



**UNIVERSIDAD TÉCNICA DEL NORTE**

**FACULTAD DE INGENIERÍA EN CIENCIAS APLICADAS**

**CARRERA DE INGENIERÍA EN MECATRÓNICA**

**INFORME FINAL DEL TRABAJO DE INTEGRACIÓN CURRICULAR,  
MODALIDAD DE PROYECTO DE INVESTIGACIÓN**

**TEMA:**

**“DISPOSITIVO DETECTOR DE OBSTÁCULOS PARA PERSONAS NO  
VIDENTES MEDIANTE VISIÓN ARTIFICIAL”**

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

**Línea de investigación:** *Biomecatrónica*

**Autor:** *William Alexander Ortega Benavides*

**Director:** *PhD, Iván Danilo García Santillán, Ing.*

**Asesor:** *PhD, David Alberto Ojeda Peña, Ing.*



## UNIVERSIDAD TÉCNICA DEL NORTE BIBLIOTECA UNIVERSITARIA

### AUTORIZACIÓN DE USO Y PUBLICACIÓN A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE

#### 1. IDENTIFICACIÓN DE LA OBRA

En cumplimiento del Art. 144 de la Ley de Educación Superior, hago la entrega del presente trabajo a la Universidad Técnica del Norte para que sea publicado en el Repositorio Digital Institucional, para lo cual pongo a disposición la siguiente información:

DATOS DE CONTACTO			
CÉDULA DE IDENTIDAD:	0401881602		
APELLIDOS Y NOMBRES:	ORTEGA BENAVIDES WILLIAM ALEXANDER		
DIRECCIÓN:	BOLIVAR		
EMAIL:	<a href="mailto:waortegab@utn.edu.ec">waortegab@utn.edu.ec</a>		
TELÉFONO FIJO:	2287559	TELÉFONO MÓVIL:	0968281927

DATOS DE LA OBRA	
TÍTULO:	Dispositivo detector de obstáculos para personas no videntes mediante visión artificial.
AUTOR (ES):	William Alexander Ortega Benavides
FECHA: DD/MM/AAAA	07/09/2023
SOLO PARA TRABAJOS DE GRADO	
PROGRAMA:	<input checked="" type="checkbox"/> PREGRADO <input type="checkbox"/> POSGRADO
TÍTULO POR EL QUE OPTA:	Ingeniero en mecatrónica
ASESOR /DIRECTOR:	Ing. Iván García, PhD

#### 2. CONSTANCIAS

El autor (es) manifiesta (n) que la obra objeto de la presente autorización es original y se la desarrolló, sin violar derechos de autor de terceros, por lo tanto la obra es original y que es (son) el (los) titular (es) de los derechos patrimoniales, por lo que asume (n) la responsabilidad sobre el contenido de la misma y saldrá (n) en defensa de la Universidad en caso de reclamación por parte de terceros.

Ibarra, a los 7 días del mes de septiembre de 2023.

#### EL AUTOR:

(Firma) 

Nombre: William Alexander Ortega Benavides

## CERTIFICACIÓN DIRECTOR DEL TRABAJO DE INTERGRACIÓN CURRICULAR

Ibarra, 7 de septiembre de 2023

PhD, Iván Danilo García Santillán, Ing.

DIRECTOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR

CERTIFICA:

Haber revisado el presente informe final del trabajo de titulación, el mismo que se ajusta a las normas vigentes de la Unidad Académica de la Universidad Técnica del Norte; en consecuencia, autorizo su presentación para los fines legales pertinentes.



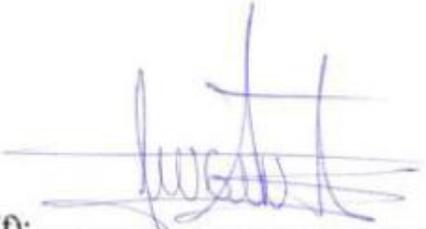
(f) .....

PhD, Iván Danilo García Santillán, Ing.

C.C.: 1002292603

## APROBACION DEL COMITE CALIFICADOR

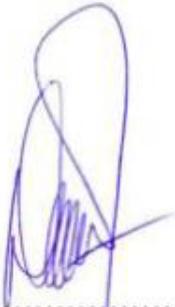
El Tribunal Examinador del trabajo de titulación "Dispositivo detector de obstáculos para personas no videntes mediante visión artificial" elaborado por William Alexander Ortega Benavides, previo a la obtención del título de Ingeniero en Mecatrónica, aprueba el presente informe de investigación en nombre de la Universidad Técnica del Norte:



(f):.....

PhD, Iván Danilo García Santillán, Ing.

C.C.: 1002292603



(f):.....

PhD, David Alberto Ojeda Peña, Ing.

C.C.:1757898489

## **DEDICATORIA**

Este trabajo de investigación es dedicado a mis padres Jorge y Rosa, por su constancia de apoyo en mi vida estudiantil y a toda mi familia en general, también es dedicado para aquellos que no pudieron ver concluir, pero siempre me dieron vos de aliento, mi abuelita Teresa.

## **AGRADECIMIENTO**

Agradezco a mi familia por apoyarme y nunca perder su confianza, en dar consejos para mejorar y aumentar mi autoestima para culminar este trabajo de investigación.

A mis amigos de grupo juvenil Sociality, quienes siempre están pendientes de los éxitos que voy logrando, también a mis compañeros de aula Anderson, Carol y David, con quienes en los últimos semestres de carrera formamos un grupo solido de trabajo, un agradecimiento especial a Carla Vallejo, una amiga incondicional que siempre está dando confianza para alcanzar mis metas.

Al señor Wilo Pabón por dar la confianza y el compromiso de realizar esta investigación y ser el quien verifique el funcionamiento correcto del dispositivo. También a la empresa Ingenious Works por impartir sus concommitos en el desarrollo e implementación del dispositivo de detección de objetos.

## RESUMEN EJECUTIVO

Las personas con discapacidad visual tienen la dificultad de ejercer su autonomía e independencia al momento de moverse de un lugar a otro, ya que requiere la implementación de un dispositivo que le ayude a realizar una caminata con normalidad, pero si el dispositivo no anticipa a tiempo un obstáculo o desnivel en el camino. esto podría ocasionar golpes o lesiones en la integridad de la persona, por lo que se ha visto indispensable la realización de un dispositivo que ayude a anticipar aquellos obstáculos en tiempo prudente y en la dirección de ubicación que se encuentran aquellas dificultades a esquivar, para que la persona no vidente esta alerta y realice un movimiento para evadir el obstáculo. Se presenta el diseño e implementación de un dispositivo que contenga algoritmos basado en la detección de objetos efectuado la programación en un entorno virtual, con la utilización de las nuevas tecnologías como es la visión artificial y software de programación que utilice un lenguaje amigable ayude al reconocimiento de objetos en un tiempo mínimo de respuesta y, sobre todo, que tenga un alcance lo suficientemente aceptable al reconocer el objeto. El dispositivo se adapta a la anatomía del cuerpo humano que presenta la persona y está ubicado en una zona donde se recolectan la mayor cantidad de información para realizar el análisis y procesamiento de datos para dar la señal de alerta.

**Palabras clave:** Python, Entorno Virtual, TensorFlow, Procesamiento de datos, Matriz de confusión, Sistemas inteligentes, Raspberry Pi.

## ABSTRACT

People with visual impairment have the difficulty to exercise their autonomy and independence when moving from one place to another, as it requires the implementation of a device that helps you to walk normally, but if the device does not anticipate in time an obstacle or unevenness in the road. This could cause blows or injuries to the integrity of the person, so it has been essential the realization of a device that helps to anticipate those obstacles in a timely manner and in the direction of location that are those difficulties to avoid, so that the blind person is alert and make a move to avoid the obstacle. We present the design and implementation of a device containing algorithms based on object detection and programming in a virtual environment, with the use of new technologies such as artificial vision and programming software that uses a friendly language to help the recognition of objects in a minimum response time and above all, that has a sufficiently acceptable range to recognize the object. The device is adapted to the anatomy of the human body that the person presents and is located in an area where most information is collected to perform the analysis and processing of data to give the warning signal.

**Keywords:** Python, Virtual Environment, TensorFlow, Data Processing, Confusion Matrix, Intelligent Systems, Raspberry Pi.

## INDICE DE CONTENIDO

INTRODUCCIÓN .....	14
<b>Problema de investigación.....</b>	<b>14</b>
<b>Objetivos .....</b>	<b>15</b>
<i>Objetivo General .....</i>	<i>15</i>
<b>Objetivos Específicos .....</b>	<b>15</b>
<b>Justificación .....</b>	<b>15</b>
<b>Alcance .....</b>	<b>16</b>
<b>CAPÍTULO I: MARCO REFERENCIAL .....</b>	<b>17</b>
<b>1.1. Antecedentes .....</b>	<b>17</b>
<b>1.2. Marco Teórico .....</b>	<b>18</b>
<i>1.2.1. La Ceguera .....</i>	<i>18</i>
<b>1.2.1.1. Discapacidad Visual. ....</b>	<b>18</b>
<b>1.2.1.2. Orientación del No Vidente. ....</b>	<b>19</b>
<b>1.2.1.3. Técnica de Movilidad con Bastón.....</b>	<b>19</b>
<b>1.2.1.4. Bastón Blanco.....</b>	<b>20</b>
<b>1.2.1.5. Percepción Auditiva. ....</b>	<b>20</b>
<i>1.2.2. Programación Orientada a Objetos .....</i>	<i>21</i>
<b>1.2.2.1. Python.....</b>	<b>22</b>
<i>1.2.3. Visión Artificial .....</i>	<i>22</i>
<b>1.2.3.1. Machine Learning.....</b>	<b>23</b>
<b>1.2.3.2. Deep Learning.....</b>	<b>24</b>
<b>1.2.3.3. Entrenamiento de Modelos de Aprendizaje Automático.....</b>	<b>24</b>
<i>1.2.4. Algoritmos en Visión Artificial .....</i>	<i>25</i>
<b>1.2.4.1. Algoritmo de Regresión.....</b>	<b>26</b>
<b>1.2.4.2. Algoritmo Basado en Instancias.....</b>	<b>26</b>
<b>1.2.4.3. Algoritmo de Árbol de Decisión. ....</b>	<b>26</b>
<b>1.2.4.4. Algoritmo de Aprendizaje Profundo. ....</b>	<b>26</b>
<b>1.2.4.5. Algoritmo De Redes Neuronales Convolucionales (CNN).....</b>	<b>26</b>
<i>1.2.5. Matriz de confusión .....</i>	<i>30</i>
<b>1.2.5.1. Estructura de una Matriz de Confusión.....</b>	<b>31</b>
<b>1.2.5.2. Curva de ROC. ....</b>	<b>33</b>
<b>CAPÍTULO II: MATERIALES Y MÉTODOS.....</b>	<b>35</b>

2.1.	Modelo de la Investigación .....	35
2.2.	Diseño de la Investigación .....	36
2.2.1.	<i>Análisis de los diferentes sistemas de detección de obstáculos con el uso de visión artificial</i> .....	36
2.2.2.	<i>Diseño del dispositivo de detección de obstáculos</i> .....	36
2.2.3.	<i>Aplicación del dispositivo de detección de obstáculos</i> .....	37
2.2.4.	<i>Validación del funcionamiento del dispositivo</i> .....	38
2.3.	Metodología Cascada.....	38
2.3.1.	Análisis .....	39
2.3.1.2.	Delimitación de Objetos. ....	41
2.3.2.	Diseño .....	41
2.3.2.1.	Planteamientos de Solución. ....	41
	Raspberry PI. ....	42
2.4.	Implementación.....	45
2.4.1.	<i>Implementación de Hardware</i> .....	46
2.4.2.	<i>Adquisición de Imágenes</i> .....	48
2.4.3.	<i>Etiquetado de Imágenes</i> .....	51
2.4.4.	<i>Configuración de la Red Neuronal Convolucional</i> .....	52
2.4.5.	<i>Entrenamiento de la Red</i> .....	56
2.5.	Preguntas de Investigación.....	58
CAPÍTULO III: RESULTADOS Y ANÁLISIS .....		59
3.1.	Verificación.....	59
3.1.1.	<i>Detección de Distancia:</i> .....	60
3.1.1.1.	Alertas e Interacción del Usuario.....	60
3.1.1.2.	Validación del sensor de distancia CJVL53LOXV. ....	61
3.1.2.	<i>Matriz de Confusión</i> .....	63
3.1.2.1.	Por Cada Variable.....	63
3.1.2.2.	Múltiples variables. ....	64
3.2.	Discusión .....	66
CONCLUSIONES.....		68
RECOMENDACIONES .....		69
REFERENCIAS.....		70
ANEXOS.....		74

## ÍNDICE DE TABLAS

<b>Tabla 1</b> <i>Denominación de objetos que interrumpen la movilidad.</i> .....	41
<b>Tabla 2</b> <i>Características del sensor VL53L0X.</i> .....	43
<b>Tabla 3</b> <i>Consumo de energía del dispositivo.</i> .....	46
<b>Tabla 4</b> <i>Estimación de precios de dispositivos electrónicos.</i> .....	48
<b>Tabla 5</b> <i>Determinación de los objetos con el respectivo simbolo Bliss.</i> .....	49
<b>Tabla 6</b> <i>Pruebas de precisión del sensor VL53L0X.</i> .....	61
<b>Tabla 7</b> <i>Tabla de resultados de cada objeto.</i> .....	63
<b>Tabla 8</b> <i>Matriz de confusión de las variables tomadas como referencia.</i> .....	65

## ÍNDICE DE FIGURAS

<b>Figura 1</b> <i>Cuadro de personas no videntes en el Ecuador.</i> .....	14
<b>Figura 2</b> <i>Esquema de discapacidad visual.</i> .....	19
<b>Figura 3</b> <i>Bastón blanco con franja roja desmontado, equipo que reconoce a las personas no videntes en su totalidad.</i> .....	20
<b>Figura 4</b> <i>Ejemplo de atributos de programación orientada a objetos.</i> .....	21
<b>Figura 5</b> <i>Aprendizaje que realiza la Inteligencia Artificial.</i> .....	23
<b>Figura 6</b> <i>Clasificación de los algoritmos de aprendizaje profundo.</i> .....	26
<b>Figura 7</b> <i>Estructura de una CNN.</i> .....	27
<b>Figura 8</b> <i>Ejemplo de cómo trabaja la capa de convolución.</i> .....	28
<b>Figura 9</b> <i>Ejemplo de una capa Pooling.</i> .....	28
<b>Figura 10</b> <i>Arquitectura del modelo MobileNet.</i> .....	29
<b>Figura 11</b> <i>Diferentes métodos de escalado frente a escalado compuesto.</i> .....	30
<b>Figura 12</b> <i>Arquitectura del modelo ResNet.</i> .....	30
<b>Figura 13</b> <i>Matriz de Confusión de 2x2.</i> .....	31
<b>Figura 14</b> <i>Ejemplos de datos que marcan la regresión operativa.</i> .....	33
<b>Figura 15</b> <i>Estructura del modelo cascada, resumido.</i> .....	39
<b>Figura 16</b> <i>Diagrama de la investigación por parte del hardware.</i> .....	42
<b>Figura 17</b> <i>Estructura Hardware de una Raspberry Pi 4.</i> .....	43
<b>Figura 18</b> <i>Sensor de distancia VL53L0X.</i> .....	43
<b>Figura 19</b> <i>Diagrama de bloque de la ejecución del software.</i> .....	44
<b>Figura 20</b> <i>Diagrama de flujo de la programación.</i> .....	45

<b>Figura 21</b> <i>Diagrama de funcionamiento del dispositivo.</i> .....	46
<b>Figura 22</b> <i>Montaje de los botones en la carcasa.</i> .....	47
<b>Figura 23</b> <i>Diseño de elemento para sostener la cámara y sensor.</i> .....	47
<b>Figura 24</b> <i>Ejemplos de Símbolos Bliss.</i> .....	49
<b>Figura 32</b> <i>Pantalla principal del programa labelImg.</i> .....	51
<b>Figura 33</b> <i>Definición de argumentos de entrada.</i> .....	52
<b>Figura 34</b> <i>Definición de la resolución de la cámara.</i> .....	54
<b>Figura 35</b> <i>Diagrama de flujo del VideoStream.</i> .....	54
<b>Figura 36</b> <i>Estructura de la programación por hilos.</i> .....	55
<b>Figura 37</b> <i>Mapa de etiquetas de los objetos.</i> .....	57
<b>Figura 38</b> <i>Perdida de rango de error.</i> .....	59
<b>Figura 39</b> <i>Grafico de perdida del detector de objetos.</i> .....	60
<b>Figura 40</b> <i>Curva de ROC de los objetos.</i> .....	64

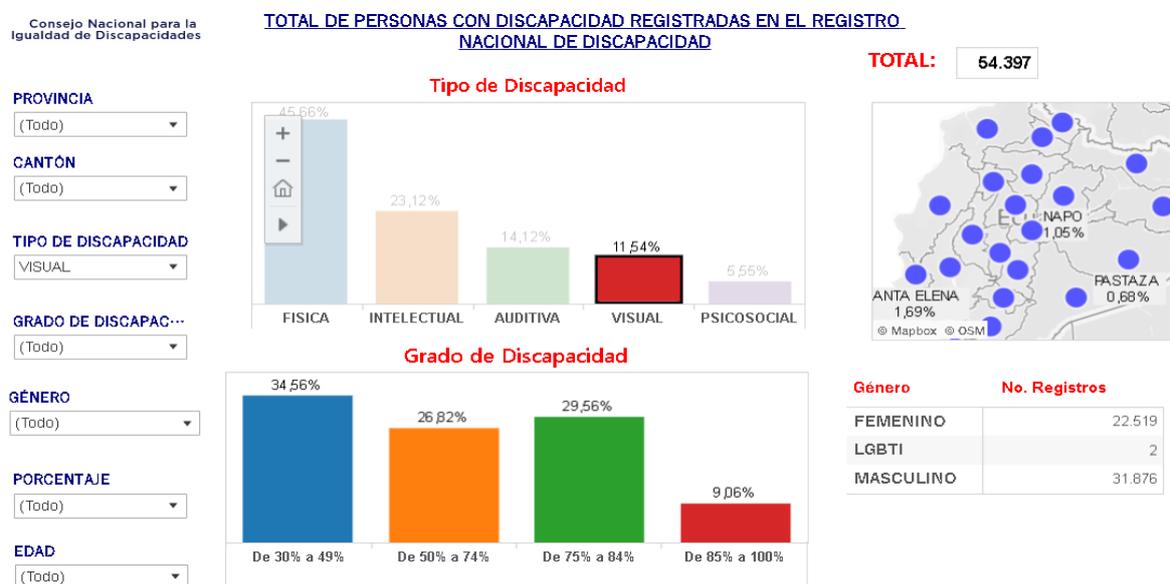
## INTRODUCCIÓN

### Problema de investigación

En el Ecuador, según los registros del Consejo Nacional para la Igualdad de Discapacidades actualizado en enero del 2022, existe un total de 54.397 personas no videntes que conforman el 11,54% de las personas con discapacidad en el país, En la Figura 1 se observa la tabla de personas con discapacidad visual registradas.

