



**Diseño e implementación de gafas inteligentes como herramienta de asistencia para  
lectura y reconocimiento del entorno para personas con discapacidad visual utilizando  
inteligencia artificial**

Arias Acosta, Marco David

Departamento de Ciencias de la Energía y Mecánica

Carrera de Ingeniería en Mecatrónica

Trabajo de titulación, previo a la obtención del título de Ingeniero en Mecatrónica

Msc. Mendoza Chipantasi, Dario José











10 de febrero 2022

Latacunga

### Document Information

Analyzed document	ARIAS-DAVID-TESIS-IA.pdf (D158552214)
Submitted	2/13/2023 8:01:00 PM
Submitted by	
Submitter email	byron.corrales@utc.edu.ec
Similarity	5%
Analysis address	byron.corrales.utc@analysis.orkund.com

### Sources included in the report

<b>SA</b>	<b>orkund_ortega_granda.pdf</b> Document orkund_ortega_granda.pdf (D26009979)	 2
<b>SA</b>	<b>TESIS BORRADOR.docx</b> Document TESIS BORRADOR.docx (D54291637)	 2
<b>W</b>	URL: <a href="http://repositorio.espe.edu.ec/handle/21000/13446">http://repositorio.espe.edu.ec/handle/21000/13446</a> Fetched: 2/13/2023 8:03:00 PM	 1
<b>W</b>	URL: <a href="https://bibdigital.epn.edu.ec/handle/15000/21431">https://bibdigital.epn.edu.ec/handle/15000/21431</a> Fetched: 2/13/2023 8:02:00 PM	 2
<b>W</b>	URL: <a href="http://repositorio.espe.edu.ec/handle/21000/11886">http://repositorio.espe.edu.ec/handle/21000/11886</a> Fetched: 2/13/2023 8:02:00 PM	 3
<b>W</b>	URL: <a href="https://www.orcam.com/es/">https://www.orcam.com/es/</a> Fetched: 2/13/2023 8:03:00 PM	 1
<b>W</b>	URL: <a href="https://pq5.pe/actualidad/tecnologia/finger-reader-el-anillo-que-facilitara-la-lectura-persona...">https://pq5.pe/actualidad/tecnologia/finger-reader-el-anillo-que-facilitara-la-lectura-persona...</a> Fetched: 2/13/2023 8:03:00 PM	 2
<b>W</b>	URL: <a href="https://www.oftalvist.es/blog/agudeza-visual-que-es-y-pruebas/">https://www.oftalvist.es/blog/agudeza-visual-que-es-y-pruebas/</a> Fetched: 2/13/2023 8:02:00 PM	 2
<b>W</b>	URL: <a href="https://www.cognifit.com/ec/ciencia/habilidad-cognitiva/campo-visual">https://www.cognifit.com/ec/ciencia/habilidad-cognitiva/campo-visual</a> Fetched: 2/13/2023 8:02:00 PM	 1
<b>W</b>	URL: <a href="https://www.martinezdecameros.com/motilidad-">https://www.martinezdecameros.com/motilidad-</a> Fetched: 2/13/2023 8:02:00 PM	 2

Firma:



Msc. Mendoza Chipantasi Darío José  
C.C.:0603110834



Departamento de Ciencias de la Energía y Mecánica

Carrera de ingeniería en Mecatrónica

### Certificación

Certifico que el trabajo de titulación: **"Diseño e implementación de gafas inteligentes como herramienta de asistencia para lectura y reconocimiento del entorno para personas con discapacidad visual utilizando inteligencia artificial"** fue realizado por el señor **Arias Acosta, Marco David**; el mismo que cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, además fue revisado y analizado en su totalidad por la herramienta de prevención y/o verificación de similitud de contenidos; razón por la cual me permito acreditar y autorizar para que se lo sustente públicamente.

Latacunga, 08 de febrero del 2022

Ing. Mendoza Chipantasi, Darío José

C.C.:0603110834



**Departamento de Ciencias de la Energía y Mecánica**  
**Carrera de Ingeniería en Mecatrónica**

**Responsabilidad de Autoría**

Yo, **Arias Acosta, Marco David**, con cédula de ciudadanía n°0504182650, declaro que el contenido, ideas y criterios del trabajo de titulación: **Diseño e implementación de gafas inteligentes como herramienta de asistencia para lectura y reconocimiento del entorno para personas con discapacidad visual utilizando inteligencia artificial** es de mi autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos, y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

**Latacunga, 08 de febrero del 2022**

**Arias Acosta, Marco David**

C.C.:0504182650



Departamento de Ciencias de la Energía y Mecánica

Carrera de Ingeniería en Mecatrónica

### Autorización de Publicación

Yo **Arias Acosta, Marco David**, con cédula de ciudadanía n°0504182650, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **Diseño e implementación de gafas inteligentes como herramienta de asistencia para lectura y reconocimiento del entorno para personas con discapacidad visual utilizando inteligencia artificial** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi/nuestra responsabilidad.

Latacunga, 08 de febrero del 2022

**Arias Acosta, Marco David**

C.C.:0504182650

### **Dedicatoria**

*El presente trabajo de titulación es dedicado con mucho amor a mis padres Marco y Alicia que con su cariño, amor en todos estos años y apoyo incondicional me han permitido cumplir un sueño más, por estar conmigo en todo momento y ser un pilar fundamental en mi vida que me han enseñado que en esta vida uno tiene que luchar por lo que se quiere, a mis hermanos Xavier y Andrea por siempre estar pendiente de mí, brindándome su cariño y aprecio, a mi sobrino Sebastián Arias, que con su existencia y amor nos alegra y nos motiva todos los días.*

*También quiero dedicar este logro importante a Santa Anita y a mi Dios por guiarme a lo largo de este proceso, ser el apoyo y fortaleza en momentos de dificultad y debilidad.*

**David Arias**

## **Agradecimiento**

*Quiero agradecer y expresar mi gratitud a Dios y Santa Anita, por todas sus bendiciones, quien me ha permitido a lo largo de mi carrera guiarme y darme las fuerzas para poder concluir con esta etapa, por permitirme alcanzar este gran sueño que tuve y hoy lo puedo ver realizado, a mis padres por ser el motor que impulsa cada uno de mis pasos, quienes estuvieron siempre a mi lado en todos los momentos y ser los mejores guías de mi vida, a mi hermano y hermana por siempre estar al pendiente en todo momento apoyándome de diferentes maneras.*

*De igual manera mi agradecimiento a mi tutor de tesis Ing. Diario Mendoza, por su colaboración y guía, por compartir sus conocimientos durante el desarrollo del prototipo.*

*Finalmente quiero expresar mi sincero agradecimiento al personal Médico del Patronato Municipal de Latacunga, por su colaboración durante el desarrollo de este proyecto, mediante su dirección y predisposición brindada.*

**David Arias**

## ÍNDICE DE CONTENIDOS

Carátula.....	1
Reporte de verificación de contenido .....	2
Certificación .....	3
Responsabilidad de autoría .....	4
Autorización de publicación .....	5
Dedicatoria.....	6
Agradecimiento .....	7
Índice de contenidos .....	8
Índice de tablas.....	13
Índice de contenidos.....	15
Resumen .....	18
Abstract.....	19
Capítulo I: Fundamento Teórico.....	20
Planteamiento del problema .....	20
Antecedentes.....	21
Patronato Municipal Gad Latacunga: Población de personas con discapacidad visual .....	25
Justificación .....	27
Objetivos .....	29
<i>Objetivo General</i> .....	29
<i>Objetivos Específicos</i> .....	29
Hipótesis.....	30
Variables de la investigación.....	30
<i>Variables Independientes</i> .....	30
<i>Variables Dependientes</i> .....	30
Deficiencia Visual .....	30



<i>Agudeza visual</i> .....	31
<i>Campo Visual</i> .....	32
<i>Motilidad ocular</i> .....	33
Clasificación de la discapacidad visual.....	34
Discapacidad visual Cotopaxi .....	35
Inteligencia Artificial .....	36
<i>Etapas del sistema aplicados en dispositivo de las gafas inteligentes.</i> .....	36
Reconocimiento óptico de caracteres OCR basado en inteligencia artificial .....	38
API de Cloud Vision .....	39
MobileNet V3.....	41
gTTS (Google Text-to-Speech) .....	43
Capítulo II: Selección y diseño de componentes .....	46
Metodología .....	46
Alternativas y selección de componentes.....	49
Selección del sistema embebido .....	50
<i>Opción 1. Raspberry Pi4 Model B</i> .....	50
<i>Opción 2. Jetson Nano</i> .....	52
<i>Opción 3. LattePanda Delta</i> .....	52
Selección del módulo de reconocimiento de voz .....	55
<i>Opción 1. Voice Recognition module V3</i> .....	55
<i>Opción 2. EasyVR3 Plus Shield Recognition</i> .....	56
<i>Opción 3. Módulo de reconocimiento de voz de Grove</i> .....	57
Arduino Nano.....	58
Selección de la pantalla .....	58
<i>Opción 1. Pantalla Nextion 5.0"</i> .....	58
<i>Opción 2. Pantalla Waveshare 5.0"</i> .....	59

<i>Opción 3. Pantalla Elecrow LCD 5.0"</i> .....	61
Selección de la cámara digital .....	62
<i>Opción 1. Módulo de cámara Raspberry Pi 2</i> .....	62
<i>Opción 2. Módulo de cámara nano IMX219-160</i> .....	63
<i>Opción 3. Mini cámara UVC Bullet</i> .....	64
Selección de fuente externa power bank.....	66
<i>Opción 1. Ldnio PR518 Power Bank</i> .....	66
<i>Opción 2. Enox 5000 KBH-I55GD</i> .....	67
<i>Opción 3. KSIX Power Bank</i> .....	68
Selección de parlantes .....	69
<i>Opción 1. Parlante Genius SP-Q160</i> .....	69
<i>Opción 2. Mini Parlante Bluetooth</i> .....	70
<i>Opción 3. Parlante Pc Mod G-101</i> .....	71
Pulsador: Pusbotton Tact & Cap.....	72
Selección del material para la impresión de la carcasa 3D .....	73
Capítulo III: Diseño e implementación de los sistemas.....	75
Diseño e implementación del reconocimiento de voz .....	75
<i>Conexión del subsistema cámara-micrófono</i> .....	78
Implementación de la pantalla .....	80
Implementación del sistema de audio.....	80
Conexión de los pulsadores en los puertos GPIO .....	81
<i>Conexión del subsistema de carga power bank</i> .....	82
<i>Conexión del sistema de expansión USB externa</i> .....	83
Implementación del software para el reconocimiento de texto y del entorno .....	84
Librerías empleadas en Python .....	84
Implementación del sistema de las gafas inteligentes.....	85

Diagrama de flujo de la estructura del programa .....	86
Creación de audios para la detección de objetos. ....	88
Diseño e implementación del soporte para las gafas inteligentes.....	89
Diseño e implementación del case para las gafas inteligentes.....	90
Resolución de la cámara UVC Bullet .....	92
Con los valores obtenidos del campo visual para el lado más largo de.....	93
Fuerza de soporte de las gafas inteligentes .....	93
<i>Cálculo de la fuerza máxima</i> .....	94
<i>Análisis de elementos finitos</i> .....	96
<i>Resultado de esfuerzo a tensión Von Mises</i> .....	97
<i>Comportamiento de caída libre del prototipo a una altura de 1.70m</i> .....	100
Cálculo del clip sujetador .....	102
<i>Análisis de fuerzas</i> .....	102
Escenas y ventanas de los modos de funcionamiento del dispositivo. ....	103
Creación del inicio .....	104
Capítulo IV: Pruebas y validación del dispositivo .....	107
Pruebas y análisis del funcionamiento del dispositivo .....	107
<i>Pruebas del modo de lectura</i> .....	107
Validación de la hipótesis .....	118
<i>Pruebas del modo reconocimiento del entorno</i> .....	123
Validación del dispositivo mediante la supervisión de un especialista .....	133
Autonomía del dispositivo .....	136
Capítulo V: Conclusiones y Recomendaciones .....	140
Conclusiones.....	140
Recomendaciones .....	143
Bibliografía .....	146

<b>Anexos .....</b>	<b>152</b>
---------------------	------------

## ÍNDICE DE TABLAS

<b>Tabla 1</b> <i>Porcentaje de acuerdo a la discapacidad visual</i> .....	35
<b>Tabla 2</b> <i>Necesidades de personas que presentan discapacidad visual</i> .....	46
<b>Tabla 3</b> <i>Características técnicas para la implementación de gafas inteligentes</i> .....	48
<b>Tabla 4</b> <i>Especificaciones técnicas del sistema embebido Raspberry Pi 4 model b</i> .....	51
<b>Tabla 5</b> <i>Especificaciones de la tarjeta Jetson Nano</i> .....	52
<b>Tabla 6</b> <i>Especificaciones técnicas LattePanda Delta</i> .....	53
<b>Tabla 7</b> <i>Parámetros que se considera para la selección de la tarjeta embebida</i> .....	54
<b>Tabla 8</b> <i>Especificaciones técnicas del módulo Voice Recognition V3</i> .....	55
<b>Tabla 9</b> <i>Especificaciones técnicas del módulo EasyVR3 Plus Shield Recognition</i> .....	56
<b>Tabla 10</b> <i>Especificaciones técnicas del módulo de reconocimiento de Grove</i> .....	57
<b>Tabla 11</b> <i>Selección del módulo del reconocimiento de voz</i> .....	57
<b>Tabla 12</b> <i>Especificaciones técnicas de la pantalla Nextion 5.0"</i> .....	59
<b>Tabla 13</b> <i>Especificaciones técnicas de la pantalla Waveshare 5.0"</i> .....	60
<b>Tabla 14</b> <i>Especificaciones técnicas de la pantalla Elecrow 5.0</i> .....	61
<b>Tabla 15</b> <i>Selección de la pantalla</i> .....	61
<b>Tabla 16</b> <i>Especificaciones técnicas del módulo de cámara de Raspberry Pi 4</i> .....	63
<b>Tabla 17</b> <i>Especificaciones técnicas de la cámara Nano IMX219-160</i> .....	64
<b>Tabla 18</b> <i>Especificaciones técnicas de la cámara mini cámara UVC Bullet</i> .....	65
<b>Tabla 19</b> <i>Selección de la cámara</i> .....	65
<b>Tabla 20</b> <i>Especificaciones técnicas power bank Ldnio</i> .....	67
<b>Tabla 21</b> <i>Especificaciones técnicas power bank Enox</i> .....	68
<b>Tabla 22</b> <i>Especificaciones técnicas KSIX power bank</i> .....	68
<b>Tabla 23</b> <i>Selección de la fuente de alimentación</i> .....	69
<b>Tabla 24</b> <i>Especificaciones técnicas parlante Genius SP-Q106</i> .....	70
<b>Tabla 25</b> <i>Especificaciones técnicas mini parlante bluetooth</i> .....	70

<b>Tabla 26</b> <i>Especificaciones técnicas parlante Pc Mod G-101</i> .....	71
<b>Tabla 27</b> <i>Selección del parlante para el sistema de audio</i> .....	71
<b>Tabla 28</b> <i>Especificaciones técnicas Pushbotton</i> .....	73
<b>Tabla 29</b> <i>Selección del material para la impresión 3D.</i> .....	74
<b>Tabla 30</b> <i>Propiedades físicas del material PLA</i> .....	93
<b>Tabla 31</b> <i>Datos usados en el Software</i> .....	97
<b>Tabla 32</b> <i>Pruebas del reconocimiento OCR con letra número 10</i> .....	109
<b>Tabla 33</b> <i>Pruebas del reconocimiento OCR con letra número 11</i> .....	110
<b>Tabla 34</b> <i>Pruebas del reconocimiento OCR con letra número 12</i> .....	111
<b>Tabla 35</b> <i>Pruebas del reconocimiento OCR con letra número 14</i> .....	113
<b>Tabla 36</b> <i>Pruebas del reconocimiento OCR ambas caras</i> .....	114
<b>Tabla 37</b> <i>Pruebas del reconocimiento OCR con texto a doble columna</i> .....	115
<b>Tabla 38</b> <i>Pruebas del reconocimiento OCR con imágenes presente en el texto</i> .....	117
<b>Tabla 39</b> <i>Pruebas del reconocimiento OCR Distribución Chi cuadrado</i> .....	118
<b>Tabla 40</b> <i>Pruebas en diferentes escenas del reconcomiendo de objetos</i> .....	125
<b>Tabla 41</b> <i>Criterios por parte del personal médico para la validación del dispositivo.</i> .....	134
<b>Tabla 42</b> <i>Autonomía de las gafas inteligentes</i> .....	136
<b>Tabla 43</b> <i>Tiempo de carga</i> .....	136
<b>Tabla 44</b> <i>Consumo de cada componente del dispositivo según su datasheet</i> .....	137

## ÍNDICE DE FIGURAS

<b>Figura 1</b> <i>Lector audible de textos para personas con discapacidad visual</i> .....	22
<b>Figura 2</b> <i>Prototipo de reconocimiento de objetos y asistente de navegación</i> .....	23
<b>Figura 3</b> <i>Dispositivo electrónico con detección de obstáculos</i> .....	24
<b>Figura 4</b> <i>Dispositivo electrónico para personas con discapacidad visual OrCam Read</i> .....	25
<b>Figura 5</b> <i>Dispositivo electrónico FingerReader</i> .....	25
<b>Figura 6</b> <i>Grados de discapacidad visual en Cotopaxi</i> .....	27
<b>Figura 7</b> <i>Test Snellen</i> .....	32
<b>Figura 8</b> <i>Campo Visual</i> .....	33
<b>Figura 9</b> <i>Motilidad Ocular</i> .....	34
<b>Figura 10</b> <i>Etapas del sistema en las gafas inteligentes</i> .....	37
<b>Figura 11</b> <i>Text Detection</i> .....	40
<b>Figura 12</b> <i>Document Text Detection</i> .....	40
<b>Figura 13</b> <i>MobileNet V3 módulo</i> .....	41
<b>Figura 14</b> <i>MobileNets</i> .....	43
<b>Figura 15</b> <i>Diagrama de flujo de texto a voz</i> .....	44
<b>Figura 16</b> <i>Traducción de texto audio mediante GTTS API</i> .....	45
<b>Figura 17</b> <i>Raspberry Pi 4 model B</i> .....	51
<b>Figura 18</b> <i>Módulo Voice Recognition module V3</i> .....	55
<b>Figura 19</b> <i>EasyVR 3 Plus Shield</i> .....	56
<b>Figura 20</b> <i>Pantalla Nextion 5.0"</i> .....	59
<b>Figura 21</b> <i>Pantalla Waveshare 5.0"</i> .....	60
<b>Figura 22</b> <i>Módulo de cámara Raspberry V2</i> .....	62
<b>Figura 23</b> <i>Módulo cámara Nano IMX219-160</i> .....	63
<b>Figura 24</b> <i>Cámara mini bullet</i> .....	64
<b>Figura 25</b> <i>Power Bank</i> .....	66

<b>Figura 26</b> <i>Power Bank Enox</i> .....	67
<b>Figura 27</b> <i>Pushbotton Tact &amp; Cap</i> .....	72
<b>Figura 28</b> <i>Conexión del módulo para el reconocimiento de voz</i> .....	76
<b>Figura 29</b> <i>Esquemático de la conexión del módulo de reconocimiento de voz hacia el arduino uno</i> .....	76
<b>Figura 30</b> <i>Comandos disponibles para grabar la voz</i> .....	77
<b>Figura 31</b> <i>Comandos load verificar las voces cargadas</i> .....	78
<b>Figura 32</b> <i>Conexión de la cámara y micrófono</i> .....	79
<b>Figura 33</b> <i>Conexión de la cámara y micrófono</i> .....	79
<b>Figura 34</b> <i>Conexión pantalla</i> .....	80
<b>Figura 35</b> <i>Conexión sistema de audio</i> .....	81
<b>Figura 36</b> <i>Conexión y configuración de los puertos GPIO</i> .....	81
<b>Figura 37</b> <i>Conexión Power Bank</i> .....	82
<b>Figura 38</b> <i>Expansión USB externa</i> .....	83
<b>Figura 39</b> <i>Captura de la imagen para lectura y detección de objetos</i> .....	86
<b>Figura 40</b> <i>Diagrama de flujo para el modo detector de objetos y lector</i> .....	87
<b>Figura 41</b> <i>Algoritmo para la creación de diferentes audios</i> .....	88
<b>Figura 42</b> <i>Conexión de soporte hacia los lentes</i> .....	89
<b>Figura 43</b> <i>Configuración de la carcasa para la instalación de los subsistemas</i> .....	91
<b>Figura 44</b> <i>Tamaño de la foto para lectura y detección de objetos</i> .....	92
<b>Figura 45</b> <i>Análisis del esfuerzo máximo en la zona de menor espesor</i> .....	94
<b>Figura 46</b> <i>Análisis del esfuerzo máximo</i> .....	96
<b>Figura 47</b> <i>Resultado de esfuerzo a tensión Von Mises</i> .....	97
<b>Figura 48</b> <i>Análisis del esfuerzo máximo a una altura de 17 m.</i> .....	99
<b>Figura 49</b> <i>Análisis del esfuerzo máximo a una altura de 1.70 m</i> .....	100
<b>Figura 50</b> <i>Análisis del esfuerzo máximo a una altura de 14 m</i> .....	101
<b>Figura 51</b> <i>Análisis para determinar la fuerza máxima</i> .....	102



<b>Figura 52</b> <i>Diagrama de cuerpo libre clip</i> .....	102
<b>Figura 53</b> <i>Procesamiento para lectura y detección de objetos</i> .....	104
<b>Figura 54</b> <i>Inicio Portada</i> .....	105
<b>Figura 55</b> <i>Modos de funcionamiento</i> .....	105
<b>Figura 56</b> <i>Disposición de la hoja en diferentes ángulos</i> .....	108
<b>Figura 57</b> <i>Pruebas del reconocimiento OCR con letra número 10</i> .....	109
<b>Figura 58</b> <i>Pruebas de validación con el reconocimiento OCR, con letra número 11</i> .....	110
<b>Figura 59</b> <i>Pruebas de validación con el reconocimiento OCR, con letra número 12</i> .....	112
<b>Figura 60</b> <i>Pruebas de validación con el reconocimiento OCR, con letra número 14</i> .....	113
<b>Figura 61</b> <i>Pruebas de validación con el reconocimiento OCR ambas caras</i> .....	114
<b>Figura 62</b> <i>Pruebas de validación con el reconocimiento OCR texto doble columna</i> .....	116
<b>Figura 63</b> <i>Pruebas de validación con el reconocimiento OCR, con presencia de imágenes en el texto</i> ....	117
<b>Figura 64</b> <i>Distribución Chi Cuadrado</i> .....	121
<b>Figura 65</b> <i>Funcionalidad del dispositivo según los datos obtenidos de las diferentes pruebas.</i> .....	122
<b>Figura 66</b> <i>Objetos identificados en un cuarto en la noche con luz cálida prueba 13</i> .....	124
<b>Figura 67</b> <i>Objetos identificados en la vía pública prueba 6</i> .....	125
<b>Figura 68</b> <i>Distribución de Chi Cuadrado con un grado de libertad 14</i> .....	131
<b>Figura 69</b> <i>Funcionalidad del dispositivo según los datos obtenidos de las diferentes pruebas.</i> .....	132
<b>Figura 70</b> <i>Pruebas en personas médico para validación del dispositivo</i> .....	135
<b>Figura 71</b> <i>Pruebas de la autonomía del dispositivo</i> .....	137

## Resumen

En el presente trabajo de titulación se pretende desarrollar un dispositivo portátil para un lector audible de textos mediante la implementación de la API de Google Cloud Vision y reconocimiento de objetos mediante la implementación y configuración de la red neuronal MobileNet V3 mediante algoritmos de inteligencia artificial, para contribuir como una herramienta de asistencia para lectura y reconocimiento del entorno que brinde autonomía a las personas con discapacidad visual permitiéndole a la persona escuchar el texto de un periódico, revista o libro, con un tamaño de fuente pequeña y mediana; el reconocimiento de objetos comunes/tradicionales, aproximados 92 objetos, pero que en simultáneo dentro del entorno se pueden reconocer aproximados 10 objetos en tiempo real y estos pueden ir cambiando en función al punto de observación de la persona con discapacidad visual, mediante el cual permitirá a los usuarios poder localizar los objetos que se encuentran en diferentes escenas. La realización de un dispositivo portátil que permita la integración de personas con discapacidad visual leve, pero que también pueda contribuir como una herramienta de asistencia para personas que presentan una discapacidad moderada y grave, mediante la asistencia y supervisión de una tercera persona, brindando autonomía y mejorando en algo la calidad de vida de las personas con discapacidad visual.

*Palabras Clave:* Google Cloud Vision, MobileNet v3, lector audible, dispositivo portátil.

### **Abstract**

In the present degree work is intend to develop a portable device for an audible text reader by implementing the Google Cloud Vision API and object recognition through the implementation and configuration of the MobileNet V3 neural network using artificial intelligence algorithms, to contribute as an assistance tool for reading and recognition of the environment that provides autonomy to people with visual impairment allowing the person to listen to the text of a newspaper, magazine or book, with a small and medium font size; the recognition of common/traditional objects, approximate 92 objects, but that simultaneously within the environment can recognize approximate 10 objects in real time and these can be changing depending on the point of observation of the visually impaired person, which will allow users to locate objects found in different scenes. The realization of a portable device that allows the integration of people with mild visual impairment, but can also contribute as an assistance tool for people with moderate and severe disability, through the assistance and supervision of a third person, providing autonomy and improving somewhat the quality of life of people with visual impairment.

*Key words:* Google Cloud Vision, MobileNet v3, audible reader, portable device.

## Capítulo I

### Fundamento Teórico

#### Planteamiento del problema

El Consejo Nacional para la igualdad de discapacidades es el organismo que cuenta con políticas para la integración y la verificación de oportunidades para las personas discapacitadas, entre estas políticas se encuentra el establecimiento de servicios sociales, información y ayuda a domicilio (CONADIS), existen 54.397 personas registradas con discapacidad visual, esta base de datos fue tomada en el mes de septiembre del 2022, lo cual en la provincia de Cotopaxi existen 1400 personas lo que representa el 2.57% del total de personas que presentan discapacidad visual en el Ecuador, ya sea moderada o grave, de acuerdo a los datos obtenidos las personas con discapacidad visual tienen acceso limitado a la información escrita. Los audiolibros y el braille son alternativas para superar estos obstáculos, pero tienen la desventaja de que deben estar preparados para un texto en particular.

Actualmente en Ecuador no existen empresas que se dediquen a desarrollar este tipo de tecnologías, por lo cual es importante incluir a las personas con discapacidad visual, es por esto que se pretende con este trabajo de titulación priorizar y dar la importancia a las personas con discapacidad, mediante el acceso a dispositivos y herramientas que les permitan adaptarse e integrarse mejor en la vida cotidiana.

A pesar que existen varios proyectos de titulación a nivel nacional e internacional, los cuales se enfocan en dispositivos para lectura de textos para bibliotecas, no existe un dispositivo portable que se adapte a los lentes de las personas que tienen discapacidad visual ya sea moderada o grave, mucho menos con reconocimiento de objetos. Por ejemplo, existen otros artículos exclusivos para personas con discapacidad visual como naipes en braille, parchís, dominó, libros de capacitación y entretenimiento, pero son difíciles de conseguir en el mercado nacional en su gran mayoría esos artículos se encuentran en inglés. Por un lado, no

se los produce, pero tampoco se los importa del extranjero debido a su alto costo. El Dispositivo OrCam MyEye Pro tiene un valor de \$3750 dólares que no está al alcance de la mayoría de personas en nuestro país. (Orcam, 2022)

Las sociedades existentes se han asegurado mediante el uso de la tecnología que la información debe ser fácilmente accesible, manipulable y reproducible, pero las personas con discapacidad visual no cuentan con esta tecnología, ni prototipos de bajo costo de ayuda y apoyo como complemento a todo su desarrollo integral en las personas. Además, existe dificultad en el manejo del sistema Braille en los centros educativos por parte de docentes y estudiantes, que deben ser desarrollados previamente, lo que conlleva a una menor comprensión del texto, menor velocidad de lectura y altos costos de traducción de textos e investigación.

Los problemas de visión afectan la vida de las personas física, mental, emocional, social, académica y profesionalmente. Estos problemas pueden dificultar el aprendizaje, la realización de actividades de la vida diaria, la comunicación y la orientación. Debido a este tipo de dificultades que presentan las personas con discapacidad visual compensan esta deficiencia utilizando la información que obtienen de sus otros sentidos más desarrollados: el oído para orientarse y el tacto para localizar objetos.

Al no tener un dispositivo con inteligencia artificial que les ayudará en el aprendizaje mediante la lectura y el reconocimiento de objetos facilitando la localización de objetos alrededor del entorno, es por tal motivo las personas con discapacidad visual no pueden relacionarse con el medio que los rodea y presentan dificultad ante un cambio de escenario.

### **Antecedentes**

Las personas con discapacidad visual se enfrentan a una serie de problemas socioculturales y entornos de vida que dificultan su desarrollo y superación personal, siendo la

lectura y el reconocimiento de objetos los principales problemas a los que se enfrentan. Aunque existen en el mercado herramientas para ayudar a las personas con discapacidad visual a leer y a reconocer los objetos, como el anillo para leer texto (FingerReader-Wearable Text-Reading Device); gafas inteligentes para personas con discapacidad visual, este proyecto busca aprovechar el potencial de estas nuevas tecnologías para contribuir y aportar en la vida de las personas con discapacidad visual de forma independiente y poder contribuir como herramienta de ayuda en sus actividades cotidianas.

