellos evidenciando de esta manera un ahorro sustancial.

CONCLUSIONES

Se desarrolló un prototipo para el reconocimiento de texto tipo imprenta con el fin de convertirlo a voz para que las personas no videntes tengan acceso a la información, se lo hizo mediante algoritmos de visión artificial, por lo cual se determinó que el más adecuado es el OCR Tesseract, debido a que reconoce el texto de la imagen mediante patrones de caracteres, para luego extraer el texto de forma digital.

Se estableció que para este prototipo es necesario el uso de un computador de placa reducida Raspberry pi 3B+, debido a las prestaciones de hardware y software que posee, ya que el microprocesador permite utilizar software libre multiplataforma, así como entornos de programación y librerías requeridas para el fin. Por otra parte, en relación al hardware se pudo determinar que este microprocesador cuenta con mejor capacidad de procesamiento tomando en cuenta la relación beneficio costo comparado con otras opciones disponibles en el medio.

Se implementó el prototipo de reconocimiento de texto utilizando los materiales y equipos establecidos previamente y que se encontraron en el mercado local, ubicando de forma táctica los componentes en la estructura diseñada para obtener los resultados deseados, así como la iluminación para que la cámara pueda captar de manera óptima todos los caracteres presentados en el texto, además, la instalación de los indicadores para que las personas del entorno al usuario no vidente, constaten el proceso que está realizando el prototipo.

Luego de realizar todas las pruebas correspondientes respecto a: iluminación, distancia de captura de imagen, tipo de fuente de texto (Calibri, Arial y Times New Roman), tamaño de texto (9,11,12 pts.), formato (normal, negrilla y cursiva), y cantidad de palabras, se puede decir que el prototipo funciona con una eficiencia superior al 95% cuando se cumplen los siguientes parámetros de entorno controlados: Iluminación igual o superior a los 180 lúmenes, distancia de 25 ± 5 cm, tipo de fuente Calibri con un tamaño de 12 pts., con formato normal con un texto de 150 ± 50 palabras.

RECOMENDACIONES

En vista que el tipo de letra Times New Roman presenta un nivel de eficiencia inferior al resto de fuentes, por sus rasgos complejos, se recomienda realizar un aprendizaje automático denominado entrenador, haciendo que el algoritmo pueda aprender por sí solo, es decir automáticamente, sin que intervenga el ser humano.

Desarrollar un algoritmo de reconocimiento de texto capaz de operar correctamente en presencia de caracteres especiales y propios del idioma español latino, en este último caso donde se presentan vocales con tildes, así como, eñes, y doble eles.

En caso de utilizar el mismo algoritmo para el sintetizador de voz, si existen los recursos económicos, se recomienda adquirir la versión pagada del mismo, con el fin de tener acceso a cualquier idioma.

Ajustar el tamaño del prototipo de acuerdo a las condiciones necesarias para su trabajo, optimizando materiales y espacio, con el fin de reducir costos y haciéndolo portátil para facilidad y comodidad del usuario.

Se recomienda evaluar diferentes opciones de software y hardware que permitan reducir el tiempo a la hora de procesar el algoritmo de reconocimiento óptico de caracteres, teniendo en cuenta que la computadora de placa reducida al tener más rendimiento y un mejor procesador, realizaría más rápido los procesos.

Al ser los algoritmos Tesseract (el conversor de imagen a texto) y GTTS (convertidor de texto a voz) en línea, donde requieren de internet para su uso, se recomienda desarrollar o buscar nuevas alternativas que no requieran conexión a internet, para que el prototipo no sea un limitante a la hora de usarlo.

BIBLIOGRAFIA

ALEXANDER, CHARLES K. Y SADIKU, MATTHEW N. O. *Fundamentos de circuitos eléctricos*. Tercera. México : McGRAW-HILL, 2006. 970-10-5606-X.

Arranz Ortega, Daniel. *Marco Para El Desarrollo De Aplicaciones Ada Sobre Microcontroladores Stm32*. Santander : Universidad De Cantabria, 2019, p. 28.