### Figura 1

*Cuadro de personas no videntes en el Ecuador.*



Fuente: Imagen extraída de [1].

El caminar de las personas no videntes ha sido uno de las grandes dificultades para ejercer su autonomía e independencia, desde hace varios años se ha vuelto indispensable la utilización del bastón para la detección de objetos o desniveles y así poder esquivarlos, también se ha visto la manera de involucrar a familiares o animales que les sirva como guía para movilizarse e idearse un mapa geográfico mental para reconocer el entorno por donde transitan, pero esta acción solo lo realizan cuando ejecutan actividades de cotidianidad, si existen modificaciones en su entorno habitual, es probable que la persona se desoriente y siga otro tipo de ruta.

Las personas no videntes, a pesar de su discapacidad, tienen la dificultad de movilizarse, de tropezar con obstáculos que se pueden encontrar en la calzada o parte superior

de las calles, ya sea por las distintas promociones que ponen las tiendas comerciales, ramas de árboles, letreros, escalones, objetos resbalosos, etc.

El uso del bastón brinda seguridad en la movilidad de las personas con discapacidad visual, ya que la oportuna anticipación en la detección de objetos, puede esquivarlos y evitar tropezarse, pero si existen cambios en el entorno o el bastón no detecta al objeto con el cual se pueda tropezar e interrumpir su movilidad, puede provocar golpes en su cuerpo o caídas las cuales provocan grandes lesiones, y el acceso a dispositivos tecnológicos o herramientas que les permitan obtener una mejora movilidad e inclusión en la cotidianidad de sus actividades es limitada o de alto costo.

## **Objetivos**

### ***Objetivo General***

- Implementar un dispositivo electrónico para la detección de obstáculos en la movilidad de las personas con discapacidad visual por medio de visión artificial.

### **Objetivos Específicos**

- Analizar los sistemas de detección de obstáculos con el uso de visión artificial para personas no videntes.
- Diseñar el dispositivo de detección de obstáculos para personas no videntes.
- Aplicar el dispositivo de detección de obstáculos.
- Validar el funcionamiento del dispositivo.

## **Justificación**

Actualmente se ha creado diversos dispositivos que ayudan a disminuir las limitaciones de movilidad en la detección de objetos, pero su excesivo costo, la poca satisfacción e incomodidad de utilizarlos en su vida cotidiana hacen que solo sea prescindible la utilización de métodos convencionales.

El presente proyecto es innovador, ya que mediante la utilización de las nuevas tecnologías se busca implementar un prototipo que ayude a la movilidad de las personas con discapacidad visual y que cumpla con los requerimientos manifestados por las personas que padecen esta discapacidad, colaborando en el ámbito social.

El desarrollo del dispositivo pretende tener una buena acogida ya que procura ofrecer confianza e independencia necesaria en el desarrollo de actividades cotidianas de las personas con discapacidad visual, con la finalidad en la investigación, de dar impulso en la creación de nuevos dispositivos que ayuden a las personas con discapacidad a superar las limitaciones.

Existen varios temas de investigación que utilizan los diferentes componentes electrónicos para la detección de objetos, pero estos son de mínimo alcance no superiores a un metro de localización de obstáculos y con un tiempo de respuesta de uno a dos segundos, lo que se plantea en este proyecto obtener un mayor rango de alcance para la verificación de objetos y un tiempo de detección más reducido, con esto se pretende mejorar el tiempo de movilidad de las personas no videntes.

### **Alcance**

El tema de investigación se desarrollará con la ayuda de un familiar que vive en la ciudad de Bolívar provincia del Carchi y el cual presenta discapacidad visual, considerando en construir un solo prototipo para validar la construcción y funcionamiento del dispositivo. Se está desarrollando diversas investigaciones con el uso de la visión artificial, por lo que se pretende desarrollar un dispositivo que ayuden a la detección de objetos, junto con el software de simulación Python en donde se desarrolla el código de programación para el funcionamiento del dispositivo.

En una de las especificaciones aportadas por la persona no vidente se considera que el dispositivo debe ser de fácil implementación y adaptación al cuerpo, con la utilización de botones de mando de encendido y apagado que se adapte al bastón para que funcione el dispositivo, con estructura y armazón resistente y una duración prolongada de la batería, se toma en consideración implementar una interfaz gráfica en el celular donde se reproduzca la señal de alerta con identificación de los objetos e indicación de la posición en cual se encuentra para ser evadido, así como también vibradores ubicados en tres regiones distintas para que emitan la señal en dirección donde se encuentra el objeto detectado.

## CAPÍTULO I: MARCO REFERENCIAL

### 1.1. Antecedentes

En la actualidad, los avances tecnológicos significan un bienestar para las personas porque dan un aporte para el desarrollo de prototipos que ayudan a mejorar el estilo de vida, en las personas con discapacidad visual, son una pieza importante para el desarrollo de nuevos proyectos tecnológicos.

En el Ecuador existen entidades que se preocupan en brindar ayuda a dichas personas, pero el acceso a dispositivos tecnológicos que le ayuden en su día a día es limitado, porque se ha comprobado que el alto costo de los aparatos tecnológicos que en verdad necesitan las personas con discapacidad visual es una limitante y en varios casos es algo imposible de adquirir.

Por eso es clave el desarrollo de nuevas técnicas de implementación de dispositivos electrónicos para el desarrollo de prototipos que se ajusten a las necesidades de las personas con discapacidad visual y que estas les ayuden a mejorar su vida, por ello el desarrollo de la tesis inspirado a ser un aporte para este tipo de personas.

La investigación realizada por [2], es una de las primeras elaboradas por una persona que padece discapacidad visual, donde trata de dar solución a su propio dilema de todos los días. Al llevar a cabo sus estudios en la carrera de ingeniería en mecatrónica, realiza un cinturón de alerta contra objetos. La investigación es tomada como base para el desarrollo del prototipo, en donde se pretende mejorar el alcance de detección de objetos con la implementación de nuevas tecnologías como es la visión artificial.

Cabe mencionar que la investigación de [3], es tomada como ejemplo para la contextualización del marco teórico, los diferentes parámetros y componentes que el prototipo debe tener.

Así como también, la investigación desarrollada por [4], en donde han desarrollado el prototipo con sensores ultrasónicos y al igual que la primera investigación tomada como referencia, tiene cierto rango mínimo de alcance, además que su infraestructura se denota con ciertos criterios de modificación. Por lo tanto, se busca mejorar estos temas de investigación y además de brindar una ayuda a las personas con discapacidad visual.

En la investigación ejecutada por [5], se analiza cómo estructurar las funciones principales de lo que es la visión artificial, así como la recolección de imágenes para la detección de objetos, esto para delimitar a qué objetos están orientados en las personas no videntes y poder esquivarlos.

## 1.2. Marco Teórico

### 1.2.1. La Ceguera

La [6] cataloga a la ceguera como una cifra total o parcial de la pérdida del sentido de la visión, esto quiere decir que, en la ceguera total, ya no puede percibir ningún tipo de luz, no pueden observar a ningún tipo de objetos, aunque ligeramente, de forma minuciosa puedan divisar un poco de luz, mientras que la ceguera parcial, pueden distinguir los objetos de forma borrosa y con gran dificultad, aquellas cosas que se interpongan en su camino a una distancia larga o corta.

También menciona que la ceguera se debe a distintos factores que pueden estar relacionados con la edad, genética, lesiones o trastornos, enfermedades donde su principal síntoma es la pérdida de visión, mientras que, aquellas enfermedades infecciosas son tratadas con gran éxito gracias a los avances clínicos que se han desarrollado para el diagnóstico preventivo o tratable en la pérdida de la visión. Pero las personas tienen la habilidad de adoptar otro tipo de estímulos, mientras pierden un sentido de su cuerpo, desarrollan otro y se ayudan con la utilización de herramientas o mecanismos que les ayude a moverse de un lugar a otro.

**1.2.1.1. Discapacidad Visual.** Como se ha mencionado, la [6] ha establecido diferentes tipos de rango en la pérdida de la visión, con la que ya se puede catalogar como una discapacidad en un grado de deterioro: deterioro de la visión distante y de la visión cercana, como se evidencia en la Figura 2 la escala que denomina el rango de discapacidad visual.

Deterioro de la visión distante:

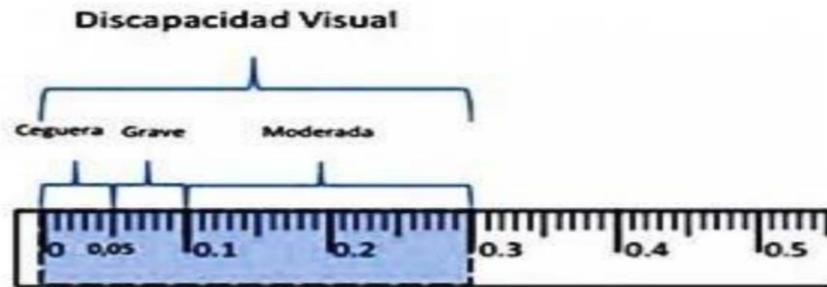
- Leve: agudeza visual inferior a 6/12 o igual o superior a 6/18.
- Moderado: agudeza visual inferior a 6/18 o igual o superior a 6/60.
- Grave: agudeza visual inferior a 6/60 o igual o superior a 3/60.
- Ceguera: agudeza visual inferior a 3/60.

Deterioro de la visión cercana:

- Agudeza visual cercana inferior a N6 o M.08 a 40 cm.

## Figura 2

*Esquema de discapacidad visual.*



Fuente: Imagen extraída de [7]

**1.2.1.2. Orientación del No Vidente.** La orientación del no vidente puede ser descrita como la capacidad que tienen las personas para desenvolverse dentro de un espacio, ya sea este abierto o cerrado, siempre está en constante actualización de establecer la posición que ocupa a través de la información sensorial, esto lo puedo sustentar [8].

Se ha determinado que las personas ciegas usan la estrategia de estar en constante contacto con la pared, entre objetos, y de objeto a entorno más cercano, todo esto para establecer un patrón o mapa geológico mental del lugar donde se encuentra, dándose una exploración periférica guiándose siempre de las paredes, describiendo los detalles más relevantes en la estructura de cada objeto por medio del tacto.

**1.2.1.3. Técnica de Movilidad con Bastón.** El uso del bastón, desde su aparición en la década de los años 1920, ha sido un instrumento de movilidad muy importante, es así como [9] lo asegura, donde este debe cumplir al menos con tres funciones principales: ofrecerle protección y apoyo al detectar objetos que podrían ocasionar golpes; orientación al tocar la superficie de un terreno transmite información de la estructura irregular; e información, ya que da un aviso a las demás personas que quien utiliza este instrumento padece de discapacidad visual., en esto también se destacan las siguientes técnicas:

- A. Sujeción o toma del bastón: funciona como un dedo índice que está completamente extendido, por lo que los dedos sujetan la parte inferior del mango, quedando el dorso de la mano hacia afuera.
- B. Técnica de toque: permite dar mayor seguridad en desplazamientos cortos, ubicando el bastón frente y al centro del cuerpo del no vidente, dando entre dos o tres toques a la superficie para saber cómo es esta en realidad.

C. Técnica más de un toque: es el tipo de técnica para desplazamientos más largos debido a que ayuda a la persona no vidente a que detecte irregularidades del lugar donde se encuentra en forma más segura.

D. Técnica para escaleras:

- Para subir: su principal propósito es brindar seguridad mientras se sube escaleras, donde el bastón mide la altura y distancia entre cada escalón.
- Para bajar escaleras: en este caso se puede contar con la ayuda de un pasamanos que le sirva de apoyo, donde en el borde de la escalera se realiza el reconocimiento de la altura y distancia que existe entre gradas.

**1.2.1.4. Bastón Blanco.** Él [10] asegura que en el año 1930 un norteamericano George A Benham, propuso el uso del bastón blanco con extremo inferior de color rojo para que sea un distintivo que identifique a las personas invidentes, esto también se dice que resulta más visible para los vehículos diferencien y reconozcan a una persona no vidente, como se describe en la Figura 3, donde se denota los colores característicos del bastón que utilizan las personas no videntes.

### Figura 3

*Bastón blanco con franja roja desmontado, equipo que reconoce a las personas no videntes en su totalidad.*



*Fuente:* Imagen extraída de [11].

**1.2.1.5. Percepción Auditiva.** La audición es un sentido en el que se perciben estímulos que pasan alrededor de una persona, al igual que la visión, este sentido favorece a la percepción de distancia, dando una idea de los acontecimientos que se están suscitando y en qué dirección se están realizando, dando a la localización de objetos y particularidades ambientales, aunque esto no especifica con exactitud las características físicas del objeto que está emitiendo el sonido, pero para poder comprender y entrenar al oído, la persona con discapacidad debe realizar una serie de tareas con las cuales entrene a su sentido auditivo, es así como lo expresa [12].

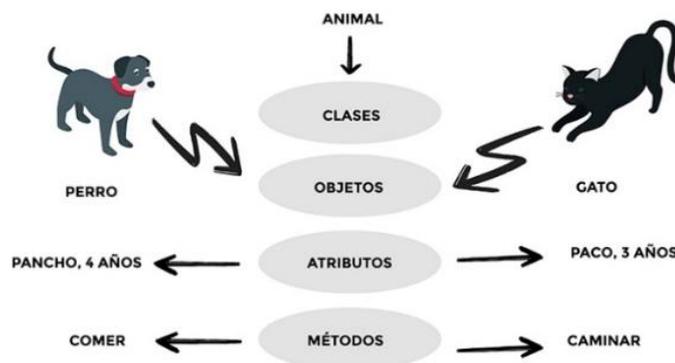
- A. **Habilidades Perceptivas:** el [12] describe que las personas con discapacidad visual, tienen que desarrollar y/o potenciar las habilidades que faciliten la información auditiva, a continuación, se describe las habilidades que se deben desarrollar:
- **Atención:** consiste en la atención que se debe dar a los sonidos del entorno.
  - **Identificación:** interpretación del área donde procede el sonido.
  - **Discriminación:** es la habilidad para seleccionar el sonido que desea escuchar, rechazando a la mezcla de los demás sonidos del entorno.
  - **Localización:** el sonido es una fuente de energía vibratoria que se diferencia en el momento de la llegada al oído.
  - **Seguimiento:** determina la dirección de la fuente sonora.
- B. **Entrenamiento auditivo:** todos los sentidos del cuerpo humano están desarrollados para recibir estimulaciones, mientras se pierde una habilidad, se desarrolla otra con más intensidad. Como toda habilidad, en el entrenamiento se pretende optimizar el oído, para percibir todas las variables de sonido y realizar un tipo de adiestramiento en un espacio interior, como lo enuncia [12], todo esto como se ha mencionado para perfeccionar cada una de las habilidades que se han denotado en las habilidades perceptivas.

### 1.2.2. Programación Orientada a Objetos

Utiliza una programación de software en piezas simples de código para la creación de pretensiones individuales de objetos, lo que se busca es dejar atrás la lógica de programación y remplazarla por el pensamiento de orientación a objetos, como se observa en la Figura 4 es así como lo define [13].

#### Figura 4

*Ejemplo de atributos de programación orientada a objetos.*



Fuente: Imagen extraída de [13].

Se emplean programas modernos como Java, C# y Python, que manejan diferentes paradigmas para definir los programas. Dejando al diseñador de software comportamientos e información relacionados en plantillas denominadas clase, en donde se crean objetos individuales a partir de esa plantilla de clase, en donde interactúan unos entre otros para crear un programa más grande que reconozca la programación de objetos designados. También menciona los principios de la programación orientada a objetos, los cuales son:

- La encapsulación define los parámetros más importantes del objeto con los que se desarrolla la programación y solo lo más importante al exterior.
- La abstracción se puede decir que es la interacción del usuario con los atributos que representa el objeto, esto facilita el mantenimiento de un código de extendida programación.
- La herencia define las relaciones jerárquicas entre las clases, de tal forma que los atributos puedan reutilizarse en otros programas, se pueden reproducir subclases en donde se puede ampliar los atributos de cada objeto.
- El polimorfismo se basa en la creación de objetos para difundir diferentes atributos, donde se puede procesar objetos en maneras diferentes, y se pueden generar con herencias.

**1.2.2.1. Python.** Es considerado como un lenguaje de programación de alto nivel de impacto, ya que su simplicidad y legibilidad es demasiado amigable para el usuario, donde permite al programador representar todas las ideas en líneas de código más simples, es por eso que es muy popular a nivel mundial, por lo que el entorno de Python puede ser compatible al ser códigos computacionalmente intensivos, como lo menciona [14].

### **1.2.3. Visión Artificial**

Para analizar lo que es visión artificial, hay que entender lo que es inteligencia artificial y dando un concepto entendible; se afirma como la capacidad que tiene las máquinas para analizar los algoritmos, aprender de los datos que se toman como información y tomar decisiones para que se ejecute una acción en pro beneficio de una actividad, como lo haría un ser humano, pero de manera autónoma, expresado en el documento de [15].