A nivel nacional se han desarrollado una serie de dispositivos para aportar en la calidad de vida de las personas que presentan discapacidad visual. Por ejemplo, Escobar Pilatasig (2017) realizó una “Investigación e implementación de un sistema de visión artificial para un lector audible de textos de fácil manejo para las personas con discapacidad visual mediante la utilización de software libre en la biblioteca de la Universidad De Las Fuerzas Armadas ESPE Extensión Latacunga”, este dispositivo puede leer de manera óptima palabras con letras de tamaño número 12, puede determinar si el documento contiene imágenes e ignorarlas para la detección por (OCR) reconocimiento óptico de caracteres, el único problema que presenta el dispositivo es cuando el documento contiene caracteres especiales.

### **Figura 1**

*Lector audible de textos para personas con discapacidad visual*



*Nota.* Sistema de visión artificial para un lector audible de textos. Tomado de (Pilatasig, 2017)

Mientras Cadena Lesly & Rogel Xavier (2021), realizó un “diseño e implementación de un prototipo electrónico de reconocimiento de objetos y asistente de navegación para personas con discapacidad visual mediante visión artificial”, este prototipo cuenta con dos modos de operación, el primero permite al usuario navegar en un entorno cerrado y el segundo permite reconocer los objetos circundantes en cualquiera de los modos de funcionamiento se transmiten al usuario mediante alertas de voz.

## Figura 2

*Prototipo de reconocimiento de objetos y asistente de navegación*



*Nota.* Visualización del prototipo de asistencia de navegación para personas. Tomado de (Cadena Lesly Stefania, 2021)

María Cristina Loayza Jaramillo (2016), realizó una investigación de un sistema electrónico con detección de obstáculos para personas con discapacidad visual que proporciona información al usuario empleando un sistema háptico compuesto por el contacto entre 4 micro servomotores y 4 dedos de la mano. En el dispositivo se utilizó un sensor de distancia ultrasónico HC SR04 para detectar obstáculos. La distancia máxima que se puede medir es de 4 metros divididos en rangos de distancias. La activación de uno o más actuadores indica que la persona está dentro de la distancia especificada y advierte a las personas que lo están empleando se detengan o continúen en su trayectoria. Existen dos modos de

funcionamiento automático: analógico y binario con 4 y 16 rangos respectivamente, modo manual para memorizar combinaciones de accionamiento. (Cristina, 2016)

### **Figura 3**

*Dispositivo electrónico con detección de obstáculos*



*Nota.* Detalles del prototipo elaboración del soporte. Tomado de (Cristina, 2016)

A nivel internacional existen dispositivos y aplicaciones que se han desarrollado, tal es el caso de la empresa OrCam Read (ver figura 4), es el primer dispositivo de asistencia para lectura para personas con reducción en la visión o dificultades de lectura, como dislexia o fatiga. OrCam Read es un dispositivo móvil capaz de detectar texto en diferentes idiomas, identificar rostros y reconocer productos. (Orcam, 2022)

El equipo de investigación de MIT Media Lab. Desarrolló el dispositivo Finger Reader (ver figura 5), para facilitar la lectura. Es un anillo asistido por cámara que lee texto impreso a través de voz sintética, ayudando a las personas con discapacidad visual a leer libros sin usar braille, argumentan que las tecnologías actuales como los lectores de pantalla, los libros electrónicos, las impresoras braille o las aplicaciones móviles se utilizan muy poco o no lo suficiente para permitir que las personas con discapacidad visual ganen más independencia en sus actividades cotidianas. (Finger Reader, 2021)



#### Figura 4

*Dispositivo electrónico para personas con discapacidad visual OrCam Read*



*Nota.* Este dispositivo presenta diversos modos de funcionamiento: lectura, identificación de billetes y reconocimiento de rostros. Tomado (Orcam, 2022)

#### Figura 5

*Dispositivo electrónico FingerReader*



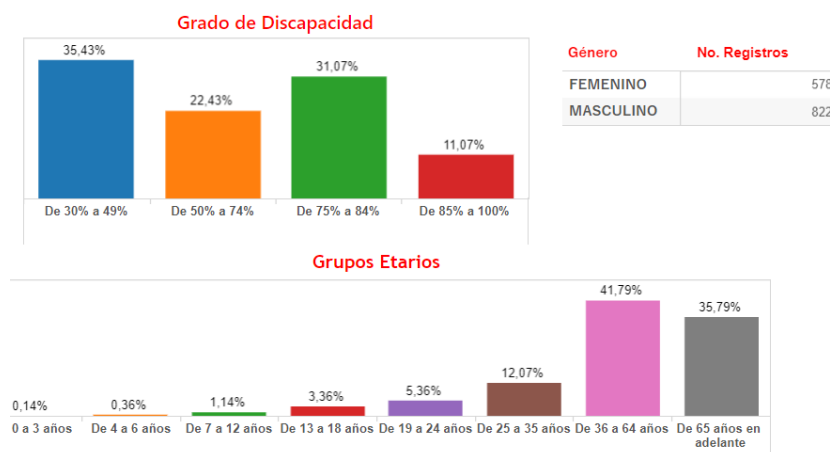
*Nota.* Este dispositivo se coloca en el dedo índice de la persona y realiza una lectura. Tomado de (EyeComitec Laz Group S.A, 2022)

**Patronato Municipal Gad Latacunga: Población de personas con discapacidad visual leve, moderada y grave de bajo recursos económicos.**

El Patronato Municipal de Amparo Social de Latacunga tiene el compromiso de prestar servicios de salud con calidad y calidez en el ámbito de la asistencia especializada, a través de su cartera de servicios cumpliendo con la responsabilidad de promoción, prevención, recuperación, rehabilitación de salud integral y acción social, conforme a las políticas del Ministerio de Salud Pública, en el marco de la justicia y la equidad social.

El Patronato Municipal de Amparo Social del Cantón Latacunga, como un centro de salud accesible que presta una atención de calidad que satisface las necesidades y expectativas de la población, bajo principios fundamentales de la salud pública y acción social, utilizando la tecnología y los recursos públicos de forma eficiente, eficaz y transparente, orientada su trabajo en el área estrictamente de servicio social, en su programa de entregas de ayudas técnicas dirigidas especialmente a los sectores más vulnerables como un modelo de inclusión social y derechos humanos enfocada a mejorar el bienestar y calidad de vida de las personas. (Patronato Municipal Gad Latacunga, 2022)

La población con discapacidad visual en Cotopaxi es de 1400 personas de las cuales el 35.43 % son personas con discapacidad moderada que equivalen a 496 personas; el 22.43 % son personas con discapacidad grave que equivalen a 314 personas; el 31.07% son personas con discapacidad muy grave que equivalen a 435 personas de las cuales en la base de datos de CONADIS se evidencia que el mayor problema de discapacidad visual se presentan en personas de 36 a 64 años lo que representa el 41,79 % que equivalen a 585 personas, de los cuales la gran mayoría son personas de bajos recursos y estas personas se sienten aisladas al no tener textos y no poder interrelacionarse con su entorno lo que provoca malestar. (Estadísticas de Discapacidad- Consejo Nacional para la Igualdad de Discapacidades, 2022)

**Figura 6***Grados de discapacidad visual en Cotopaxi*

*Nota.* Estadísticas de personas que presentan discapacidad visual. Tomado de (Estadísticas de Discapacidad- Consejo Nacional para la Igualdad de Discapacidades, 2022)

**Justificación**

Este proyecto tiene como beneficiario al Patronato de Amparo Social de Latacunga para personas con discapacidad visual, los cuales contempla en su plan de acción la prestación de asistencia técnica dirigidos a los grupos vulnerables (niños, niñas, personas con discapacidad y adultos mayores), a través de la inclusión social y los derechos humanos para garantizar una mejor protección social. Mediante las gafas inteligentes como herramienta para lectura y el reconocimiento de objetos que permitan a las personas que presentan una discapacidad permanente, temporal y debido a la edad disminuir la desventaja social, promover su bienestar y su autonomía.

Al implementar el dispositivo para personas que presentan discapacidad visual, parte de la necesidad de una fácil operación en áreas como la lectura y el reconocimiento de objetos, el desarrollo del presente proyecto facilitará la lectura y contribuirá como una herramienta de asistencia y reconocimiento de objetos sea más accesible para personas con discapacidad

visual, permitiéndoles acceder a la información impresa, lo cual permitirá a los usuarios llevar una vida más independiente y mejorar su estado emocional.

Los dispositivos existentes disponibles en el exterior, que se utilizan para la lectura de texto a partir de imágenes están limitados al idioma español y no corresponden al vocabulario y dialectos que se utilizan en el país. El desarrollo del proyecto tiene como objetivo ayudar a las personas con discapacidad visual mediante el desarrollo de un dispositivo cómodo, pequeño y fácil de usar que pueda leer texto y reconocer objetos, con el fin de mejorar en gran medida la calidad de vida del usuario, esto ayudará a todas las personas con discapacidad visual tengan más opciones y puedan lograr una independencia.

Este dispositivo es una herramienta necesaria para las personas con discapacidad visual, mediante la implementación de inteligencia artificial, para la lectura de texto y reconocimiento de objetos como nuevas tendencias en la educación en la que las tecnologías de la información y la comunicación contribuyen un papel fundamental como fuentes de apoyo a la enseñanza y el aprendizaje. Esto surge a partir de la inclusión para personas con discapacidad visual con una variedad de propósitos cotidianos, económicos y sociales, pero en el ámbito educativo buscan satisfacer las necesidades básicas de aprendizaje que surgen dentro y fuera del aula gracias a su capacidad multifuncional, las TIC también posibilitan la creación de nuevas metodologías y estrategias pedagógicas, las que son utilizadas como herramientas para contribuir en los procesos de enseñanza-aprendizaje, mediante la implementación de las gafas inteligentes convirtiéndose en herramientas de apoyo para facilitar la autonomía y participación de las personas con discapacidad visual.

En el Ecuador actualmente en el ministerio de telecomunicaciones y de la sociedad de la información (Mintel), promueve el desarrollo de proyectos tecnológicos e innovadores mediante nuevas tecnologías de desarrollo a través de inteligencia artificial para solventar las necesidades reales de la población, como es el caso para personas con discapacidad visual de

bajos recursos, mediante los conocimientos adquiridos en la carrera de ingeniería mecatrónica se pretende aplicar capacidades e intercambiar conocimientos con el objeto de garantizar herramientas tecnológicas, para ayudar a mitigar en algo las necesidades de las personas.

## **Objetivos**

### ***Objetivo General***

Diseño e implementación de gafas inteligentes como herramienta de asistencia para lectura y reconocimiento del entorno para personas con discapacidad visual utilizando inteligencia artificial.

### ***Objetivos Específicos***

- Realizar una revisión bibliográfica de los diferentes dispositivos y sistemas electrónicos comerciales para lectura de textos y reconocimiento de objetos para personas con discapacidad visual moderada o grave.
- Diseñar el modelo 3D del dispositivo y seleccionar los materiales para su elaboración para la estructura y la carcasa, que brinden soporte y protección de los componentes internos, que permita comodidad y ergonomía al usuario.
- Desarrollar e implementar algoritmos para el reconocimiento de texto y objetos mediante una cámara web aplicando diferentes técnicas y operaciones de visión artificial.
- Implementar el sistema de inteligencia artificial mediante la configuración de la red neuronal convolucional para el reconocimiento de objetos.
- Diseñar e implementar los circuitos electrónicos para la transmisión de datos hacia el sistema embebido.
- Implementar el sistema de convertidor de texto a voz a través de la comunicación entre el microcontrolador y el entorno mediante una interfaz gráfica amigable que

permita la supervisión del dispositivo para el monitoreo del reconocimiento del texto y objetos al usuario.

- Determinar si las gafas inteligentes contribuyen como una herramienta de asistencia para lectura y reconocimiento del entorno mediante diferentes pruebas de funcionamiento del dispositivo con personas que presentan discapacidad visual moderada o grave, mediante la supervisión de un especialista.

### **Hipótesis**

¿El diseño e implementación de gafas inteligentes podrá ser considerada como una herramienta de asistencia para lectura y reconocimiento del entorno en el que se desempeñan las personas con discapacidad visual?

### **Variables de la investigación**

#### ***Variables Independientes***

Gafas inteligentes

#### ***Variables Dependientes***

Herramienta de asistencia para lectura y reconocimiento del entorno para personas con discapacidad

### **Deficiencia Visual**

Para la Dra. Ananya Mandal (2012) el deterioro se define como “cualquier pérdida o anomalía en una estructura anatómica o una función fisiológica o psicológica”, de manera similar, una discapacidad es “cualquier restricción de la capacidad para realizar una actividad dentro del rango considerado normal para un ser humano”. Esto coloca a las personas en desventaja ante la sociedad debido a un impedimento o discapacidad.

También se puede definir la deficiencia visual la cual desencadena limitaciones en la ejecución de actividades cotidianas, acciones y funciones del sistema visual, por lo cual el Instituto Nacional del Ojo hace referencia que la baja visión parte como un impedimento visual,

que no se puede corregir con lentes estándares, lentes de contacto, medicamentos o cirugía que interfiere con la capacidad de realizar actividades de la vida diaria. (Ananya, 2012)

Mediante un análisis de un especialista en oftalmología se puede determinar la capacidad de la visión a través de diferentes análisis; agudeza visual, el campo visual y la motilidad ocular, conceptos que a continuación se explica:

### ***Agudeza visual***

Es la capacidad del sistema visual para distinguir los objetos con nitidez a determinada distancia y bajo ciertas condiciones, en este aspecto, cada ojo puede ser diferente del otro, se refiere a la capacidad de ver los detalles de un objeto contra un fondo uniforme o de ver dos objetos que están muy cerca, por ello es recomendable realizar el test de Snellen para conocer la agudeza visual de cada ojo con y sin lentes correctores. (Calderón, 2012)

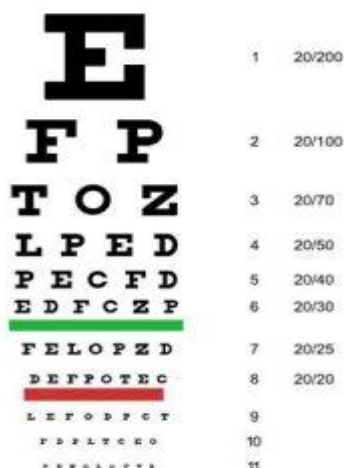
Test de Snellen consiste que la persona observe filas de letras que van reduciendo su tamaño, el más habitual que se practica en este test es el de la letra E, cuantas más filas sea capaz de determinar la persona mejor será la evaluación correspondiente agudeza visual.

Tres factores físicos y neurológicos que se determinan mediante la agudeza visual.

- La precisión con que la córnea y el cristalino del ojo enfocan la luz sobre la retina.
- La sensibilidad de los nervios de la retina y de los centros de la visión en el cerebro.
- La capacidad del cerebro para interpretar la información recibida de los ojos.

## Figura 7

### Test Snellen



*Nota.* Disposición de las letras para la realización del test Snellen. Tomado de (Oftalvist, 2019)

### Campo Visual

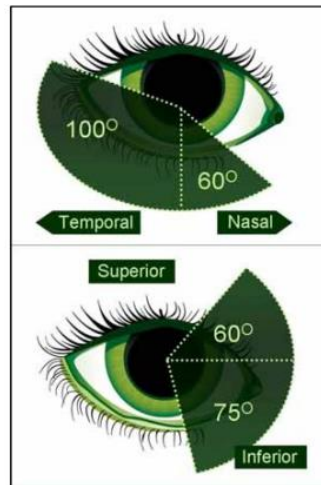
El campo visual se define como el área del espacio en la que los objetos son visibles, durante la fijación constante de la mirada en una dirección. Incluye las cosas que se ven de manera directa mediante su visión central, así como lo que se puede ver alrededor de ese punto: arriba, abajo y a los lados, comúnmente se conoce como visión periférica. (Randhawa, 2023)

- **Porción Nasal:** Espacio del centro del campo de visión hacia la nariz, el límite normal es de 60°(Eje Horizontal).
- **Porción Temporal:** Espacio que va del centro del campo de visión hacia la oreja, el límite normal es de 100°(Eje Horizontal).
- **Porción Superior:** Espacio que va del centro del campo de visión hacia arriba, el límite normal es de 60°(Eje Vertical).
- **Porción Inferior:** Espacio que va del centro del campo de visión hacia abajo, el límite normal es de 75°(Eje Vertical).



## Figura 8

### *Campo Visual*



*Nota.* Rangos visuales presentes en las personas. Tomado de (Bonafonte, 2022)

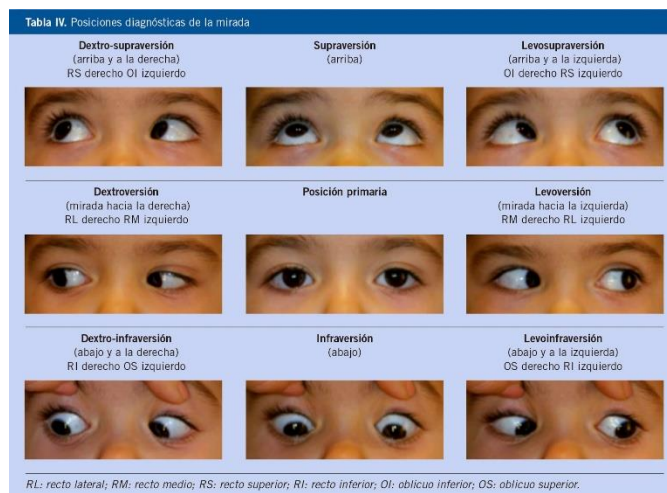
### ***Motilidad ocular***

Hace referencia al control del movimiento, lo cual aumenta el campo visual efectivo tanto monocular como binocular, permite mantener la imagen enfocada en la fovea y mantener la alineación de los ojos, evitando diplopías o visión doble. (Carneros, 2019)

- Ducciones: movimientos monoculares (referencia del ojo con respecto a los ejes)
- Versiones: movimientos binoculares asociados
- Vergencias: movimiento de los dos ojos de manera sincrónica y simétrica en igual dirección y sentido contrario.
- Convergencia existe una afinidad fisiológica cuando miras de cerca

## Figura 9

### Motilidad Ocular



*Nota.* Posiciones diagnosticadas mediante motilidad ocular. Tomado de (Motilidad ocular-Clínica Oftalmológica Halcon, 2022)

### Clasificación de la discapacidad visual

Para la Organización Mundial de la Salud (OMS), se puede clasificar en cuatro niveles: visión normal, discapacidad visual moderada, discapacidad visual grave y ceguera. Mediante la OMS se clasifica de diferentes grados la discapacidad visual, en función de dos parámetros, la agudeza visual y el campo visual de las personas. (OMS, 2012)

- Visión normal: agudeza visual 0.8 o mayor
- Baja visión moderada: agudeza visual < 0.3
- Baja visión grave: agudeza visual < 0.12
- Baja visión profunda: agudeza visual < 0.05
- Ceguera casi total: agudeza visual < 0.02
- Ceguera total: sin percepción de luz

## Discapacidad visual Cotopaxi

Mediante el Consejo Nacional para la igualdad de discapacidades (CONADIS), las personas con discapacidad visual registradas en el Ecuador existen 54.397 personas que presentan discapacidad visual, se registra en la provincia de Cotopaxi 1400 personas lo que representa el 2.57% del total de personas que presentan discapacidad visual en el Ecuador, esta ya sea moderada, grave, muy grave y discapacidad completa o ceguera.

**Tabla 1**

*Porcentaje de acuerdo a la discapacidad visual registrada en Consejo Nacional para la igualdad CONADIS.*

Grado de discapacidad visual	Número de personas con discapacidad
De 30% a 49%	496
De 50% a 74%	314
De 75% a 84%	435
De 85% a 100%	155
Total	1400

*Nota.* Personas con discapacidad visual Registradas en el Consejo Nacional para la igualdad CONADIS. Tomada del (Consejo Nacional para igualdad de discapacidades, 2021)

Existiendo 496 personas que presentan una discapacidad moderada que son secuelas permanentes que presentan limitaciones a la hora de realizar actividades de la vida cotidiana; una discapacidad grave siendo deficiencias permanentes que afectan a las personas en su totalidad e imposibilidad la realización de sus actividades cotidianas y necesitan de ayuda de una tercera persona para realizar sus actividades, siendo estos los grupos con mayor número de personas que necesitan una atención prioritaria.

## **Inteligencia Artificial**

La inteligencia artificial (IA) es un conjunto de tecnologías que permiten el uso de computadores con diversas capacidades avanzadas para ver, comprender y traducir palabras habladas, escritas y analizar datos.

La IA es el pilar sobre el cual se implementan innovaciones informáticas modernas. Automatiza los procesos y proporciona información a partir de grandes conjuntos de datos para ofrecer valor a las personas y a las empresas. El aprendizaje automático es un subconjunto importante de la IA, donde máquinas construyen modelos a partir de datos de entrenamiento, garantizando predicciones más precisas. (Google Cloud, 2022)

La finalidad de implementar un dispositivo de gafas inteligentes aplicando inteligencia artificial (IA), que permita ayudar en la vida de las personas que presentan discapacidad visual, auditiva, intelectual y psicosocial, por tal motivo con estos avances tecnológicos puedan facilitar el proceso de comunicación, brindar mayor independencia y obtengan oportunidades de realizar sus tareas cotidianas de su diario vivir. Por ejemplo, funciones de reconocimiento de texto para lectura y reconocimiento del entorno.

La inteligencia artificial tiene influencia en los campos de la medicina y salud con el propósito de manejar información de manera más eficiente con patrones que se pueda obtener un diagnóstico para tratamientos más efectivos. (Mascaró, 2020)

### ***Etapas del sistema aplicados en dispositivo de las gafas inteligentes.***

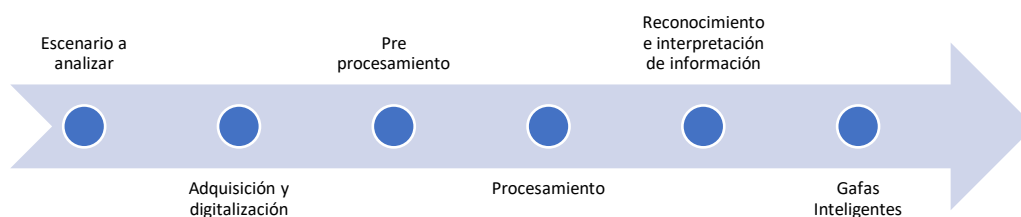
Los procesos de inteligencia artificial buscan imitar y simular las acciones de los humanos, como obtener imágenes a través de niveles de luz, enviarlas al cerebro y luego tomar decisiones sobre una acción en particular, como los procesos de inteligencia artificial basada en el aprendizaje automático supervisado.

El reconocimiento del texto, idioma y de objetos en diferentes escenas, que trasmita a una computadora y procesé la información de la imagen mediante una toma de decisión, sobre

una acción particular que debe realizar en este caso la detección de texto y reconocimiento del entorno; este proceso se divide en cinco etapas que se muestran en la figura 10.

### Figura 10

*Etapas del sistema en las gafas inteligentes*



*Nota.* Diferentes etapas del sistema en las gafas inteligentes.

**Escenario a analizar:** Es el área que queremos capturar, donde se encuentra la información que buscamos procesar.

**Adquisición y digitalización:** Consiste en analizar la imagen capturada a través de la cámara para localizar las características más relevantes que se presenta en la imagen y pasarla a algún formato digital compatible.

**Procesamiento previo:** La preparación de la imagen puede implicar la conversión de imágenes en color a escala de grises para reducir la complejidad computacional, la estandarización de las imágenes que tengan anchos y altos idénticos, la rotación y otras transformaciones típicas, dependiendo del modo de funcionamiento del dispositivo. (Team, 2020)

**Procesamiento:** Permite la transformación de datos en información necesaria para identificar y asociar el sistema mediante los datos de la imagen digital y el uso de software para realizar funciones específicas que luego se utilizan para evaluar. Estas evaluaciones son

condiciones preprogramadas que definen los criterios de aceptación en función del tipo de dato y del requerimiento del proceso. (Solución Ingenieril, 2022)

**Reconocimiento e interpretación de información:** En esta etapa se procesa la información obtenida en la etapa previa, se da una interpretación de las características necesarias para la toma de decisiones; dicha acción será uno de los dos modos de funcionamiento del dispositivo ya sea como lector de texto o detector de objetos, mediante el análisis de patrones para poder predecir y categorizar qué clase de objeto se está analizando.

**Aplicación:** Dispositivo gafas inteligentes utilizando inteligencia artificial para lectura de textos y reconocimientos de objetos.

### **Reconocimiento óptico de caracteres OCR basado en inteligencia artificial**

El reconocimiento óptico de caracteres (OCR), es un proceso especializado que permite buscar, extraer, indexar y optimizar datos en un formato legible por máquina, este conjunto de datos del documento escaneado de manera confiable, permite extraer información de documentos impresos, escritos, facturas y letreros etc.

El software OCR extrae las letras en la imagen, las convierte en palabras y luego las convierte en oraciones, lo que permite el acceso y la edición del contenido original, los sistemas OCR utilizan una combinación de hardware y software para convertir documentos físicos impresos en texto legible por máquina. El hardware como un escáner óptico o una placa de circuito especializada, copia o lee texto; luego el software generalmente maneja el procesamiento avanzado.

El software OCR puede aprovechar la inteligencia artificial (IA), para implementar métodos más avanzados de reconocimiento inteligente de caracteres (ICR), como identificar idiomas o estilos de escritura a mano. El proceso de OCR se usa más comúnmente para convertir documentos legales o históricos en papel en documentos PDF, para que los usuarios

puedan editar, formatear y buscar los documentos como si se hubiera creado con un procesador de textos.

En general, esta técnica tiene las siguientes ventajas:

- Reconocimiento y lectura en entornos complejos, oscuros, borrosos o de alta velocidad.
- Múltiples lecturas al mismo tiempo, pueden identificar diferentes idiomas en el mismo registro.
- Reduzca los costos de hardware y mantenimiento, independientemente de la gran infraestructura.
- Precisión cercana al 100% superando el estándar de las soluciones existentes.
- Integración automática de datos en los sistemas de gestión de información existente. (Read, 2022)

### **API de Cloud Vision**

La API Application Programming Interface (Interfaz de Programación de Aplicaciones) de Visión, ofrece modelos de aprendizaje automático preentrenados y muy potentes a través de las APIs REST y RPC, que generan etiquetas a las imágenes para poderlas clasificar de una manera más sencilla y rápida en millones de categorías predefinidas. Permite la detección de textos impresos, escritos a mano y extraen valiosos metadatos para su base de datos.

La API de Cloud Vision, es capaz de detectar y extraer texto de las imágenes. Existen dos funciones de anotación que admiten el reconocimiento óptico de caracteres (OCR).

TEXT\_DETECTION detecta y extrae el texto de cualquier imagen. Por ejemplo, una fotografía en la cual exista presencia de texto. El archivo. JSON incluye toda la string extraída, palabras individuales y sus cuadros de límite. (Google Cloud, 2022)

**Figura 11***Text Detection*

*Nota.* Reconocimiento mediante Text Detection. Tomado de (Google Cloud, 2022)

DOCUMENT\_TEXT\_DETECTION también extrae el texto de una imagen, pero la respuesta se optimiza para textos y documentos densos. El archivo JSON incluye información de la página, el bloque, el párrafo, la palabra y la división.

**Figura 12***Document Text Detection*



*Nota.* Detalles en una imagen mediante Document Text Detection. Tomado de (Google Cloud, 2022)

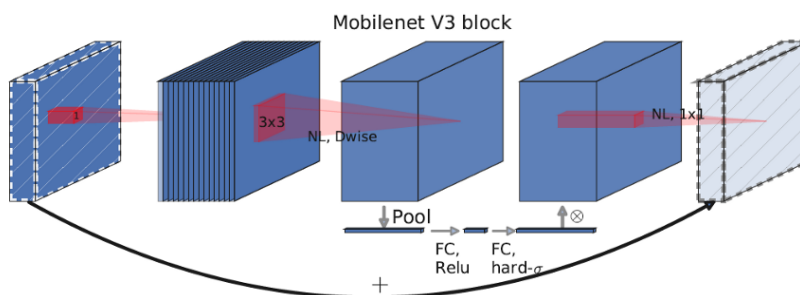
### MobileNet V3

MobileNet está basada en la combinación de técnicas de búsqueda de arquitectura de red (NAS), consciente del hardware, es decir MnasNET complementa el algoritmo NetAdapt, se usa simplemente en el modelo de esta versión Squeeze and Excitation (SE).

NetAdapt, adapta automáticamente una red previamente entrenada a una plataforma móvil mediante la optimización de la red de NetAdapt de forma automática para reducir gradualmente el consumo de recursos de una red preentrenada mientras se maximiza la precisión. (Sik-Ho, 2020)

### Figura 13

*MobileNet V3 módulo*



*Nota.* Módulo de la red MobileNet V3. Tomado de (Sik-Ho, 2020)

La biblioteca Coco (objeto común en contexto), es un conjunto de datos de subtítulos, segmentación y detección de objetos a gran escala. El modelo mobilenet-ssd es una red de detección de caja múltiple (SSD) de disparo único, diseñada para realizar la detección de objetos, este archivo es un modelo de Tensorflow, entrenado previamente en conjunto con la base de datos de COCO.

MobileNet es una clase de modelos eficientes llamados para aplicaciones de visión integradas y móviles. Esta clase de modelos se basa en una arquitectura simplificada que utiliza convoluciones separables en profundidad para construir redes neuronales profundas ligeras.

El modelo MobileNet V3 está conformado en convoluciones separables en profundidad, de forma de convoluciones factorizadas. Factoriza mediante la convolución estándar en una convolución en profundidad y una convolución  $1 \times 1$  conocida como convolución puntual. La convolución en profundidad aplica un solo filtro a cada canal de entrada en el caso de MobileNets. Luego la convolución puntual genera una convolución  $1 \times 1$  que permite combinar las salidas desde la convolución profunda. Una convolución estándar tiene un solo paso para filtrar y combinar entradas hacia el conjunto de salidas. Pero la convolución separable en profundidad divide esto en dos capas, una capa separada para filtrar y una capa separada para combinar. Esta factorización reduce drásticamente el cálculo y garantiza un mayor rendimiento en el tamaño del modelo. (Datapeaker, 2022)

Los modelos de MobileNet se pueden aplicar a varias tareas de reconocimiento para una inteligencia eficiente en el dispositivo. La biblioteca utilizada permite que la Raspberry Pi pueda identificar 91 objetos/animales únicos y proporcionar una calificación de confianza que se actualiza constantemente.