Asociación D.O.C.E. KNFB READER, UN ESCÁNER QUE LEE EL TEXTO IMPRESO EN VOZ ALTA | DISCAPACIDAD VISUAL D.O.C.E. (DISCAPACITADOS OTROS CIEGOS DE ESPAÑA). *KNFB READER, UN ESCÁNER QUE LEE EL TEXTO IMPRESO EN VOZ ALTA*. [En línea] 03 de Noviembre de 2015. <https://asociaciondoce.com/2015/11/03/knfb-reader-un-escaner-que-lee-el-texto-impreso-en-voz-alta/>.

Baliga, B. Jayant. *Fundamentals of Power Semiconductors Devices*. New York : Springer Science+Business Media, 2008. 978-0-387-47313-0.

Balogh, Laszlo. *A Practical Introduction to Digital Power Supply Control*. Texas : Texas Instruments Incorporated, 2005.

Barreno, Franklin. *Diseño de Prototipo Doméstico de Video Vigilancia con Cámaras IP por internet*. Quito : UNIVERSIDAD SAN FRANCISCO DE QUITO, 2013, p. 19.

Barrón Ruiz, Mariano. *USO DIDÁCTICO DEL SOFTWARE DE AYUDA AL DISEÑO ELECTRÓNICO "PROTEUS"*. País Vasco : Universidad del País Vasco, 2004, p. 28.

Bermúdez Cabra, Antonio Manuel, Muñoz Sevilla, José Antonio y Rodríguez Hernandez, Antonio. *Tecnologías específicas para personas ciegas y deficientes visuales*. 2012, p. 17.

Boylestad, Robert y Nashelsky, Louis. *Electrónica Teoría de Circuitos y Dispositivos Electrónicos*. México : Prentice Hall, 2009.

Brown, Marty. *Practical Switching Power Supply Design*. The United States Of America : Academic Press, 1990. 0-12-137030-5.

Brunete, Alberto, San Segundo, Pablo y Herrero, Rebeca. *Introducción a la Automatización Industrial*. Madrid - España : Universidad Politécnica de Madrid, 2020, p. 25. ISBN: 978-84-09-22291-9.

Canacuan Ipiiales, Cristian German. *DISEÑO DE UN SISTEMA DE VIDEO VIGILANCIA IP Y ALARMA BASADA EN MOVIMIENTO, UTILIZANDO SOFTWARE LIBRE SOBRE UN*

COMPUTADOR DE PLACA REDUCIDA, PARA LA EMPRESA COLOR 2000 DE LA CIUDAD DE IBARRA. Ibarra - Ecuador : Universidad Técnica del Norte, 2018.

Castillo, Diana. Definición de sistema operativo by diana castillo - issuu. *Definición de sistema operativo.* [En línea] 26 de Octubre de 2015, p. 12. https://issuu.com/dianacastillo07/docs/definici__n_de_sistema_operativo.

Chinchero Iza, Jorge Luis. *DESARROLLO DE UN DISPOSITIVO QUE MEDIANTE VISIÓN ARTIFICIAL PERMITA ADQUIRIR IMÁGENES CON PALABRAS PARA LA CONVERSIÓN A AUDIO, ORIENTADO A LA AYUDA DE PERSONAS INVIDENTES.* QUITO - ECUADOR : UNIVERSIDAD TECNOLÓGICA ISRAEL, 2019, pp. 14-28 .

CONADIS. Estadísticas de Discapacidad – Consejo Nacional para la Igualdad de Discapacidades. *Estadísticas de Discapacidad.* [En línea] 2020. <https://www.consejodiscapacidades.gob.ec/estadisticas-de-discapacidad/>.

Concepción, Ricardo. Tarjetas de Desarrollo (Episodio #8) - rjconcepcion. *Tarjetas de Desarrollo (Episodio #8).* [En línea] 19 de Diciembre de 2019. <https://www.rjconcepcion.com/podcast/tarjetas-de-desarrollo-episodio-8/#:~:text=Una%20tarjeta%20de%20desarrollo%20es,instrucciones%20de%20un%20programa%20suministrado..>

Corporation, Measurement Computing. *Data Acquisition Hand Book.* Estados Unidos : s.n., 2012.