El primer término de visión artificial surgió en el año de 1959, la principal función era conectar una cámara de video hacia un ordenador. Pocos años atrás, un ordenador tenía solo la capacidad de almacenar imágenes, reproducirlas y cambiar ciertos parámetros según sea la

conveniencia, donde se muestra una gran cantidad de información que se representa en píxeles de almacenamiento de luz o color.

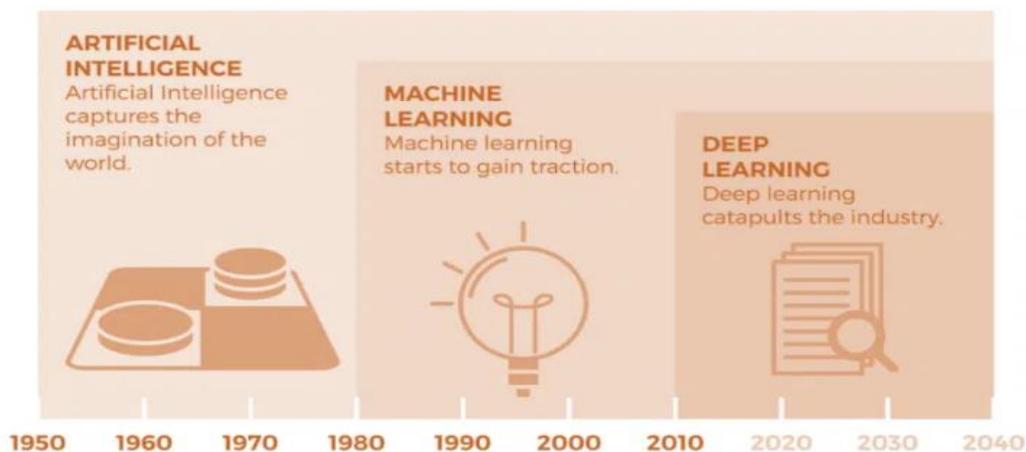
La visión artificial es un método científico cuyo objetivo principal es interpretar la información que aporta una imagen, se utilice diferentes algoritmos para que se realice una acción previamente programada y establezca un patrón esperado, pero esta acción debe ser realizada en tiempo real.

**1.2.3.1. Machine Learning.** Es un procedimiento de la Inteligencia Artificial, donde la principal función básica es que las máquinas aprendan de manera autónoma la toma de disposiciones por sí solas, para luego ser capaces de tomar predicciones o enviar sugerencias en cada elemento real que se analiza, también permite detectar y clasificar nuevos datos que sean correspondientes a los que se han ingresado, como lo manifiesta [16]. En la Figura 5 se evidencia como es la utilización de la inteligencia artificial para que los algoritmos sean entrenados y luego tomen las decisiones.

“En definitiva, el ‘Machine Learning’ es un maestro del reconocimiento de patrones, y es capaz de convertir una muestra de datos en un programa informático capaz de extraer inferencias de nuevos conjuntos de datos para los que no ha sido entrenado previamente” es así como lo expresa [17], donde la principal función es la mejora de sí mismo en la utilización de algoritmos.

### Figura 5

*Aprendizaje que realiza la Inteligencia Artificial.*



Fuente: imagen extraída de [16].

### Fases de la Machine Learning.

- Adquisición de datos — imágenes en base de datos.
- Creación de dataset — etiquetado de los datos obtenidos.
- Entrenamiento del modelo — se entrenan los datos después de la etiqueta.
- Evaluación del modelo — se analiza el comportamiento del modelo con nuevos datos que no han sido descritos en el entrenamiento.

**1.2.3.2. Deep Learning.** Es un tipo de método que se desglosa de una de las ramas de la Inteligencia Artificial, donde se pueden efectuar tareas que las personas la realizarán de forma automática, pero que para una máquina estas tareas serían complejas de realizar, se basa principalmente en el uso de las redes neuronales artificiales, que están inspiradas en las redes neuronales que tiene el cerebro humano. Estas redes neuronales artificiales reciben y transmiten impulsos, cuya información principal es transmitida hasta la red neuronal artificial final, la cual lleva a cabo una acción previamente establecida en función a los datos que se han ingresado, y toma las decisiones necesarias para ejecutar acciones, como lo manifiesta [18].

El principal funcionamiento de una Deep Learning, es la utilización de múltiples transformadas no lineales, consisten en alterar la posición de los píxeles que se encuentran en una imagen, pero manteniendo la intensidad de los mismos, por lo que permite verificar la posición de un objeto que está en movimiento.

Con lo que se logra que la Deep Learning se aplique en diversos objetos que se encuentran en diferentes ámbitos, por lo que se puede afirmar que este tipo de algoritmos están hechos para trabajar en entornos complejos y cambiantes, donde los algoritmos tradicionales no pueden obtener la información.

**1.2.3.3. Entrenamiento de Modelos de Aprendizaje Automático.** Estos modelos son para que se dé el entrenamiento del algoritmo, este paso es crucial para que el modelo aprenda y realice las predicciones. Para ejecutar las predicciones se necesita tener una gran cantidad de información y deben estar debidamente etiquetados, en cada entrenamiento se alimenta y refuerza los parámetros de los objetos de importancia.

Existen diferentes bibliotecas y software con entorno virtual para entrenar los modelos de aprendizaje automático, por las que se destacan Keras y TensorFlow.

A. Keras: es una biblioteca de código abierto que se puede utilizar en Python, su

principal objetivo es acelerar la creación de redes neuronales, por ello, Keras trabaja como una interfaz de uso intuitivo, proporciona bloques modulares sobre los que se acoplan modelos complejos, es más utilizado en operaciones de alto nivel.

Las principales ventajas de emplear código abierto es la aportación que sirve para crear redes neuronales, ya que simplifican este proceso. Los modelos producidos son fáciles de manejar y en diferentes plataformas como lo afirma [19].

- B. TensorFlow: es una biblioteca de computación numérica en código abierto para Machine Learning, basado en redes neuronales profundas, que ha sido construida por Google Brian Team, para realizar investigación y desarrollo en sistemas de producción, es fácil comenzar a construir, difundir y entrenar modelos con acceso a la API de Keras, como lo declara [20].

Es una plataforma que facilita a la creación e implementación de desarrollo que generan aprendizaje autónomo del modelo de manera fácil, que busca automatizar estos procesos, ya sea en servidores o en la web, presentando soportes para lenguaje C++ y Java. “Los desarrolladores no quieren perder más tiempo en el mantenimiento de la infraestructura o desarrollo de los procesos y lo que buscan es generar valor”, lo dice [21] en su página web.

#### ***1.2.4. Algoritmos en Visión Artificial***

Los algoritmos se definen como el conjunto de instrucciones o reglas previamente definidas que permiten delimitar y solucionar un problema, procesar diferentes datos que son programables para el reconocimiento de objetos previamente definidos, como lo manifiesta [16]. Se intenta que las máquinas lleguen a responder un problema empleando un conjunto de conocimientos previamente establecidos, llamados algoritmos de aprendizaje automático, como se especifica en la Figura 6.

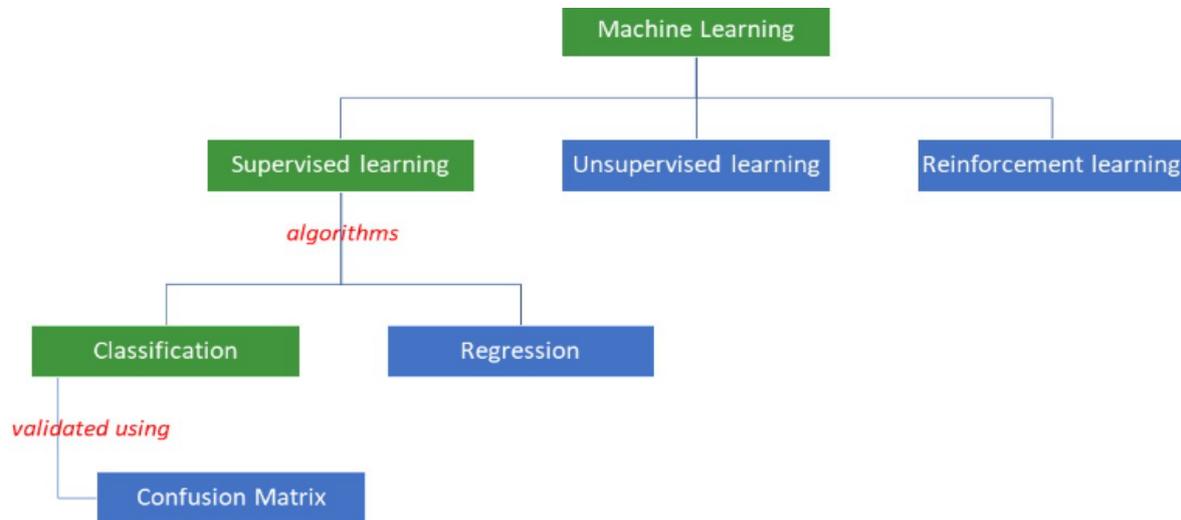
Los algoritmos se clasifican en: supervisados, no supervisados o de refuerzo. Los algoritmos supervisados, se tiene conocimiento de la respuesta y se entrena a las máquinas, para que lleguen a esa respuesta. Los algoritmos de aprendizaje automático supervisados se clasifican en: clasificación y regresión, como lo menciona [22].

Los algoritmos de aprendizaje profundo, supervisados de clasificación, son los que ordenan los datos de un conjunto con base en los criterios que el usuario le especifica. En el

documento de [23] se analizan algunos de los algoritmos de clasificación utilizados en visión artificial.

### Figura 6

*Clasificación de los algoritmos de aprendizaje profundo.*



Fuente: Imagen tomada de [22]

**1.2.4.1. Algoritmo de Regresión.** Es un tipo de algoritmo que utiliza una medida de error, la cual se intentara establecer como la mínima posible, para que se realice las predicciones lo más verdadero a la realidad, en un proceso que pretende compensar la respuesta más óptima en la implementación de un pequeño error.

**1.2.4.2. Algoritmo Basado en Instancias.** Este tipo de algoritmos son modelos que se ajustan en problemas de decisión con datos que son sometidos a entrenamiento, estos datos están basados en memoria, donde se compran con nuevos datos para verificar la similitud que existen entre esta combinación y poder ejecutar la predicción.

**1.2.4.3. Algoritmo de Árbol de Decisión.** Es empleado para la clasificación de información a partir de la toma de decisiones que tienen los valores reales de un modelo, de los atributos que se la asignan, para que se tomen los posibles caminos en mejorar la precisión del modelo.

**1.2.4.4. Algoritmo de Aprendizaje Profundo.** Es un tipo de algoritmo que maneja el amplio espacio de la tecnología, para exportar la cantidad de datos necesarios para la ejecución de un trabajo, colocar la información en varias redes neuronales que están interconectadas en diversas capas que se ejecutan en paralelo para hacer una acción.

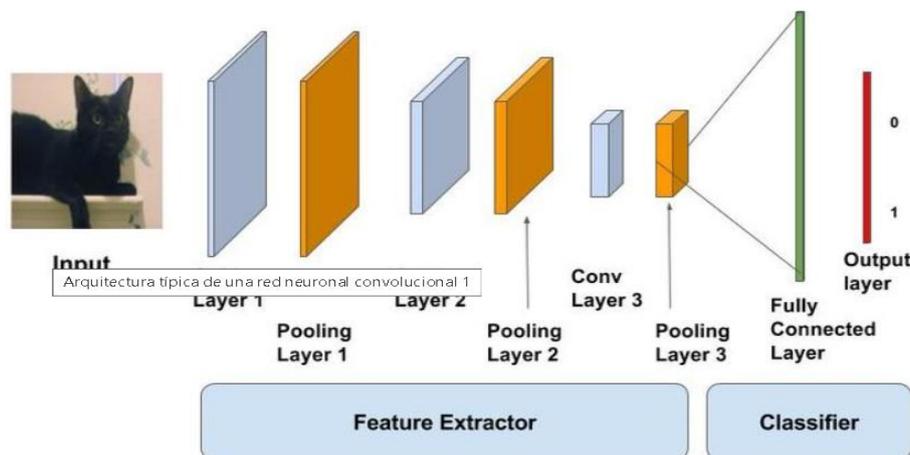
**1.2.4.5. Algoritmo De Redes Neuronales Convolucionales (CNN).** Son algoritmos que se basan en la funcionalidad de las redes neuronales de los humanos y son un

subconjunto de la Machine Learning. Aplica los conceptos de álgebra lineal, para identificar patrones específicos que se encuentran en una imagen, necesita de unidades de procesamiento gráfico (GPU) para entrenar cada modelo y de esta manera puede diferenciar entre cada imagen y catalogarla, como se observa en la Figura 7.

Trabaja en diferentes capas y nodos y al menos una de ellas debe ser de convolución, la primera capa es de entrada que contiene las características relevantes de un objeto en importancia, se introduce en las capas ocultas, ya trabaja en reconocer las formas específicas del objeto, hasta que la capa de salida mostrara una predicción de lo que se espera obtener de acuerdo a un valor de umbral detallado, como lo especifica [24].

### Figura 7

*Estructura de una CNN.*



Fuente: Imagen tomada de [25]

#### A. Tipos de capas

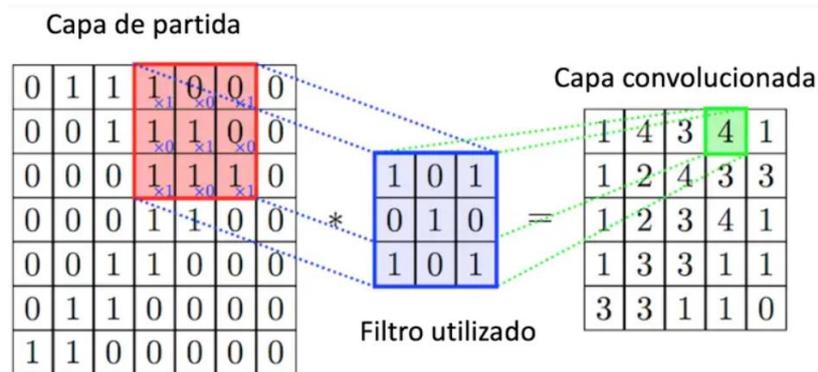
**Capa Convolución.** Después de la capa de entrada, en la convolución se encarga de enlazar cada nodo con una región específica de la entrada, para extraer los rasgos específicos de la imagen base, luego realiza operaciones de suma y producto entre la imagen de entrada y el filtro (kernel) que forma un mapa de particularidades correspondiente al posible lugar del objeto. La principal ventaja que ofrece esta capa es que el mismo filtro sirve para extraer las mismas peculiaridades en cualquier posición donde se encuentre el objeto de importancia de la imagen de entrada, como se evidencia en la Figura 8.

**Capa de Reducción – Pooling.** Cuando se haya completado la convolución, sigue la capa de reducción que acorta la cantidad de parámetros y solo se queda los más

relevantes. Esto se logra, mediante la extracción de padrones como el promedio o el máximo de una región fija del mapa, para ello se utiliza un nuevo filtro, pero manteniendo la principal información como se observa en la Figura 9, al reducir características se disipa precisión, sin embargo, mejora la compatibilidad, como lo asegura [26].

### Figura 8

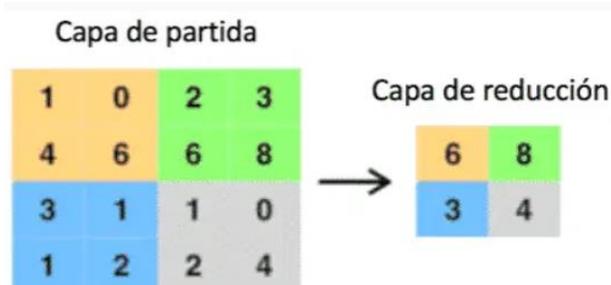
*Ejemplo de cómo trabaja la capa de convolución.*



Fuente: Imagen extraída de [26].

### Figura 9

*Ejemplo de una capa Pooling.*



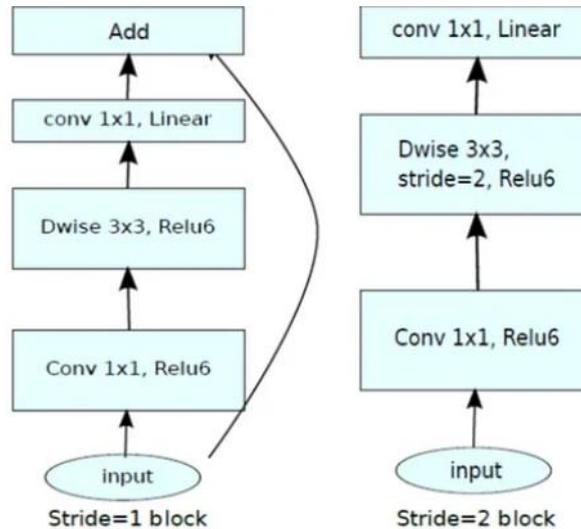
Fuente: Imagen extraída de [26].

## B. Tipos de arquitecturas

**MobileNet.** Es una serie de modelos creadas por Google que se basan en TensorFlow, para que sea empleado en dispositivos móviles, pero con recursos limitados, de baja latencia y poco consumo de energía, este modelo emplea arquitecturas simplificadas con el uso de convoluciones separadas con el fin de construir redes neuronales profundas y ligeras que comprimen la cantidad de operaciones por lo que hace al modelo más rápido y eficiente, como se muestra en la Figura 10. MobileNet ha demostrado comprarse con modelos más grandes y complejos que clasifican imágenes, como lo afirma [27].

**Figura 10**

Arquitectura del modelo MobileNet.



Fuente: Imagen extraída de [27].

**EfficientNet.** Es una de las arquitecturas más poderosas que se aplican para lograr mejores precisiones cuando exista más recursos. Usa una técnica de coeficiente compuesto para escalar modelos, donde escala uniformemente la dimensión y se ajusta para obtener la máxima precisión, como se muestra en la

, está constituida por una serie de bloques que emplean la convolución, el pooling y la activación, para tomar características distintivas de una imagen.