Cada vez que se identifique un objeto, dibujara un cuadro a su alrededor en la transmisión en vivo, lo etiquetará y proporcionará una calificación de confianza, se puede modificar el valor del porcentaje de umbral, aumentar esto significa que el software solo dibujara un cuadro alrededor de un objeto cuando esté absolutamente seguro, el valor porcentual nms, se podrá modificar con la finalidad de limitar la cantidad de objetos identificados simultáneamente. (Core Electronics , 2022)

## Figura 14

### MobileNets



*Nota.* Modelo MobileNet reconocimiento de objetos. Tomado de (Programador clic, 2022)

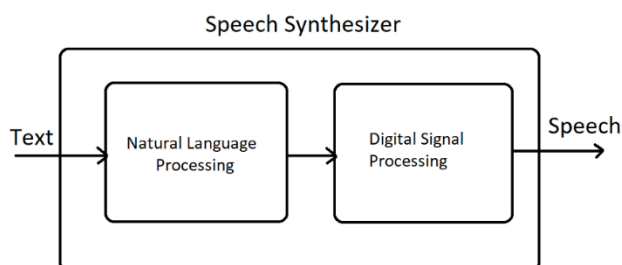
### gTTS (Google Text-to-Speech)

gTTS (Google Text-to-Speech API), una biblioteca de Python y una herramienta CLI (Interfaz de línea de comandos), para interactuar con la API de texto a voz de Google Translate, que permite convertir texto a voz de una forma simple y efectiva, cuenta con preprocesamiento y tokenización flexibles.

Implementa tecnologías innovadoras de Google que permite la generación de voces con entonación similar a las de los seres humanos. La API se desarrolló mediante la experiencia DeepMind en la síntesis de voz y ofrece voces de calidad idénticas a las humanas.

## Figura 15

Diagrama de flujo de texto a voz



*Nota.* Especificación de la conversión de texto a voz que realiza esta herramienta. Tomado de (Brarath, 2022)

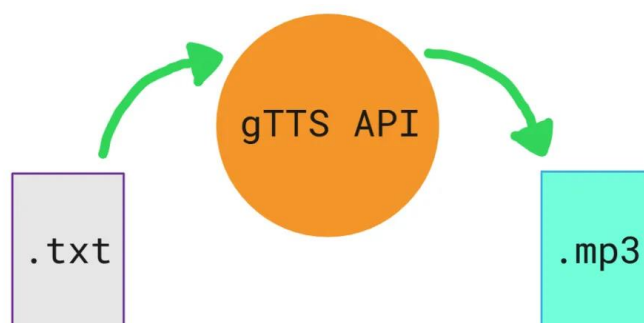
El texto a voz es el proceso de convertir palabras en una forma de audio vocal. El programa, herramienta o software toma un texto de entrada del usuario, utilizando métodos de procesamiento de imágenes, comprende la lingüística del idioma que se utiliza y realiza inferencias lógicas en el texto. Este texto procesado pasa al siguiente bloque donde se realiza el procesamiento de señales digitales en el texto procesado. Utilizando algoritmos y transformaciones, este texto procesado finalmente se convierte en un formato de voz.

La API gTTS (Google Text-to-Speech), admite varios idiomas, incluidos inglés, hindi, tamil, francés, y alemán etc. Se puede pronunciar en cualquiera de las dos velocidades de audio disponibles, rápida o lenta.

Esto es extremadamente útil cuando hay una barrera de comunicación y el usuario no puede transmitir sus mensajes a las personas. La conversión de texto a voz es de gran ayuda para las personas con discapacidad visual o con otras discapacidades, permitiendo la traducción de texto a voz.

**Figura 16**

*Traducción de texto audio mediante GTTS API*



*Nota.* Conversión de texto a voz mediante gTTS API. Tomado de (Artturi, 2022)

## Capítulo II

### Selección y diseño de componentes

#### Metodología

La metodología de diseño e implementación aplicada en el trabajo de titulación se basa en un dispositivo mecatrónico como herramienta de asistencia para lectura y reconocimiento del entorno para personas con discapacidad visual, las cuales presentan los siguientes requerimientos a considerar: diseño, autonomía, fiabilidad, ergonomía del dispositivo, selección de componentes electrónicos, costo en el mercado y concepto final del dispositivo.

En la tabla 2, se considera de forma general las necesidades principales de mayor influencia para la implementación del dispositivo: Diseño, algoritmo mediante IA, funcionalidad, las cuales se establecen en necesidades: Por parte del usuario (A), por parte de CEAPAT (Centro Estatal de Autonomía Personal y Ayudas Técnicas) (B), factores de análisis de conceptos y dispositivos existentes (C).

#### Tabla 2

*Necesidades de personas que presentan discapacidad visual para la implementación de dispositivos de ayuda técnica.*

#	Tipo	Necesidades
1	A	Sistema inteligente debe capturar imágenes
2	C	Sistema inteligente debe reconocer diferentes textos
3	C	Sistema inteligente debe identificar objetos en diferentes escenas
4	A	Sistema inteligente debe tener diferentes modos de funcionamiento
5	A	Sistema inteligente debe tener interfaz gráfica para asistencia
6	A	Sistema inteligente debe convertir texto a voz
7	B	Sistema inteligente debe controlar el sistema de audio

#	Tipo	Necesidades
8	A/B	Gafas inteligentes deben ser de fácil acceso al usuario
9	A/B	Sistema inteligente para el soporte de gafas debe ser ergonómico, liviano y pequeño
10	A/B	Gafas inteligentes debe soportar caídas (Case, soporte)
11	A/C	Gafas inteligentes debe tener avisos con el usuario:(Inicialización, Batería y conexión a internet)
12	A	Gafas inteligentes no debe tener un alto costo en el mercado

### **Definición de especificaciones**

Mediante la consideración de las necesidades de los usuarios, en este apartado se debe generar soluciones a las necesidades mediante métricas las cuales intervienen de manera directa en busca de establecer las características más relevantes para la implementación del dispositivo, funcionamiento del algoritmo mediante IA y el diseño del dispositivo que se adapte fácilmente para las personas con discapacidad visual.

El desarrollo del dispositivo de gafas inteligentes, comprende desde la obtención de conceptos, ideas iniciales, verificación y el desarrollo el cual se parte de la identificación de las necesidades que presentan los usuarios, mediante la selección de materiales, componentes electrónicos e implementación de nuevas tecnologías mediante inteligencia artificial, implementando metodologías innovadoras que sean de aporte para personas con discapacidad.

**Tabla 3**

*Características técnicas para la implementación de gafas inteligentes para personas con discapacidad visual*

#Necesidades	Métricas	Unidades
1,2,3	Resolución de la cámara digital	Píxel/in
1,2,3,5,6	Escalamiento de la imagen	Píxel
1,2,5,6	Cantidad de palabras reconocidas	# de palabras
2,5	Distancia máxima de detección de texto	cm
3,5	Cantidad de objetos a identificar	# de objetos
1,3	Distancia máxima de detección de objetos	cm
2,3,4	Modos de funcionamiento	Unidad
2,3,5	Resolución de la pantalla HMI	Píxel
2,3,7	Frecuencia del convertidor texto a voz	Hz
2,3,6,7,8	Control de volumen de salida de audio	Watios
2,3,4,5,6,7	Autonomía de las gafas inteligentes	Minutos
7,8	Modularidad	Unidad
8,9,10	Peso del dispositivo	kg



#Necesidades	Métricas	Unidades
9,10	Resistencia del dispositivo (Impresión 3D PLA)	MPa
2,3,4,6,7,	Notificaciones en el dispositivo	Unidad
12	Costo de elaboración (Manufactura)	USD\$

Los parámetros e indicadores que se establecen para la implementación del dispositivo de las gafas inteligentes: eficiencia del reconocimiento OCR y del detector de objetos, autonomía del dispositivo, diseño ergonómico, fiabilidad y el costo de manufactura siendo aspectos secundarios, el peso del dispositivo y su modularidad que presenta case y soporte para las gafas inteligentes.

Es necesario considerar los siguientes aspectos primordiales antes mencionados y etapas de funcionamiento que tendrá el dispositivo; software directamente con el sistema operativo, herramientas de programación de Google API y librerías para la implementación del dispositivo, en la parte de hardware el soporte físico que brindan los sistemas electrónicos.

### **Alternativas y selección de componentes**

Este tipo de método cualitativo por puntos, se deben incluir los principales factores para la ubicación, para personalizar los valores de peso de acuerdo con los criterios de los investigadores, se consideran diferentes alternativas con características de cada elemento.

Al comparar dos o más localizaciones opcionales, se procede a asignar una calificación a cada factor en una localización de acuerdo a una escala determinada por el evaluador de cero a diez, la suma de las calificaciones que acumulé el mayor puntaje se establece diferentes criterios de análisis y selección para lo cual se sugieren los siguientes pasos para clasificar los factores específicos. (Urbina, 2010, pág. 87)

1. Se identifican y se seleccionan factores de localización más influyentes
2. Asignar un peso a cada factor que indique su importancia relativa, la suma de los pesos debe sumar 1 y el peso asignado por el evaluador dependerá del criterio.
3. Asignar una escala común a cada factor de análisis (0 a 10) y elegir cuál es el máximo, calificar a cada sitio potencial y cuál es el mínimo.
4. Calificar a cada sitio potencial de acuerdo con la escala designada y multiplicar la calificación por el peso.
5. Sumar la puntuación de cada sitio y elegir la máxima puntuación (Urbina, 2010, pág. 87)

### **Selección del sistema embebido**

#### ***Opción 1. Raspberry Pi4 Model B***

El Raspberry Pi Model B ofrece mejoras innovadoras en la velocidad de procesamiento, el rendimiento multimedia, presenta un mayor almacenamiento en la memoria RAM de 8Gb; destinado para el desarrollo de dispositivos de control de bajo costo con capacidades de proporcionar herramientas necesarias para aplicaciones de inteligencia artificial, dispone de diversos tipos de conexiones (HDMI, USB, video y sonido), que conectan diferentes dispositivos periféricos, mantiene la compatibilidad con versiones anteriores y un consumo de energía similar, cuál posee una gran versatilidad y su plataforma está desarrollada para código abierto.

**Tabla 4**

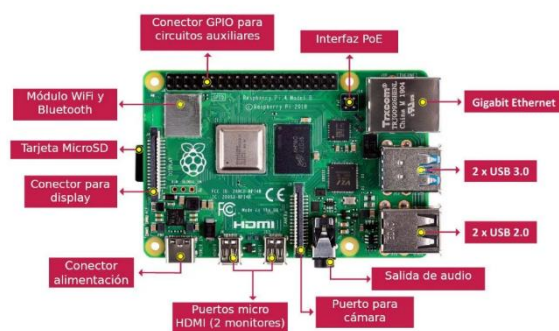
*Especificaciones técnicas del sistema embebido Raspberry Pi 4 model b*

Procesador	Broadcom BCM2711, CORTEX-A72 de cuatro núcleos (ARM v8) 64 bits SoC @ 1.5GHz
GPU	Video Core VI (OpenGL)
Multimedia	H.265 (decodificación 4Kp60); H.264 (codificación 1080p60, codificación 1080p30); OpenGL ES, gráficos 3.0
Video y sonido	2 x puertos micro HDMI (admite hasta 4Kp60) Puerto de pantalla MIPI DSI Puerto de cámara MIPI CSI Puerto de audio estéreo y video
Conectividad	5,0 GHz - 2,4 GHz, inalámbrico LAN, Bluetooth 5.0, BLE Gigabit Ethernet 2 x puertos USB 3.0 2 x puertos USB 2.0.
GPIO	40 pines

*Nota.* Disposición de la Raspberry Pi 4 model B. Tomado de (Geek Factory, 2022)

**Figura 17**

*Raspberry Pi 4 model B*



*Nota.* Disposición de la Raspberry Pi 4 model B. Tomado de (Geek Factory, 2022)

### **Opción 2. Jetson Nano**

Jetson Nano es una computadora pequeña y poderosa para aplicaciones integradas y sistemas de inteligencia artificial IoT, que brinda el poder de la IA moderna en un módulo. El JetPack SDK integral con bibliotecas aceleradas para aprendizaje profundo, visión por computadora, gráficos, multimedia y más. Jetson Nano tiene el rendimiento y las capacidades que necesita para ejecutar cargas de trabajo de IA modernas, lo que le brinda una manera rápida y fácil de agregar, incluye periféricos necesarios para el desarrollo de un sistema embebido con IA y redes neuronales. (NVIDIA Developer, 2019)

**Tabla 5**

*Especificaciones de la tarjeta Jetson Nano*

Procesador	Quad-core ARM Cortex-A57 MPCore de cuatro núcleos
GPU	NVIDIA MAXWELL- NVIDIA CUDA con 128 núcleos
Multimedia	H.265 (4Kp60); H.264 (4kp30)
Video y sonido	4x USB 3.0, USB 2.0, Micro-B HDMI 2.0 o DP 1.2; eDP 1.4; DSI (1 x 2)
Conectividad	Gigabit Ethernet, M.2 Clave E
GPIO	260 pines

*Nota.* Especificaciones técnicas de la tarjeta Jetson Nano. Tomado de (NVIDIA Developer, 2019)

### **Opción 3. LattePanda Delta**

LattePanda Delta es una SBC (computadora de placa única) de alto rendimiento, con bajo consumo de energía que ejecuta el sistema operativo completo Windows 10 o Linux. Es ampliamente utilizado en computación perimetral, máquinas expendedoras, máquinas

publicitarias, automatización industrial, desarrollador de IoT (Internet de las cosas), integrador de sistemas o proveedor de soluciones, LattePanda es una poderosa placa de desarrollo que puede potenciar la creación y acelerar su productividad. Toda esta potencia informática utiliza solo 6 W.

**Tabla 6**

*Especificaciones técnicas LattePanda Delta*

Procesador	Intel Celeron N4100
GPU	Gráficos Intel HD 600, 200-700 MHz
Multimedia	H.265 (decodificación 4Kp60); H.264 (codificación 1080p60, codificación 1080p30); OpenGL ES, gráficos 3.0
Video y sonido	DP 1.4 (40850x2160, a 60 Hz) 4K Ultra HD
Conectividad	WIFI 802.11 AC, 2.4G y 5G de doble banda Bluetooth 5G Gigabit Ethernet Puertos USB: 3x USB 3.0 Tipo A USB Tipo C
GPIO	2 GPIO de 50 pines, incluidos I2C, I2S, USB, TTL, UART, RTC, administración de energía.

*Nota.* Especificaciones técnicas de la tarjeta LattePanda Delta. Tomado de (LattePanda, 2022)

Se procede a la selección de la tarjeta embebida, considerando las especificaciones técnicas y soportes en el sistema operativo que brinda cada una de las tarjetas embebidas, la velocidad de procesamiento que facilite la implementación de la red neuronal MobileNet V3 y la API de Google Vision y la accesibilidad en el mercado.

**Tabla 7**

*Parámetros que se considera para la selección de la tarjeta embebida para las gafas*

*inteligentes*

Aspecto Importante	Peso Asignando	Raspberry Pi 4		Jetson Nano		LattePanda Delta 432	
		Califi.	Ponde.	Califi.	Ponde.	Califi.	Ponde.
Procesador	0,3	8	2,4	9	2,7	10	3
Velocidad del procesamiento	0,2	8	1,6	9	1,8	10	2
RAM	0,2	8	1,6	9	1,8	9	1,8
Sistema operativo	0,1	10	1	6	0,6	5	0,5
Conexión audio/video	0,1	10	1	6	0,6	5	0,5
Accesibilidad en el mercado	0,1	10	1	5	0,5	4	0,4
Total	1		8,6		8		8,2
Selección			Si		No		No

En referencia a lo presentado en la tabla 7, la adquisición de la tarjeta embebida será la Raspberry Pi 4 model B de 8GB, la cual ofrece las prestaciones suficientes para el desarrollo del dispositivo de las gafas inteligentes: presenta una velocidad superior en el procesamiento, fluidez en la decodificación de video a 60 fps sin compatibilidad HDR (High Dynamic Range) la Raspberry Pi 4 soporta una variedad de sistemas operativos en las cuales se puede adaptar a la necesidad del usuario, presenta conectores de audio y video para la visualización y la reproducción del sonido, lo que garantiza que la tarjeta embebida está disponible en el mercado y es la idónea para el desarrollo del dispositivo.

## Selección del módulo de reconocimiento de voz

### ***Opción 1. Voice Recognition module V3***

El módulo de reconocimiento de voz Voice Recognition module V3. Permite un registro de hasta 80 comandos de voz en total. Este no convertirá los comandos en texto, pero lo compara con un conjunto de voces ya grabadas por el usuario. Cualquier sonido podría ser entrenado como comando. El usuario debe entrenar al módulo antes de que pueda reconocer los comandos de voz, por lo que es posible que no se reconozca los comandos pronunciados por otros usuarios. (botnroll.com, 2022)

### **Tabla 8**

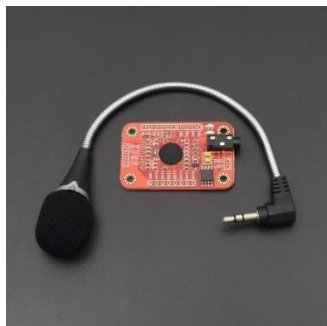
#### *Especificaciones técnicas del módulo Voice Recognition V3*

Rango de Voltaje	4,5 a 5 V
Consumo de corriente	<40mA
Tiempo de grabación	150(s)
Interfaz digital	5V TTL UART Y GPIO
Tamaño	31 mm x 50 mm
Precisión de reconocimiento	99 % (bajo condiciones controladas)

*Nota.* Especificaciones técnicas del módulo Voice Recognition V3. Tomado de (botnroll.com, 2022)

### **Figura 18**

#### *Módulo Voice Recognition module V3*



*Nota.* Dispositivo para reconocimiento de voz. Tomado de (botnroll.com, 2022)

### **Opción 2. EasyVR3 Plus Shield Recognition**

El módulo EasyVR 3 Plus es un módulo de reconocimiento de voz multipropósito diseñado para agregar fácilmente capacidades de reconocimiento de voz versátiles, robustas y rentables a cualquier aplicación. EasyVR 3 Plus permite hasta 256 comandos dependientes del hablante (SD) o Verificación del hablante (SV), definidos por el usuario que se pueden entrenar en cualquier idioma. (Mouser Electronics, 2022)

#### **Tabla 9**

*Especificaciones técnicas del módulo EasyVR3 Plus Shield Recognition*

Rango de Voltaje	3,3 a 5 V
Consumo de corriente	<40mA
Tiempo de grabación	100(s)
Interfaz digital	UART estándar
Tamaño	25.4 mm x 56.4 mm
Precisión de reconocimiento	95 %

*Nota.* Especificaciones técnicas del módulo Voice Recognition V3. Tomado de (botnroll.com, 2022)

#### **Figura 19**

*EasyVR 3 Plus Shield*



*Nota.* Módulo de reconocimiento de voz. Tomado de (Mouser Electronics, 2022)



### Opción 3. Módulo de reconocimiento de voz de Grove

“El motor de reconocimiento de voz de Grove está diseñado para aplicaciones controladas por voz, hogares inteligentes, juguetes inteligentes y robots controlados por voz y cualquier cosa que se desee controlar a través de voz” (RobotShop Europe, 2023).

**Tabla 10**

*Especificaciones técnicas del módulo de reconocimiento de Grove*

Rango de Voltaje	3 a 5 V
Consumo de corriente	25 mA – 130 mA
Tiempo de grabación	120(s)
Interfaz digital	JST2.0
Tamaño	25.4 mm x 56.4 mm
Precisión de reconocimiento	95 % (bajo condiciones normales)

*Nota.* Especificaciones técnicas del módulo reconocimiento de voz Grove V3. Tomado de (RobotShop Europe, 2023)

**Tabla 11**

*Selección del módulo del reconocimiento de voz*

Aspecto Importante	Peso Asignando	EasyVR3 Plus Shield		Reconocimiento de voz de Grove		Voice Recognition module V3	
		Califi.	Ponde.	Califi.	Ponde.	Califi.	Ponde.
Comandos de voz	0,1	9	0,9	7	0,7	8	0,8
Tiempo grabación	0,2	8	1,6	8	1,6	9	1,8
Precisión reconocimiento	0,3	8	2,4	8	2,4	8	2,4
Compatibilidad	0,1	10	1	9	0,9	9	0,9
Interfaz digital	0,2	8	1,6	7	1,4	9	1,8

Aspecto Importante	Peso Asignando	Voice Recognition module V3					
		EasyVR3 Plus Shield		Reconocimiento de voz de Grove		Voice Recognition module V3	
		Califi.	Ponde.	Califi.	Ponde.	Califi.	Ponde.
Accesibilidad en el mercado	0,1	7	0,7	8	0,8	10	1
Total	1	8,2		7,8		8,7	
Selección		No		No		Si	

En referencia a lo presentado en la tabla 11, la alternativa ideal para la implementación es el módulo de reconocimiento de voz Voice Recognition V3, garantizando una precisión del 99 % en el reconocimiento de voz bajo condiciones controladas, proporcionando un mayor tiempo para grabar, presentando diferentes protocolos de comunicación: UART y GPIO los cuales garantizan una comunicación en serie de manera síncrona como asíncrona, lo que certifica que es un módulo eficiente y compatible, permitiendo que la relación entre persona y dispositivo sea fluida para establecer comandos de voz y siendo accesible en el mercado.

### **Arduino Nano**

Para el funcionamiento del módulo Voice Recognition se necesita la conexión de una tarjeta de desarrollo: una placa pequeña que presente un menor consumo de energía, compatibilidad, protocolo de comunicación UART, la velocidad de su microcontrolador para el procesamiento de datos con el módulo de reconocimiento de voz es aceptable por tal motivo, la tarjeta que cumple con todas estas prestaciones es el arduino nano, siendo accesible en el mercado por su costo y funcionalidad.

### **Selección de la pantalla**

#### ***Opción 1. Pantalla Nextion 5.0"***

Nextion es una solución perfecta de interfaz hombre-máquina (HMI), que proporciona una interfaz de control y visualización entre humanos y procesos, aplicaciones o

dispositivos. Nextion se utiliza principalmente en los campos del internet de las cosas IoT o en las industrias de electrónica de consumo. El software proporciona una manera fácil de crear una interfaz de usuario táctil sorprendente e intuitiva. Nextion es la mejor solución HMI rentable para los usuarios con una curva de aprendizaje baja y reducida. (SandoRobotics, 2022)

## Figura 20

*Pantalla Nextion 5.0"*



*Nota.* Pantalla Nextion 5.0" High performance hardware. Tomado de (Nextion, 2022)

## Tabla 12

*Especificaciones técnicas de la pantalla Nextion 5.0"*

Resolución	800 x 480
Área visual:	108 mm (L) x 64.8 mm (W)
Soporte S.O.	Raspberry Pi A+, B+, Arduino
Tamaño	133.5 mm x 84 mm
Consumo de energía	5V 410mA
Pantalla TFT	panel táctil resistivo
Operación sin controlador	UART

*Nota.* Especificaciones técnicas de la pantalla Nextion. Tomado de (Nextion, 2022)

### ***Opción 2. Pantalla Waveshare 5.0"***

La pantalla táctil de 5" permite proyectos integrados, sistemas de información y entretenimiento. La pantalla táctil de 800 x 480 píxeles permite una interacción perfecta con el

usuario, es ideal para mostrar imágenes en aplicaciones con interfaces gráficas, esta pantalla LCD puede admitir sistemas Raspberry Pi OS / Ubuntu / Kali / Retropié; cuando la pantalla LCD funciona en sistemas como Raspberry Pi, la resolución debe configurarse manualmente, de lo contrario provocará una visualización anormal. (Waveshare , 2022)

**Tabla 13**

*Especificaciones técnicas de la pantalla Waveshare 5.0"*

Resolución	800 x 480
Área visual:	108 mm (L) x 64.8 mm (W)
Soporte S.O.	Raspberry Pi 4, Jetson Nano, etc.
Tamaño	12.1 x 89.4 cm
Consumo de energía	5V 410mA
Pantalla TFT	panel táctil resistivo
Operación sin controlador	plug and play

*Nota.* Especificaciones técnicas de la pantalla Nextion. Tomado de (Nextion, 2022)

**Figura 21**

*Pantalla Waveshare 5.0"*



*Nota.* Pantalla Waveshare compatible con todos los modelos de Raspberry Pi 4. Tomado de (Waveshare , 2022)

### **Opción 3. Pantalla Elecrow LCD 5.0”**

Esta pantalla táctil TFT de 5 pulgadas es un monitor mini HDMI de montaje en panel. Con una resolución estándar de 800 x 480 para Raspberry Pi, pero es muy pequeña y simple. Sin embargo, puede usar este monitor con cualquier dispositivo que tenga una salida HDMI, su forma de conexión hacia dispositivos electrónicos es sencilla.

**Tabla 14**

*Especificaciones técnicas de la pantalla Elecrow 5.0*

Resolución	800 x 480
Área visual:	108 mm (L) x 64.2 mm (W)
Soporte S.O.	Raspberry Pi B+, Jetson Nano
Tamaño	12.1 x 78 cm
Consumo de energía	5V 410mA
Pantalla TFT	panel táctil resistivo
Operación sin controlador	plug and play

*Nota.* Especificaciones técnicas de la pantalla Elecrow. Tomado de (RobotShop Europe, 2022)

**Tabla 15**

*Selección de la pantalla*

Aspecto Importante	Peso Asignando	Pantalla Nextion		Pantalla Waveshare		Pantalla Elecrow	
		Califi.	Ponde.	Califi.	Ponde.	Califi.	Ponde.
Compatibilidad S.O	0,2	7	1,4	9	1,8	7	1,4
Controlador	0,1	5	0,5	10	1	10	1
Precio	0,3	7	2,1	8	2,4	8	2,4
Panel táctil	0,1	10	1	10	1	10	1
Tamaño	0,2	8	1,6	8	1,6	8	1,6
Accesibilidad en el mercado	0,1	7	0,7	9	0,9	7	0,7



*Nota.* Módulo de cámara de Raspberry Pi 4. Tomado de (The Pi Hut, 2022)

**Tabla 16**

*Especificaciones técnicas del módulo de cámara de Raspberry Pi 4*

Resolución	3280 x 2464p
Soporte S.O.	Raspbian
Tamaño	25 mm x 23 mm x 9.1 mm
Conexión	Conector plano (CSI)
Sensor	Sony IMX219
Frecuencia	30fps

*Nota.* Especificaciones técnicas del módulo cámara de Raspberry Pi 4 Tomado de (The Pi Hut, 2022)

### **Opción 2. Módulo de cámara nano IMX219-160**

El módulo de la cámara nano IMX219-160 es compatible para visión nocturna, principalmente este tipo de módulos de cámara se utilizan en proyectos de IA, reconocimiento facial, detección OCR, que permite capturar imágenes y grabar videos en alta calidad. El módulo de la cámara incorpora un sensor Sony IMC219 de 8 megapíxeles, es compatible con los siguientes modos de video: 1080p/30fps, 720p/60fps y 480p/90fps. (Waveshare, 2023)

**Figura 23**

*Módulo cámara Nano IMX219-160*



*Nota.* Cámara Nano IMX219-160. Tomado de (Waveshare, 2023)

**Tabla 17**

*Especificaciones técnicas de la cámara Nano IMX219-160*

Resolución	3280 x 2464p
Soporte S.O.	NVIDIA Jetson
Tamaño	25 mm x 24 mm x 15 mm
Conexión	FFC 15 pines
Sensor	CMOS
Frecuencia	60fps

*Nota.* Especificaciones técnicas de la cámara nano IMX219-160 Tomado de (Waveshare, 2023)

### **Opción 3. Mini cámara UVC Bullet**

La mini cámara UVC Bullet son dispositivos pequeños y portátiles, el cual incorpora un lente estenopeica de 3,7 mm garantizando unas excelentes imágenes en alta resolución y videos en 1080 píxeles, lo cual proporciona dimensiones pequeñas que se ajustan a la aplicación de desarrollo, su conexión a la Raspberry es sencilla que se conectan mediante el puerto USB 2.0 el cual no necesita de instalación de drivers; una ventaja que presenta la cámara digital es la incorporación de un dial en los cuales se puede mostrar los diferentes modos de exposición de la cámara y ajustarlos a nuestra necesidad, dichas características son las que convierten este dispositivo en el ideal para desarrollo de la aplicación de las gafas inteligentes. (Alibaba.com, 2022)

**Figura 24**

*Cámara mini bullet*





*Nota.* Cámara con dial incorporado. Tomado de (Alibaba.com, 2022)

**Tabla 18**

*Especificaciones técnicas de la cámara mini cámara UVC Bullet.*

Resolución	1090 x 1080p
Soporte S.O.	Windows 10, macOS 10.14, Linux, Android
Tamaño	19 mm x 19 mm x 19 mm
Conexión	Cable USB – 2.0
Sensor	CMOS
Frecuencia	30fps

*Nota.* Especificaciones técnicas de la cámara mini cámara UVC Bullet. Tomado de (Alibaba.com, 2022)

**Tabla 19**

*Selección de la cámara*

Aspecto Importante	Peso Asignando	Módulo Raspberry V2		Nano IMX219- 160		Mini Cámara UVC Bullet	
		Califi.	Ponde.	Califi.	Ponde.	Califi.	Ponde.
Resolución	0,3	8	2,4	9	2,7	8	2,4
Tipo de conexión	0,2	7	1,4	6	1,4	9	1,8
Tamaño	0,2	7	1,4	6	1,2	8	1,6
Sistema Operativo	0,1	6	0,6	5	0,5	9	0,9
Precio	0,1	8	0,8	6	0,6	8	0,8
Accesibilidad en el mercado	0,1	8	0,9	5	0,5	8	0,8
Total	1	7,5		6,9		8,3	
Selección		No		No		Si	

En referencia a lo presentado en la tabla 19, la cámara digital que se selecciona es la Mini cámara UVC Bullet, la cual ofrece las prestaciones suficientes para el desarrollo del dispositivo de las gafas inteligentes: presenta una resolución de 1920 x 1080 píxeles que proporciona unas excelentes imágenes en alta resolución y videos con una fluidez de 30 fps, teniendo influencia en aspectos como el tipo de conexión mediante USB 2.0, las dimensiones del dispositivo para su adaptación en el diseño del soporte, asegurando su funcionalidad.