Coughlin, Robert F. y Driscoll, Frederick F. *Amplificadores operacionales y circuitos integrados lineales.* México : Prentice-Hall Inc, 2000. 968-880-284-0.

Fairchild, Industries. *Reliability Stress Test Descriptions.* 2003.

Granda Miguel, Mercedes y Mediavilla Bolado, Elena. *Intrumentacion electrónica: Transductores y Acondicionadores de señal.* Cantabria : s.n., 2010, p. 14.

Gutierrez, Richard, Frydson, María Fernanda y Vintimilla, Boris. *Aplicación de visión por computador para el reconocimiento automático de placas vehiculares utilizando ocr's convencionales.* Guayaquil : ESPOL, 2011, pp. 22-23.

Guzmán Arreola, Martín Alejandro. *Sintetizador de voz para la enseñanza de la lectura a niños mexicanos.* Cholula, Puebla, México : Universidad de Las Americas Puebla, 2004, p. 17.

Hart, Daniel W. *Electrónica de Potencia.* Madrid : Prentice Hall , 2001. 84-205-3179-0.

Hernandez, Jorge Luis. *New Methodologies and Instrumentations for Power.* Rende : s.n., 2015.

Hernandez, José Román. *10 sintetizadores de voz TTS para Android para cambiar las voces de nuestro móvil y leer textos, ebooks o notificaciones.* España, 2015.

Lutz, Josef. *Semiconductor Power Devices, Physics, Characteristics, Reliability .* Berlin : Springer, 2011. 978-3-642-11125-9.

Manosalvas Salazar, Carlos Arturo. *DISEÑO DE UN SISTEMA EMBEBIDO PARA EL CONTROL DE INGRESO Y SALIDA DE VEHÍCULOS A TRAVÉS DE INTERNET, EN EL ACCESO PRINCIPAL DE LA ESPOCH.* Riobamba : ESPOCH, 2017, p. 15.

Merino, María y Pérez Porto, Julián. *Definición de cámara web - Qué es, Significado y Concepto. DEFINICIÓN DE CÁMARA WEB. [En línea] 2016, p. 27.* <https://definicion.de/camara-web/>.

Mohan, Ned, Underland, Tore y Robbins, Willam. *ELECTRÓNICA DE POTENCIA: Convertidores, aplicaciones y diseño.* México : MCGRAW-HILL, 2009. 978-0-471-22693-.

Nogué, Albert y Antiga, Jordi. *Aplicación práctica de la visión artificial en el control de procesos industriales.* España : Gobierno de España (Ministerio de Educación), 2012, pp. 10-12.

Ochoa Duarte, Alexei. *Tabla comparativa entre modelos de Frambuesa Fuente: Tomado de [11] / Descargar diagrama científico, 2018.*

Ogata, Katsuhiko. *Ingeniería de control moderna.* Madrid : PEARSON EDUCACIÓN, 2010. 978-84-8322-660-5.

ORGANIZACION MUNDIAL DE LA SALUD. *Ceguera y discapacidad visual. Ceguera y discapacidad visual.* [En línea] 11 de Octubre de 2018. <https://www.who.int/es/news-room/fact-sheets/detail/blindness-and-visual-impairment>.

Organización Mundial de la Salud. *Ceguera y discapacidad visual. Ceguera y discapacidad visual.* [En línea] 11 de Octubre de 2018. <https://www.who.int/es/news-room/fact-sheets/detail/blindness-and-visual-impairment>.

Organización Nacional de Ciegos Españoles. *Discapacidad visual y autonomía personal.* Madrid : Departamento de Promoción Cultural y Braille, 2011.