Esta arquitectura ha demostrado ser precisa en una diversidad de trabajos con visión por computadora, creando nuevos patrones de precisión en modelos de redes neuronales convolucionales, pero se desajusta si la red consume bastantes recursos vistos desde el punto computacional, como lo expresa [28].

**ResNet.** Es un modelo producido por Microsoft Research en el año 2015, se utiliza para resolver el problema del gradiente de fuga, este problema se deriva porque cuando la red es demasiado profunda, los gradientes a partir de la cual se calcula alguna función se reducen a cero, este resultado nunca actualiza los valores y, por lo tanto, no se realiza ningún aprendizaje.

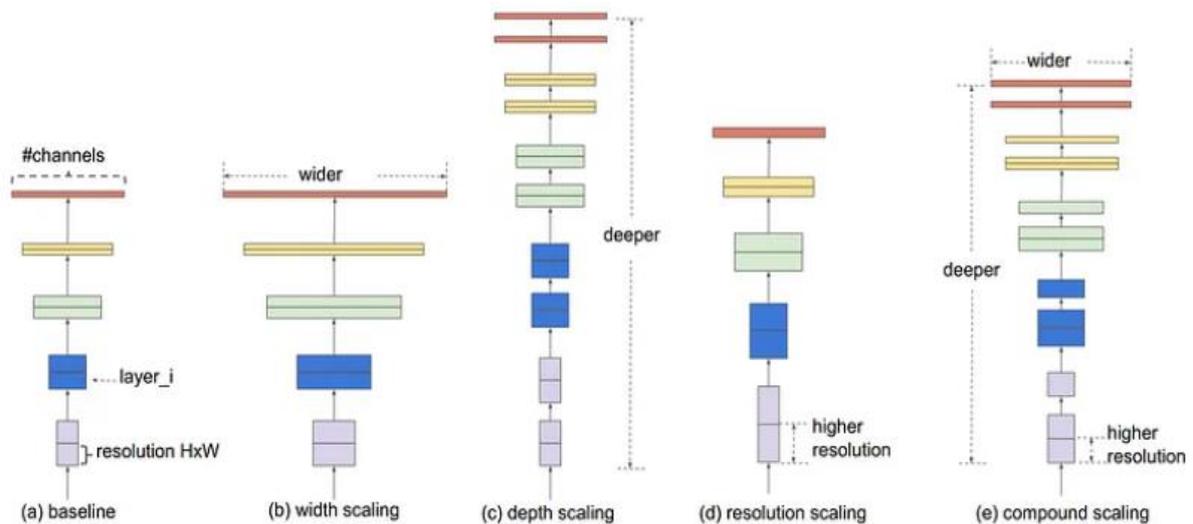
Los gradientes pueden fluir a través de las conexiones de salto hacia atrás desde las primeras capas, ya que utiliza bloques que contienen atajos, para saltar una o más capas,

evitando que se desvanezcan los gradientes a medida que pasan por capas profundas, como se evidencia en la Figura 12.

La principal ventaja que tiene este tipo de conexión es que, si alguna capa daña el rendimiento de la arquitectura, la regularización omitirá, por lo que da como resultado el entrenamiento de la red neuronal sin los problemas causados por el gradiente de desaparición/explosión, como lo asegura [29].

**Figura 11**

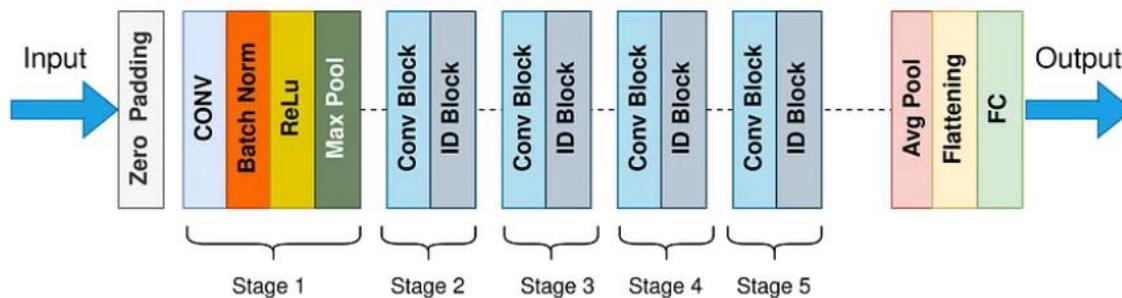
*Diferentes métodos de escalado frente a escalado compuesto.*



Fuente: Imagen tomada de [28].

**Figura 12**

*Arquitectura del modelo ResNet.*



Fuente: Imagen extraída de [29].

### 1.2.5. Matriz de confusión

Es también distinguida como una matriz de error, donde se plantea una tabla que se utiliza para evaluar el rendimiento del modelo que tiene la imagen, la cual se está analizando, donde se muestra las predicciones correctas e incorrectas en un número exacto de valores que

se extraen en diferentes clases [30]. En la Figura 13, se muestra un ejemplo de una matriz de confusión de 2x2.

En la tabla se muestra un número de filas que representan las clases reales que deberían ser resueltas, mientras que las columnas representan las posibles predicciones que se han propuesto, por lo que esta tabla, verifica fácilmente las predicciones erróneas que se han planteado.

### Figura 13

*Matriz de Confusión de 2x2.*

		Actual Values	
		Yes	No
Predicted Values	Yes	True Positive	False Positive
	No	False Negative	True Negative

*Fuente:* Imagen extraída de [30].

#### 1.2.5.1. Estructura de una Matriz de Confusión.

Positivo (P): la observación es positiva.

Negativo (N): la observación no es positiva.

Verdadero Positivo (TP): resultado en el que el modelo predice correctamente la clase positiva.

Verdadero Negativo (TN): resultado donde el modelo predice correctamente la clase negativa.

Falso Positivo (FP): también llamado error de tipo 1, resultado donde el modelo predice incorrectamente la clase positiva cuando en realidad es negativa.

Falso Negativo (FN): también llamado error de tipo 2, un resultado en el que el modelo predice incorrectamente la clase negativa cuando en realidad es positiva [30].

También en el documento se muestra las siguientes propiedades que debe tener una matriz de confusión con su respectiva ecuación, que son consideradas de gran importancia.

Exactitud — Es similar al pronóstico que el modelo clasificó correctamente, visto aplicando la Ecuación  $ACC = exactitud = \frac{\# \text{ predicciones correctas}}{\text{total \# predicciones}} = \frac{TP + TN}{TP + TN + FP + FN}$  (1).

$$ACC = exactitud = \frac{\# \text{ predicciones correctas}}{\text{total \# predicciones}} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precisión — Es el número de predicciones positivas que son instancias relevantes entre las instancias recuperadas, se puede apreciar aplicando la Ecuación  $PPV = precision = \frac{TP}{TP + FP}$  (2).

$$PPV = precision = \frac{TP}{TP + FP} \quad (2)$$

Sensibilidad — Es la tasa de aciertos o tasa positiva real, o también dicho son la cantidad de instancias pertinentes que se recuperaron realmente, que se evidencia con la ejecución de la Ecuación  $sensibilidad = \frac{TP}{TP + FN}$  (3).

$$sensibilidad = \frac{TP}{TP + FN} \quad (3)$$

Especificidad — Es la tasa negativa real, la cual se encarga de medir las instancias de negativos reales, es lo opuesto a la sensibilidad, visto lo antes dicho en la  $especificidad = \frac{TN}{TN + FP}$  (4).

$$especificidad = \frac{TN}{TN + FP} \quad (4)$$

Puntuación F1 — Denominada como la media armónica de precisión y sensibilidad, que puede tomar valores de puntuación máxima de 1 y una mínima de 0, se calcula mediante  $lapuntuacion \quad F1 = \frac{2TP}{2TP + FP + FN}$  (5).

$$\text{puntuacion } F1 = \frac{2TP}{2TP+FP+FN} \quad (5)$$

**1.2.5.2. Curva de ROC.** La palabra curva ROC proviene del significado (Curva de Características Operativas del Receptor), la cual es un método estadístico con representación gráfica que se ajusta a un modelo de regresión cuando la variable es en respuesta binaria, con el que se evalúa que tan bien se adapta un modelo de regresión logística contra un conjunto de datos, como lo manifiesta [31] y se pueden apreciar las siguientes métricas:

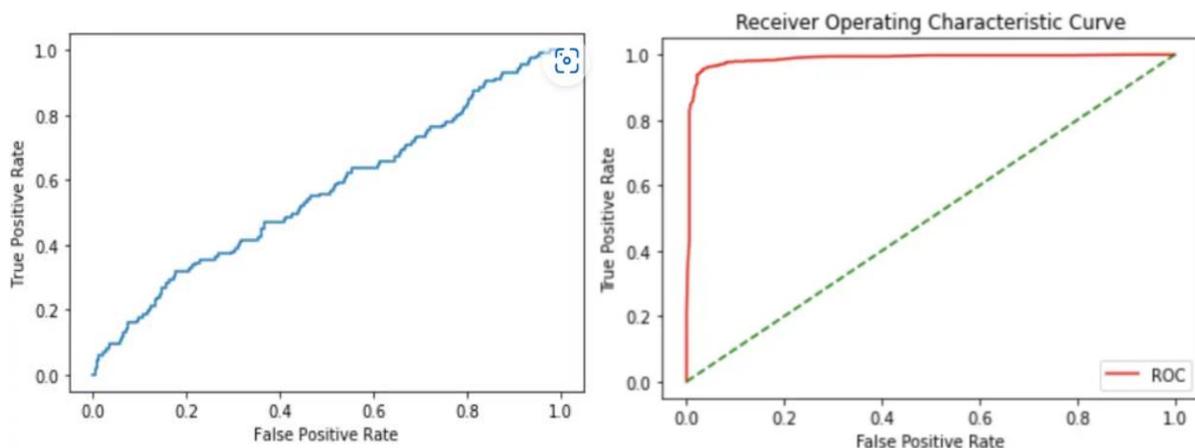
Sensibilidad — la probabilidad de que el modelo prediga un resultado positivo en una realidad de resultados positivos.

Especificidad — la probabilidad de que el modelo prediga un resultado negativo en una observación cuando la realidad de resultados es negativa.

Las dos métricas son puestas en gráfica con los respectivos datos, donde la sensibilidad se reproduce en el eje de las ordenadas, mientras que la especificidad se reproduce en el eje de las abscisas, como se mira en la Figura 14.

#### Figura 14

*Ejemplos de datos que marcan la regresión operativa.*



a) La imagen muestra una curva pobre  
Fuente: Imagen extraída de [31]

b) La imagen muestra una curva de ROC perfecta

Por lo que se puede afirmar que mientras la curva se acerque a la esquina superior izquierda, la categorización de los datos se realiza de mejor manera, como se aprecia en el segmento b), donde muestra que los verdaderos positivos se asemejan a datos verdaderos mientras que las predicciones o falsos positivos no se asemejan a la realidad, como es el caso de la imagen a). Para determinar la curvatura es necesario conocer el área bajo la curva con

forme se da la Sensibilidad y Especificidad, con el que se cuente con un estimador de variables que cuente con parámetros ajustables, en donde se definen los conceptos de *Tasa de Verdaderos Positivos (TVP) = TP / (TP + FN)* (6) y *Tasa de Falsos Positivos (TFP) = FP / (FP + TN)* (7):

$$\text{Tasa de Verdaderos Positivos (TVP)} = TP / (TP + FN) \quad (6)$$

$$\text{Tasa de Falsos Positivos (TFP)} = FP / (FP + TN) \quad (7)$$

## CAPÍTULO II: MATERIALES Y MÉTODOS

### 2.1. Modelo de la Investigación

La detección de objetos se piensa que es una tarea sencilla de realizar, pero en la realidad es una tarea difícil, es uno de los grandes retos que se tiene en el área de la visión artificial por computadora, ya que el principal objetivo es encontrar e identificar uno o varios objetos en una imagen, donde se debe combinar grandes tareas importantes como son, la localización y la clasificación de los objetos. Es por ello la importancia de la utilización de métodos tradicionales, así también la utilización de nuevas técnicas de aprendizaje, que están permitiendo llevar a cabo una mejor precisión y exactitud en detectar objetos que son de interés, a pesar de las dificultades o limitaciones que se presenten, como pueden ser el ambiente, el tamaño y posición del objeto, la iluminación o condición climática y varias restricciones más.

La investigación que se desarrolló tiene como propósito resolver un problema social en particular, ejecutando la investigación aplicada, que trata de plantear un dispositivo electrónico que reconozca diferentes objetos que puedan obstruir la movilidad de las personas que tengan discapacidad visual.[32] Para la recolección de información, se planteó una entrevista [33], donde el principal objetivo de cada una de las preguntas, es asegurarse que el dispositivo a desarrollar tenga la eficiencia y comodidad que la persona no vidente pretende obtener, además de establecer algunos datos a través de la observación para analizar como la persona con discapacidad efectúa sus actividades y que el dispositivo a implementar no altere su vida cotidiana.

Se procedió con la investigación documental, ya que se lleva un proceso de análisis de búsqueda de información, ya sea esto en exploraciones anteriores sobre el tema de interés, artículos científicos, libros [34], y se puede mencionar que la investigación es de campo, por el hecho de que el dispositivo a desarrollar es para una persona en específico, donde se busca satisfacer los requerimientos que la persona establece.

Se puede mencionar que esta investigación es de tipo experimental, puesto que se pretende implementar un dispositivo que no es invasivo, pero que tenga funcionalidad para ejecutar pruebas y que cada vez se mejore su funcionalidad [34]. También, se aplicó el método inductivo para buscar las falencias del dispositivo en la dificultad de la movilización por

pérdida de la vista, a su vez en el método deductivo se desarrollan diagnósticos de donde sería más eficaz la implementación del dispositivo en el cuerpo de la persona no vidente [35].

## **2.2. Diseño de la Investigación**

En el capítulo se trata de redactar cada una de las actividades que se realizan para dar con el cumplimiento de cada uno de los objetivos específicos, dando un detalle de manera secuencial del proceso para la implementación del dispositivo reconocedor de objetos para las personas no videntes.

### ***2.2.1. Análisis de los diferentes sistemas de detección de obstáculos con el uso de visión artificial***

#### **Actividad 1: “búsqueda de información acerca de los sistemas que utilicen visión artificial para la detección de objetos”**

Se considera diferentes documentos que contengan información sobre la construcción de dispositivos que evaden obstáculos, también sobre el lenguaje de programación para el reconocimiento de objetos en visión artificial en software de simulación libre.

#### **Actividad 2: “entrevista”**

Realización de una entrevista a la persona no vidente que colabora en el desarrollo de la investigación para escuchar cuáles son los inconvenientes que tiene al movilizarse de un lugar a otro dentro del lugar donde más desarrolla sus actividades.

#### **Actividad 3: “recopilación de requerimientos”**

Establecida la entrevista, se analizan los requerimientos más relevantes que la persona no vidente desearía poseer en la construcción del diseño del dispositivo en la detección de objetos.

### ***2.2.2. Diseño del dispositivo de detección de obstáculos***

#### **Actividad 1: “análisis de sistema según los requerimientos”**

Entendido cada uno de los requerimientos que la persona no vidente desea tener en el dispositivo, se realiza una investigación bibliográfica de los diferentes sistemas para la detección de objetos, tanto en hardware como en software.

### **Actividad 2: “flujograma y diagrama de bloques”**

Dada la revisión de las diferentes investigaciones sobre la detección de objetos, se plantea como se debe efectuar la secuencia de construcción del hardware y del software.

### **Actividad 3: “Implementación de metodologías”**

Se analiza el entorno virtual donde se pueda realizar proceder la programación y se dé la recolección de datos necesarios para que el sistema los analice y se ejecute el comando necesario para esquivar a objetos determinados.

### **Actividad 4: “obtención de datos”**

Se llevan a cabo simulaciones para discriminar los falsos positivos y negativos para cada uno de los objetos con que la persona no vidente se enfrenta día a día en sus actividades.

## ***2.2.3. Aplicación del dispositivo de detección de obstáculos***

### **Actividad 1: “Selección de elementos”**

Dada la comprobación de las diferentes investigaciones sobre la detección de objetos, se analiza los dispositivos electrónicos que mejor rango de alcance tenga y verificar la disponibilidad del mercado local.

### **Actividad 2: “Implementación de dispositivos”**

Se analiza implementar una tarjeta electrónica que ejecute los comandos de programación y obtenga los datos para el procesamiento de los mismos, además de la inclusión de los diferentes dispositivos electrónicos para la detección de objetos.

### **Actividad 3: “realización de simulaciones”**

Considerado el ensamblaje electrónico de cada uno de los dispositivos, se comprueba el funcionamiento de los mismos y se procede a realizar pequeñas simulaciones en el reconocimiento de objetos.

#### **2.2.4. Validación del funcionamiento del dispositivo**

Es aquí donde la investigación experimental tendrá mayor relevancia porque se dará las pruebas de funcionamiento del dispositivo con la persona no vidente que ayuda en esta investigación.

##### **Actividad 1: “pruebas de funcionamiento”**

Con la implementación completa de todos los dispositivos electrónicos, se procede a realizar diferentes pruebas de reconocimiento de objetos en diferentes escenarios, para establecer el rango de eficiencia en el reconocimiento de objetos.

##### **Actividad 2: “modificaciones de funcionamiento”**

Llevadas a cabo las pruebas de simulación, se considera el rango de alcance de los dispositivos electrónicos y se es necesario realizar modificaciones, para que esta eficiencia alcance un rango considerado se las realiza.

##### **Actividad 3: “ajustes de funcionamiento”**

Lo que se busca en esta actividad es corregir los errores que presenta el dispositivo para que el reconocimiento de objetos se asemeje a una eficiencia óptima y garanticen la movilidad de la persona no vidente.

### **2.3. Metodología Cascada**

Es un modelo de procedimiento lineal que se especifica por dividir un todo en procesos de desarrollo sucesivos, donde cada fase solo se ejecuta una sola vez y los resultados obtenidos de una fase sirven como hipótesis para la siguiente fase, como lo menciona [36].