### **Selección de fuente externa power bank**

#### ***Opción 1. Ldnio PR518 Power Bank***

La fuente externa Ldnio PR518 es uno de los tipos más comunes de cargadores portátiles, los cuales brindan un método muy conveniente y fácil para cargar dispositivos, el cual presenta una forma y tamaño que se adaptan al requerimiento para la implementación del sistema. La power bank PR518 constituye una batería de polímero de litio que garantiza una vida útil al dispositivo que no sufra el envejecimiento en la misma medida, la cual puede resultar más costosa. (Botech, 2022)

### **Figura 25**

#### *Power Bank*



*Nota.* Fuente externa de 5000mAh. Tomado de (Botech, 2022)

**Tabla 20***Especificaciones técnicas power bank Ldnio*

Capacidad	5000 mAh/18,5 Wh
Entradas	Micro/Tipo-C :5V/2A máx.
Tamaño	118 mm x 44,5 mm x13 mm
Salidas	USB-ADC5V/2.1 A max.
Batería	Litio
Indicador	Estado de batería luz led

*Nota.* Especificaciones técnicas power bank Ldnio. Tomado de (Botech, 2022)

### **Opción 2. Enox 5000 KBH-I55GD**

Enox5000 es un banco de energía de carga rápida que ofrece la máxima productividad sobre la marcha al mantener su dispositivo funcionando tanto tiempo, es capaz de disipar el exceso de temperatura de manera efectiva; es una fuente de carga rápida, compatible con la mayoría de los dispositivos portátiles recargables. (Klipxtreme, 2022)

**Figura 26***Power Bank Enox*

*Nota.* Power bank Enox de 5000mAh. Tomado de (Klipxtreme, 2022)

**Tabla 21***Especificaciones técnicas power bank Enox*

Capacidad	5000 mAh
Entradas	Micro USB
Tamaño	125 mm x 68,8 mm x10 mm
Salidas	USB 2.0
Batería	Litio
Indicador	Estado de batería luz led

*Nota.* Especificaciones técnicas de la fuente externa power bank Enox. Tomado de (Klipxtreme, 2022)

**Opción 3. KSIX Power Bank**

El banco de energía KSIX power Bank es un banco de energía portátil que es capaz de recargar varios dispositivos electrónicos alimentados por la batería con el propósito de extender la vida útil. Luego se puede recargar el banco de energía a su máxima capacidad desde un puerto micro-USB.

**Tabla 22***Especificaciones técnicas KSIX power bank*

Capacidad	3000 mAh
Entradas	Micro USB
Tamaño	120 mm x 48 mm x9,5 mm
Salidas	USB 2.0
Batería	Litio
Indicador	Estado de batería no registra

*Nota.* Especificaciones técnicas KSIX. Tomado de (LDLC-High Tech Experience, 2023)

**Tabla 23***Selección de la fuente de alimentación*

Aspecto Importante	Peso Asignando	Enox 5000 KBH-I55GD		Ldnio PR518 Power Bank		KSIX BXBA 3000 U01	
		Califi.	Ponde.	Califi.	Ponde.	Califi.	Ponde.
Puertos de conexión	0,3	5	1,5	7	2,1	4	1,2
Capacidad	0,2	9	1,8	9	1,8	5	1
Precio	0,2	8	1,6	8	1,6	9	1,8
Tipo de batería	0,1	10	1	10	1	10	1
Diseño	0,1	5	0,5	6	0,6	5	0,5
Accesibilidad en el mercado	0,1	7	0,7	8	0,8	8	0,8
Total	1	7,1		7,9		6,3	
Selección		No		Si		No	

De acuerdo con la Tabla 23, la fuente externa que se selecciona es la power bank Ldnio PR518, que presente mayores prestaciones en su funcionamiento. La capacidad de la batería permite tener una autonomía aceptable para el dispositivo que es compatible con la mayoría de los dispositivos portátiles recargables, al ser una batería de litio nos garantiza una carga rápida y una duración mayor, incorporando puertos de entrada y salida para el correcto funcionamiento del dispositivo, la accesibilidad en costo y mercado del dispositivo.

**Selección de parlantes*****Opción 1. Parlante Genius SP-Q160***

El parlante Genius SP-Q160 es un dispositivo que da una salida de audio, que la finalidad de implementar en el dispositivo de las gafas inteligentes es escuchar la traducción de

texto a voz mediante el altavoz, de esta manera puede interactuar con el usuario por medio del sonido generado.

**Tabla 24**

*Especificaciones técnicas parlante Genius SP-Q106*

Potencia	6 W
Rango de frecuencia	150 – 20000 Hz
Amplificador	Integrado
Conector	USB Type A; 3,5 mm
Dimensiones	102mm x 85mm x 81 mm

*Nota.* Especificaciones técnicas del parlante Genius SP-Q106. Tomado de (Cemco, 2022)

### ***Opción 2. Mini Parlante Bluetooth***

El mini parlante Bluetooth puede vincularse con cualquier dispositivo que presente tecnología por bluetooth compatible con el sistema operativo de Raspbian, facilitando la reproducción mp3, limitando su autonomía debido a su capacidad de batería.

**Tabla 25**

*Especificaciones técnicas mini parlante bluetooth*

Potencia	3 W
Rango de frecuencia	280 – 16KHz
Amplificador	Integrado
Conector	Bluetooth
Dimensiones	60 mm x 60 mm x 50 mm

*Nota.* Especificaciones técnicas del parlante mini bluetooth. Tomado de (Tecno Trade, 2022)

### **Opción 3. Parlante Pc Mod G-101**

El parlante Pc Mod G-101 presenta un diseño compacto ofreciendo un sonido envolvente y eficiente, el cuál incorpora tecnología de frecuencia digital para un sonido más suave.

**Tabla 26**

*Especificaciones técnicas parlante Pc Mod G-101*

Potencia	5 W
Rango de frecuencia	100 – 20000 Hz
Amplificador	Incorporado
Conector	USB, Jack 3,5 mm
Dimensiones	70 mm x 70 mm x 68 mm

*Nota.* Especificaciones técnicas del parlante G-101. Tomado de (SIJ Distribuciones SaS, 2022)

**Tabla 27**

*Selección del parlante para el sistema de audio*

Aspecto Importante	Peso Asignando	Genius SP- Q106		Parlante mini bluetooth		Parlante Pc Mod G-101	
		Califi.	Ponde.	Califi.	Ponde.	Califi.	Ponde.
Potencia	0,3	8	2,4	5	1,5	6	1,8
Rango de Hz	0,2	7	1,4	9	1,8	7	1,4
Amplificador	0,1	10	1	10	1	10	1
Conector	0,2	9	1,8	5	1	8	1,6
Diseño	0,1	7	0,7	8	0,8	7	0,7
Dimensiones	0,1	6	0,6	8	0,8	8	0,8
Total	1		7,9		6,9		7,3
Selección			Si		No		No

De acuerdo con la Tabla 27, la implementación del sistema de audio se opta por el parlante Genius SP-Q106 que es el dispositivo idóneo para la implementación en el sistema de las gafas inteligentes que permite reproducir mediante audio la traducción de texto a voz, por medio de los altavoces del parlante, mediante una potencia de salida de audio aceptable para que pueda interactuar con el usuario, un control de volumen que permita ajustar: subir o bajar dependiendo de la necesidad.

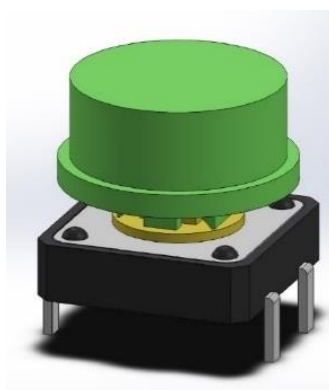
### **Pulsador: Pushbotton Tact & Cap**

El pulsador Tact & Cap, es un dispositivo que permite o impide el paso de corriente, lo cual permite acceder a activar o desactivar un modo de funcionamiento del dispositivo.

Según la norma DIN 33 402 .2: Las dimensiones de las manos son de gran importancia para el diseño de dispositivo de mando herramienta, según ese criterio se estableció el tamaño del pulsador, se opta que los botones deben ser grandes con la finalidad de una manipulación más sencilla con cualquiera de sus dedos de la mano siendo un promedio el tamaño del pulsador. (Melo, 2016)

### **Figura 27**

*Pushbotton Tact & Cap*



*Nota.* Pushbotton para personas con discapacidad visual. Tomado del Autor



**Tabla 28***Especificaciones técnicas Pushbotton*

Ciclos de vida	100 000 operaciones
Estilo interruptor	Botón táctil
Material	Plástico y metal
Pines	4
Dimensiones	12 mm x 12 mm x 7,3 mm

*Nota.* Especificaciones técnicas del pulsador Pushnotton. Tomado de (Novatronic, 2022)

De acuerdo con la Tabla 28, el pulsador Tact & Cap, permite la interacción con personas con discapacidad visual, con el propósito de adaptar un entorno de fácil acceso, se optó que el tamaño del pulsador debe tener botones grandes con la finalidad de una manipulación más sencilla en las personas con discapacidad visual, permitiendo ganar en independencia al momento que el usuario interactúe con el dispositivo.

### **Selección del material para la impresión de la carcasa 3D**

La selección del material para la carcasa y del soporte se basan en especificaciones técnicas que presentan los diferentes materiales, que tienen influencia en las aplicaciones en las cuales son recomendables. El material para la impresión 3D se establecerá de acuerdo a parámetros de cada material; características mecánicas, químicas, ópticas y térmicas que determinan el rendimiento de las diferentes piezas impresas en 3D, conforme a la importancia y ergonomía del prototipo.

**Tabla 29***Selección del material para la impresión 3D.*

Especificaciones	Peso Asignado	PLA		ABS		ABSi	
		Califi.	Ponde.	Califi.	Ponde.	Califi.	Ponde.
Resistencia	0.3	8	2.4	9	2.7	8	2.4
Dureza	0.2	8	1.6	8	1.6	8	1.6
Rendimiento moderado	0.1	9	0.9	7	0.7	6	0.6
Facilidad de impresión	0.1	9	0.9	8	0.8	6	0.6
Costo de impresión	0.2	10	2	7	1.4	5	1
Disponibilidad de impresión	0.1	10	1	8	0.8	0.7	0.7
Total			8.8		8		6.9
Selección			Si		No		No

De acuerdo con la Tabla 29, los resultados de las características del material para la impresión en 3D se realizará mediante la impresión en PLA, presentando ventajas en la impresión con respecto a los demás, la disponibilidad de impresión y materiales de consumo de bajo costo, que permiten la creación de prototipos, el cual presenta un rendimiento moderado para la elaboración de la carcasa y del soporte para las gafas inteligentes, presentando una resistencia considerable para impactos lo cual garantiza y valida la seguridad para el prototipo.

## **Capítulo III**

### **Diseño e implementación de los sistemas**

En este capítulo, se plantea el diseño y la implementación de cada uno de los sistemas, módulos, interfaces, considerando estudios existentes de dispositivo para personas con discapacidad visual, mediante el centro estatal de autonomía personal y ayudas técnicas, con la finalidad de implementar un dispositivo portable, ergonómico y eficiente mediante el desarrollo de algoritmos y configuración de herramientas: Api Cloud Vision, Google Text-To-Speech y de la red neuronal MobileNet V3.

Norma ISO 9241-420:2011. Ergonomía de la interacción hombre-sistema.

Norma ISO 9241-420:2011. Brinda orientación para la selección de dispositivos de entrada para sistemas interactivos, en función de factores ergonómicos, considerando las limitaciones y capacidades de los usuarios y las tareas específicas y el contexto de uso. (Instituto Nacional de Seguridad, 2022)

### **Diseño e implementación del reconocimiento de voz**

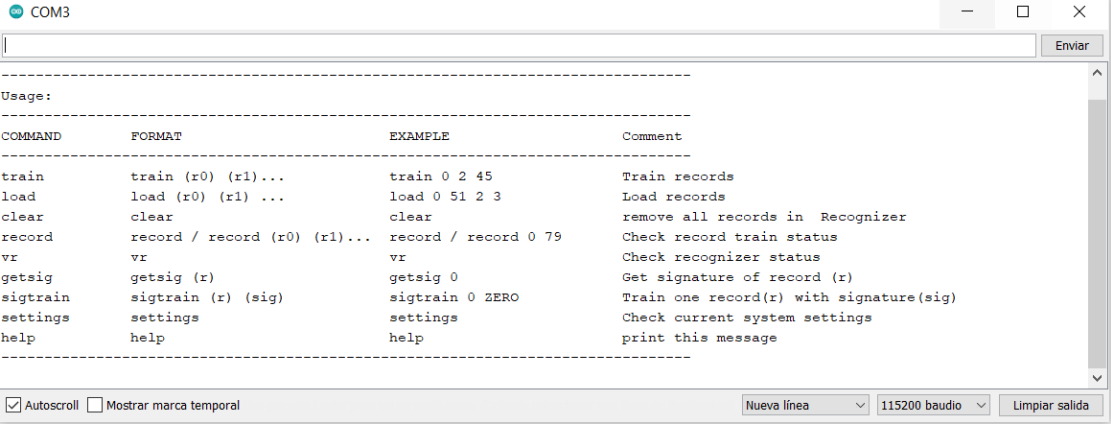
La implementación y diseño del módulo de reconocimiento de voz tiene la misión de establecer una comunicación entre humano y el sistema de reconocimiento, que sea capaz de captar voces lejanas/cercanas mediante comandos de palabras simples que el usuario puede emitir, dictar u ordenar al dispositivo.



Para grabar, cargar o modificar los comandos que previamente se grabaron en el módulo, se debe ejecutar un programa que viene incorporado en la librería: `vr_sample_train` el cual se ejecuta y se abre en el monitor serie y previamente se debe configurar a 115.500 baudios.

**Figura 30**

*Comandos disponibles para grabar la voz*



The screenshot shows a serial monitor window titled 'COM3'. It displays the usage information for the 'vr\_sample\_train' program, which is a table with four columns: COMMAND, FORMAT, EXAMPLE, and Comment. The table lists various commands such as 'train', 'load', 'clear', 'record', 'vr', 'getsig', 'sigtrain', 'settings', and 'help', each with its corresponding format, an example command, and a brief description of its function. At the bottom of the window, there are checkboxes for 'Autoscroll' (checked) and 'Mostrar marca temporal' (unchecked), along with dropdown menus for 'Nueva línea' and '115200 baudio', and a 'Limpiar salida' button.

```

Usage:
-----
COMMAND      FORMAT      EXAMPLE      Comment
-----
train        train (r0) (r1)...  train 0 2 45  Train records
load        load (r0) (r1) ...  load 0 51 2 3  Load records
clear       clear                          remove all records in Recognizer
record     record / record (r0) (r1)... record / record 0 79  Check record train status
vr         vr                          Check recognizer status
getsig     getsig (r)                getsig 0        Get signature of record (r)
sigtrain   sigtrain (r) (sig)        sigtrain 0 ZERO  Train one record(r) with signature(sig)
settings   settings                  Check current system settings
help       help                          print this message
-----
 Autoscroll  Mostrar marca temporal
Nueva línea 115200 baudio Limpiar salida

```

*Nota.* Parámetros para grabar los diferentes comandos de voz. Tomado de Autor

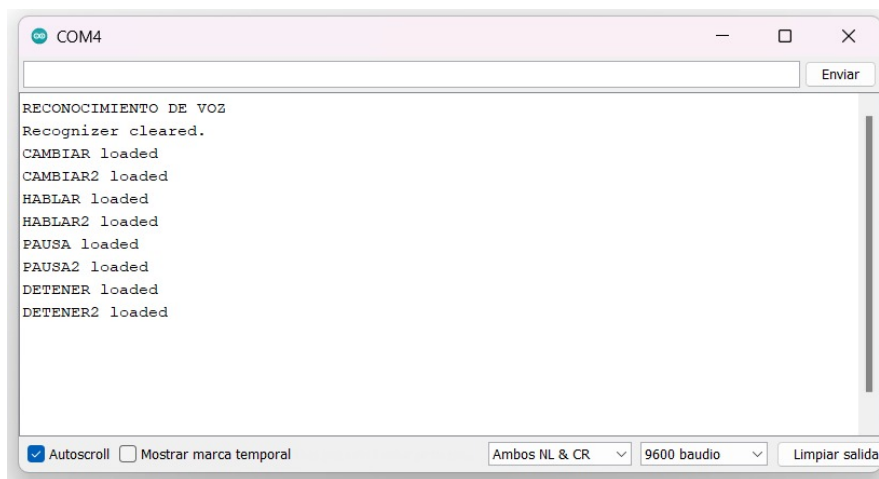
El primer comando que se debe utilizar para grabar la voz es el comando `train`, en el cual se debe establecer la posición en la que se quiere guardar y el nombre que se asigna, en este caso las voces que se graban son las siguientes: Cambiar, Hablar, Pausa y Detener.

El comando `load` se utiliza para cargar las voces que se grabó, donde se indican los números de registros que se quiere cargar, en este caso; "load 0 y 1" en el cual 0 vendría a ser "cambiar", grabada con una voz cerca al micrófono y 1 con una voz grabada lejos del micrófono, este proceso se realizará para todas las voces grabadas previamente.

Para comprobar que esté correctamente grabadas las voces se ejecuta el comando `record`, para verificar el estado de registro se localiza mediante el nombre y su posición 0 al 7, donde nos indica las voces grabadas previamente para su validación.

## Figura 31

Comandos load verificar las voces cargadas



```
COM4
RECONOCIMIENTO DE VOZ
Recognizer cleared.
CAMBIAR loaded
CAMBIAR2 loaded
HABLAR loaded
HABLAR2 loaded
PAUSA loaded
PAUSA2 loaded
DETENER loaded
DETENER2 loaded
```

Autoscroll  Mostrar marca temporal  Ambos NL & CR 9600 baudio Limpiar salida

*Nota.* Comandos de voz establecidos para la implementación en las gafas inteligentes. Tomado de Autor

El sistema de reconocimiento de voz es capaz de entender palabras, frases cortas, en condiciones normales, el cual limitará el reconocimiento en caso de presencia de ruido, voces más agudas, hombres/mujeres debido a la diferencia de tonalidad en su voz.

### **Conexión del subsistema cámara-micrófono**

Se debe realizar una configuración en los puertos de entrada del módulo del reconocimiento de voz, hacia la cámara que es un puerto USB 2.0 y del micrófono que es un puerto Jack, con la finalidad de unificar y optimizar espacio en el diseño del soporte de la cámara, mediante la conexión hacia un puerto HDMI, el cual garantice la transferencia de audio y video a través de un único cable, como se puede ver en la figura 32.

### Figura 32

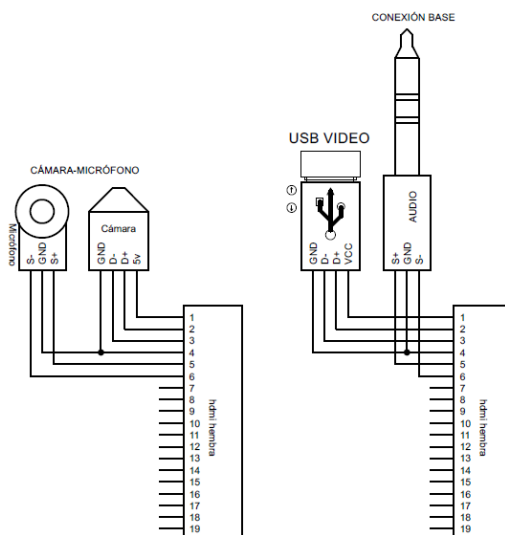
Conexión de la cámara y micrófono



*Nota.* Esquema de conexión de un conector unificado para el audio y video mediante la configuración de un cable HDMI. Tomado del Autor.

### Figura 33

Conexión de la cámara y micrófono



*Nota.* Esquema de conexión de la cámara con el micrófono. Tomado del Autor

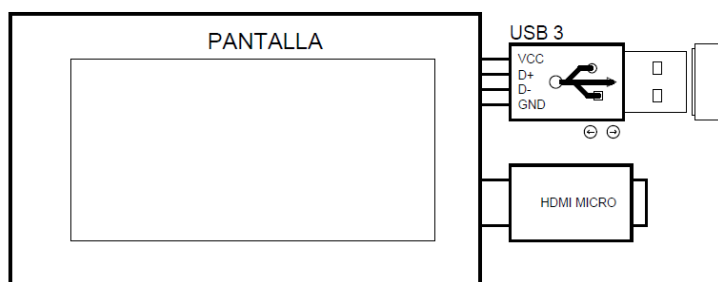
## Implementación de la pantalla

La implementación de la pantalla Waveshare de 5 pulgadas, mediante una interfaz (HMI), como asistente para el proceso de detección de texto y la visualización de la identificación de los objetos mostrando información en tiempo real.

Una vez que se conectan los puertos HDMI, micro USB en la pantalla Waveshare, se debe conectar las salidas a los puertos USB y HDMI mini en la tarjeta Raspberry Pi, como se muestra en la figura 34.

### Figura 34

*Conexión pantalla*



*Nota.* Esquema de conexión de la pantalla. Tomado del Autor

## Implementación del sistema de audio

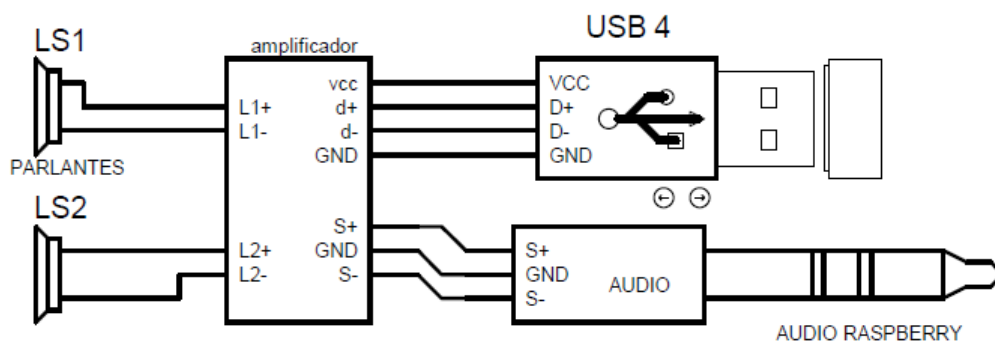
La implementación del sistema de audio en el dispositivo de gafas inteligentes garantiza una salida de audio por el altavoz a partir de la detección de texto a voz, el reconocimiento de objetos a través de audios de cada objeto identificado, de esta manera puede asistir e interactuar con el usuario.

Se establece la conexión del sistema de audio mediante la conexión de los puertos USB y Jack, en sus puertos destinados a la tarjeta Raspberry Pi.



## Figura 35

Conexión sistema de audio



Nota. Esquema de conexión del sistema de audio. Tomado del Autor

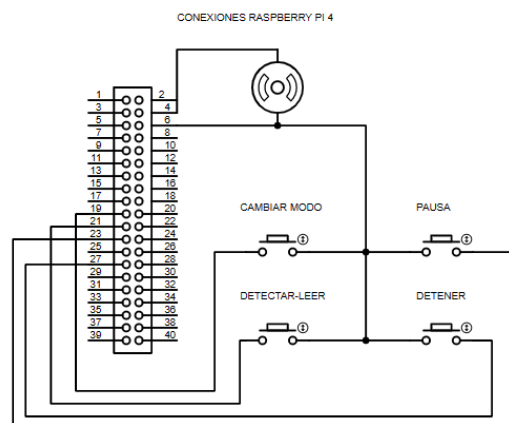
El sistema de audio establece un amplificador para controlar el volumen: subir/bajar el porcentaje del volumen por las personas, esto dependerá si el dispositivo se encuentra en un ambiente abierto o cerrado.

## Conexión de los pulsadores en los puertos GPIO

Para establecer la conexión de los pulsadores se debe realizar la conexión en las entradas digitales asignadas en el GPIO.

## Figura 36

Conexión y configuración de los puertos GPIO



*Nota.* Esquema de conexión y configuraciones de los pines de la Raspberry Pi 4 Model B.

Tomado del Autor.

Se conectan las entradas digitales de la Raspberry Pi mediante la configuración Pull-Up, que establece la conexión hacia los pulsadores, uno de los pines del pulsador se debe conectar al GND y el segundo pin se debe conectar a los 3.3 V de la fuente, en este caso cuando el pulsador no está presionado el pin de la señal tendrá un valor de 3.3 V, cuando el pulsador esté presionando el pin de la señal tendrá un valor de 0 V.

La configuración Pull-Up, cuando el pulsador está pulsado la raspberry lee 0 y cuando el pulsador está sin ser pulsado lee 1 lógico, permitiendo activar o desactivar los diferentes modos de funcionamiento del dispositivo.

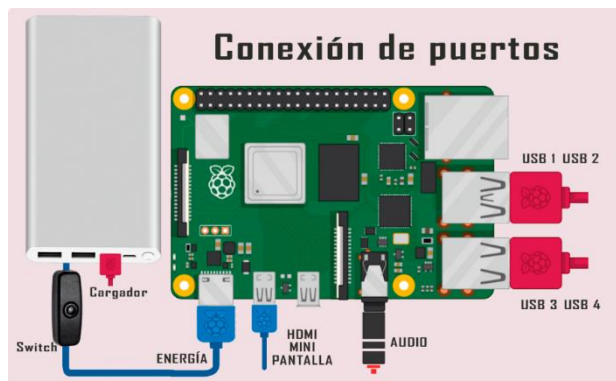
### ***Conexión del subsistema de carga power bank***

La conexión de la batería power bank, se conecta a un switch que se encarga de habilitar el encendido/apagado, que funciona como una fuente de alimentación para cada uno de los componentes electrónicos que están conectados en la Raspberry Pi.

Para que la batería pueda recargarse se debe conectar el cargador de la power bank que incorpora una alta velocidad de carga hacia el puerto USB Type C.

### **Figura 37**

*Conexión Power Bank*



*Nota.* Conexión de la fuente externa power bank hacia la tarjeta. Tomado del Autor

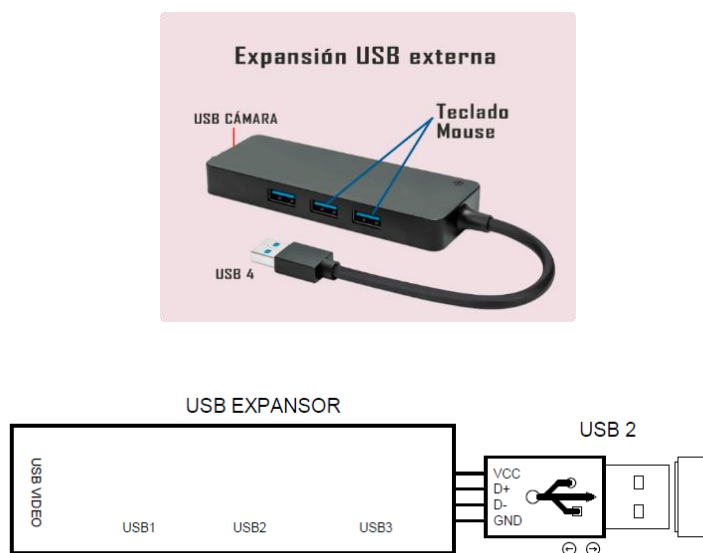
La capacidad de la batería es de 5000 mA. Brindando una autonomía aceptable al dispositivo, garantizando el correcto funcionamiento de los sistemas conectados hacia el dispositivo de las gafas inteligentes.

### ***Conexión del subsistema de expansión USB externa***

Este sistema tiene la finalidad que se puedan conectar diferentes puertos adicionales: mouse y teclado que permita interactuar y transferir información hacia la Raspberry Pi, permitiendo acceder a recursos de conexión a internet; mediante la conexión VNC (Virtual Network Computing) que es un programa de escritorio remoto que permite controlar y acceder a la Raspberry Pi por un ordenador.

### **Figura 38**

#### *Expansión USB externa*



*Nota.* Dispositivo expansor USB externo para la conexión de diferentes dispositivos adicionales mouse, teclado. Tomado del Autor

## Implementación del software para el reconocimiento de texto y del entorno

El sistema operativo que se instala para el desarrollo es Raspbian que es un sistema operativo libre, optimizado para el hardware de Raspberry Pi, la implementación mediante software se debe instalar previamente diferentes librerías, archivos adicionales que son necesarios para la ejecución del programa.

Para instalar las diferentes librerías mediante el gestor (PIP) Python Package Index, el siguiente comando desde consola.

```
pip3 install opencv-python
```

Este proceso se hace para cada una de las librerías que a continuación se detallan, para la implementación del algoritmo se utilizan algunas librerías que necesitan de conexión a internet para su funcionamiento.

### Librerías empleadas en Python

**Pickle:** Permite almacenar fácilmente colecciones y objetos en ficheros separando toda la parte de escritura y lectura que permita una serialización del objeto.

**CV2:** Es una biblioteca de visión artificial que permite trabajar con imágenes.

**Numpy:** Herramienta que permite trabajar con matrices.

**Pygame:** Es una librería para cargar, crear eventos y reproducir sonidos.

**Serial:** Permite comunicarse a través del puerto serial.

**Time:** Es una librería que tiene funciones relacionadas con la medición del tiempo. Por ejemplo, la librería permite establecer un intervalo de tiempo luego de alguna función o proceso utilizado.