Pérez Porto, Julián y Merino, María. *Definición de cámara digital - Qué es, Significado y Concepto.* *DEFINICIÓN DE CÁMARA DIGITAL.* [En línea] 2016, p. 22. <https://definicion.de/camara-digital/>.

PROCER TECNOLOGIA INCLUSIVA. *PROCER Tecnología inclusiva | PROCER 2. PROCER 2.* [En línea] 2021. <https://www.procertecnologias.com/products/4#:~:text=Gracias%20al%20esc%C3%A1ner%20port%C3%A1til%20adaptado,conectando%20un%20teclado%20al%20dispositivo..>

Rashid, Muhammad H. *Power Electronics: Circuits, Devices and Applications.* England : Pearson Education Limited, 2014. 978-0-273-76908-8.

Rossano, Victor. *Proteus VSM.* Buenos Aires : USERSHOP, 2013, p. 18.

Rouhiainen, Lasse. *Inteligencia Artificial.* Barcelona : Planeta S.A, 2018, p. 5.

Salcedo Tovar, Marco Luis. *Minicomputador educacional de bajo costo Raspberry Pi: Primera parte.* Miranda - Venezuela : Revista Ethos Venezolana, 2015, p. 14. ISSN 1856-9862.

Sánchez Fernández, Carlos Javier y Sandonís Consuegra, Victor. *Reconocimiento Óptico de Caracteres (OCR).* Madrid : Universidad Carlo, 2008, pp. 11-16.

Sierra, Antonio. *Lenguajes de programación.* Guayaquil : Escuela Superior Politécnica del Litoral, 2009, p. 26.

Solano Aguilar, Gabriela Del Cisne. *DISEÑO Y CONSTRUCCIÓN DE UN DISPOSITIVO PORTATIL PARA LA IDENTIFICACIÓN DE BILLETES ORIENTADO A PERSONAS CON DISCAPACIDAD VISUAL MEDIANTE EL USO DE VISIÓN ARTIFICIAL.* Riobamba : ESPOCH, 2018, pp. 17-43.

Terven Salinas, Juan Ramon, Salas, Joaquín y Raducanu, Bogdan. *Estado del Arte en Sistemas de Visión Artificial para Personas Invidentes.* Mexico : Sociedad Mexicana de Inteligencia Artificial , 2013, pp. 32-36.

The Henry Ford. *Xerox "Kurzweil Reading Edge Optical Recognition Speech Synthesizer," 1992 - The Henry Ford. Xerox "Kurzweil Reading Edge Optical Recognition Speech Synthesizer," 1992.* [En línea] 19 de Octubre de 2019. <https://www.thehenryford.org/collections-and-research/digital-collections/artifact/476405/#slide=gs-464333>.

Toulson, Rob y Wilmshurst, Tim. *Fast and Effective Embedded Systems Design.* Oxford : Elsevier, 2017. 978-0-08-100880-5.

Vélez Caicedo, Juan David. *Diseño y desarrollo de dispositivo de lectura de textos para personas en situación de discapacidad visual, basado en redes neuronales profundas.* Santiago de Cali : Universidad Autónoma de Occidente, 2019.

Vinueza Escobar, Cristhian Patricio. *DISEÑO Y CONSTRUCCIÓN DE UNA IMPRESORA BRAILLE DE BAJO COSTO, UTILIZANDO PLATAFORMA HARDWARE DE CÓDIGO ABIERTO.* Riobamba - Ecuador : ESPOCH, 2017, p. 17.

Xiao, Perry. *Designing Embedded Systems and the Internet of Things (IoT) with the ARM® Mbed™.* Ucrania : John Wiley & Sons, 2018. ISBN 9781119363996.