Cabe recalcar que el modelo se le atribuye al teórico de la informática Wiston W. Royce. Aunque no es el creador del modelo, lo presento como una alternativa de un modelo interactivo creciente en el que cada fase se basa en la anterior. Royce plantea un modelo combinado por siete fases:

1. Requisitos de sistema
2. Requisitos de software
3. Análisis

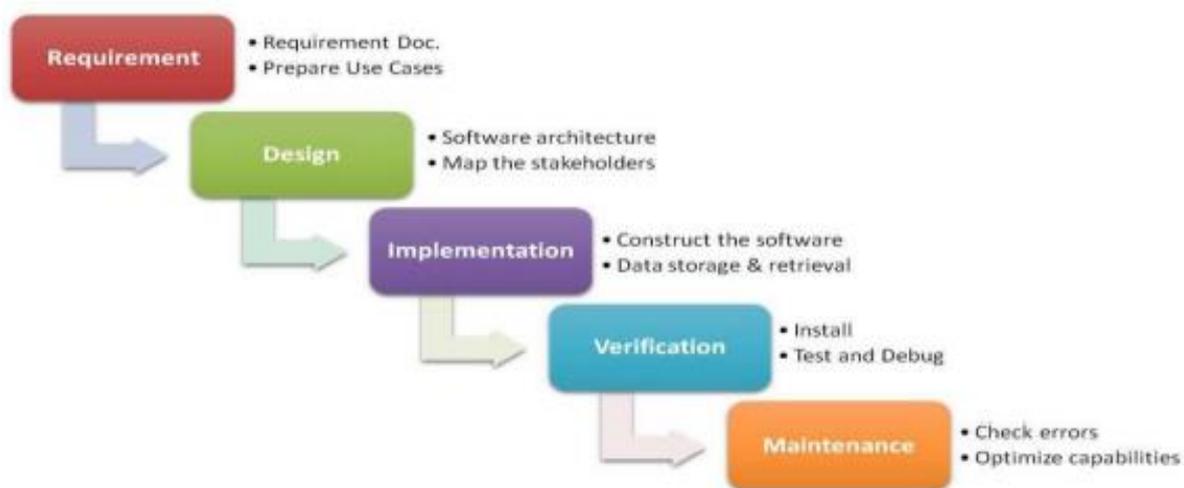
4. Diseño
5. Implementación
6. Prueba
7. Servicio

Para una investigación, las fases mencionadas por Royce 1, 2 y 3 se fusionan en una sola fase a modo de análisis de requisitos, como se explica en la siguiente Figura 15, solo se concentran en el análisis que necesita la investigación.

1. Análisis: fase de análisis, planificación y definición de requisitos.
2. Diseño: etapa de diseño, arquitectura y especificación del dispositivo.
3. Implementación: programación y pruebas de funcionamiento.
4. Verificación: validación y pruebas del dispositivo.
5. Mantenimiento: entrega, mejora y mantenimiento del dispositivo.

**Figura 15**

*Estructura del modelo cascada, resumido.*



*Fuente:* La imagen fue extraída de [36].

### 2.3.1. Análisis

En este apartado, se busca coleccionar la mayor cantidad de información sobre los requisitos para diseñar el dispositivo, teniendo en cuenta siempre la opinión de la persona no vidente, realizando una investigación de campo con visita técnica donde se plantea una entrevista con preguntas puntuales y concretas, para definir correctamente qué parámetros son

los que se deben tomar en cuenta para realizar la arquitectura del dispositivo y la programación del modelo.

**2.3.1.1. Especificaciones del Sistema.** Se presenta un listado de criterios que el usuario desearía que se implemente en el dispositivo, con un diseño sencillo, fácil de manipular y de utilizar, obtenidos a través del análisis cualitativo, que se especifican en la entrevista realizada. Criterios que aporten al diseño y construcción del dispositivo para el reconocimiento de objetos.

**Componentes electrónicos.** Dispositivo que tengan un rango de alcance de entre 1 a 3 metros de distancia. Con la utilización de tarjeta electrónica u ordenador para el procesamiento y ejecución de datos.

**Tiempo de respuesta.** El reconocimiento de objetos, el análisis de datos y la respuesta debe ser la menor posible para que la persona no vidente pueda saber la ubicación del objeto a esquivar y efectúe una acción evasiva.

**Diseño.** Se plantea un diseño sencillo para a la vez versátil y fácil de comprender su funcionamiento, además que debe acoplarse a la zona torácica superior.

**Ergonomía.** Se debe acoplar a la anatomía de la persona, con la ideología de acoplarse en la parte torácica superior.

**Peso.** El dispositivo debe ser lo más ligero posible, la persona que lo va a utilizar es un adulto mayor y no debe sobrepasar los 400 g.

**Geometría.** Las dimensiones del dispositivo se acoplarán al ensamblaje de los dispositivos electrónicos.

**Precio.** El dispositivo debe ser lo más económico posible, tomando como referencia los precios comerciales y no debe sobrepasar los \$550 .00 en dispositivos similares.

**Mecanismo.** La infraestructura del dispositivo debe tener diferentes ajustes.

**Innovación.** Se trata de crear un dispositivo que no sea común a los de tipo comercial, pero con mayor funcionalidad.

**Interfaz.** El dispositivo tendrá que ser fácil de utilizar, acoplándose con una tarjeta de mando.

**Materiales.** Para el recubrimiento de los dispositivos eléctricos, se realizará una carcasa plástica en impresión 3D que emplee un plástico resistente y donde se incorporen todos los elementos.

**Arquitectura.** El dispositivo se compondrá de sensores de distancia y con cámara de óptima resolución.

**2.3.1.2. Delimitación de Objetos.** De acuerdo con la entrevista a la persona con discapacidad visual, que es el usuario final del dispositivo, se ha determinado una lista de objetos con los que, con mayor frecuencia, le llegan a interrumpir la movilidad dentro de su hogar, se muestran en la Tabla 1, están descritos de acuerdo con el grado de importancia que manifiesta la persona.

**Tabla 1**

*Denominación de objetos que interrumpen la movilidad.*

<i>Número Objeto</i>	<i>Denominación</i>
1	Cama
2	Silla
3	Basurero
4	Quebrada
5	Puerta
6	Baño
7	Cocina

### **2.3.2. Diseño**

Una vez establecidos los requerimientos que la persona no vidente desea tener en el dispositivo a diseñar, se procede plantear como realizar el desarrollo del módulo, tanto en la parte física, la arquitectura que va a tener y como se va a acoplar al cuerpo de la persona no vidente, así como también, la programación y los diferentes entornos a utilizar.

**2.3.2.1. Planteamientos de Solución.** Planteados los criterios impartidos por la persona no vidente que ayuda a la realización de esta investigación, se establece una propuesta de solución, con la implementación de los dispositivos electrónicos y entornos virtuales de simulación. A continuación, se detallan cada uno de los pasos a seguir, pero subdividiendo en

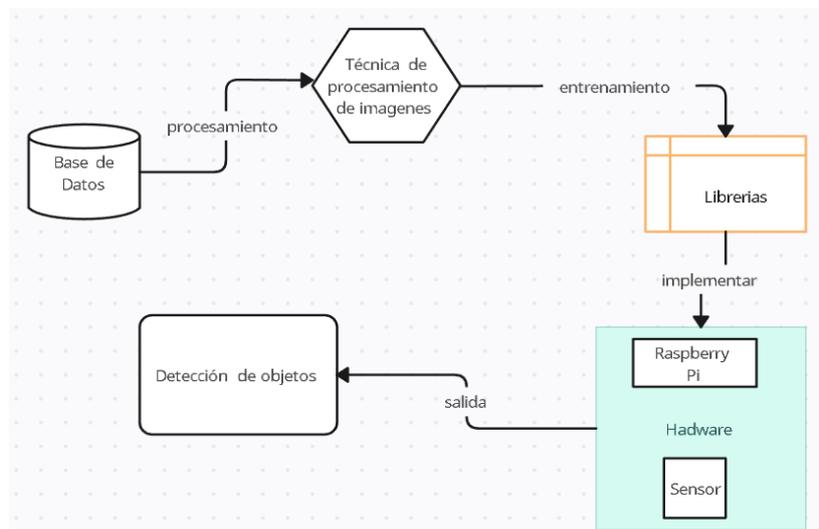
dos categorías; la construcción de hardware y software.

### A. Construcción de Hardware

La cimentación del hardware se lleva a cabo mediante la inserción de los dispositivos electrónicos, el sensor de distancia, la cámara, la tarjeta de mando o Raspberry Pi 4 y el reproductor de audio que deben estar conectados, para que el dispositivo dé una buena obtención de datos y cumpla con las funciones y especificaciones requeridas, como se observa en la Figura 16.

**Figura 16**

*Diagrama de la investigación por parte del hardware.*



*Nota:* Imagen realizada con base a investigaciones anteriores.

Los elementos electrónicos utilizados en el desarrollo de esta investigación, son componentes que ofrecen alta capacidad de funcionamiento, facilidad de implementación y que se encuentren en el medio, por lo que se ha visto indispensable la utilización de:

**Raspberry PI.** Es un tipo de ordenador, pero de forma reducida desarrollada por la Fundación Raspberry Pi. Fue construida como una herramienta de enseñanza para los inicios de programación y computación en los centros de estudio primario y se quería poner en el alcance de todos, dispone de diferentes puertos en el que se pueden adaptar aparatos electrónicos para el desarrollo de proyectos e inclusive se puede conectar un teclado o un mouse para la introducción de información y recepción de datos, es como se puede apreciar en la Figura 17, y como lo declara [37].

**Sensor de Distancia.** La importancia de este tipo de sensor es para verificar la distancia que se encuentra el objeto a ser medido, para comprobar la distancia máxima y mínima de cada objeto, comprobar aquellas cosas que pueda afectar la movilidad de la persona no vidente y despreciar aquellos que no interrumpen la caminata, se puede apreciar en la Figura 18 la infraestructura del sensor.

### Figura 17

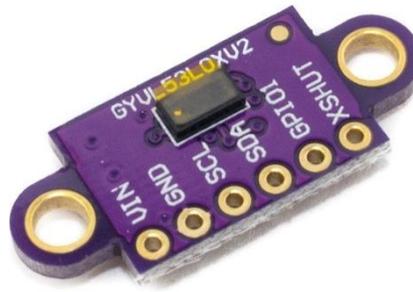
*Estructura Hardware de una Raspberry Pi 4.*



*Fuente:* Imagen extraída de [37].

### Figura 18

*Sensor de distancia VL53L0X.*



*Nota:* Imagen real del sensor extraída de [38].

Se ha visto indispensable la utilización del sensor VL53L0X, que tiene la capacidad de medir con exactitud el tiempo que tarda la luz en viajar desde el sensor hasta el objeto tomado como referencia, como se muestra en la **¡Error! No se encuentra el origen de l a referencia.**, dicho tiempo se transforma en distancia con independencia de la reflectividad de cada objeto, las características se muestran en la Tabla 2.

### Tabla 2

*Características del sensor VL53L0X.*

Características	Detalle
Tamaño	4,40 x 2,40 x 1,00 mm

Tensión de funcionamiento	2,6 a 5 V
Temperatura funcionamiento	-20 a 70 ° C
Emisor infrarrojo	940 nm
Distancia	1200 mm
Precisión	+/- 30 mm
Consumo corriente	10 mA

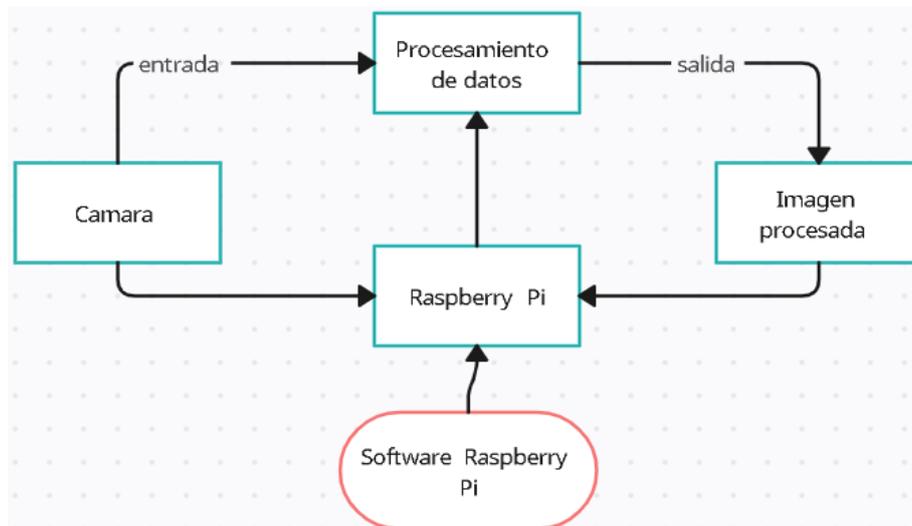
*Fuente:* Características específicas del sensor tomadas de [38].

## B. Construcción de Software

Esta es una de las partes más importantes, ya que la programación bien estructurada permitirá la recolección de datos para ser procesados y que se dé una respuesta rápida y eficiente, que se debe tener un plan de diseño detallando la ejecución de cada una de las líneas de código debe satisfacer el reconocimiento de objetos que se delimiten a ser esquivados, como se muestra en la Figura 19.

**Figura 19**

*Diagrama de bloque de la ejecución del software.*

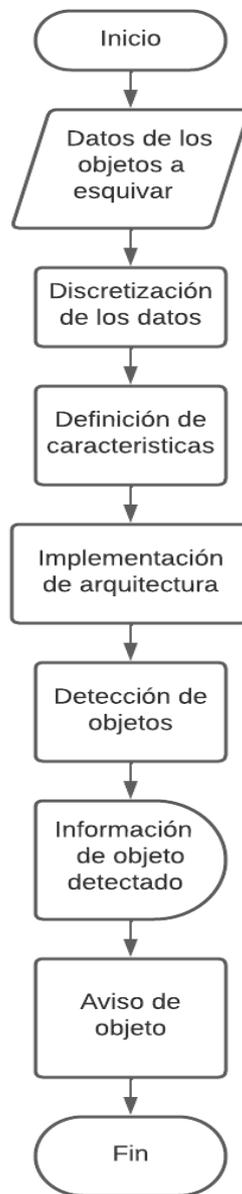


*Nota:* Diagrama construido en base anteriores investigaciones.

El sistema operativo a utilizar es Raspbian, se utilizará TensorFlow, una biblioteca de código abierto para aprendizaje automático, para entrenar un modelo de detección de símbolos Bliss. Este modelo se basará en MobileNet V2, una arquitectura de red neuronal convolucional eficiente y ligera diseñada para dispositivos móviles y aplicaciones integradas, para más entendimiento se muestra en la Figura 20 el diagrama de flujo de la programación empleada.

**Figura 20**

*Diagrama de flujo de la programación.*



*Nota:* Diagrama implementado con base en otras investigaciones.

#### **2.4. Implementación**

Ya podemos desarrollar la programación en el entorno virtual descrito, tomando en cuenta la información requerida y basándose en los lineamientos de programación, para que la implementación del código se realice en secuencia de pequeños pasos, con el fin de ir minimizando los errores y realizar pruebas unitarias que verifiquen el correcto funcionamiento del mismo.

### 2.4.1. Implementación de Hardware

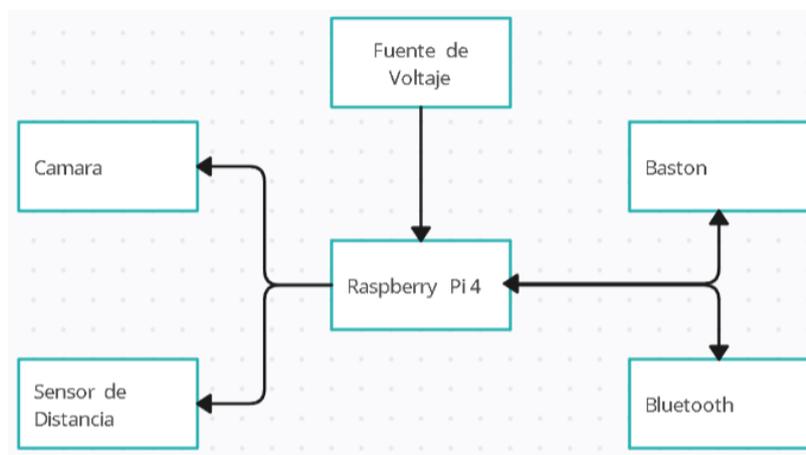
En la Figura 21, se nota el diagrama de flujo sobre la implementación de los componentes electrónicos utilizados en la investigación, para hacer las pruebas de funcionamiento del dispositivo detector de objetos.

Se implementa tan solo un botón para encendido y apagado del dispositivo, se ejecuta un modelo de carcasa donde se incorpora la fuente de voltaje, donde se reúne un par de pilas de 8000 mAH, en la Tabla 3, se muestra el consumo de energía de los componentes electrónicos para ver las horas de duración del dispositivo.

Cabe mencionar que el elemento que más consume energía es la Raspberry Pi 4, que trabaja a un voltaje entre 4,7 a 5,5 Voltios, además en la Figura 22, se muestra el montaje de la carcasa, impresa en 3D, con el montaje del botón de mando para el paso y discontinuo de energía, el funcionamiento del botón es de fácil entendimiento ya que con un pulso se activa el paso de energía, con otra pulsada se desconecta el circuito, se implementa solo este sencillo funcionamiento ya que la persona solo quiere que el dispositivo sea de fácil maniobra.

#### Figura 21

Diagrama de funcionamiento del dispositivo.



#### Tabla 3

Consumo de energía del dispositivo.