**RPi.GPIO as Gpio:** Es una librería que permite controlar la interfaz GPIO de la Raspberry Pi.

**Os:** Este módulo consta de funciones para crear y administrar diferentes procesos en ejecución y contenidos de archivos, permite acceder a la estructura de directorios leer y escribir archivos. Por ejemplo, la reproducción de audio.

**Io:** Este módulo interpreta operaciones de entrada/salida en los archivos.

**Google.cloud:** Biblioteca cliente de Cloud, garantiza una experiencia optimizada en el desarrollo de uso de convenciones y estilos de lenguaje, recomendadas para la API de Google Cloud.

**Gtts:** Incorpora el servicio de Google Translate que permite convertir texto a voz.

**Mixer:** Controla la reproducción mediante eventos: detener, reanudar, etc.

**Wave:** Esta librería sirve para leer y escribir archivos.

**Screeninfo:** Obtiene la ubicación y tamaño de la pantalla.

**Socket:** Envía mensajes a través de la red, mediante una interfaz de red de bajo nivel. Por ejemplo, verifica la disponibilidad de internet para la detección mediante OCR.

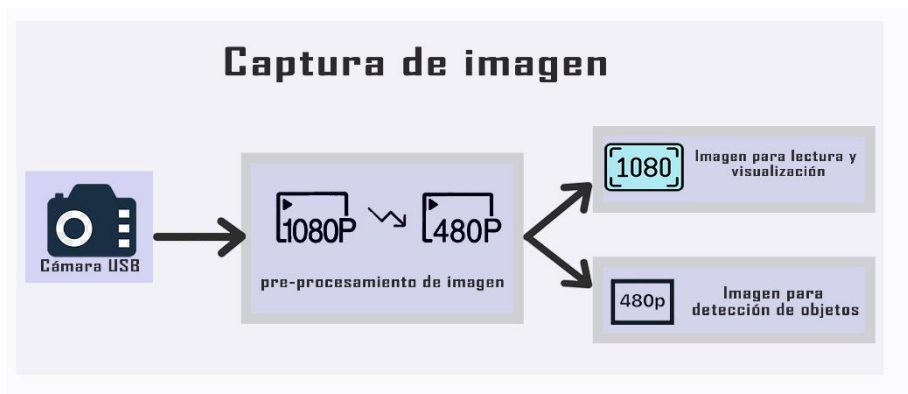
**Pil:** Edición de imágenes de manera directa desde Python, soporta diferentes formatos.

### **Implementación del sistema de las gafas inteligentes**

Para la implementación de las gafas inteligentes se contemplan dos funciones que realiza el dispositivo la primera que es la lectura, visualización y la segunda la detección de objetos en diferentes escenas, en la figura 41 se representa el funcionamiento general del dispositivo.

### Figura 39

*Captura de la imagen para lectura y detección de objetos*



*Nota.* Pre-procesamiento de imagen. Tomado del Autor

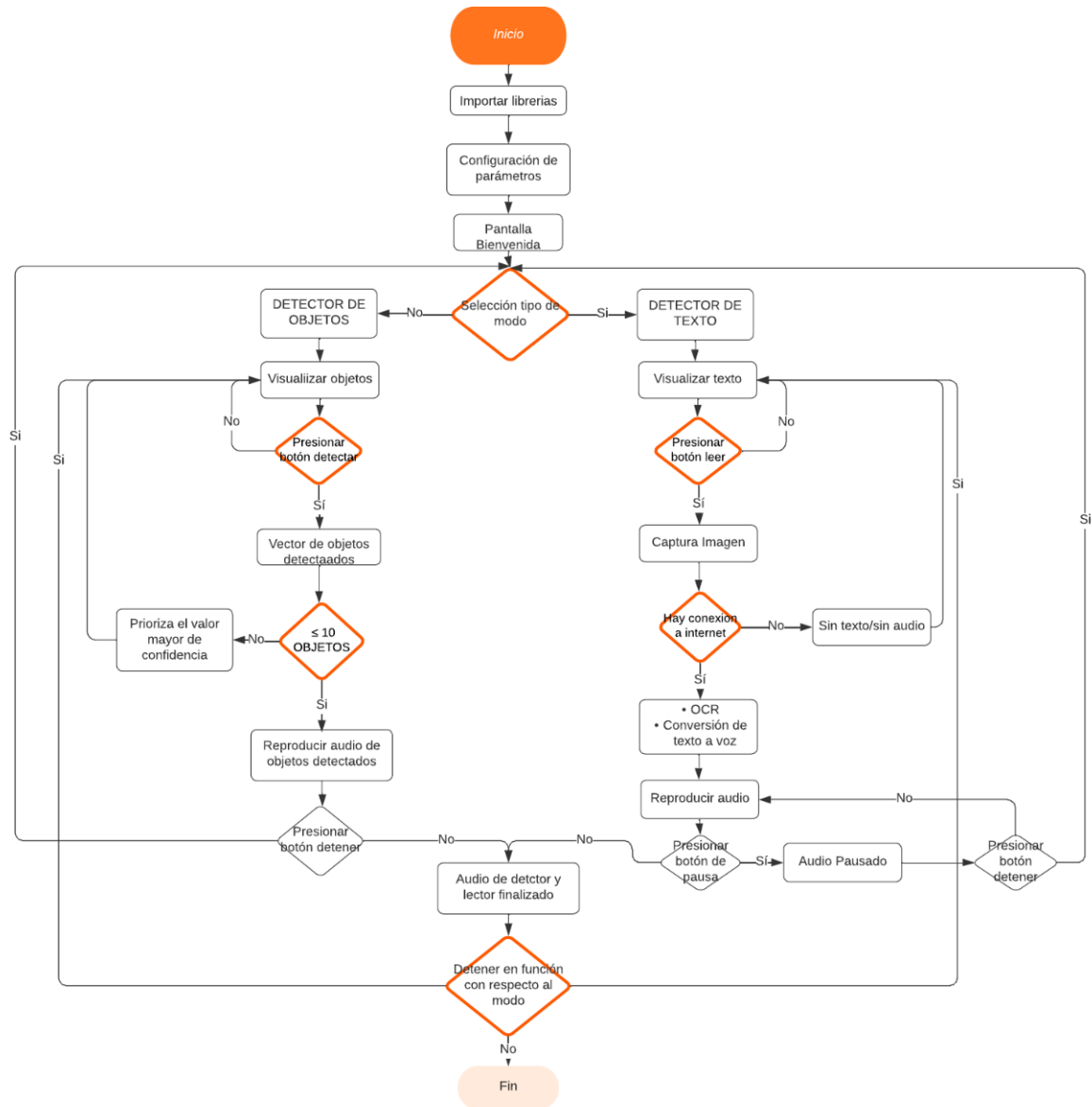
### Diagrama de flujo de la estructura del programa

El diagrama de flujo del dispositivo de las gafas inteligentes describe el proceso y los pasos que se implementó para poder alcanzar la solución del problema.

La implementación del diagrama de flujo del dispositivo para que funcione como lector o detector se indica en la figura 40, donde se detallan todos los procesos previos que se deben cumplir, selección de diferentes modos que el usuario puede utilizar para su correcto funcionamiento.

Figura 40

Diagrama de flujo para el modo detector de objetos y lector



Nota. Diagrama funcional del dispositivo gafas inteligentes. Tomado Autor

Para el funcionamiento del modo lector debe cumplir con configuraciones de parámetros: la ruta de las direcciones donde se van a guardar las imágenes, credenciales de la API Google visión y la configuración del ancho y alto de las imágenes y operaciones morfológicas en las imágenes.

Para obtener la credencial se debe registrar en Google Cloud para tener acceso de los recursos de la API de Google Vision, lo cual genera un archivo token.Json.

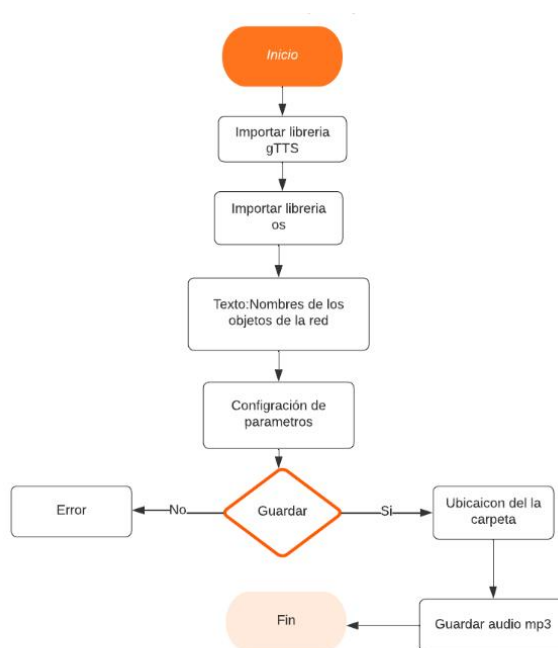
Para la implementación del algoritmo para la detección de objetos se debe configurar los parámetros de la red MobileNet V3; se debe descargar los modelos previamente entrenados con la base de datos de Coco.names, se descargan dos documentos, el primer archivo el modelo de la red MobileNet V3 y el segundo los pesos pertinentes de cada uno de los objetos.

### Creación de audios para la detección de objetos.

En esta parte se realiza la creación de 92 audios para cada uno de los objetos de la red MobileNet V3, con la finalidad que se guarden en un archivo mp3, permitiendo una reproducción de audio sin conexión a internet, la dirección de estos archivos se guarda en la carpeta principal del programa.

### Figura 41

*Algoritmo para la creación de diferentes audios*





*Nota.* Creación de audios de los diferentes objetos de la red MobileNet V3. Tomado del Autor

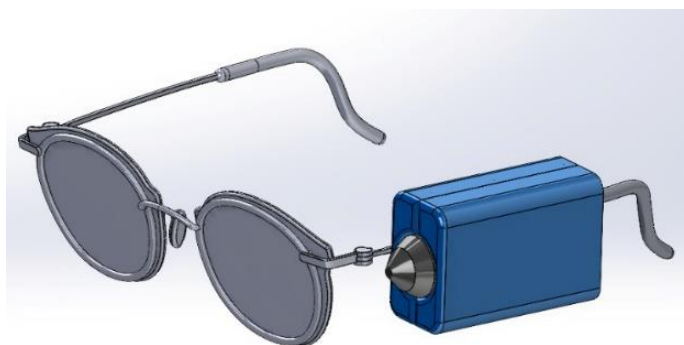
Para la reproducción del audio de cada objeto, el sistema podrá reproducir hasta 10 objetos en simultáneo optimizando el recurso computacional, la reproducción del audio se define mediante: `pygame.mixer.Sound`; es una biblioteca de audio que permite controlar la reproducción del audio a partir de los audios grabados en mp3, se reproduce mediante la activación del modo detectar.

### **Diseño e implementación del soporte para las gafas inteligentes.**

El diseño del soporte para las gafas inteligentes se realiza mediante la revisión de dispositivos para personas con discapacidad visual que detallan aspectos para un diseño accesible, dimensiones, ubicación del dispositivo al lado izquierdo del marco de los lentes, los componentes funcionales del soporte de las gafas inteligentes, la cámara estará acoplada en el soporte brindando funcionalidad y adaptación en cualquier tipo de lentes o gafas como se puede ver en la figura 42, brindando comodidad hacia las personas con discapacidad visual para su funcionamiento.

### **Figura 42**

*Conexión de soporte hacia los lentes*



*Nota.* Diseño del soporte para la cámara. Tomado del Autor

En el diseño para el soporte se consideran las dimensiones de las varillas que presentan los lentes, en este caso se determina que el dispositivo será más eficiente en varillas normales, permitiendo una sujeción más fija con respecto a otras varillas más pequeñas, se establece que para ocupar el dispositivo correctamente los lentes deben ser de la medida correcta de la persona con la finalidad que, al momento de utilizar, inclinar la cabeza este permanezca fijo y no tienda a caerse.

### **Diseño e implementación del case para las gafas inteligentes.**

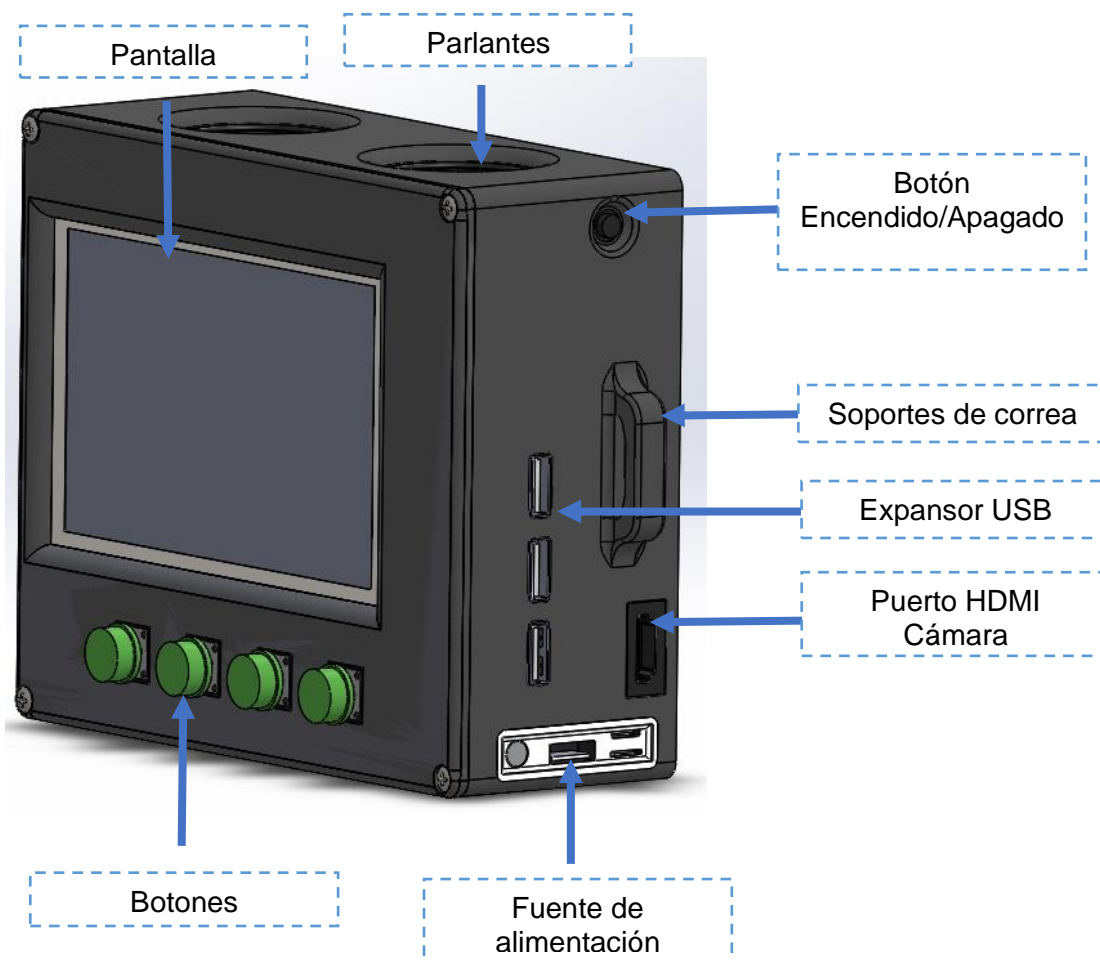
El diseño del case para las gafas inteligentes se realiza mediante la revisión de guías técnicas que detallan aspectos concretos para un diseño accesible, las dimensiones, ubicación de cada componente electrónico, ergonomía y la accesibilidad de las personas con discapacidad visual para su funcionamiento.

En donde se define la arquitectura del producto mediante los componentes funcionales de las gafas inteligentes, estos están acoplados en el case brindando la adaptación y accesibilidad mediante una arquitectura modular, como se puede ver en la figura 43.

La integración del sistema pretende la optimización y accesibilidad integral para el desarrollo de equipos tecnológicos de ayudas técnicas para personas con discapacidad visual, por tal motivo se incorporan pulsadores grandes con la finalidad de una manipulación sencilla con cualquiera de sus dedos de la mano, siendo un promedio el tamaño del dedo pulgar.

**Figura 43**

*Configuración de la carcasa para la instalación de los subsistemas*



*Nota.* Diseño del soporte para el sistema embebido

En la parte posterior del diseño se implementa rejillas de ventilación con la finalidad de reducir el calor que se genera dentro del dispositivo para obtener una temperatura adecuada, la ubicación de los parlantes se sugiere colocar en la parte superior para garantizar una salida de audio eficiente y clara.

## Resolución de la cámara UVC Bullet

La resolución de una fotografía digital es la cantidad de píxeles por unidad de longitud de ancho x alto, que es capaz la cámara de capturar en una imagen.

De esta manera con las especificaciones de la cámara para la implementación del dispositivo de las gafas inteligentes se tiene que la resolución del sensor que es de 1920 píxeles de largo y 1080 píxeles de ancho, se determina la resolución final que se puede obtener de la cámara digital mediante la siguiente ecuación.

$$\text{Resolución Real} = (\text{Píxeles de largo}) * (\text{Píxeles de ancho})$$

$$\text{Resolución Real} = (1920\text{píxeles}) * (1080\text{píxeles})$$

$$\text{Resolución Real} = 2073600$$

La resolución real que presenta la cámara UVC Bullet es de 2073600 píxeles lo cual no varía con respecto a las especificaciones de resolución de la cámara digital de 2 megapíxeles.

### Figura 44

*Tamaño de la foto para lectura y detección de objetos*

Sensor Megapíxeles y Resolución	Tamaño de la foto									
	2x3"	3x5"	4x6"	5x7"	6x8"	8x10"	11x14"	13x19"	16x20"	24x36"
1MP 1280 x 960	Profesional 427	Excelente 256	Bueno 213	Bueno 183	Bueno 160	Aceptable 120	Malo 87	Malo 67	Malo 60	Malo 36
2MP 1600 x 1200	Profesional 533	Excelente 320	Excelente 267	Bueno 229	Bueno 200	Aceptable 150	Aceptable 109	Malo 84	Malo 75	Malo 44
3MP 2048 x 1536	Profesional 683	Profesional 410	Excelente 341	Excelente 293	Excelente 256	Bueno 192	Aceptable 140	Aceptable 108	Malo 96	Malo 57
4MP 2464 x 1632	Profesional 816	Profesional 493	Profesional 408	Excelente 326	Excelente 272	Bueno 204	Aceptable 148	Aceptable 126	Aceptable 102	Malo 68
6MP 3008 x 2000	Profesional 1000	Profesional 602	Profesional 500	Profesional 400	Excelente 333	Excelente 250	Bueno 182	Bueno 154	Aceptable 125	Malo 83
8MP 3504 x 2336	Profesional 1168	Profesional 701	Profesional 584	Profesional 467	Profesional 389	Excelente 292	Bueno 212	Bueno 180	Aceptable 146	Malo 97
10MP 3872 x 2592	Profesional 1291	Profesional 774	Profesional 645	Profesional 518	Profesional 432	Excelente 324	Bueno 236	Bueno 199	Bueno 162	Aceptable 108
12MP 4288 x 2848	Profesional 1424	Profesional 858	Profesional 712	Profesional 570	Profesional 475	Profesional 356	Excelente 259	Bueno 219	Bueno 178	Aceptable 119
16MP 4992 x 3328	Profesional 1664	Profesional 998	Profesional 832	Profesional 666	Profesional 555	Profesional 416	Excelente 303	Excelente 256	Bueno 208	Aceptable 139
18MP 5232 x 3516	Profesional 1744	Profesional 1046	Profesional 872	Profesional 703	Profesional 586	Profesional 440	Excelente 320	Excelente 270	Bueno 220	Aceptable 145
21MP 5616 x 3744	Profesional 1872	Profesional 1123	Profesional 936	Profesional 749	Profesional 624	Profesional 468	Excelente 340	Excelente 288	Bueno 234	Bueno 156

*Nota.* Tamaño de la foto según su resolución de la cámara. Tomado de (Rodríguez, 2022)

Para determinar el campo visual de la cámara, se debe utilizar la siguiente fórmula que se deberá emplear con la resolución de la cámara y la resolución recomendada por la figura 44.

$$C.V = \frac{\text{Resolución del sensor}}{\text{Resolución sugerida}}$$

Se analiza el campo visual para el lado más largo del sensor.

$$C.V = \frac{1920 \text{ Píxeles}}{200 \text{ Píxeles/Pulg}} * \frac{2.54 \text{ cm}}{1 \text{ Pulg}} = 24.384 \text{ cm.}$$

Se analiza el campo visual para el lado más corto del sensor.

$$C.V = \frac{1080 \text{ Píxeles}}{200 \text{ Píxeles/Pulg}} * \frac{2.54 \text{ cm}}{1 \text{ Pulg}} = 13.716 \text{ cm.}$$

Con los valores obtenidos del campo visual para el lado más largo de 24.384 cm y del campo visual para el lado más corto de 13.716 cm, se determina que la cámara capturará fotografías aceptables para el pre-procesamiento de imágenes, con la finalidad que los objetos a analizar y el texto se encuentren dentro del campo visual para una detección aceptable en las dimensiones sugeridas (24.384 cm \* 13.716 cm).

### Fuerza de soporte de las gafas inteligentes

El material que ofrece altas prestaciones en resistencia es el PLA, que proporciona una calidad superficial aceptable y accesibilidad en la impresión en el mercado para prototipos.

**Tabla 30**

*Propiedades físicas del material PLA*

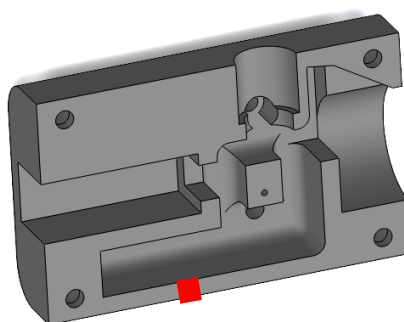
<b>Propiedades generales</b>	<b>Valores</b>
Módulo de Young	3,6 GPa
Coefficiente de Poisson	0,38
Límite elástico	72 MPa
Resistencia a la tracción	70 MPa

Propiedades generales	Valores
Resistencia a la compresión	86 MPa
Elongación	3-6 %
Resistencia a fatiga	27,7 MPa

### **Cálculo de la fuerza máxima**

#### **Figura 45**

*Análisis del esfuerzo máximo en la zona de menor espesor*



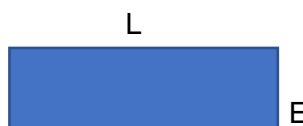
*Nota.* Análisis en el case del dispositivo

Para determinar la carga máxima que pueda deformar el componente se ubicará en la zona con menor espesor, de manera de ver cuánto es lo máximo que pueda resistir.

Cálculo de esfuerzos a compresión:

$$\sigma_{\text{Compresión}} = \frac{F_{\text{máxima}}}{A}$$

Donde A: es el área transversal perpendicular a la carga, para este caso sería un rectángulo a estudiar.



$$L = 1 \times 10^{-3} m$$

$$E = 1 \times 10^{-3} m$$

$$A = 1 \times 10^{-3} m * 1 \times 10^{-3} m$$

$$A = 1 \times 10^{-6} m^2$$

Para el esfuerzo de compresión se usará el esfuerzo admisible del material el cual es 86 MPa.

Sustituyendo en la ecuación:

$$\sigma_{Compresión} = \frac{F_{máxima}}{A}$$

$$86 \times 10^6 Pa = \frac{F_{máxima}}{1 \times 10^{-6} m^2}$$

$$F_{máxima} = 86 N$$

$$F_{máxima} = 86 N \approx 8.7 kg$$

Cálculo de la altura de caída:

$$F_{impacto} = \frac{1}{2} * \frac{mv^2}{D}$$

Donde:

M=masa

V= velocidad final

D= altura de frenado para este caso se considerará 1

Para Determinar la velocidad se aplicará la siguiente ecuación:

$$v = \sqrt{2gh}$$

Combinando las 2 ecuaciones:

$$F_{\text{impacto}} = \frac{1}{2} * \frac{m(\sqrt{2gh})^2}{D}$$

Sustituyendo:

$$8.7 \text{ kg} = \frac{1}{2} * \frac{0.05 \text{ Kg} \left( \sqrt{2 \left( \frac{9.8 \text{ m}}{\text{s}} \right) h} \right)^2}{1}$$

Despejando H:

$$h = 17.2 \text{ m}$$

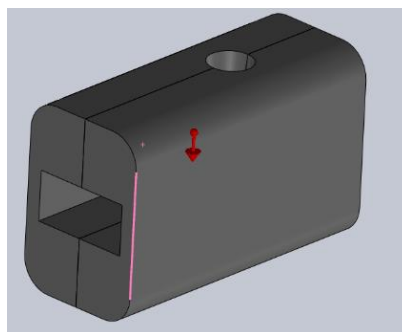
h es la altura máxima donde el prototipo es más propenso a fallar.

### **Análisis de elementos finitos**

Mediante la herramienta computacional CAE, se presentan los análisis de elementos finitos, el cual se estudia el impacto en caída libre del prototipo con la finalidad de simular el rendimiento del prototipo.

### **Figura 46**

*Análisis del esfuerzo máximo*





Nota. Análisis en el case del dispositivo

**Tabla 31**

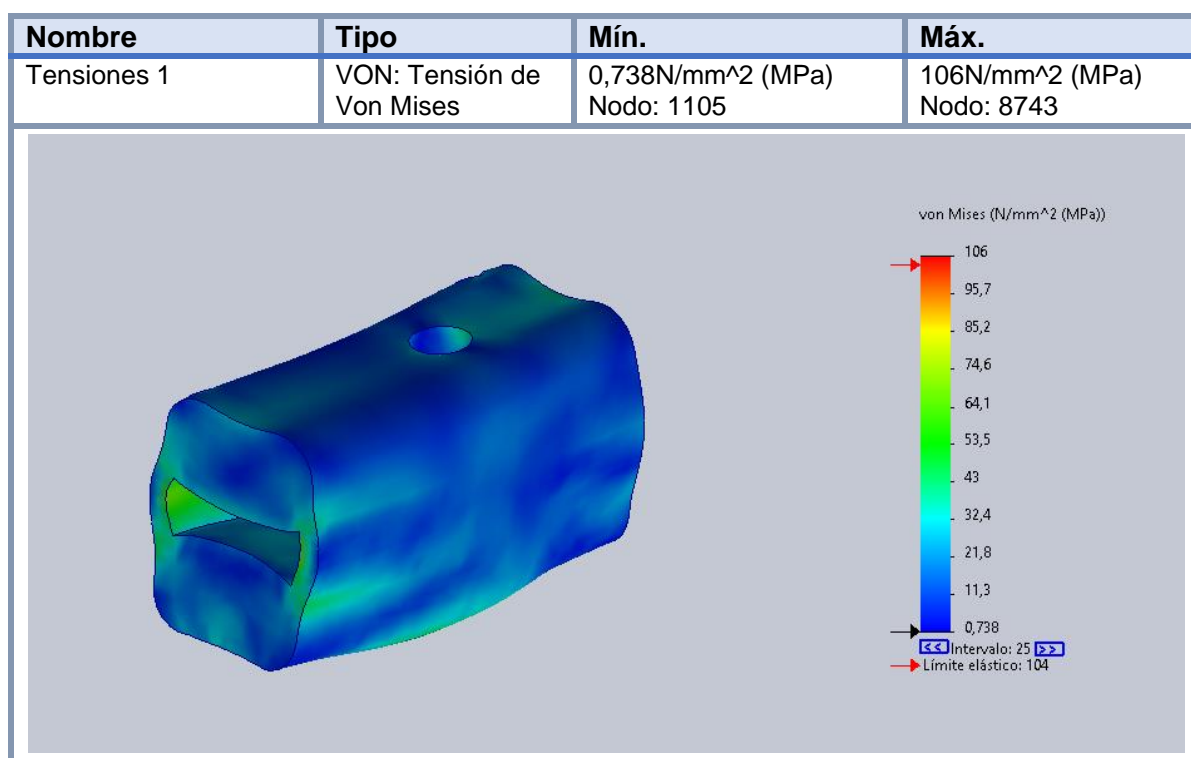
*Datos usados en el Software*

Altura de caída	14 m
Gravedad	9.81m/s
Tipo de suelo	Suelo rígido

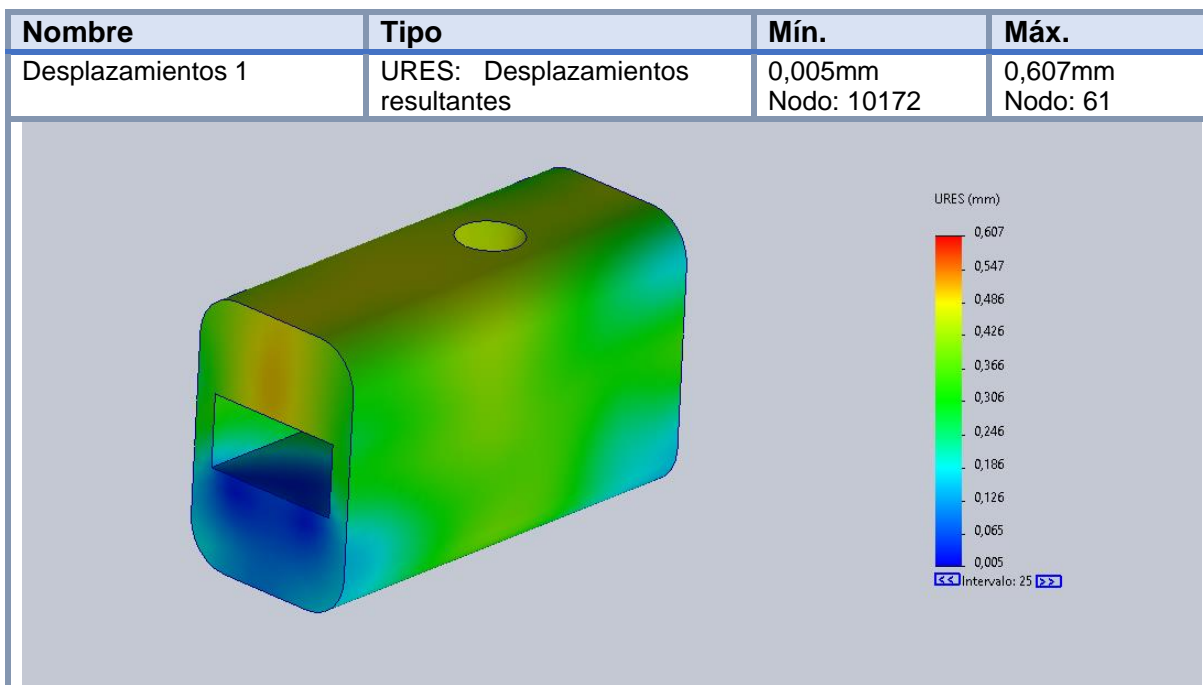
**Resultado de esfuerzo a tensión Von Mises**

**Figura 47**

*Resultado de esfuerzo a tensión Von Mises*



## Resultado Desplazamientos

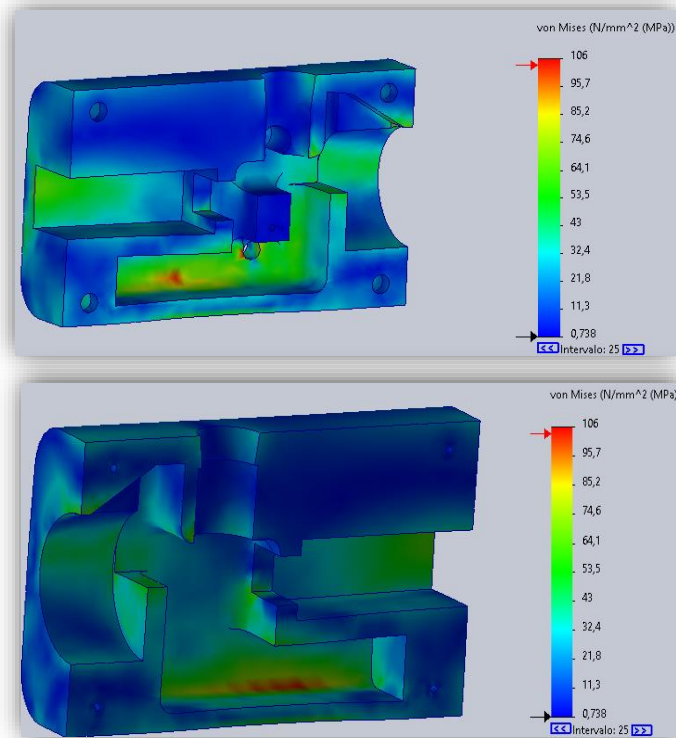


## Análisis de resultado de la simulación

En la figura 48, se muestra un esfuerzo máximo encontrado de 106 MPa el cual supera el esfuerzo de fluencia del PLA de 86 MPa, esto quiere decir que si el prototipo se cae a una altura de 17 m es muy probable que falle. Sin embargo, la figura 49 muestra zonas de color azul, donde existe muy poca concentración de esfuerzo, haciendo que el material no falle, por lo tanto, se puede concluir que existen zonas propensas a fallar a la caída mientras que las zonas donde se encuentre mayor área no fallará, para corresponder a lo dicho anteriormente se oculta una de las tapas del prototipo para ver como se encuentran los esfuerzos en la parte interna del prototipo.