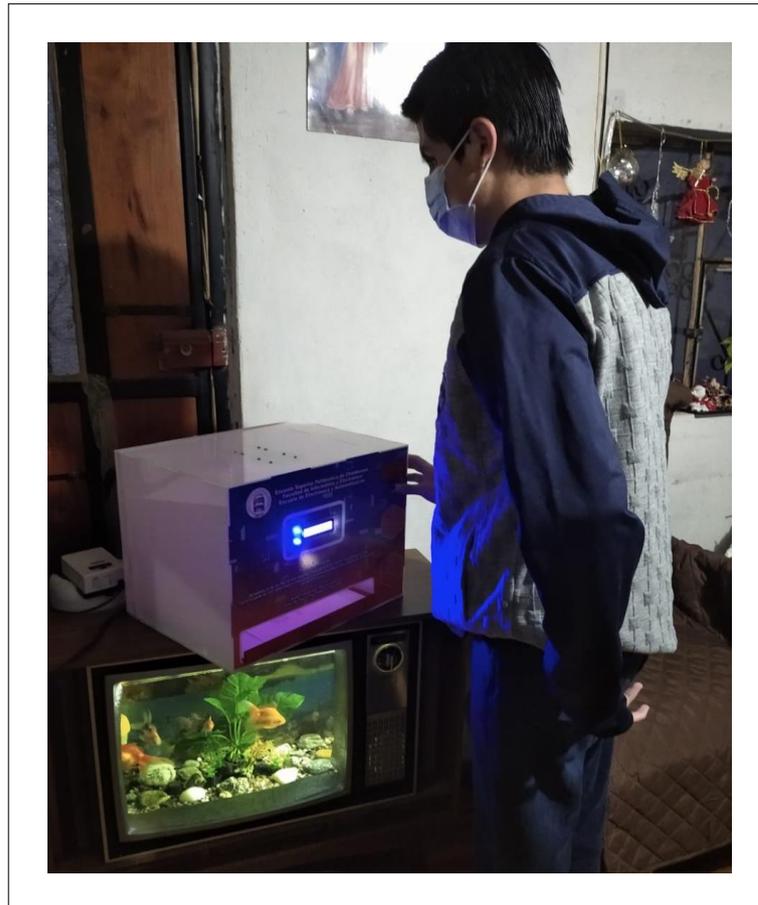
Zambrano Sonroza, Johanna María, Toala, Ramon Alfredo y Bolívar Chávez, Oscar Elías. *Inclusión educativa virtual en estudiantes con discapacidades visuales. 2017. FFLCE. INCLUSION EDUCATIVA VIRTUAL EN ESTUDIANTES CON DISCAPACIDADES VISUALES.pdf.* [En línea] Marzo de 2017.
<https://www.utm.edu.ec/investigacion/phocadownload/publicaciones/Publicaciones-Regionales/2017/FLCE/2017.%20FFLCE.%20INCLUSION%20EDUCATIVA%20VIRTUAL%20EN%20ESTUDIANTES%20CON%20DISCAPACIDADES%20VISUALES.pdf>.

ANEXOS

Anexo A: Prototipo finalizado y operando



Anexo B: Persona no vidente iniciando el sistema

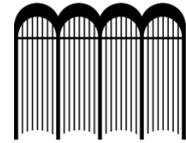


Anexo C: Persona no vidente insertando la bandeja con el documento de texto





ESCUELA SUPERIOR POLITÉCNICA DE
CHIMBORAZO



DIRECCIÓN DE BIBLIOTECAS Y RECURSOS DEL
APRENDIZAJE

DBRA

UNIDAD DE PROCESOS TÉCNICOS

REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y BIBLIOGRAFÍA

Fecha de entrega: 26 / 01 / 2022

INFORMACIÓN DEL AUTOR/A (S)
Nombres – Apellidos: DANILO ORLANDO LUNA SORIA
INFORMACIÓN INSTITUCIONAL
Facultad: INFORMÁTICA Y ELECTRÓNICA
Carrera: ELECTRÓNICA Y AUTOMATIZACIÓN
Título a optar: INGENIERO EN ELECTRÓNICA Y AUTOMATIZACIÓN
f. Analista de Biblioteca responsable: Lcdo. Holger Ramos, MSc.



Firmado electrónicamente por:
**HOLGER GERMAN
RAMOS UVIDIA**

1664-DBRA-UPT-2021