Componente	Consumo
Raspberri Py 4	Máximo 3 AH
VL53L0X	10 mAH
Cámara	40 mAH

Auricular

40 mAH

---

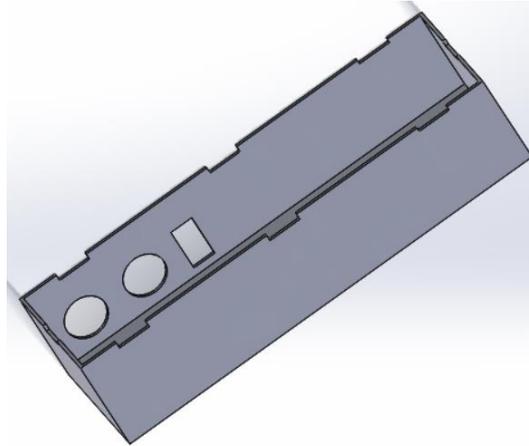
**TOTAL** 3.1 AH

---

*Nota:* Descripción del consumo energético de los componentes eléctricos, donde la duración del dispositivo se estima en 5 horas en implementación de dos pilas de 8000 mAH.

**Figura 22**

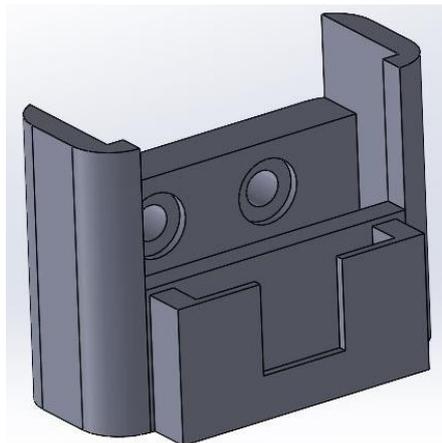
*Montaje de los botones en la carcasa.*



Además, se diseñó un elemento para que se sostenga la cámara y el sensor, como se mira en la Figura 23, este elemento también es impreso en 3D, diseño que se realizó en el software SolidWorks, las especificaciones y mediciones se redactan en el apartado de Anexos.

**Figura 23**

*Diseño de elemento para sostener la cámara y sensor.*



Se ha visto indispensable que se construya un elemento donde se implemente los elementos electrónicos, en el diseño, la cámara se ubica en la parte superior y debajo está el sensor de distancia, se ha visto mediante las pruebas de funcionamiento, que, para tener una buena medición del sensor, se lo ponga a trabajar sin tanta iluminación que afecte a la emisión de la luz infrarroja hacia el objeto a detectar.

### Estimación.

En la Tabla 4, se detalla el presupuesto de cada uno de los componentes eléctricos utilizados para la construcción del dispositivo, se ha visto conveniente ilustrar la tabla para dar cumplimiento a uno de los requerimientos sobre el costo de cimentación.

**Tabla 4**

*Estimación de precios de dispositivos electrónicos.*

DESCRIPCIÓN	COSTO REAL (\$)	COSTO ACTUAL (\$)
Raspberri Py 4	300.00	250.00
Sensor VL53L0X	30.00	24.00
Cámara	30.00	26.50
Auriculares	5.00	3.00
Bastón	19.00	16.00
Botones	3.00	2.00
Chaleco	20.00	10.00
Pilas	16.00	9.00
Hojas impresas con simbología	7.00	3.50
<b>TOTAL</b>	<b>437.00</b>	<b>344.00</b>

*Nota:* El costo real hace referencia al precio que se encuentra en tiendas web, mientras que el costo actual es el precio que se consiguió en el mercado.

#### 2.4.2. Adquisición de Imágenes

Como primera etapa de construcción del dispositivo para reconocimiento de objetos, se tiene la adquisición de imágenes, que luego de delimitar los objetos, se ha visto conveniente la utilización de la simbología Bliss, cada símbolo se ubica en una posición específica del objeto para que el reconocimiento se realice con mayor eficacia. En la Tabla 5 se muestra la simbología correspondiente que es utilizada para cada uno de los objetos mencionados en la **¡Error! No se encuentra el origen de la referencia..**

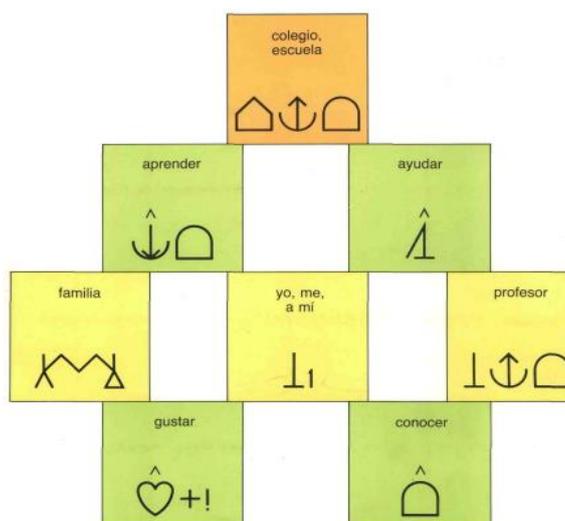
#### Simbología Bliss

Es un sistema que utiliza una serie de símbolos compuesto por las formas geográficas simples tales como círculos, cuadrados, triángulos, líneas horizontales y verticales, con el principal objetivo principal es utilizarlo en personas que presentan limitaciones de comunicación y expresión, como lo expresa [39] en su página web.

Fue creado por Charles Bliss en 1965, pero no tuvo la acogida suficiente sino hasta el año de 1971 en un centro de niños que tenían parálisis cerebral en la ciudad de Ontario. Estos símbolos son objetos o hechos que simbolizan una realidad y deben de seguir una normativa y procedimiento adecuado en su estructura para ser entendidos, donde se emitan los pensamientos, emociones, ideas y sentimientos que la personas tenga en su mente, pero no puede exponer en forma de sonido mediante la voz, como se muestra en la Figura 24.

### Figura 24

*Ejemplos de Símbolos Bliss.*



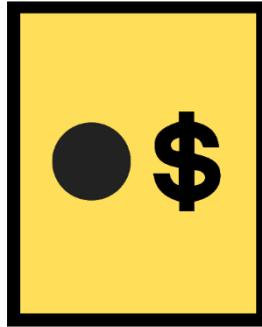
*Fuente:* Imagen tomada de [40].

### Tabla 5

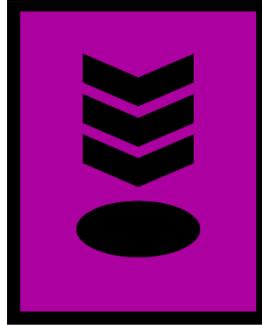
*Determinación de los objetos con el respectivo simbolo Bliss.*

OBJETO	SÍMBOLO BLISS
Cama	

Silla



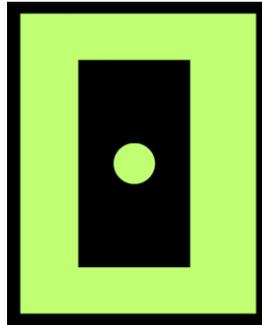
Basurero



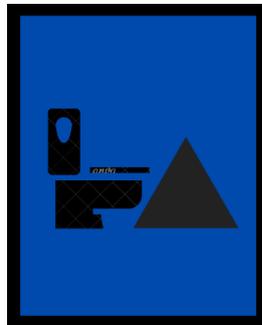
Quebrada



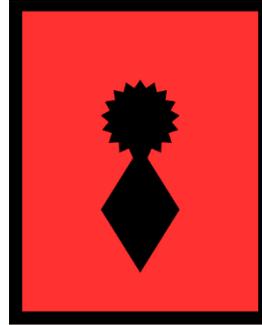
Puerta



Baño



Cocina



---

### 2.4.3. *Etiquetado de Imágenes*

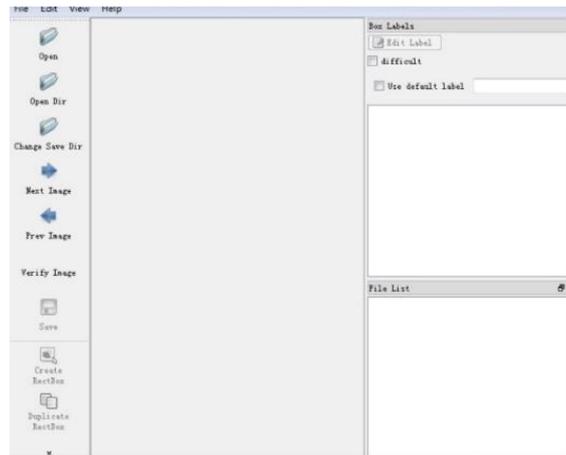
Se tomaron 500 imágenes de cada uno de los símbolos con los respectivos objetos en diferentes escenarios, tanto en forma positiva, así como también en forma negativa, en diferente distancia y con ángulo de posición distinta, para la captura de imágenes de entrenamiento se utiliza el celular que tiene una cámara de 16 + 2 Mp, y con la menor resolución en megapíxeles, se debe tener en constancia que las imágenes no deben ser demasiado grandes en peso KB, ya que si supera los 200 KB tardará más tiempo en entrenar el modelo.

Para realizar el etiquetado de cada una de las imágenes, donde se recortan solo el símbolo de importancia que se encuentra dentro de la imagen, los demás objetos que contenga se desprecian y no es tomada en cuenta. La imagen recortada se guarda en formato .mlx, esto se logra con la utilización del programa “labelImg”, que es un entorno virtual de Python, como se muestra en la Figura 25.

Este tipo de programa puede etiquetar la topografía que tiene el objeto de importancia, sin importar el contorno que tenga la imagen, pues se puede ampliar la imagen, delimitar y seguir el perímetro mediante líneas consecutivas, donde representan correctamente el dibujo de interés.

### **Figura 25**

*Pantalla principal del programa labelImg.*



Hay que tener en cuenta que el etiquetado de las imágenes es un proceso largo, por lo que se debe tener paciencia y realizarlo de forma organizada, dividiendo en carpetas cada una de las clases con las imágenes, carpeta de muestreo que contiene el 60% de las imágenes a etiquetar, carpeta de prueba con un 20% de imágenes y la carpeta de validación con el otro 20% de imágenes con los dataset obtenidos y las etiquetas de esas imágenes.

#### 2.4.4. Configuración de la Red Neuronal Convolutiva

Se inicia la configuración del modelo con la generación de los TFRecords, los cuales son los datos de entrada en TensorFlow, estos parámetros son de importancia, ya que se especifican las determinaciones que tiene el modelo de detección. Los datos de entrada se ubican en las primeras capas de la arquitectura MobileNet.

El código utiliza el módulo `argparse` como se muestran en la Figura 26, para analizar los argumentos de línea de comandos proporcionados al programa. Los argumentos que se pueden proporcionar son:

**Figura 26**

*Definición de argumentos de entrada.*

```

77 # Definir y analizar argumentos de entrada
78 parser = argparse.ArgumentParser()
79 parser.add_argument('--modeldir', help='Folder the .tflite file is located in',
80                    required=True)
81 parser.add_argument('--graph', help='Name of the .tflite file, if different than detect.tflite',
82                    default='detect.tflite')
83 parser.add_argument('--labels', help='Name of the labelmap file, if different than labelmap.txt',
84                    default='labelmap.txt')
85 parser.add_argument('--threshold', help='Minimum confidence threshold for displaying detected objects',
86                    default=0.9)
87 parser.add_argument('--resolution', help='Desired webcam resolution in WxH. If the webcam does not support the resolution entered, errors may occur.',
88                    default='640x360')
89 parser.add_argument('--edgetpu', help='Use Coral Edge TPU Accelerator to speed up detection',
90                    action='store_true')

```

`modeldir`: este directorio especifica la ruta del directorio que contiene el modelo de detección (TFLite).

`graph`: Nombre del archivo TFLite que contiene el modelo de detección, que se especifica cuando se ejecuta el programa.

`labels`: Nombre del archivo que contiene el mapa de las etiquetas de clase del modelo si es diferente del valor predeterminado.

`threshold`: establece el umbral de confianza mínimo para mostrar los objetos detectados, todo esto para no proporcionar respuestas erróneas si no superan el 90% de exactitud.

`resolution`: se configura la resolución deseada de la cámara web en formato WxH, donde W es el ancho y H es el alto de la resolución, es importante tener en cuenta que configurar una resolución que la cámara web no admite puede provocar errores.

`edgetu`: Se utiliza para permitir a los usuarios habilitar o deshabilitar el uso del acelerador Coral Edge TPU (Unidad de procesamiento tensorial) para acelerar la detección de objetos.

Es relevante definir en la importación de librerías que son compatibles con Python, poniéndole mayor énfasis en “board” que es un módulo que proporciona acceso a los pines GPIO en la Raspberry Pi, y también para poder reproducir los archivos de formato de audio MP3 se utiliza “pydub, AudioSegment, play”, así como también “busio” que facilita la comunicación I2C en la Raspberry Pi. Para que se acople el sensor con el ordenador principal se emplea “adafruit\_vl53l0x”, un módulo que permite interactuar con el sensor de distancia VL53L0X mediante I2C.

Para iniciar la detección de objetos se crea una instancia de la clase VideoStream para comenzar a capturar los fotogramas del flujo de video de la cámara. Luego, se pone en marcha un bucle infinito que realiza la captura del último fotograma disponible desde la cámara, que utiliza una resolución de 320 x 320 como se muestra en la Figura 28, y con MobileNet V2 permite ajustar la resolución de entrada de las imágenes. Una resolución más baja, disminuyendo la cantidad de píxeles en la imagen de entrada, lo que reduce aún más la cantidad de cálculos necesarios, el fotograma se redimensiona al tamaño requerido por el modelo de detección y se realiza la detección de objetos utilizando el modelo cargado y se obtienen las

coordenadas y confianza de los objetos detectados, como se muestra en la Figura 28, el diagrama de flujo de la recepción de la clase del video en directo.

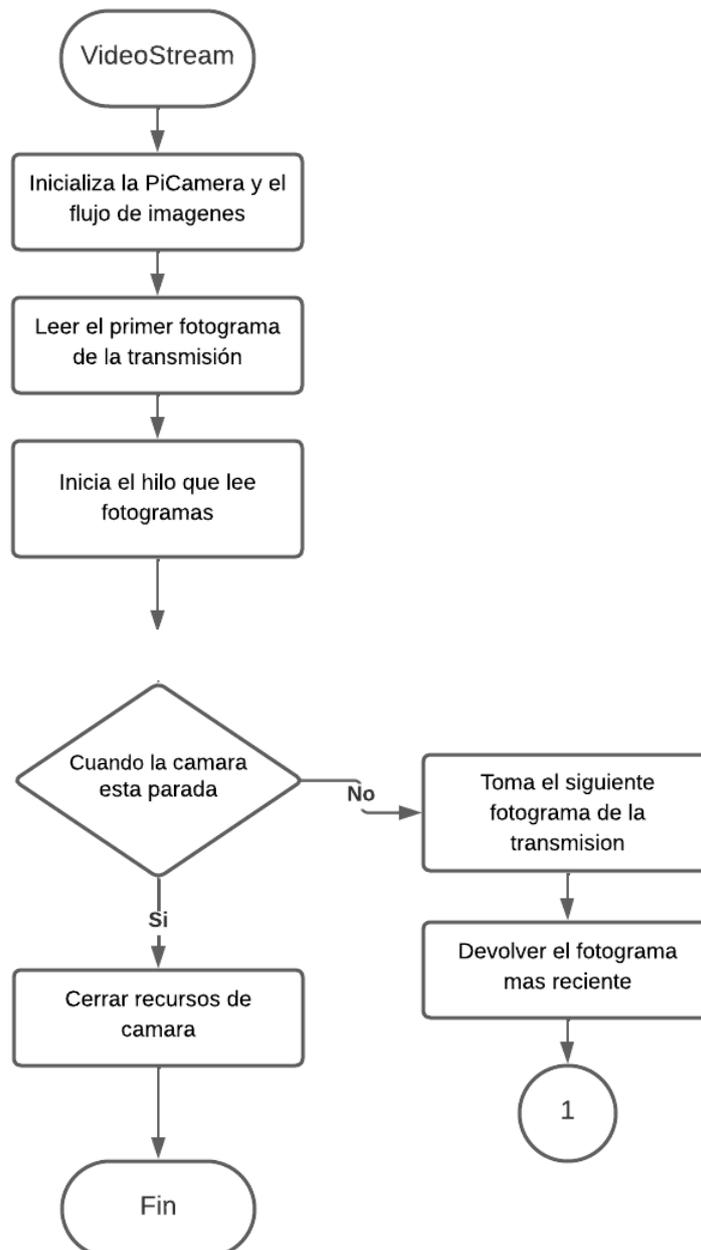
### Figura 27

*Definición de la resolución de la cámara.*

```
38 """Camera object that controls video streaming from the Picamera"""
39 def __init__(self,resolution=(320,320),framerate=30):
```