**Figura 48**

Análisis del esfuerzo máximo a una altura de 17 m.



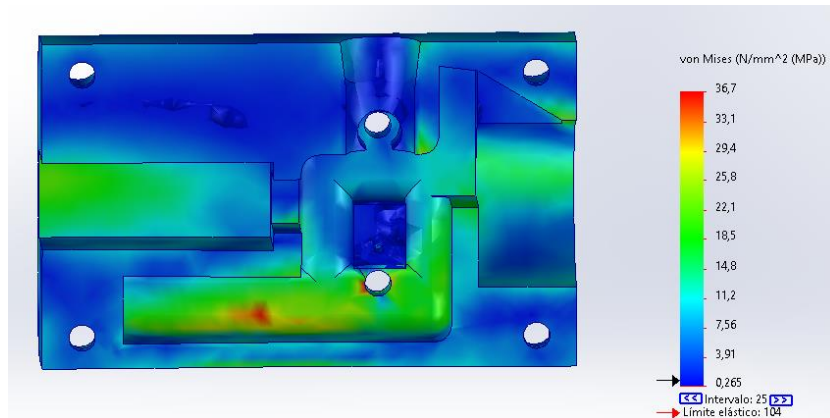
*Nota.* Análisis en el case del dispositivo a 17m. Tomado del Autor

En la figura 48, se puede apreciar que se encuentra de color rojo el esfuerzo de mayor magnitud conocido como concentradores de esfuerzos, esto quiere decir que esa zona deberá ser mejorada agregando más material PLA o eliminando aristas vivas.

## Comportamiento de caída libre del prototipo a una altura de 1.70m

**Figura 49**

*Análisis del esfuerzo máximo a una altura de 1.70 m*



*Nota.* Análisis en el case del dispositivo a 1.70 m. Tomado del Autor

Como se puede observar, el esfuerzo resultante máximo no supera el esfuerzo de fluencia del PLA a una altura de 1.7 m, el cual es la altura promedio de una persona donde se encuentra ubicado el prototipo garantizando su funcionalidad y resistencia.

Es importante resaltar que se tienen que hacer modificaciones de geometría para disminuir el concentrador de esfuerzo.

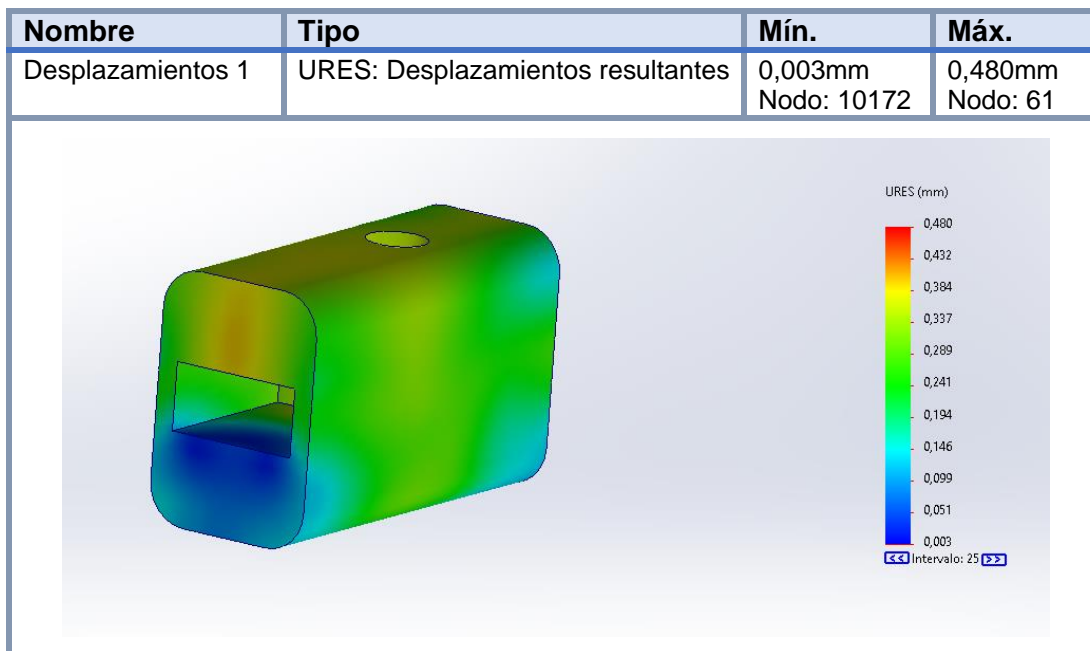
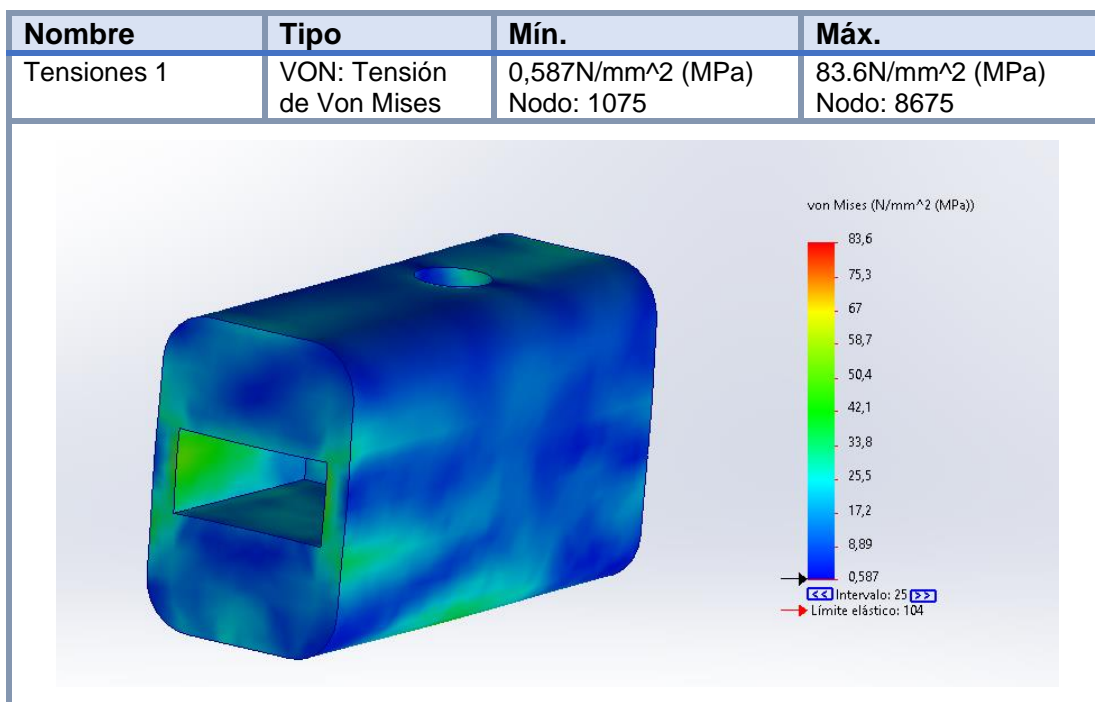
### **Altura real 14m donde se forma el esfuerzo máximo del material**

Como se puede evidenciar, el esfuerzo resultante máximo no supera el esfuerzo de fluencia del PLA a una altura de 14 m, el cual es la altura promedio que puede soportar el dispositivo haciendo que el material no falle, por lo tanto, se puede concluir que no existirán zonas propensas a fallar a la caída mientras que las zonas donde se encuentre mayor área no fallarán.

## Resultado de esfuerzo a tensión Von Mises

**Figura 50**

*Análisis del esfuerzo máximo a una altura de 14 m*



## Cálculo del clip sujetador

### Figura 51

*Análisis para determinar la fuerza máxima*



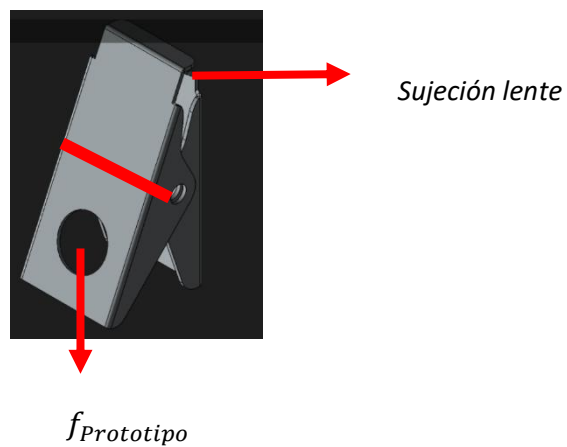
*Nota.* Análisis fuerza máxima soporta el clip. Tomado del Autor

A continuación, se presentarán los cálculos para determinar la fuerza máxima que puede soportar el clip, tiene un pasador que se someterá a cortante, el diámetro del pasador es de 2 mm, por lo tanto, se determina el esfuerzo máximo a cortante del eje de aluminio 6061 con esfuerzo de fluencia de 78 MPa.

### **Análisis de fuerzas**

### Figura 52

*Diagrama de cuerpo libre clip*



*Nota.* Análisis diagrama de cuerpo libre clip. Tomado del Autor

Cálculo de esfuerzos a cortante:

$$\sigma_{Cortante} = \frac{F_{m\acute{a}xima}}{A}$$

Donde el área se determina de la siguiente manera:

$$A = \frac{\pi * \phi^2}{4}$$

$$A = \frac{\pi * (2x10^{-3})^2}{4}$$

$$A = 3.14x10^{-6}$$

Sustituyendo:

$$78x10^6 Pa = \frac{F_{m\acute{a}xima}}{3.14x10^{-6}}$$

$$F_{m\acute{a}xima} = 244.92 N$$

La carga máxima que puede soportar el clip es de 24 Kg a cortante, garantizando una sujeción fija, siendo la carga del dispositivo menor que la calculada.

### **Escenas y ventanas de los modos de funcionamiento del dispositivo.**

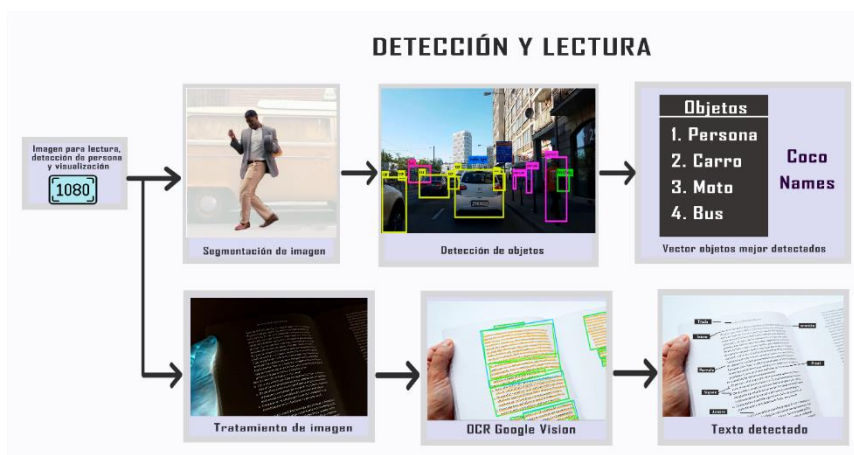
Se establece dos modos de funcionamiento independientes, el primer modo para la detección de objetos por medio de la red neuronal MobileNet V3, en la cual se realizan previamente la segmentación para la clasificación de los diferentes objetos que están disponibles en la red neuronal, lo cual permite la detección de cada uno de los objetos de la red y los puede mostrar en un cuadro con una etiqueta y un valor de confianza, se establece un vector de reconocimiento hasta 10 objetos, que tengan los valores más altos los cuales se ordenan de mayor a menor según su valor de confianza.

Para el segundo modo para la lectura se realiza un tratamiento de imágenes, que internamente se ejecutan en la API Google Cloud Vision, mediante un archivo token. Json; incluyen información de la página, el bloque, el párrafo, las palabras y las divisiones del texto y el idioma que se puede presentar en un texto.

La herramienta API de Google Cloud Vision, proporciona modelos preentrenados para la detección de textos impresos y escritos a mano, el cual presenta un recurso computacional más eficiente y una precisión cercana al 100%.

### Figura 53

*Procesamiento para lectura y detección de objetos*



### Creación del inicio

- **Portada**

La presentación de la portada donde se define el título del proyecto, se muestra la carátula un aproximado de 5 segundos.

Se establecen avisos de audio para que el usuario pueda usar el dispositivo.

1. Iniciando sistema por favor espere un momento
2. Sistema listo



Figura 54

Inicio Portada



- **Modos de Funcionamiento**

El dispositivo presenta dos modos de funcionamiento: Modo Detector de objetos/Modo lector, mediante 4 botones que permiten acceder a las siguientes funciones del dispositivo.

Figura 55

Modos de funcionamiento



En el presente panel se presentan 4 funciones del dispositivo.

1. “Modo”, se refiere al modo de funcionamiento del dispositivo como: detector/lector.
2. “Detectar”, reproduce a través de audio los objetos detectados en ese instante.  
“Leer”, reproduce el texto identificado en ese instante a través de audio”
3. “Pausa”, esta opción permite al usuario detener en cualquier punto la lectura y poder continuar en cualquier momento.
4. “Detener”, esta opción permite al dispositivo dejar de ejecutar cualquiera de los dos modos de funcionamiento.
5. Nota: Esta opción o modos también se pueden realizar por comandos de voz, los cuales el usuario puede ejecutar a través de comandos de voz por medio del usuario.

## Capítulo IV

### Pruebas y validación del dispositivo

#### Pruebas y análisis del funcionamiento del dispositivo

Con la finalidad de determinar el alcance del dispositivo de las gafas inteligentes, se aplican diversas pruebas de funcionamiento los cuales se detallan a continuación.

Para el modo de reconocimiento de objetos se establece diferentes escenas con diferentes objetos con la finalidad de estimar cuántos objetos es capaz de reconocer el dispositivo.

Para el modo de lectura se aplican diversas pruebas con diferentes textos: periódicos, revistas, documentos impresos, en diferentes formatos A4, A3; con el propósito de estimar las palabras correctas y las incorrectas, al momento de realizar el reconocimiento OCR.

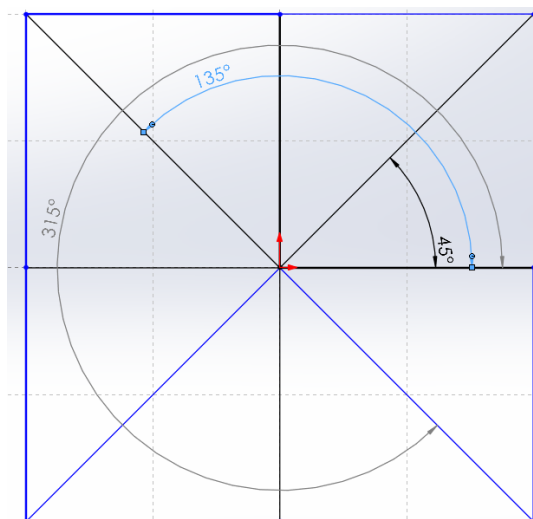
La autonomía del dispositivo se determina mediante el rendimiento del dispositivo: en sus modos de funcionamiento conectado al cargador y desconectado para determinar las horas de uso por el usuario, determinar el tiempo de carga/descarga del dispositivo para su correcto funcionamiento.

#### ***Pruebas del modo de lectura.***

Se establecen diferentes pruebas con diferentes organizaciones del texto; tamaño de fuente, tipo de fuente y la distancia del reconocimiento con el propósito de determinar la cantidad de palabras correctas e incorrectas, de esta manera identificar el porcentaje de error que genera las gafas inteligentes, en las siguientes tablas se observan las pruebas realizadas.

## Figura 56

*Disposición de la hoja en diferentes ángulos*



*Nota.* Referencia de ángulos para diferentes pruebas

### Explicación prueba 1:

El tipo de texto que se tiene para realizar las diferentes pruebas, es un periódico que presentan las siguientes características: hoja A3, tamaño de fuente 10, tipo de letra Century Old Style, gráficos presentes en el documento, presenta 10 párrafos en el documento. A una distancia de enfoque para el reconocimiento de 35 cm y 40cm.

Por lo general este tipo de documentos se presentan en un tamaño de fuente entre 10 y 13 que es un tamaño promedio, en su gran parte estos documentos se encuentran con letra Times New Roman y Century Old Style.

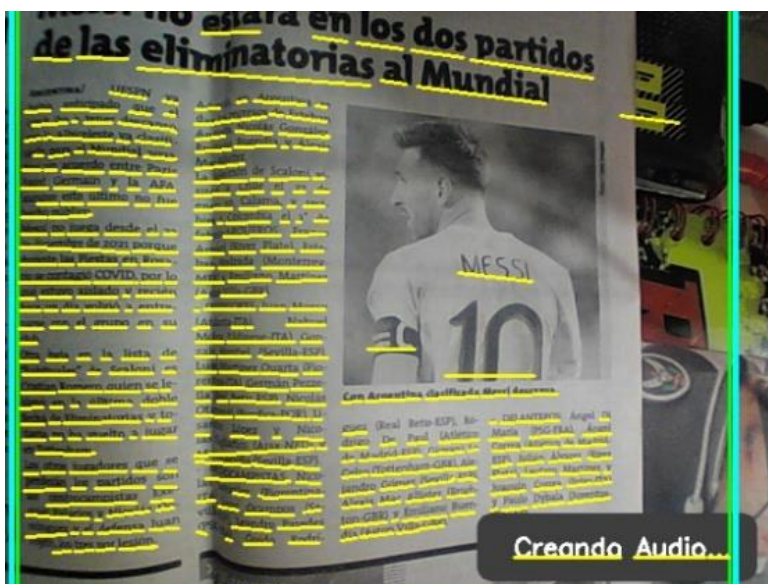
Tabla 32

Pruebas del reconocimiento OCR con letra número 10

Prueba	Palabras			% de error
	Totales	Detectadas	No Detectadas	
Prueba Vertical	463	447	16	3,45572354
Prueba Horizontal	463	449	14	3,0237581
Prueba 45°	463	438	25	5,39956803
Prueba 135°	463	430	33	7,12742981
Prueba 315°	463	450	13	2,80777538
Total	2315	2214	101	4,36285097

Figura 57

Pruebas del reconocimiento OCR con letra número 10



Realizadas diferentes pruebas se determina que el detector presenta un error de detección del 4,36 %. A una distancia mayor a 40 cm, debido al tamaño de fuente número 10, omitiendo algunas palabras para la detección y también presentado una mala interpretación entre las letras I, L; no interpreta los acentos en las palabras con tilde.

## Explicación prueba 2:

El tipo de texto que se tiene para realizar las diferentes pruebas, es un documento impreso que presentan las siguientes características: hoja A4, tamaño de fuente 11, tipo de letra Times New Roman, presenta gráficos en el documento en la parte superior derecha, presenta 5 párrafos en el documento. A una distancia de enfoque de 36 cm.

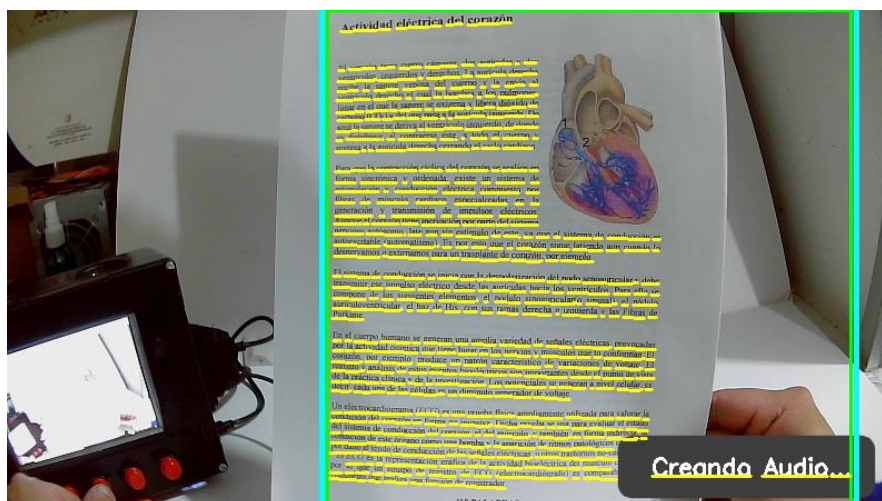
**Tabla 33**

*Pruebas del reconocimiento OCR con letra número 11*

Prueba	Palabras			% de error
	Totales	Detectadas	No detectadas	
Prueba Vertical	418	406	12	2,8708134
Prueba Horizontal	418	403	15	3,58851675
Prueba 45°	418	390	28	6,69856459
Prueba 135°	418	397	21	5,02392344
Prueba 315°	418	401	17	4,06698565
Total	2090	1997	93	4,44976077

**Figura 58**

*Pruebas de validación con el reconocimiento OCR, con letra número 11*



Realizadas diferentes pruebas se determina una buena detección en el reconocimiento de texto a una distancia de 40 cm, debido al tamaño de fuente número 11, existiendo una mala interpretación en signos de puntuación que fueron evidentes en estas pruebas ya que se las realizo en una hoja impresa donde el texto no se encuentra justificado.

### Explicación prueba 3:

El tipo de texto que se tiene para realizar las diferentes pruebas, es un documento impreso que presentan las siguientes características: hoja A4, tamaño de fuente 12, tipo de letra Times New Roman, no presenta gráficos en el documento, un solo párrafo con texto seguido. A una distancia de enfoque de 25 cm y 35 cm.

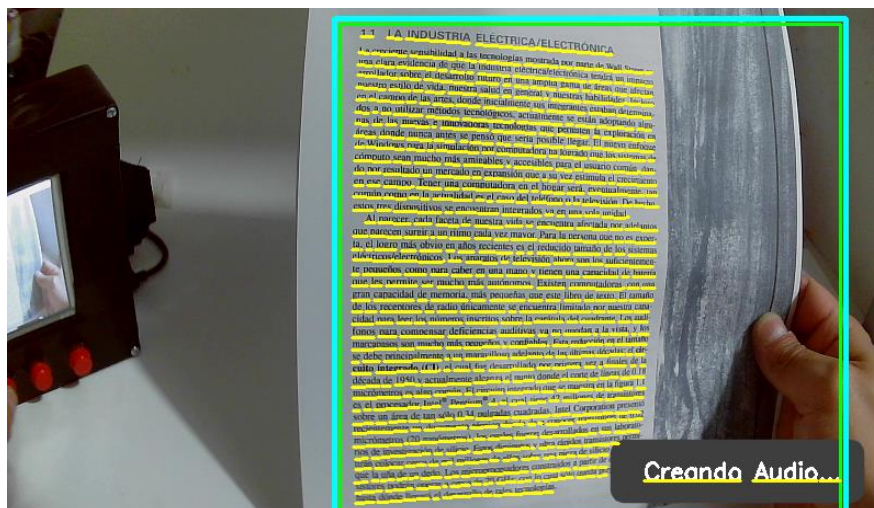
**Tabla 34**

*Pruebas del reconocimiento OCR con letra número 12*

Prueba	Palabras			% de error
	Totales	Detectadas	No Detectadas	
Prueba Vertical	450	435	15	3,33333333
Prueba Horizontal	450	437	13	2,88888889
Prueba 45°	450	435	15	3,33333333
Prueba 135°	450	435	15	3,33333333
Prueba 315°	450	439	11	2,44444444
Total	2250	2181	69	3,06666667

**Figura 59**

*Pruebas de validación con el reconocimiento OCR, con letra número 12*



Realizadas diferentes pruebas se determina una buena detección del texto de interés para su análisis, debido al tamaño de fuente número 12, el detector presenta limitaciones en la detección de palabras con separación de guion (-), ya que detecta como palabras separadas debido a la disposición del texto que se encuentra justificado.

#### **Explicación prueba 4:**

El tipo de texto que se tiene para realizar las diferentes pruebas, es un libro que presentan las siguientes características: hoja A3, disposición de texto ambas caras, tamaño de fuente 14, tipo de letra Times New Roman cursiva, presenta gráficos en centro del documento y presentes 4 párrafos en el documento. A una distancia de enfoque de 35 cm y 40 cm.

El tipo de tipografía que se presentan en libros impresos y revistas es mediante la fuente Times New Roman.



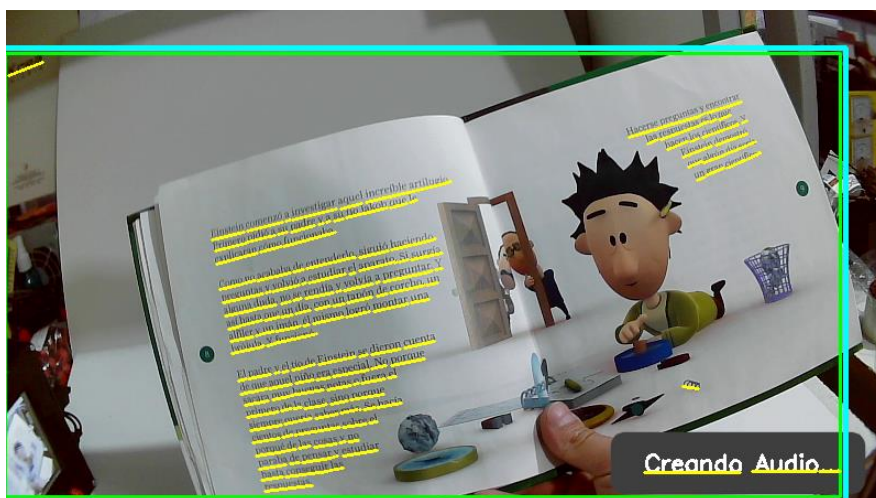
Tabla 35

Pruebas del reconocimiento OCR con letra número 14

Prueba	Palabras			% de error
	Totales	Detectadas	No Detectadas	
Prueba Vertical	164	160	4	2,43902439
Prueba Horizontal	164	160	4	2,43902439
Prueba 45°	164	161	3	1,82926829
Prueba 135°	164	159	5	3,04878049
Prueba 315°	164	157	7	4,26829268
Total	820	797	23	2,80487805

Figura 60

Pruebas de validación con el reconocimiento OCR, con letra número 14



Realizadas diferentes pruebas se determina que el detector funciona de manera eficiente, presentando una buena detección del 100 % del texto, omitiendo las imágenes presentes en el texto, debido al tamaño de fuente número 14, el error es mínimo debido que reconoce una figura como letra (detecta el resorte y la interpreta como L), el detector interpreta palabras que no se encuentran justificas, priorizando la lectura mediante nuestro alfabeto de izquierda a derecha.

## Explicación prueba 5:

El tipo de texto que se tiene para realizar las diferentes pruebas, es un libro que presentan las siguientes características: hoja A3, disposición de texto ambas caras, tamaño de fuente 16, presenta gráficos en centro del documento, con presencia de caracteres especiales y presentes 6 párrafos en el documento. A una distancia de enfoque de 35 cm y 40 cm.