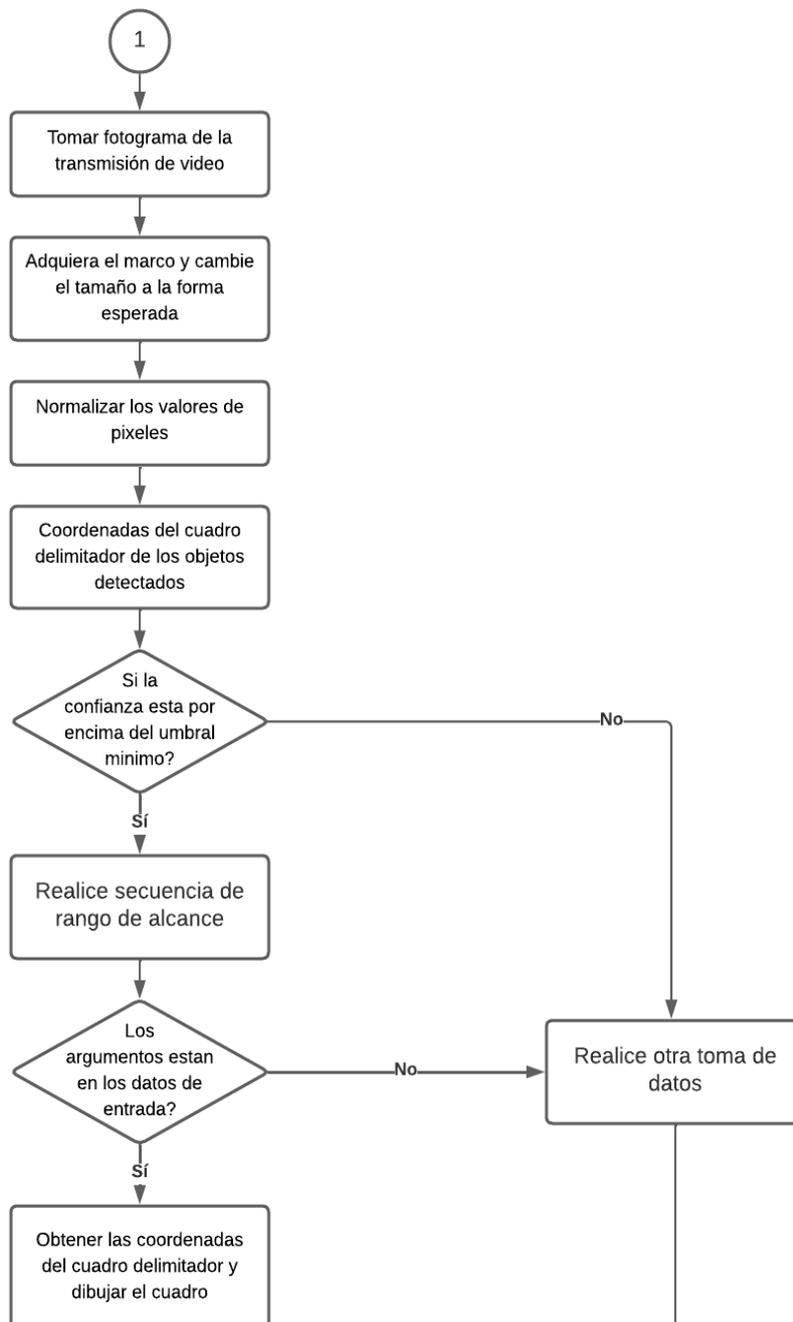
### Figura 28

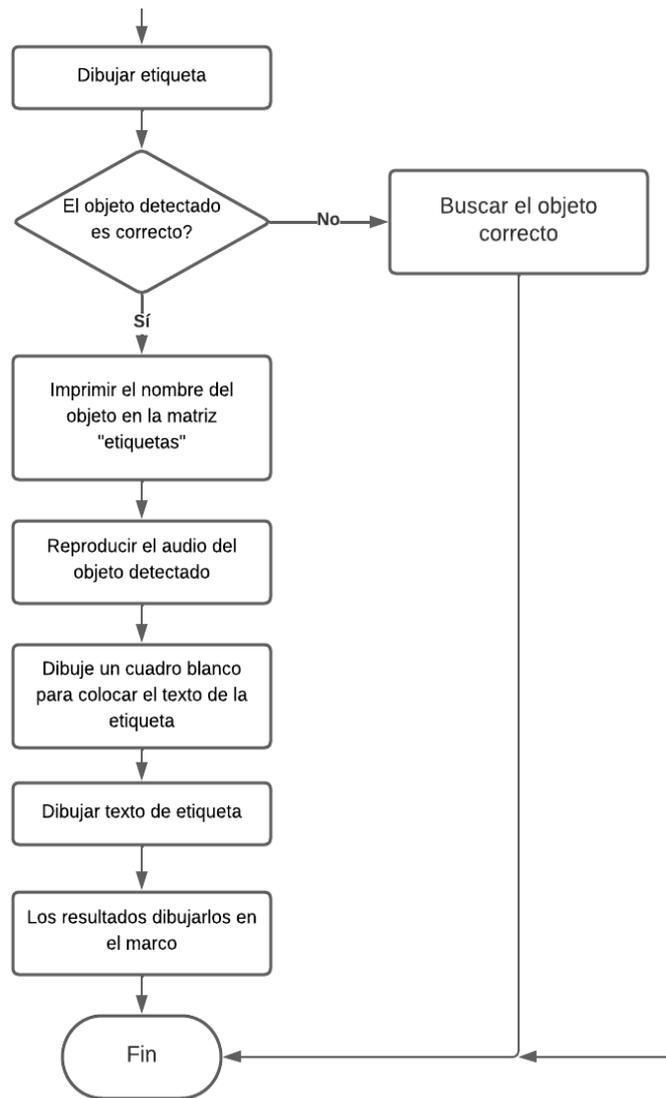
*Diagrama de flujo del VideoStream.*



Una vez configurado el modelo, se convertirá a un formato TensorFlow Lite más ligero y eficiente, que luego se ejecutará en la Raspberry Pi para la detección en tiempo real. Cuando el sistema detecte un símbolo Bliss a través del sensor CJVL53L0Xv2, emitirá una señal de alerta a través de una salida de audio o vibración, como se muestra en el diagrama de flujo de la Figura 29. Si el objeto de interés está dentro del umbral de confianza y en el rango de alcance del sensor VL53L0X, se realiza una acción específica, como reproducir un archivo de audio asociado al objeto.

**Figura 29** Estructura de la programación por hilos.





Se muestra el resultado de la detección en tiempo real, junto con la tasa de fotogramas por segundo (FPS), la estructura de programación que se realiza en secuencia por hilos, que puede ser beneficioso para mantener la captura de video y la detección para un rendimiento más suave. Sin embargo, es importante asegurarse de que los recursos compartidos entre los hilos estén sincronizados adecuadamente para evitar posibles problemas de concurrencia.

#### 2.4.5. Entrenamiento de la Red

Terminado la configuración del modelo de detección de objetos, se empieza el entrenamiento del modelo primero, configurando las imágenes que están con extensión .xml para adaptarlos en extensión .csv, con cada una de las carpetas, para ello se debe ejecutar el siguiente código:

```
(tensorflow1) C:\tensorflow1\models\research\object_detection> python xml_to_csv.py
```

Ahora se debe configurar el mapa de etiquetas de cada uno de los objetos donde se le asigna un número específico de identificación. Para ello se puede crear un editor de texto que se generen las etiquetas como se mira en la Figura 30 siguiente.

### Figura 30

*Mapa de etiquetas de los objetos.*

```
Archivo Edición Formato Ver Ayuda
item {
  id: 1
  name: 'cama'
}

item {
  id: 2
  name: 'silla'
}

item {
  id: 3
  name: 'basurero'
}

item {
  id: 4
  name: 'quebrada'
}

item {
  id: 5
  name: 'puerta'
}

item {
  id: 6
  name: 'baño'
}

item {
  id: 7
  name: 'cocina'
}
```

Posteriormente, se debe establecer el canal de entrenamiento para realizar la detección de objetos, con la estimación planteada se tiene que por cada objeto se demora 7 horas en realizar el entrenamiento, utilizando una máquina con procesador Intel i7 4.8 GHZ, procesador de onceava generación y utilizando una tarjeta gráfica de DDR4 Nvidia RTX 3060 y con 4 TB de almacenamiento.

## **2.5. Preguntas de Investigación**

Para recolectar información relevante por parte de la persona no vidente que está ayudando a la investigación, se realiza una entrevista a la persona para establecer los requerimientos necesarios que debe tener el dispositivo para que el reconocimiento de objetos se de en un porcentaje de aceptación con un mayor rango de alcance y reconozca a objetos que en verdad puede alterar la movilidad de la persona.

Las respuestas impartidas por la persona no vidente en la entrevista son descritas en el apartado de Anexos y son de gran importancia ya que ayudan a delimitar la planificación en el desarrollo de la programación y solo tomar en cuenta aquellas bibliotecas compatibles con el entorno virtual que en verdad se necesita utilizar.

## CAPÍTULO III: RESULTADOS Y ANÁLISIS

A continuación, se presenta la construcción del dispositivo para el reconocimiento de objetos, qué elementos electrónicos son los más apropiados para que se den la recolección de datos, presentaremos y analizaremos los resultados obtenidos durante las pruebas de nuestro dispositivo de detección de obstáculos para personas con discapacidad visual. Nuestro objetivo es desarrollar un dispositivo que mejore la movilidad y autonomía de las personas no videntes mediante el uso de visión artificial. Para evaluar la efectividad de nuestro dispositivo, hemos realizado una serie de pruebas que nos han proporcionado datos valiosos sobre su rendimiento.

### 3.1. Verificación

Una vez efectuado el entrenamiento del modelo para la detección de objetos, se empieza con la fase de realización de pruebas del modelo, por lo que se tiene como prioridad la distancia de trabajo comprando la del sensor con lo expresado en la hoja de datos, para verificar que la detección se realice con total normalidad en las distancias establecidas, además también para medir la precisión que tiene el modelo y el rendimiento que se tendrá.

Se empieza verificando los datos que nos deja el modelo con forme se realiza el entrenamiento, cada vez que pasa el tiempo y se realiza el entrenamiento se realiza la información de la perdida de errores en el modelo.

Como se ha mencionado en este entrenamiento se utiliza la arquitectura MobileNet V2, que utiliza alrededor de 53 capas que son de convolución y de acumulación, la perdida comenzó en 20 y antes de terminar el entrenamiento se mostró una perdida que se acerca a 1, como se muestra en la Figura 31, y en la Figura 32, se evidencia la gráfica cada 5 minutos que se guarda periódicamente mientras se realiza el entrenamiento de la perdida conjunto con el muestreo, que para la arquitectura utilizada es muy efectiva y que recalca que fue una buena opción de utilizar esta arquitectura.

#### **Figura 31**

*Perdida de rango de error.*

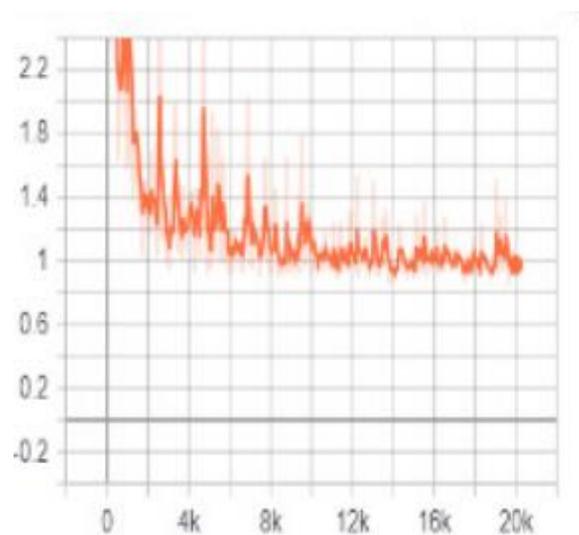
```

Windows PowerShell x Anaconda - python train.py - x + v
I0804 09:38:41.113283 16772 learning.py:507] global step 7993: loss = 1.4049 (6.810 sec/step)
INFO:tensorflow:global step 7994: loss = 2.0634 (6.809 sec/step)
I0804 09:38:47.925750 16772 learning.py:507] global step 7994: loss = 2.0634 (6.809 sec/step)
INFO:tensorflow:global step 7995: loss = 1.0842 (5.676 sec/step)
I0804 09:38:53.602364 16772 learning.py:507] global step 7995: loss = 1.0842 (5.676 sec/step)
INFO:tensorflow:global step 7996: loss = 1.5863 (5.671 sec/step)
I0804 09:38:59.276793 16772 learning.py:507] global step 7996: loss = 1.5863 (5.671 sec/step)
INFO:tensorflow:global step 7997: loss = 1.4594 (5.627 sec/step)
I0804 09:39:04.905450 16772 learning.py:507] global step 7997: loss = 1.4594 (5.627 sec/step)
INFO:tensorflow:global step 7998: loss = 1.4224 (5.616 sec/step)
I0804 09:39:10.523548 16772 learning.py:507] global step 7998: loss = 1.4224 (5.616 sec/step)
INFO:tensorflow:global step 7999: loss = 0.9841 (5.719 sec/step)
I0804 09:39:16.244331 16772 learning.py:507] global step 7999: loss = 0.9841 (5.719 sec/step)
INFO:tensorflow:global step 8000: loss = 1.3469 (5.687 sec/step)
I0804 09:39:21.932704 16772 learning.py:507] global step 8000: loss = 1.3469 (5.687 sec/step)
INFO:tensorflow:global step 8001: loss = 1.4523 (5.513 sec/step)
I0804 09:39:27.448667 16772 learning.py:507] global step 8001: loss = 1.4523 (5.513 sec/step)
INFO:tensorflow:global step 8002: loss = 1.4595 (5.556 sec/step)
I0804 09:39:33.006380 16772 learning.py:507] global step 8002: loss = 1.4595 (5.556 sec/step)
INFO:tensorflow:global step 8003: loss = 4.4861 (5.690 sec/step)
I0804 09:39:38.698826 16772 learning.py:507] global step 8003: loss = 4.4861 (5.690 sec/step)
INFO:tensorflow:global step 8004: loss = 1.5981 (5.671 sec/step)
I0804 09:39:44.372274 16772 learning.py:507] global step 8004: loss = 1.5981 (5.671 sec/step)
INFO:tensorflow:global step 8005: loss = 1.2853 (5.562 sec/step)
I0804 09:39:49.936476 16772 learning.py:507] global step 8005: loss = 1.2853 (5.562 sec/step)
INFO:tensorflow:global step 8006: loss = 2.0382 (5.928 sec/step)
I0804 09:39:55.866453 16772 learning.py:507] global step 8006: loss = 2.0382 (5.928 sec/step)
INFO:tensorflow:global step 8007: loss = 1.1059 (5.675 sec/step)
I0804 09:40:01.542748 16772 learning.py:507] global step 8007: loss = 1.1059 (5.675 sec/step)

```

**Figura 32**

*Gráfico de pérdida del detector de objetos.*



*Nota:* La imagen es extraída desde TensorBoard.

### 3.1.1. Detección de Distancia:

Uno de los componentes del dispositivo es la detección de distancia, que se realiza utilizando un sensor de proximidad CJVL53L0Xv2. Este sensor puede detectar objetos que se encuentran a una distancia de hasta 2 metros, lo que permite al dispositivo advertir a los usuarios de obstáculos próximos que pueden representar un riesgo.

**3.1.1.1. Alertas e Interacción del Usuario.** Cuando el dispositivo detecta uno

de los objetos predefinidos o un obstáculo a través del sensor de proximidad, genera una señal de alerta. Esta señal puede ser un sonido emitido a través de una salida de audio o una vibración, según las preferencias del usuario.

Además, el dispositivo incluye botones físicos para encender y apagar el sistema. Cabe recalcar que la persona no vidente que es beneficiaria de este proyecto es un adulto mayor que no tiene mucho conocimiento de la tecnología, por lo que con botones de encendido y apagado son recursos más que suficientes de interacción.

**3.1.1.2. Validación del sensor de distancia CJVL53LOXV.** Un componente esencial en nuestro dispositivo de detección de obstáculos para personas con discapacidad visual es el sensor de distancia CJVL53LOXV2. Este sensor juega un papel crucial al permitirnos identificar y cuantificar la proximidad de los obstáculos en el entorno inmediato del usuario, proporcionando una medida vital de seguridad y funcionalidad para el dispositivo.

La eficacia de nuestro sistema depende, en gran medida, de la precisión y fiabilidad de este sensor en particular. Por lo tanto, es imperativo llevar a cabo un proceso de validación exhaustivo para asegurarnos de que el sensor de distancia CJVL53LOXV2 está operando a su máxima capacidad y que su rendimiento está en línea con nuestras expectativas y necesidades.

**Calibración y Pruebas de Precisión.** Para probar la precisión del sensor, se realiza una serie de mediciones a diferentes distancias conocidas y se compara las lecturas del sensor con los valores reales. Esta prueba se llevó a cabo en diferentes condiciones para asegurarse de que el sensor puede dar lecturas precisas en diversas situaciones.

Tomando en cuenta esto se tomó rangos de medida desde los 5 cm hasta los 100 cm, que es la métrica deseada para la detención de obstáculos donde actuaría la identificación del objeto, para evitar un desbordamiento auditivo de reconocimientos, en las métricas tengo la siguiente Tabla 6.

**Tabla 6**

*Pruebas de precisión del sensor VL53LOX.*

Distancia Real	VL53LOX.	E. Absoluto	E. Relativo %
5	5,3	0,30	6,00
10	10,9	0,90	9,00
15	15,8	0,80	5,33

20	19,9	-0,10	-0,50
30	32,3	2,30	7,67
40	41,1	1,10	2,75
50	53	3,00	6,00
60	62,5	2,50	4,17
70	72,1	2,10	3,00
80	82,4	2,40	3,00
90	93,1	3,10	3,44
100	99,45	-0,55	-0,55

Nota: La tabla que proporcionaste muestra la comparación entre la distancia real y la medida

El sensor VL53LOX obtuvo los resultados mostrados en la tabla, así como el error absoluto y el error relativo asociados con cada medición, en este caso se tiene un análisis general de los resultados:

**Error Absoluto:** Los valores del error absoluto indican cuánto se desvía la medida del sensor de la distancia real. Estos valores parecen ser relativamente bajos, lo que sugiere que las mediciones del sensor son razonablemente precisas. Sin embargo, es importante notar que el sensor tiende a sobreestimar la distancia, especialmente para distancias más largas. La excepción ocurre a las distancias de 20 y 100 cm, donde el sensor subestima ligeramente la distancia.

**Error Relativo:** El error relativo te da una idea del error en relación con el tamaño del objeto que estás midiendo. Se observa que el error relativo es bastante bajo para la mayoría de las mediciones, lo que recomienda que el sensor es bastante preciso. Sin embargo, hay un aumento en el error relativo en distancias más cortas (5, 10 y 15 cm) y en una medida a 30 cm.

Una posibilidad para estos errores mayores en distancias cortas podría ser que el sensor tenga una precisión reducida a distancias más cortas. En cuanto al error de 30 cm, podría ser una anomalía o podría proponer que el sensor tiene problemas particulares con esa distancia.

En general, los resultados sugieren que el sensor VL53LOX tiene una precisión razonable a lo largo de una serie de distancias. Sin embargo, tiende a sobreestimar las distancias más largas y tiene un rendimiento menos preciso a distancias más cortas. Estos detalles podrían ser útiles para ajustar o calibrar el sensor en el uso práctico.

### 3.1.2. Matriz de Confusión

Para la construcción de la matriz de confusión se utilizan las ecuaciones descritas en el apartado del Capítulo I, que especifica la matriz que debe contener la matriz para analizar si el modelo utilizado se puede utilizar en la detección de objetos.