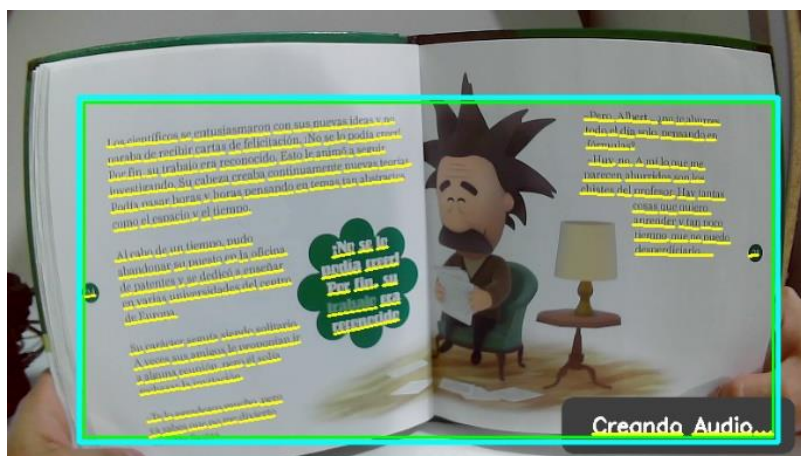
**Tabla 36**

*Pruebas del reconocimiento OCR ambas caras*

Prueba	Palabras			% de error
	Totales	Detectadas	No Detectadas	
Prueba Vertical	170	163	7	4,11764706
Prueba Horizontal	170	161	9	5,29411765
Prueba 45°	170	160	10	5,88235294
Prueba 135°	170	163	7	4,11764706
Prueba 315°	170	162	8	4,70588235
Total	850	809	41	4,82352941

**Figura 61**

*Pruebas de validación con el reconocimiento OCR ambas caras*



Realizadas diferentes pruebas se determina que el detector funciona de manera eficiente, presentando una detección del 100 % del texto omitiendo las imágenes presentes en el texto, debido al tamaño de fuente número 16, el error es mínimo debido a la interpretación y reconocimiento del número de la página, el detector interpreta palabras que no se encuentran justificadas y signos de puntuación, priorizando la lectura mediante nuestro alfabeto de izquierda a derecha.

### Explicación prueba 6:

El tipo de texto que se tiene para realizar las diferentes pruebas, es un documento impreso que presentan las siguientes características: hoja A4 a doble columna, tamaño de fuente 11,14 y 24, tipo de letra Times New Roman cursiva, presenta gráficos en la parte superior izquierda de documento y presenta 10 párrafos en el documento. A una distancia de enfoque de 30 cm y 35 cm.

**Tabla 37**

*Pruebas del reconocimiento OCR con texto a doble columna*

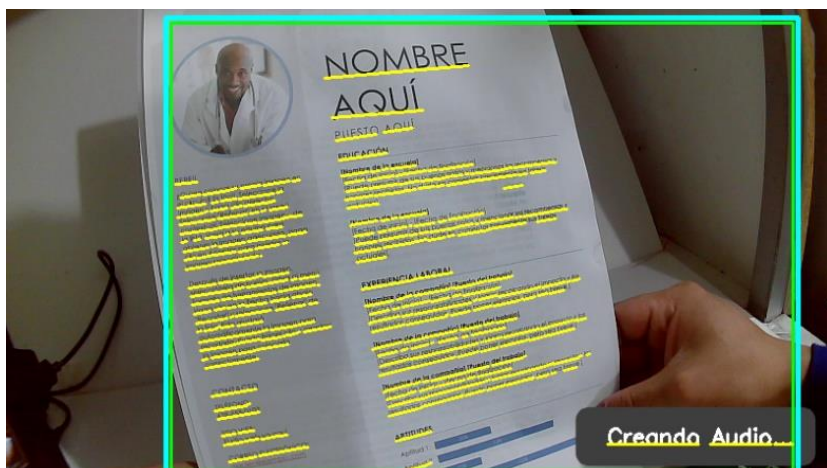
Prueba	Palabras			% de error
	Totales	Detectadas	No Detectadas	
Prueba Vertical	326	311	15	4,60122699
Prueba Horizontal	326	312	14	5,21472393
Prueba 45°	326	312	14	4,29447853
Prueba 135°	326	309	17	4,29447853
Prueba 315°	326	310	16	4,90797546
Total	1630	1554	76	4,66257669

Realizadas diferentes pruebas se determina que el detector, presenta ciertos errores de reconocimiento del texto, debido a diferentes tamaños de fuentes, el detector presenta cierta

dificultad al momento de identificar tablas, produciendo un error del 4,66% cuando el texto se ubica con ángulos de 135 y 315 grados en el reconocimiento de texto, con una mala interpretación de las siguientes letras i y l; el detector es capaz de identificar texto a doble columna y priorizar la lectura mediante nuestro alfabeto de izquierda a derecha.

## Figura 62

*Pruebas de validación con el reconocimiento OCR texto doble columna*



## Explicación prueba 7

El tipo de texto que se tiene para realizar las diferentes pruebas, es una revista que presenta las siguientes características: hoja A3 a doble cara, tamaño de fuente 12,14,16,20 y 24, tipo de letra Times New Roman cursiva y Bell Mt, presenta gráficos en diferentes partes del documento, presenta 5 párrafos en el documento, el texto no se encuentra justificado y con una distancia de enfoque de 35 cm y 40 cm.

El tipo de tipografía que se presentan en libros impresos y revistas es mediante la fuente Times New Roman y Bell Mt.

Tabla 38

Pruebas del reconocimiento OCR con imágenes presente en el texto y letra de diferente tamaño

Prueba	Palabras			% de error
	Totales	Detectadas	No Detectadas	
Prueba Vertical	214	205	9	4,20560748
Prueba Horizontal	214	203	11	5,14018692
Prueba 45°	214	207	7	3,27102804
Prueba 135°	214	205	9	4,20560748
Prueba 315°	214	206	8	3,73831776
Total	1070	1026	44	4,11214953

Figura 63

Pruebas de validación con el reconocimiento OCR, con presencia de imágenes en el texto



Realizadas diferentes pruebas se determina que el detector, presenta una eficiencia del 100 % en la detección del texto, debido a diferentes tamaños de fuentes facilita la detección, se puede visualizar que si una imagen presenta texto el detector es capaz de reconocer de una manera eficiente, se puede detectar diferentes tipos de letras de una manera óptima y priorizar la lectura de izquierda a derecha.

## Validación de la hipótesis

Para validar las diferentes pruebas mediante chi cuadrado, donde se determina si la hipótesis aplicada en la investigación cumple o no, la hipótesis alternativa H1 y la hipótesis nula H0.

H1: ¿El diseño e implementación de gafas inteligentes podrá ser considerada como una herramienta de asistencia para lectura y reconocimiento del entorno en el que se desempeñan las personas con discapacidad visual?

H0: Mediante la implementación de gafas inteligentes, no se considera como una herramienta de asistencia para lectura y reconocimiento del entorno en el que se desempeñan las personas con discapacidad visual.

Para la validez de la hipótesis mediante el método de validación probabilístico chi cuadrado con el propósito de cuantificar el porcentaje de aciertos de las palabras detectadas, si influye el tamaño de fuente, organización y distribución del contenido que se puede encontrar en los diferentes textos como herramienta de asistencia para lectura y estimar el error de palabras que el dispositivo no detecta, así determinar si la hipótesis cumple o no.

**Tabla 39**

*Pruebas del reconocimiento OCR Distribución Chi cuadrado*

Prueba	Palabras			%Error
	Totales	Detectadas	No Detectadas	
Lectura letra n°10	2315	2214	101	4,36285097
Lectura letra n°11	2090	1997	93	4,44976077
Lectura letra n°12	2250	2181	69	3,06666667
Lectura letra n°14	820	797	23	2,80487805
Lectura y organización del texto a ambas caras	850	809	41	4,82352941
Lecturas a doble columna	1630	1554	76	4,66257669
Lectura diferente tamaños de fuentes de letras	1070	1026	44	4,11214953

TOTAL	11025	10578	447	4,04034458
-------	-------	-------	-----	------------

Distribución del chi cuadrado. Lo cual determinan las frecuencias teóricas, utilizando la siguiente ecuación.

$$ft = \frac{(Total\ de\ columnas) * (Total\ de\ filas)}{Total\ de\ palabras}$$

Frecuencias teóricas que se estima para los datos obtenidos para la lectura con letra número 10.

$$ft\ 2214 = \frac{(10578) * (2315)}{11025} = 2221.1401$$

$$ft\ 101 = \frac{(447) * (2315)}{11025} = 93.8598$$

Frecuencias teóricas estimadas para lectura con letra número 11.

$$ft\ 1997 = \frac{(10578) * (2090)}{11025} = 2005.2625$$

$$ft\ 93 = \frac{(447) * (2090)}{11025} = 84.7374$$

Frecuencias teóricas estimadas para lectura con letra número 12.

$$ft\ 2181 = \frac{(10578) * (2250)}{11025} = 2158.7755$$

$$ft\ 69 = \frac{(435) * (2250)}{11025} = 91.2244$$

Frecuencias teóricas estimadas para lectura con letra número 14.

$$ft\ 797 = \frac{(10578) * (820)}{11025} = 7826.7537$$

$$ft\ 23 = \frac{(447) * (820)}{11025} = 33.2462$$

Frecuencias teóricas estimadas para una distribución a ambas caras.

$$ft_{809} = \frac{(10578) * (850)}{11025} = 815.5374$$

$$ft_{41} = \frac{(447) * (850)}{11025} = 34.4625$$

Frecuencias teóricas estimadas para lectura a doble columna.

$$ft_{1554} = \frac{(10578) * (1630)}{11025} = 1563.9129$$

$$ft_{76} = \frac{(447) * (1630)}{11025} = 66.0870$$

Frecuencias teóricas estimadas para lectura de diferente tamaño de fuente 12,14,16,20 y 24,

$$ft_{1026} = \frac{(10578) * (1070)}{11025} = 1026.6176$$

$$ft_{44} = \frac{(447) * (1070)}{11025} = 43.3823$$

Se debe determinar el grado de libertad de la tabla 16, de los datos mediante la distribución chi cuadrado, se utiliza la siguiente fórmula donde,  $V$  es el grado de libertad.

$$V = (\#de\ filas - 1) * (\#de\ columnas - 1)$$

Con los datos de la tabla 16, se obtiene que existen 7 filas de datos por 2 columnas de datos, el grado de libertad es el siguiente.

$$V = (7 - 1) * (2 - 1) = 6$$

Mediante los datos obtenidos se determina el chi cuadrado con la siguiente fórmula.

$$x^2 = \sum \frac{(f - ft)^2}{ft}$$



$$\begin{aligned}
 x^2 = & \frac{(2214 - 2221.1401)^2}{2221.1401} + \frac{(101 - 93.8598)^2}{93.8598} + \frac{(1997 - 2005.2625)^2}{2005.2625} + \frac{(93 - 84.7374)^2}{84.7374} \\
 & + \frac{(2181 - 2158.7755)^2}{2158.7755} + \frac{(69 - 91.2244)^2}{91.2244} + \frac{(797 - 786.7537)^2}{786.7537} \\
 & + \frac{(23 - 33.2462)^2}{33.2462} + \frac{(809 - 815.5374)^2}{815.5374} + \frac{(41 - 34.4625)^2}{34.4625} \\
 & + \frac{(1554 - 1563.9129)^2}{1563.9129} + \frac{(76 - 66.0870)^2}{66.0870} + \frac{(1026 - 1026.6176)^2}{1026.6176} \\
 & + \frac{(44 - 43.3823)^2}{43.3823}
 \end{aligned}$$

$$\begin{aligned}
 x^2 = & 0.02295 + 0.54317 + 0.03405 + 0.80567 + 0.22880 + 5.41442 + 0.13344 + 3.15782 \\
 & + 0.05240 + 1.24012 + 0.06283 + 1.48692 + 0.00037 + 0.00879
 \end{aligned}$$

$$x^2 \text{ calculada} = 13.19176$$

**Figura 64**

*Distribución Chi Cuadrado*

v/p	0,001	0,0025	0,005	0,01	0,025	0,05	0,1	0,15	0,2	0,25	0,3	0,35
1	10,8274	9,1404	7,8794	6,6349	5,0239	3,8415	2,7055	2,0722	1,6424	1,3233	1,0742	0,8735
2	13,8150	11,9827	10,5965	9,2104	7,3778	5,9915	4,6052	3,7942	3,2189	2,7726	2,4079	2,0996
3	16,2660	14,3202	12,8381	11,3449	9,3484	7,8147	6,2514	5,3170	4,6416	4,1083	3,6649	3,2831
4	18,4662	16,4238	14,8602	13,2767	11,1433	9,4877	7,7794	6,7449	5,9886	5,3853	4,8784	4,4377
5	20,5147	18,3854	16,7496	15,0863	12,8325	11,0705	9,2363	8,1152	7,2893	6,6257	6,0644	5,5731
6	22,4575	20,2491	18,5475	16,8119	14,4494	12,5916	10,6446	9,4461	8,5581	7,8408	7,2311	6,6948
7	24,3213	22,0402	20,2777	18,4753	16,0128	14,0671	12,0170	10,7479	9,8032	9,0371	8,3834	7,8061
8	26,1239	23,7742	21,9549	20,0902	17,5345	15,5073	13,3616	12,0271	11,0301	10,2189	9,5245	8,9094

*Nota.* Tabla de distribución Chi cuadrado. Tomado de (Cátedra: Probabilidad y Estadística , 2023)

Calculado el valor de chi cuadrado que es  $x^2 \text{ calculada} = 13.19176$ , se compara con la tabla de distribución de chi cuadrado de la figura 64, este valor se obtiene del grado de libertad calculado que es  $V = 6$ , con un margen de error de 0.1.

Con estos dos valores encontrados del chi calculado de 13.19176 y el valor de chi cuadrado de tabla de 10.6446, estos valores se comparan para establecer cuál de las dos hipótesis cumple, alternativa o nula para lo cual debe cumplir la siguiente condición, que el chi calculado debe ser mayor al valor del chi cuadrado de la tabla, la hipótesis alternativa cumple, caso contrario se establece que la hipótesis nula es valedera.

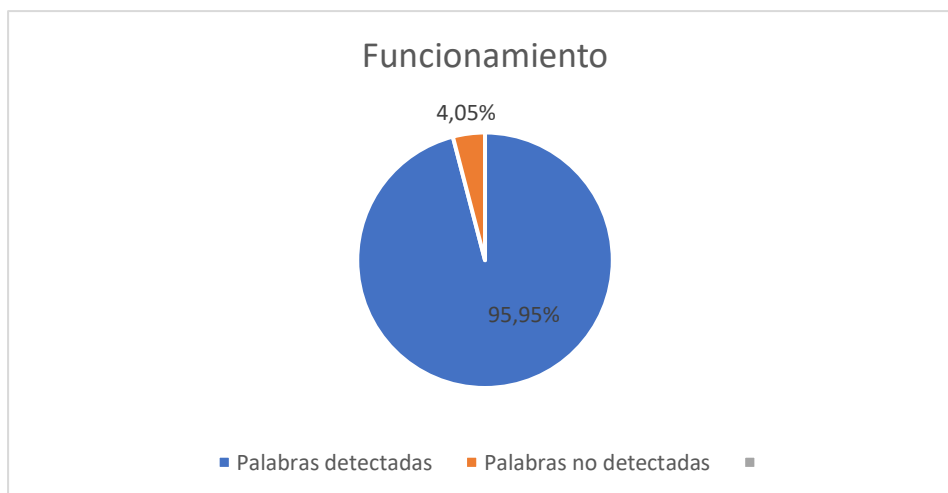
$$x^2_{calculada} > x^2_{tabla}$$

$$x^2_{13.19176} > x^2_{10.6446}$$

De acuerdo al valor obtenido del chi cuadrado calculado se puede determinar que es mayor que el chi cuadrado de la tabla, por lo cual la hipótesis alternativa del proyecto de investigación cumple y se da validez al dispositivo como una herramienta de asistencia para lectura para las personas con discapacidad visual.

### Figura 65

*Funcionalidad del dispositivo según los datos obtenidos de las diferentes pruebas.*



Una vez que se realizaron diferentes pruebas con diferentes tipos de textos; periódicos, folletos, libros y hojas impresas en diferentes tamaños de fuente se puede determinar que el dispositivo tiene una eficacia del 95.94 % en la detección mediante OCR y existe un error del

4.05 % el mismo que se produce en la detección de caracteres especiales, signos de puntuación o cuando las palabras se encuentran incompletas y presentan un salto de línea, lo cual el detector las detecta como palabras por separado, generalmente las palabras que se encuentran con un acento ortográfico, no las reconoce ni las interpreta a una distancia considerable.

### ***Pruebas del modo reconocimiento del entorno***

Para el reconocimiento del entorno se realizan diferentes pruebas con diferentes objetos cotidianos, en diferentes escenas; en un cuarto, en la vía pública y en la noche.

Para determinar si el detector funciona de una manera eficiente se realizaron diversas pruebas con diferentes condiciones de luz, diferentes escenarios para poder determinar la eficiencia del detector y cuales vendrían hacer las limitaciones que pueden presentar; garantizar que el dispositivo pueda reconocer aleatoriamente 10 objetos de una base de datos de 92 objetos.

### **Explicación prueba dormitorio, vía pública y noche:**

Esta prueba se realiza en 3 escenarios.

Primer escenario en un cuarto de 2,5 m de ancho por 3 m de largo, luz natural, objetos medianos: Tv, botella, silla, escritorio, persona, laptop, cama, maleta, libro y mouse.

Segundo escenario en la vía pública, luz natural. Objetos a identificar: Bus, auto, personas, moto, perro, semáforos, maceta, banca y bicicletas.

Tercer escenario en un cuarto de 2,5 m de ancho por 3 m de largo, luz artificial (noche). Objetos a identificar: Tv, botella, silla, escritorio, persona, laptop, cama, maleta, libro y mouse.

**Figura 66**

*Objetos identificados en un cuarto en la noche con luz cálida prueba 13*

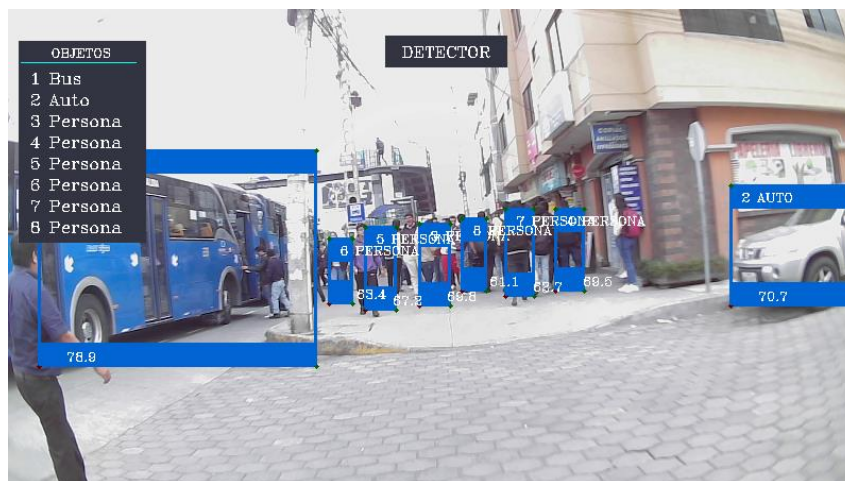


El resultado de la detección de objetos se realiza en la vía pública donde se puede visualizar que el detector reconoce de manera eficiente los objetos que encuentra en análisis, el detector es capaz de identificar 10 objetos en simultáneo, en la figura 67 se evidencia la detección del 80 % de objetos con un de error de 20% en la detección.

Debido que en la vía pública resulta difícil poder encontrar 10 objetos diferentes y en simultáneo se considera realizar pruebas con presencia de objetos repetitivos como se puede ver en la figura 67.

**Figura 67**

*Objetos identificados en la vía pública prueba 6*



El reconocimiento de objetos en las diferentes escenas se considera pertinente combinar objetos grandes y pequeños; la ubicación de los objetos varía dependiendo de las pruebas, distancia que se encuentra los objetos, con la finalidad que el detector sea capaz de detectar de una manera eficiente.

**Tabla 40**

*Pruebas en diferentes escenas del reconcomiendo de objetos*

Objeto	Clasificación			% de error
	Total, de objetos	Objetos identificados	Objetos sin identificar	
Prueba 1: Escenario 1	10	7	3	30
Prueba 2: Escenario 1	10	9	1	10
Prueba 3: Escenario 1	10	7	3	30
Prueba 4: Escenario 1	10	8	2	20
Prueba 5: Escenario 1	10	6	4	40
Prueba 6: Escenario 2	10	8	2	20
Prueba 7: Escenario 2	10	7	3	30
Prueba 8: Escenario 2	10	7	3	30
Prueba 9: Escenario 2	10	5	5	50
Prueba 10: Escenario 2	10	8	2	20
Prueba 11: Escenario 3	10	4	6	60

Objeto	Clasificación			% de error
	Total, de objetos	Objetos identificados	Objetos sin identificar	
Prueba 12: Escenario 3	10	4	6	60
Prueba 13: Escenario 3	10	3	7	70
Prueba 14: Escenario 3	10	4	6	60
Prueba 15: Escenario 3	10	4	6	60
Total	150	91	59	39,3333333

Distribución chi cuadrado, determinamos las frecuencias teóricas utilizando la siguiente ecuación.

$$ft = \frac{(Total\ de\ columnas) * (Total\ de\ filas)}{Total\ de\ objetos}$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en una habitación de 3m de largo por 2.50 de ancho

$$ft\ 7 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 3 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en una habitación de 3m de largo por 2.50 de ancho

$$ft\ 9 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 1 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en una habitación de 3m de largo por 2.50 de ancho

$$ft\ 7 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 3 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en una habitación de 3m de largo por 2.50 de ancho

$$ft\ 8 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 2 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en una habitación de 3m de largo por 2.50 de ancho

$$ft\ 6 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 4 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la vía pública.

$$ft\ 8 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 2 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la vía pública.

$$ft\ 7 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 3 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la vía pública.

$$ft\ 7 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 3 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la vía pública.

$$ft\ 5 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 5 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la vía pública.

$$ft\ 8 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 2 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la noche en una habitación y en la vía pública.

$$ft\ 4 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 6 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la noche.

$$ft\ 4 = \frac{(91) * (10)}{150} = 6.0666$$



$$ft\ 6 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la noche.

$$ft\ 3 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 7 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la noche.

$$ft\ 4 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 6 = \frac{(59) * (10)}{150} = 3.9333$$

Frecuencias teóricas que se estima para los datos obtenidos para el reconocimiento de objetos en la noche.

$$ft\ 4 = \frac{(91) * (10)}{150} = 6.0666$$

$$ft\ 6 = \frac{(59) * (10)}{150} = 3.9333$$

Se debe determinar el grado de libertad de la tabla 18, de los datos mediante la distribución chi cuadrado, se utiliza la siguiente fórmula donde  $V$ , es el grado de libertad.

$$V = (\#de\ filas - 1) * (\#de\ columnas - 1)$$

Con los datos de la tabla 18, se obtiene que existen 15 filas de datos por 2 columnas de datos, el grado de libertad es el siguiente.

$$V = (15 - 1) * (2 - 1) = 14$$

Con los datos obtenidos se determina el chi cuadrado con la siguiente fórmula.

$$x^2 = \sum \frac{(f - ft)^2}{ft}$$

$$\begin{aligned} x^2 = & \frac{(7 - 6.0666)^2}{6.0666} + \frac{(3 - 3.9333)^2}{3.9333} + \frac{(9 - 6.0666)^2}{6.0666} + \frac{(1 - 3.9333)^2}{3.9333} + \frac{(7 - 6.0666)^2}{6.0666} \\ & + \frac{(3 - 3.9333)^2}{3.9333} + \frac{(8 - 6.0666)^2}{6.0666} + \frac{(2 - 3.9333)^2}{3.9333} + \frac{(6 - 6.0666)^2}{6.0666} + \frac{(4 - 3.9333)^2}{3.9333} \\ & + \frac{(8 - 6.0666)^2}{6.0666} + \frac{(2 - 3.9333)^2}{3.9333} + \frac{(7 - 6.0666)^2}{6.0666} + \frac{(3 - 3.9333)^2}{3.9333} + \frac{(7 - 6.0666)^2}{6.0666} \\ & + \frac{(3 - 3.9333)^2}{3.9333} + \frac{(5 - 6.0666)^2}{6.0666} + \frac{(5 - 3.9333)^2}{3.9333} + \frac{(8 - 6.0666)^2}{6.0666} + \frac{(2 - 3.9333)^2}{3.9333} \\ & + \frac{(4 - 6.0666)^2}{6.0666} + \frac{(6 - 3.9333)^2}{3.9333} + \frac{(4 - 6.0666)^2}{6.0666} + \frac{(6 - 3.9333)^2}{3.9333} + \frac{(3 - 6.0666)^2}{6.0666} \\ & + \frac{(7 - 3.9333)^2}{3.9333} + \frac{(4 - 6.0666)^2}{6.0666} + \frac{(6 - 3.9333)^2}{3.9333} + \frac{(4 - 6.0666)^2}{6.0666} + \frac{(6 - 3.9333)^2}{3.9333} \end{aligned}$$

$$\begin{aligned} x^2 = & 0.14361 + 0.22145 + 1.41840 + 2.18754 + 0.14361 + 0.22145 + 0.61617 + 0.95026 \\ & + 0.00073 + 0.00113 + 0.61617 + 0.95026 + 0.14361 + 0.22145 + 0.14361 \\ & + 0.22145 + 0.18752 + 0.28929 + 0.61617 + 0.95026 + 0.70399 + 1.08592 \\ & + 0.70399 + 1.08592 + 1.55013 + 2.39103 + 0.70399 + 1.08592 + 0.70399 \\ & + 1.08592 \end{aligned}$$

$$x^2 \text{ calculada} = 21.34495$$

## Figura 68

Distribución de Chi Cuadrado con un grado de libertad 14

v/p	0,001	0,0025	0,005	0,01	0,025	0,05	0,1	0,15	0,2	0,25	0,3
1	10,8274	9,1404	7,8794	6,6349	5,0239	3,8415	2,7055	2,0722	1,6424	1,3233	1,0742
2	13,8150	11,9827	10,5965	9,2104	7,3778	5,9915	4,6052	3,7942	3,2189	2,7726	2,4079
3	16,2660	14,3202	12,8381	11,3449	9,3484	7,8147	6,2514	5,3170	4,6416	4,1083	3,6649
4	18,4662	16,4238	14,8602	13,2767	11,1433	9,4877	7,7794	6,7449	5,9886	5,3853	4,8784
5	20,5147	18,3854	16,7496	15,0863	12,8325	11,0705	9,2363	8,1152	7,2893	6,6257	6,0644
6	22,4575	20,2491	18,5475	16,8119	14,4494	12,5916	10,6446	9,4461	8,5581	7,8408	7,2311
7	24,3213	22,0402	20,2777	18,4753	16,0128	14,0671	12,0170	10,7479	9,8032	9,0371	8,3834
8	26,1239	23,7742	21,9549	20,0902	17,5345	15,5073	13,3616	12,0271	11,0301	10,2189	9,5245
9	27,8767	25,4625	23,5893	21,6660	19,0228	16,9190	14,6837	13,2880	12,2421	11,3887	10,6564
10	29,5879	27,1119	25,1881	23,2093	20,4832	18,3070	15,9872	14,5339	13,4420	12,5489	11,7807
11	31,2635	28,7291	26,7569	24,7250	21,9200	19,6752	17,2750	15,7671	14,6314	13,7007	12,8987
12	32,9092	30,3182	28,2997	26,2170	23,3367	21,0261	18,5493	16,9893	15,8120	14,8454	14,0111
13	34,5274	31,8830	29,8193	27,6882	24,7356	22,3620	19,8119	18,2020	16,9848	15,9839	15,1187
14	36,1239	33,4262	31,3194	29,1412	26,1189	23,6848	21,0641	19,4062	18,1508	17,1169	16,2221

Nota. Tabla de distribución Chi cuadrado. Tomado de (Cátedra: Probabilidad y Estadística , 2023)

Calculado el valor de chi cuadrado;  $x^2_{calculada} = 21.34495$ , se compara con la tabla de distribución de chi cuadrado de la figura 57, este valor se obtiene del grado de libertad calculado que es  $V = 14$ , con un margen de error de 0.1.

Con estos dos valores encontrados del chi calculado de 21.34495 y el valor de chi cuadrado de tabla es de 21.0641, se comparan para establecer si la hipótesis alternativa cumple o la hipótesis nula, para lo cual debe cumplir la siguiente condición que el chi calculado debe ser mayor al valor del chi cuadrado de la tabla, la hipótesis alternativa cumple, caso contrario se establece que la hipótesis nula es valedera.

$$x^2_{calculada} > x^2_{tabla}$$

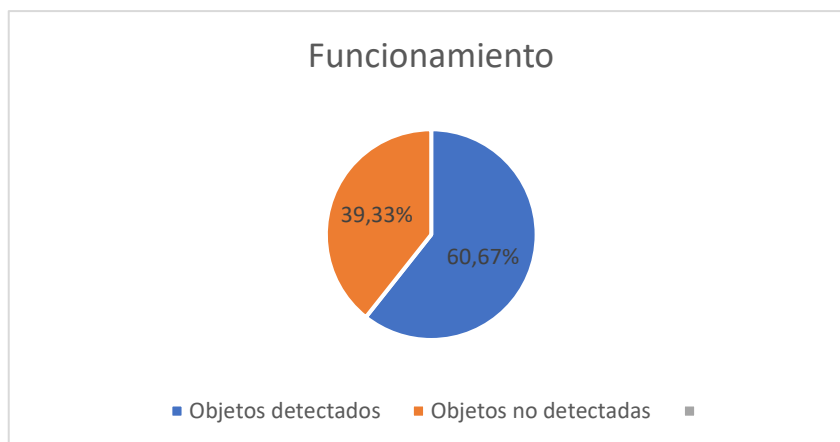
$$x^2_{21.34495} > x^2_{21,0641}$$

De acuerdo al valor obtenido del chi cuadrado calculado, se determina que es mayor que el chi cuadrado de la tabla, por lo cual la hipótesis alternativa del proyecto de investigación

cumple y se da validez al dispositivo como una herramienta de asistencia para el reconocimiento del entorno en el que se desempeñan las personas con discapacidad visual.

### Figura 69

*Funcionalidad del dispositivo según los datos obtenidos de las diferentes pruebas.*



Una vez que se realizaron diferentes pruebas con diferentes tipos de objetos en diferentes escenas, se puede determinar que el dispositivo tiene una eficacia del 60,67 % en la detección de objetos de las cuales se puede determinar que el reconocimiento de objetos influye directamente en los siguientes aspectos; tamaño del objeto, la ubicación del objeto el fondo en el que se encuentra y la distancia en la que se requiere reconocer ya que algunos objetos son capaces de detectar a una distancia de 3 m y otros objetos deben ser detectados a una distancia menor de 1 m, lo cual genera un error del 39.33 % el mismo que se produce en el reconocimiento por la noche ya que debido a diferentes condiciones de luz es ineficiente el reconocimiento, en las pruebas que se realizaron con la ayuda de luz cálida el reconocimiento de objetos es mínimo debido a la sombra que genera la misma persona, lo cual genera un reconocimiento erróneo y a su vez en ciertos objetos no reconoce nada.

Mediante la validación de la hipótesis se demostró que hay un rendimiento menor en la precisión de la red MobileNet V3, generando un error en la detección del % 14.98, pero que al

final fueron buenas selecciones en la detección de objetos en diferentes escenarios que pueden contribuir en las personas con discapacidad visual como una herramienta de asistencia para el reconocimiento del entorno.

### **Validación del dispositivo mediante la supervisión de un especialista**

El dispositivo de las gafas inteligentes se destinará como un dispositivo de asistencia para la lectura y el reconocimiento del entorno en personas que presentan una discapacidad leve y moderada; en donde se estableció y analizó diferentes criterios para su validación por parte del personal médico del Patronato de Latacunga, donde se establecen criterios primordiales para el correcto funcionamiento y las limitaciones que tendrá el mismo.

Para establecer la funcionalidad y la validación del dispositivo se realizaron diferentes preguntas al personal médico que pudo ocupar el dispositivo.

El criterio más importante que se tomó en cuenta fue la portabilidad del dispositivo, en donde se analiza el peso, la comodidad que tiene este al ser ubicado en lentes o gafas de las personas, que tan molesto es utilizar un cinturón, esto permite obtener una respuesta de satisfacción o de molestia por las personas que lo vayan a utilizar.

El segundo criterio que se analiza es la utilidad del dispositivo el cual se buscó conocer en cuanto ayudaría esto a una persona que presente discapacidad visual, si este serviría o no como ayuda asistencial para las personas.

El tercer criterio que se analiza es la funcionalidad en donde se pretende saber en las personas si tienen alguna dificultad para la operación del dispositivo, si su forma de uso es de una manera fácil o difícil.

El cuarto criterio de análisis fue la implementación donde se determina si las personas que ocuparon las gafas inteligentes están conformes con la calidad del sintetizador de voz, si

este dispositivo como herramienta de asistencia para lectura y reconocimiento de objetos es de beneficio para las personas que presentan discapacidad visual.

Con los criterios mencionados se pretende establecer un valor de ponderación del 1 a 5, en donde 5 excelente, 4 muy bueno, 3 bueno, 2 regular y 1 insuficiente.

Persona externa 1, es una persona que presenta el 43 % de discapacidad visual moderada.

Persona externa 2, se realizó pruebas con una persona joven que no presenta discapacidad visual.

#### **Tabla 41**

*Criterios por parte del personal médico para la validación del dispositivo.*

Personal Médico/ Externas	Portabilidad	Utilidad	Funcionalidad	Implementación	Aceptación
Médico 1	4	5	4	4	85%
Médico 2	3	4	5	5	85%
Médico 3	5	4	4	5	90%
Externa 1	4	4	3	4	75%
Externa 2	5	4	5	4	90%

Para la persona con discapacidad visual moderada el dispositivo no resulta ser intuitivo presentando ciertas complicaciones en su operación de funcionamiento con respecto a las personas que no presentan problemas de visión.

Debido a que las personas con discapacidad visual necesitan de mayor tiempo para su adaptación para desarrollar habilidades y acostumbrarse al dispositivo mediante la compensación de otros sentidos como el tacto.

Mediante los diferentes criterios por parte del personal médico del patronato se establece que el dispositivo tendrá limitaciones y se enfocará en un grupo prioritario para personas con discapacidad leve y moderada, en las cuales presentan secuelas y una disminución muy importante en la capacidad de la persona para poder realizar en mayor parte sus actividades cotidianas, lo cual el dispositivo puede utilizarse con mayor facilidad a este grupo de personas, pero también el dispositivo podrá ser utilizado para personas con discapacidad grave lo que dificulta a la persona poder realizar la mayoría de sus actividades de una manera independiente lo cual requiere necesariamente de una tercera persona que asista y supervise el dispositivo.

Este tipo de proyecto que se implementó se ve como un aporte de gran utilidad hacia las personas que presentan una discapacidad visual lo cual les permitirá una inclusión social y una accesibilidad a textos y al reconocimiento de objetos en diferentes escenas.

### **Figura 70**

*Pruebas en personas médico para validación del dispositivo*



### Autonomía del dispositivo

La autonomía del dispositivo de las gafas inteligentes se determinará por el tiempo de uso que le pueden dar los usuarios en relación a horas de funcionamiento del dispositivo con respecto a la capacidad de la batería, que dependerá de los diferentes dispositivos conectados al sistema.

**Tabla 42**

*Autonomía de las gafas inteligentes*

Descarga	Tiempo
Reposo	3,2 horas
Sistema en uso	2,3 horas

Las gafas inteligentes se pueden ocupar de diferentes maneras por parte del usuario, de una manera continua el dispositivo alcanza una autonomía de 2,30 horas de manera frecuente el dispositivo alcanza una autonomía de 3,2 horas.

**Tabla 43**

*Tiempo de carga*

Carga	Tiempo
Funcionamiento	2 horas
Apagado	1,40 horas

La carga del dispositivo se puede realizar de dos formas: cuando el dispositivo se encuentra en estado de reposo se estima un tiempo de carga de 1.40 horas hasta completar su



carga máxima, cuando el dispositivo se encuentra en estado funcional se estima un tiempo de carga de 2 horas hasta completar su carga.

**Tabla 44**


*Consumo de cada componente del dispositivo según su datasheet*

Cargas	Medidas Referenciales (mAh)
Raspberry Pi4	1000
Pantalla Táctil	90
Arduino	15
Web Cam	400
Parlantes	800
Total	2305
Batería	5000
Autonomía	2,30 horas

Se determina el cálculo de la autonomía mediante medidas referenciales (mA), con respecto a la capacidad de la batería (5000 mAh).

**Figura 71**

Pruebas de la autonomía del dispositivo

Capacidad Disponible:	<input type="text" value="5000"/>	<input type="text" value="mA·h"/>
Consumo: 	<input type="text" value="2305"/>	<input type="text" value="mA"/>
Duración:	<input type="text" value="2,16919"/>	Horas

*Nota.* Autonomía de las gafas inteligentes

Para determinar el consumo de corriente en el dispositivo se establecen pruebas mediante un medidor de consumo de cada uno de los componentes electrónicos conectados en el sistema.

$$P = (5.05 + 0,45 + 0,075 + 2 + 4)$$

$$P = 11,575 \text{ w}$$

Donde:

$$V_b \times I_b = W_b$$

$V_b$ , voltaje de la batería

$I_b$ , capacidad de la batería

$W_b$ , potencia de la batería power bank Ldnio

$$5V \times 5000 \text{ mAh} = 25 \text{ Wh}$$

$W_c$ , potencia consumida de cada componente conectado en el sistema

$$W_c = 11,575 \text{ W}$$

Consumo total

$$\frac{W_b}{W_c} = h$$

$$\frac{25 \text{ wh}}{11,525 \text{ w}} = 2,20 \text{ horas}$$

Conversión de horas a minutos

$$2,20 \text{ horas} \times 60 \text{ min}$$

$$132 \text{ minutos}$$

El dispositivo de las gafas inteligentes brindará una autonomía al usuario para su funcionamiento de 132 minutos de manera interrumpida, el cual permita funcionar con normalidad en sus dos modos: en el reconocimiento de texto y objetos.

## Capítulo V

### Conclusiones y recomendaciones

#### Conclusiones

- Realizada la revisión bibliografía de los diferentes dispositivos y sistemas electrónicos, se determina que actualmente en el país no se comercializa ni se desarrollan dispositivos de gafas inteligentes como herramienta de asistencia para lectura y el reconocimiento del entorno mediante inteligencia artificial para personas con discapacidad visual, lo que limita a las personas a una integración en su entorno.
- EL diseño y modelo 3D del dispositivo de las gafas inteligentes se determinó mediante una revisión bibliografía de guías técnicas de dispositivos existentes de ayuda para personas con discapacidad visual, que detallan parámetros de selección, dimensiones para los componentes electrónicos debido que consta de varias partes: Raspberry Pi 4, cámara, micrófono, pantalla, batería, botones y parlantes para un diseño óptimo y funcional.
- Se creó un diseño CAD para la estructura y el soporte de los lentes el cual se concluye que la mejor opción para la impresión es mediante PLA que es un material que ofrece mejores propiedades físicas y mecánicas; garantizando su funcionalidad y protección en el dispositivo, brindando comodidad y ergonomía en el usuario.
- Se determinó un diseño de arquitectura modular que permite dividir en dos apartados el diseño: case para el soporte de la cámara y audio; estructura donde se encuentran los diferentes componentes electrónicos que permiten la protección, flexibilidad y adaptación en las personas.
- Dentro del desarrollo e implementación del algoritmo para el reconocimiento de texto se emplea la herramienta de Google API Cloud Vision que optimiza algoritmos de procesamiento de imágenes, algoritmos basados en aprendizaje profundo y redes

neuronales para el reconocimiento óptico de caracteres (OCR), mejorando en si el rendimiento computacional y su eficiencia en el reconocimiento de texto.

- Se implementó la interfaz del dispositivo de las gafas inteligentes, con la finalidad que pueda interactuar con el dispositivo la persona que asista y supervise a las personas con discapacidad visual grave y pueda brindar asistencia para la lectura y el reconocimiento de entorno.
- Se implementó el sistema de inteligencia artificial mediante la configuración de modelo de la red SSD MobileNet V3, por su precisión y fluidez en el reconocimiento de objetos, que permite emplear menos recursos computacionales.
- Se determinó que el modelo de red SSD MobileNet V3 emplean modelos de TensorFlow preentrenados previamente en conjunto a la base de datos de Coco que conllevan errores de detección debido a los entornos, tipo de objetos y diversas características que no se toman en cuenta en el entrenamiento.
- Se determinó que no es pertinente la creación de circuitos electrónicos para la transmisión de datos hacia el sistema embebido, debido que los componentes electrónicos ofrecen una conexión plug & play, que permite la conexión de manera directa al sistema embebido Raspberry Pi 4.
- Se implementó el sistema convertidor de texto a voz mediante la herramienta de API Google Translate, que emplea la librería de gTTS (Text To Speech), que brinda mayores prestaciones en la reproducción del audio y permite la configuración de la frecuencia para una voz de forma más natural, precisa y con soporte en varios idiomas.
- Se determinó por parte del personal médico del Patronato de Latacunga que el dispositivo de gafas inteligentes, para contribuir como una herramienta de asistencia para lectura y reconocimiento del entorno, lo cual ayuda a las personas con

discapacidad visual leve o moderada, a eliminar barreras en la detención de objetos, garantizando una inclusión social, accesibilidad y funcionalidad del dispositivo.

- El dispositivo de gafas inteligentes también tiene como grupo de atención a las personas que presentan discapacidad visual grave, pero presenta limitaciones en su funcionalidad debido que este grado de discapacidad necesita de una tercera persona para su supervisión y funcionamiento.
- La validación del dispositivo como herramienta de asistencia para lectura y reconocimiento del entorno para personas con discapacidad visual utilizando inteligencia artificial, se estableció mediante diferentes pruebas de texto, libros, revistas, folletos, hojas impresas en diferente tamaño de fuente y diferentes tipos de letras con una eficiencia en el detector del % 95.94 lo cual garantiza un reconocimiento aceptable en su funcionamiento.
- La validación del dispositivo como herramienta de asistencia para el reconocimiento del entorno para personas con discapacidad visual, se realizó la validación con diferentes objetos en diferentes escenas con una eficiencia del 60,67 % en el reconocimiento presentando limitaciones en el reconocimiento de objetos debido a dimensiones de los objetos, distancia a identificar, condiciones no favorables lo cual interviene directamente en el funcionamiento del dispositivo.
- Se determinó que el dispositivo para personas con discapacidad visual grave, resulta ser poco intuitivo presentando limitaciones en su modo funcional por parte del usuario, hasta su adaptación para el correcto funcionamiento.
- Se determinó que la API de Cloud Vision de Google, mediante la herramienta Text Detection proporciona una precisión del 95.40% una vez realizadas diversas pruebas de reconocimiento de texto en diferentes condiciones, con respecto a la precisión

recomendada por la API que en condiciones normales que dice ofrecer una precisión cercana al 100 %.

- Después que se realizó la detección de objetos mediante la red MobileNet V3, que proporciona una precisión de la red recomendada del 75.6 %, lo cual, realizadas diversas pruebas de detección de objetos en diferentes escenarios, se concluye que la red MobileNet ofrece una precisión real del 60.66 %, presentando buenas detecciones de objetos en condiciones normales.
- Se determinó que el dispositivo de gafas inteligentes presenta un bajo costo de implementación con respecto a dispositivos comerciales existentes con un alto valor de costo e inexistentes en el país.

## **Recomendaciones**

- Es pertinente el uso de un acelerador para aumentar los fps en el proceso de detección de objetos por tal motivo se recomienda el uso de Google coral mediante un hardware acelerador para el proceso de detección de objetos, se puede optar por el uso de modelos que demanden más carga computación como TensorFlow y Yolo.
- Considerando los recursos computaciones que genera la detección de objetos mediante la red MobileNet, se recomienda el uso de una tarjeta de desarrollo que preste mayores prestaciones computaciones, mejorando el desempeño en el procesador, tarjeta gráfica y capacidad que garantice una detección más fluida Enel reconocimiento de los objetos.
- Es recomendable el uso de una cámara web más pequeña con micrófono incorporado que sea compatible con Raspberry, para mejorar la ergonomía del dispositivo.
- Se recomienda el uso de una batería de mayor capacidad para aumentar la autonomía del dispositivo y extender su rendimiento.

- El acceso de las herramientas de Google API está limitado en su funcionamiento, debido a que se necesita una conexión estable a internet.
- Para el correcto funcionamiento de la herramienta de Google API, es necesario la creación de un usuario que permita acceder a los recursos y archivos. Json para detección de texto, si se supera la cantidad solicitudes gratuitas de detección de texto es necesario el pago de una suscripción mensual para acceder a todas las herramientas.
- Para la detección de texto se recomienda establecer una distancia visual considerable de menos de 35 cm, cuando exista documentos con una distribución ambas caras y presencia de tamaño de fuente inferior a 10 puntos.
- Para la detección de objetos grandes y medianos es recomendable ubicarse a una distancia mayor o igual a 3 m, debido al campo visual que ofrece la cámara.
- Para la detección de objetos pequeños es recomendable ubicarse a una distancia menor o igual a 1m para una detección aceptable.
- El reconocimiento de objetos en la red neuronal MobileNet V3, presenta limitaciones debido a la resolución de la cámara; pre-procesamiento que se realiza en la imagen, mediante un reajuste en la resolución de las imágenes limitando el reconocimiento de objetos en la noche.
- La resolución de la cámara limita en el reconocimiento de objetos, debido al pre-procesamiento que se realiza en las imágenes lo cual se establece un reajuste en la resolución de la imagen para la detección de objetos.
- La reproducción de texto a audio en condiciones no favorables presenta limitaciones hacia el usuario debido a la potencia de salida del audio.
- Se recomienda que la configuración para el soporte del case se deba implementar adicionalmente tirantes que brinden una mayor sujeción fija al dispositivo.



- Se sugiere que el soporte para los lentes se emplee un clip, con una protección para evitar rayones en las varillas de los lentes o gafas.
- La implementación de la interfaz del dispositivo se puede modificar el tamaño de la pantalla a 3.5 pulgadas, debido que la pantalla es de asistencia para la persona que supervisa el dispositivo.
- Evitar el uso en ambientes húmedos, calurosos y con presencia de polvo debido a que perjudica el proceso de disipación de calor y funcionamiento de los componentes que conforman el dispositivo.

## Bibliografía

- (19 de 12 de 2021). Obtenido de Finger Reader: <https://pqs.pe/actualidad/tecnologia/finger-reader-el-anillo-que-facilitara-la-lectura-personas-invidentes/>
- (25 de 10 de 2022). Obtenido de RobotShop Europe:  
<https://eu.robotshop.com/es/products/elecrow-rc070-7-inch-1024x600-hdmi-lcd-display-w-touch-screen>
- (19 de 01 de 2023). Obtenido de Waveshare: <https://www.waveshare.com/imx219-160-camera.htm>
- Alibaba.com. (10 de 08 de 2022). Obtenido de [https://www.alibaba.com/product-detail/Best-sale-2MP-1080P-Mini-Spy\\_62451413472.html?spm=a2700.wholesale.0.0.620b6dc3m5leJi](https://www.alibaba.com/product-detail/Best-sale-2MP-1080P-Mini-Spy_62451413472.html?spm=a2700.wholesale.0.0.620b6dc3m5leJi)
- Ananya, M. (27 de 06 de 2012). Noticias-Medical.net. Obtenido de <https://www.news-medical.net/health/Types-of-visual-impairment.aspx>
- Artturi, J. (29 de 09 de 2022). codingem.com. Obtenido de <https://www.codingem.com/python-text-to-speech/>
- Av Electronics. (10 de 12 de 2022). Obtenido de <https://avelectronics.cc/producto/interruptor-deslizante-spdt/>
- bcnvision. (07 de 01 de 2022). Obtenido de Lectura OCR basada en Deep Learning:  
<https://www.bcnvision.es/blog-vision-artificial/lectura-ocr-deep-learning/>
- Bonafonte, D. (04 de 08 de 2022). Campo visual normal- Centro de Oftalmología Bonafonte.  
Obtenido de <https://www.youtube.com/watch?v=SzNxgyRDE-s>
- Botech. (29 de 09 de 2022). Obtenido de <https://btech.com/en/ldnio-power-bank-white-pr518.html>
- botnroll.com. (18 de 08 de 2022). Obtenido de <https://www.botnroll.com/en/sound/3991-voice-recognition-module-v3-with-microphone.html>

- Brarath, K. (27 de 08 de 2022). Medium. Obtenido de <https://towardsdatascience.com/how-to-get-started-with-google-text-to-speech-using-python-485e43d1d544>
- Cadena Lesly Stefania, R. X. (22 de 02 de 2021). Repositorio Digital EPN Ingenieria en Electronica y control. Obtenido de <https://bibdigital.epn.edu.ec/handle/15000/21431>
- Calderón, E. L. (2012). Educación pertinente e inclusiva. La discapacidad en educación indígena. México: Secretaria de Educación Publica.
- Carneros, M. D. (04 de 09 de 2019). Oftalmologos Martines De Carneros Madrid. Obtenido de <https://www.martinezdecarneros.com/motilidad-ocular/#:~:text=La%20motilidad%20ocular%20es%20una,evitando%20diplop%C3%ADAs%20o%20visi%C3%B3n%20doble.>
- Cemco. (08 de 10 de 2022). Obtenido de <https://cemco.com.ec/producto/parlantes-genius-rojo-sp-q160/>
- CIE9MC. (2014). Clasificación Internacional de Enfermedades. 9a Clasificación Internacional de Enfermedades. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad. ISBN: 978-84-7670-721-0.
- Cognex. (2019). Introducción a la visión artificial una guia para la automatización de procesos y mejoras de calidad. Estados Unidos: Corporate Headquarters.
- CogniFit. (04 de 08 de 2022). Campo Visual - Habilidades Cognitivas. Obtenido de <https://www.cognifit.com/ec/ciencia/habilidad-cognitiva/campo-visual>
- Contaval. (08 de 01 de 2022). Contaval automatismos y componentes electronicos SI. Obtenido de <https://www.contaval.es/que-es-la-vision-artificial-y-para-que-sirve/>
- Core Electronics . (12 de 05 de 2022). Obtenido de <https://core-electronics.com.au/guides/object-identify-raspberry-pi/>
- Cristina, J. L. (23 de 08 de 2016). Obtenido de <http://repositorio.espe.edu.ec/handle/21000/11886>

Daniel, B. (30 de 08 de 2022). Trenton Systems. Obtenido de

<https://www.trentonsystems.com/blog/what-are-embedded-systems>

Datapeaker. (27 de 10 de 2022). Obtenido de Detección de objetos mediante la API de

TensorFlow: <https://datapeaker.com/big-data/deteccion-de-objetos-mediante-la-api-de-tensorflow/>

Estadísticas de Discapacidad- Consejo Nacional para la Igualdad de Discapacidades. (12 de 08

de 2022). Obtenido de <https://www.consejodiscapacidades.gob.ec/estadisticas-de-discapacidad/>

EyeComitec Laz Group S.A. (07 de 01 de 2022). Obtenido de [https://eyecomtec.com/es/3136-](https://eyecomtec.com/es/3136-EyeTalk-Eye-Tracker)

[EyeTalk-Eye-Tracker](https://eyecomtec.com/es/3136-EyeTalk-Eye-Tracker)

Geek Factory. (30 de 08 de 2022). Obtenido de

<https://www.geekfactory.mx/tienda/tarjetas/raspberry-pi/raspberry-pi-4-modelo-b/>

Google Cloud. (12 de 08 de 2022). Obtenido de

[https://cloud.google.com/vision/docs/ocr?hl=es\\_419#:~:text=batch%20image%20annotation,-,Optical%20Character%20Recognition%20\(OCR\),street%20sign%20or%20traffic%20sign.](https://cloud.google.com/vision/docs/ocr?hl=es_419#:~:text=batch%20image%20annotation,-,Optical%20Character%20Recognition%20(OCR),street%20sign%20or%20traffic%20sign.)

Google Cloud. (27 de 10 de 2022). Obtenido de [https://cloud.google.com/learn/what-is-artificial-](https://cloud.google.com/learn/what-is-artificial-intelligence)

[intelligence](https://cloud.google.com/learn/what-is-artificial-intelligence)

Guia de juegos y juguetes Aiju. (28 de 11 de 2022). Obtenido de <https://www.guiaaiju.com/>

IBM Reconomiento optico OCR. (12 de 08 de 2022). Obtenido de

<https://www.ibm.com/cloud/blog/optical-character-recognition>

Instituto Nacional de Seguridad, y. S. (10 de 09 de 2022). Herramientas Prevención Riesgos

Laborales - INSST. Obtenido de <https://herramientasprl.insst.es/ergonomia/distancia-visual/contenido/194>

- Klipxtreme. (11 de 08 de 2022). Obtenido de <https://www.klipxtreme.com/st-products/internal/KBH-155/>
- Kvetonovad, L., & Rehurek, J. (2011). Tratamiento Educativo de la Diversidad de Tipo Visual. Madrid: ISBN: 974-84-362-6201-8.
- LattePanda. (6 de 11 de 2022). Obtenido de <http://www.lattepanda.com:2006/lattepanda-delta>
- LDLC-High Tech Experience. (10 de 01 de 2023). Obtenido de <https://www.ldlc.com/es-es/ficha/PB00280675.html>
- LIÉBANA, M., & SIMÓN, V. (2000). MARTÍNEZ LIÉBANA, Ismael. ISBN: 8448402227, 9788448402228.
- Logitech. (06 de 08 de 2022). Obtenido de <https://www.logitechstore.com.ar/camara-logitech-streamcam-plus/p>
- Mascaró, D. B. (2020). Inteligencia Artificial y Decisiones Clínicas. Madrid: Gaspar Casal.
- Melo, J. L. (2016). Pequeño Manual de Higiene y Seguridad en el Trabajo. Quilmes.
- Motilidad ocular- Clínica Oftalmológica Halcon. (04 de 08 de 2022). Obtenido de <https://clinicaoftalmologicahalcon.com/motilidad-ocular>
- Mouser Electronics. (16 de 03 de 2022). Obtenido de <https://www.mouser.ec/new/fortebit/fortebit-easyvr-3-plus-module/>
- Novatronic. (10 de 12 de 2022). Obtenido de <https://novatronic.ec.com/index.php/product/pulsador-de-boton-cuadrado-6x6x7-3mm-con-tapa-plastica/>
- NVIDIA Developer. (06 de 03 de 2019). Obtenido de <https://developer.nvidia.com/embedded/jetson-nano#:~:text=Jetson%20Nano%20is%20a%20small,graphics%2C%20multimedia%2C%20and%20more.>
- Oftalvist, G. (10 de 09 de 2019). Blog Oftalvist:Lo Ultimo en Salud Ocular. Obtenido de <https://www.oftalvist.es/blog/agudeza-visual-que-es-y-pruebas/>

- OMS. (2012). Change the Definition of Blindness. Definitions of blindness and visual impairment. Ginebra, Organización Mundial de la Salud.
- OMS. (2014). Ceguera y Discapacidad Visual Organización Mundial de la Salud. N° 282,OMS.
- OMS. (2019). ASAMBLEA MUNDIAL DE LA SALUD. Proyecto de plan de acción para la prevención de la ceguera y la discapacidad visual evitables 2014 – 2019. Salud ocular universal: un plan de acción mundial, A66/11 , punto 13.4 del orden del día.
- Orcam. (10 de 01 de 2022). Obtenido de <https://www.orcam.com/es/>
- Patrick. (23 de 08 de 2022). Nextion. Obtenido de <https://nextion.tech/datasheets/nx8048t050/>
- Patronato Municipal Gad Latacunga. (15 de 03 de 2022). Obtenido de <https://patronatomunicipallatacunga.gob.ec/mision-vision-y-transparencia/>
- Pilatasig, A. V. (07 de 08 de 2017). Repositorio Institucional de la Universidad de las Fuerzas Armadas ESPE. Obtenido de <http://repositorio.espe.edu.ec/handle/21000/13446>
- Programador clic. (07 de 10 de 2022). Obtenido de <https://programmerclick.com/article/2166891983/>
- Randhawa, D. (08 de 02 de 2023). Legacy Eye Care. Obtenido de <https://legacyeyecare.ca/what-is-a-visual-field-test-why-is-it-important/>
- Raspberry Pi 4 Computer Model B. (2022). Raspberry Pi, 8.
- Read, A. (26 de 05 de 2022). AllRead. Obtenido de <https://www.allread.ai/es/ocr-tradicional-vs-software-inteligencia-artificial/>
- Robot Electronica. (10 de 12 de 2022). Obtenido de <https://robot.com.ve/product/kit-de-2-pbs-11b-interruptores-boton-pulsador-push-button-rojo/>
- RobotShop Europe. (08 de 01 de 2023). Obtenido de <https://eu.robotshop.com/es/products/grove-speech-recognition-module>
- Rodriguez, H. (27 de 09 de 2022). Crehana. Obtenido de <https://www.crehana.com/blog/estilo-vida/megapixeles/>

Sabina, N. &. (2022). Objetc Detection using Yolo and Mobilene SSD:A Comparative Study.

International Journal of Engineering Research & Technology, 5.

SandoRobotics. (25 de 08 de 2022). Obtenido de

<https://sandorobotics.com/producto/im150416006/>

SIJ Distribuciones SaS. (16 de 12 de 2022). Obtenido de

<https://sijdistribuciones.com/producto/parlante-economico-para-pc-mod-g-101/>

Sik-Ho, T. (15 de 11 de 2020). MobileNet V3. MobileNet V3 Clasificacion de imagenes, pág. 12.

Solución Ingenieril. (12 de 01 de 2022). Obtenido de Etapas de un sistema de visión - Visión

artificial: [http://solucioningenieril.com/vision\\_artificial/etapas\\_de\\_un\\_sistema\\_de\\_vision#](http://solucioningenieril.com/vision_artificial/etapas_de_un_sistema_de_vision#)

Team, Q. T. (20 de 08 de 2020). Machine Vision Process Flow. Obtenido de

<https://qualitastech.com/blog/quality-control-insights/machine-vision-process-flow/>

Tecno Trade. (28 de 11 de 2022). Obtenido de

<https://www.tecnotrade.com.ec/producto/parlante-bluetooth-con-radio-koa073a/>

The Pi Hut. (20 de 07 de 2022). Obtenido de [https://thepihut.com/products/raspberry-pi-camera-](https://thepihut.com/products/raspberry-pi-camera-module)

[module](https://thepihut.com/products/raspberry-pi-camera-module)

Urbina, G. B. (2010). Evaluación De Proyectos. Mexico: Mc Graw Hill Educacion.

UTN, F. R. (06 de 01 de 2023). Cátedra: Probabilidad y Estadística . Obtenido de

<https://mat.uda.cl/hsalinas/cursos/2010/eyp2/Tabla%20Chi-Cuadrado.pdf>

Waveshare . (20 de 08 de 2022). Obtenido de

[https://www.waveshare.com/wiki/5inch\\_HDMI\\_LCD\\_\(B\)](https://www.waveshare.com/wiki/5inch_HDMI_LCD_(B))

Waveshare. (15 de 01 de 2023). Obtenido de [https://www.waveshare.com/imx219-160-](https://www.waveshare.com/imx219-160-camera.htm)

[camera.htm](https://www.waveshare.com/imx219-160-camera.htm)

## Anexos