**3.1.2.1. Por Cada Variable.** Para evaluar el rendimiento del modelo de visión artificial en la detección de obstáculos y símbolos Bliss, se utiliza una matriz de confusión. Esta es una herramienta poderosa que permite desglosar las predicciones del modelo. Al analizar estos resultados, podemos obtener una visión detallada de cómo se está comportando nuestro modelo en la práctica.

Además, a partir de la matriz de confusión, podemos calcular varias métricas, como la precisión, la sensibilidad, la especificidad y el F1 score. Estas métricas nos ayudarán a cuantificar el rendimiento de nuestro modelo en cada una de las variables u objetos descritos de manera más precisa, por lo que se utiliza la matriz de confusión de 2x2 de forma manual, con la utilización del modelo implementado y se muestra los resultados en la Tabla 7.

**Tabla 7**

*Tabla de resultados de cada objeto.*

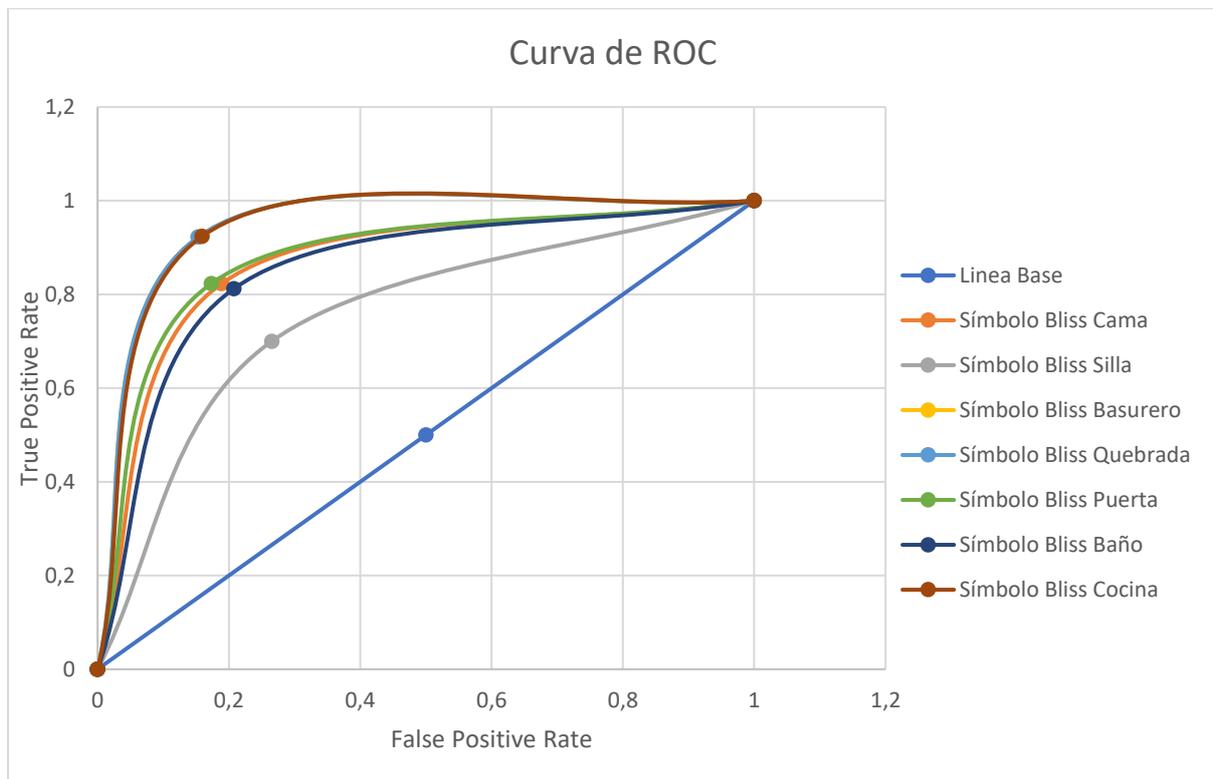
Objetos	Exactitud (%)	Precisión (%)	Sensibilidad (%)	Especificad (%)	F1 (%)
Cama	81,7	80,8	82,3	81,1	81,5
Silla	71,6	75,6	70	73,5	72,7
Basurero	88,4	83,8	92,2	85,1	87,8
Quebrada	89,9	73	88,8	77,1	80
Puerta	82,5	82,8	82,3	82,7	82,6
Baño	80,2	78,6	81,2	79,3	79,9
Cocina	87,8	82,4	92,4	84,1	87,1
<b>TOTAL</b>	<b>83,16</b>	<b>79,57</b>	<b>84,17</b>	<b>80,41</b>	<b>81,66</b>

Donde se observa que la precisión en cada una de las variables supera el 70% de precisión, por lo que se puede asegurar que el modelo es óptimo para trabajar. Además, el puntaje F1 balancea precisión y sensibilidad, ofreciendo una visión global del rendimiento del modelo. Todos los objetos tienen un puntaje F1 superior al 70%, lo que indica un rendimiento consistente y equilibrado del modelo.

**Curva de ROC.** Mientras que en la Figura 33, se muestra la construcción de la curva de ROC de los objetos tomados en consideración, que es una herramienta gráfica utilizada para evaluar la capacidad de un modelo de clasificación para discriminar entre clases positivas y negativas a todos los niveles de umbral de clasificación. Cuanto más cerca esté la curva ROC del borde superior izquierdo, mejor será el rendimiento general del modelo.

**Figura 33**

*Curva de ROC de los objetos.*



La grafica nos muestra que los modelos utilizados en cada uno de los objetos son los más óptimos, esto se puede afirmar ya que el punto de cruce entre las métricas de Verdaderos Positivos y Falsos Positivos, están por encima de la línea tomada como referencia y se acerca a la esquina superior izquierda. Es importante tener en cuenta que el rendimiento del programa puede variar según el modelo de detección y la complejidad de las operaciones realizadas en tiempo real.

**3.1.2.2. Múltiples variables.** Al tener  $n \times n$  variables en el modelo, se construye la matriz de confusión, donde la diagonal principal contiene los valores de Verdaderos Positivos que detecto el modelo utilizado para detectar los objetos, y se completa las demás

celdas con los valores consignados como errores. La matriz fue construida tomando en cuenta el 80% de las imágenes que se tienen de muestra tanto en las carpetas de etiquetar y prueba, para que el modelo clasifique a que tipo de variable es la imagen que se procesa, dejando como resultado lo que se muestra en la Tabla 8, donde se determinan las métricas de progresión global y el estadístico de Kappa, métricas que hacen entender que tan versátil es la utilización del modelo construido.

**Tabla 8**

*Matriz de confusión de las variables tomadas como referencia.*

MATRIZ CONFUSIÓN VARIAS VARIABLES								
	Cama	Silla	Basurero	Quebrada	Puerta	Baño	Cocina	total
Cama	404	0	1	1	0	0	0	406
Silla	2	378	2	0	0	0	0	382
Basurero	0	1	416	0	1	0	0	418
Quebrada	0	0	1	365	0	0	1	367
Puerta	0	0	0	2	414	1	0	417
Baño	0	3	0	0	0	393	1	397
Cocina	0	0	0	1	0	0	412	413
TOTAL	406	382	420	369	415	394	414	2800

Como lo especifica [41], se calcula las métricas especificadas, la progresión global donde se aprecia la precisión y sensibilidad que tiene el modelo en todas las variables para la detección de objetos y la ecuación para calcular el estadístico Kappa, es un tipo de estadístico que busca encontrar la concordancia entre el conjunto de datos observado, con la que podría suceder en realidad, como se describe en la Ecuación Progresión global =  $\sum_{i=1}^k n_{ixi} / \sum n_{ixi}$

$$(8) \text{ y Estadístico Kappa} = K^{\wedge} = \frac{n \sum_{i=1}^k n_{ii} - \sum_{i=1}^k n_i + n_i}{n^2 - \sum_{i=1}^k n_i + n_i} \quad ($$

9).

$$\text{Progresión global} = \sum_{i=1}^k n_{ixi} / \sum n_{ixi} \quad (8)$$

$$\text{Progresión global} = \frac{2782}{2813} = 0,98$$

$$\text{Estadístico Kappa} = K^{\wedge} = \frac{n \sum_{i=1}^k n_{ii} - \sum_{i=1}^k n_i + n_i}{n^2 - \sum_{i=1}^k n_i + n_i} \quad (9)$$

$$K^{\wedge} = \frac{6693199}{6780402} = 0,98$$

*Presision* = 0,795

*Sensibilidad* = 0,841

*F1* = 0,816

### **3.2. Discusión**

Aunque hay variaciones en el rendimiento entre los diferentes objetos, el modelo muestra una capacidad sólida y consistente para identificar todos los objetos. Este éxito sugiere que con ajustes adicionales y posiblemente más datos de entrenamiento, en particular para la "silla", el rendimiento puede mejorar aún más. Sin embargo, es importante celebrar el logro que representa este alto nivel de rendimiento en todas las categorías.

Al obtener en el objeto "silla", uno de los valores bajos en sensibilidad, con un 70%, lo que demuestra que este tipo de objetos tienden a ser más vulnerables en su detección, y se puede afirmar ya que existen de diferentes formas, tamaños tanto en altura como en el ancho, por lo que en cierto periodo de tiempo podría mostrar verdaderos negativos.

También se tiene que para el objeto "quebrada", al no tener suficientes imágenes de referencia en sitios web, se tomaron imágenes de prueba, por lo que, como muestra el resultado de su precisión, con un 73%, uno de los datos más bajos en esta categoría, se puede afirmar que, en ambientes con demasiada luminosidad, el dispositivo tenga que acercarse al rango mínimo para ser detectado el objeto.

Para dar solución a este inconveniente se ha optado por aplicar colores de fondo en cada uno de los símbolos, como lo manifiesta [39], con el fin de establecer un dato de entrada adicional, para que el modelo realice con mayor eficiencia la detección de objetos. Se llegó a esta decisión ya que en pruebas de validación del dispositivo es el único objeto que se encuentra fuera del hogar, pero que es de vital importancia para la persona no vidente porque, necesita saber la ubicación del objeto, ya que siempre realiza trabajos de agricultura y utiliza el canal de riego que pasa por su casa, todas estas actividades lo hacen sentir una persona activa.

La base de datos obtenida debe estar bien clasificada y dividirlas en 7 categorías bien establecidas, también se ayudó a tener información de los dataset de cada categoría para establecer mejor a que simbología se adapta cada objeto. Se debe tener suficiente entrenamiento del modelo, para que se lo utilice en cualquier ambiente.

La utilización del modelo de aprendizaje automático de supervisión con base en la clasificación de modelos MobileNet V2, el dispositivo tiene mejor desempeño conforme más cantidad de datos analizar se ostenten, ya que actúan de acorde al entrenamiento y asignación de los datos que se les dan a las redes neuronales, por lo que se afirma que son óptimos en la utilización de modelos con gran cantidad de datos o con gran afluencia de ellos. También el modelo edgetpu, ayuda a comprimir el TensorFlow en un TensorFlow Lite, para que no consuma recursos de la Raspberri Py 4.

Al realizar las pruebas de funcionamiento del dispositivo, se debe tener en cuenta que la persona quien va a utilizarlo, debe sentirse conforme y satisfecha con lo que a diario va a utilizar, por lo que se ha visto innecesario colocar elementos que en teoría ayudan a detectar con mayor eficiencia los objetos, pero que producen inconformidad en el usuario y hacen perder el interés en utilizar el dispositivo, por lo que es conveniente realizar un modelo sencillo pero eficaz, a implementar un dispositivo robusto en elementos pero que no tenga una apreciación del usuario final.

En resumen, todos los modelos tienen un buen rendimiento, siendo particularmente destacados los modelos para "basurero" y "cocina". Sin embargo, el modelo de "silla" podría necesitar algunos ajustes para reducir su FPR y mejorar su rendimiento general. Recuerda que idealmente, queremos que nuestro modelo tenga un alto TPR y un bajo FPR. Esto se ve reflejado gracias a la aplicación de la matriz de confusión, que evalúa si el modelo desarrollado es idóneo para trabajar y cumpla con las expectativas tanto del investigador como de la persona quien lo va a utilizar.

Se puede dar conformidad que se ha construido un dispositivo detector de objetos que cumple con todas las expectativas de la persona no vidente que lo va a utilizar, aunque es un modelo sencillo, se adapta perfectamente a cumplir con el propósito principal, detectar los objetos descritos y emitir la señal respectiva para que la persona no vidente tenga idea de la ubicación exacta del objeto a esquivar y realice su movilidad dentro del hogar con mayor confianza, resaltando también que es más factible su implementación en futuras investigaciones a diferencia de los dispositivos analizados.

## CONCLUSIONES

Se puede afirmar que existen varios algoritmos para detectar objetos que están en constante actualización con forme lo hace la tecnología, algunos de estos, con la correcta utilización ayudan a la movilidad de las personas que presentan discapacidad visual, para tener una correcta recolección de información, debe someterse a un entrenamiento bastante riguroso, para que no emita señales erróneas de objetos que no alteran la movilización de la persona.

Con la implementación de un entorno virtual denominado Python la programación se desarrolló de manera más eficiente y de alta precisión, ya que el lenguaje de programación que utiliza el software es amigable y de fácil entendimiento, así también la utilización de dispositivos electrónicos eficientes, con fácil alcance en el mercado, realizan la recolección de datos de alta calidad y con mayor eficiencia.

El análisis y procesamiento de datos de los objetos que mayor incomodidad tiene la persona no vidente al momento de moverse dentro de su hogar, se examinan con la ejecución de la matriz de confusión, donde se explora los patrones de interés que presenta en una sola imagen, y con la implementación de la simbología Bliss, la detección se realiza en un tiempo mínimo y con alta validez de virtud.

Se realizó un diseño accesible y de fácil instalación para que el dispositivo detector de objetos se incorpore al cuerpo humano de la persona, y sobre todo se analizó el área de ubicación del dispositivo para que se tome la mayor recolección de datos, donde se presenta un mínimo errores en las detecciones, todo esto gracias a los requerimientos que impartió la persona no vidente.

## RECOMENDACIONES

Se recomienda adoptar la ayuda de un experto, que ayude en la planificación y desarrollo de del dispositivo en todas las etapas, ya que, al utilizar una programación amigable, se debe tener el conocimiento de la estructura de la misma para que al momento de recolectar información, solo aporte con la de mayor importancia y no asigne datos erróneos.

El algoritmo debe tener el suficiente entrenamiento para reconocer los objetos que interrumpen la movilidad de las personas, ya que, si no se lo realiza, la detección de objetos lo realiza de manera errónea y reconoce objetos falsos con interpretación de positivos.

Son varios los objetos que interrumpen la movilidad de las personas no videntes, pero para la implementación de todos estos casos, se necesita un componente electrónico de mayor complejidad, pero es de alto costo y necesita mayor tiempo de investigación, por lo que se recomienda investigar cuales son los objetos que producen interrupción de movilidad y delimitarse a los requerimientos que dice el beneficiario.

La persona con discapacidad visual, al momento de que el dispositivo reconoce el objeto y escucha el audio al cual el objeto pertenece, con tal solo el audio es más que suficiente el aviso para saber en qué posición se encuentra el objeto, por lo que se observó incomodidad que además del audio, la implementación del vibrador este actuando. Se recomienda delimitar el diseño que tendrá el objeto junto con los elementos necesarios para que el aviso o señales de alerta no alteren al temperamento de la persona.

## REFERENCIAS

- [1] Consejo Nacional para la igualdad de Discapacidades, «Estadísticas de discapacidad», enero de 2022. <https://www.consejodiscapacidades.gob.ec/estadisticas-de-discapacidad/> (accedido 26 de octubre de 2022).
- [2] H. M. Grijalva Rivadeneira, «Cinturón y manillas vibratoras ultrasónicas para no videntes», Universidad Técnica del Norte, Ibarra, 2017.
- [3] A. J. Villacreses Córdova, «Desarrollo de un prototipo de gafas electrónicas con sensores ultrasónicos orientado a personas no videntes para el laboratorio de robótica en la carrera de Ingeniería en Computación y Redes», Universidad Estatal del Sur de Manabí, Manabí, 2019.
- [4] D. A. Espinoza Moncayo y C. D. Peña Mendoza, «Diseño e implementación de un prototipo de gafas electrónicas con comunicación bluetooth a un celular para la detección de objetos circundantes que servirá como ayuda en personas no videntes», Universidad Politécnica Salesiana, Guayaquil, 2017.
- [5] F. A. Quintana Loaiza, A. D. Herrera Manzano, y E. L. Salazar Múnera, «Sistema de visión artificial para conteo de objetos en movimiento», Universidad Autónoma de Occidente, Cali, 2017.
- [6] Organización Mundial de la Salud, «Ceguera y Discapacidad Visual», 13 de octubre de 2022. <https://www.who.int/es/news-room/fact-sheets/detail/blindness-and-visual-impairment> (accedido 3 de noviembre de 2022).
- [7] A. F. Pérez Freire, «Implementación de un dispositivo inteligente de movilidad usando IoT para personas con discapacidad visual», Universidad de Guayaquil, Guayaquil, 2019.
- [8] S. González Zurita, «Influencia de la formación en fisioterapia en el desarrollo de competencias de rotación mental, representación espacial, orientación y movilidad en personas videntes y con ceguera», Master, Universidad Autónoma de Madrid, Madrid, 2019.
- [9] FESOCE, «Guía de orientación y movilidad para personas con sordoceguera», 26 de junio de 2018. [https://www.saludinforma.es/portalsi/documents/10179/2793801/Guia\\_de\\_orientacion\\_y\\_movilidad\\_sin\\_clave.pdf/134eea73-3596-48ed-942b-94aa57dbffdb](https://www.saludinforma.es/portalsi/documents/10179/2793801/Guia_de_orientacion_y_movilidad_sin_clave.pdf/134eea73-3596-48ed-942b-94aa57dbffdb) (accedido 28 de octubre de 2022).
- [10] EL COMERCIO, «Cómo nació el bastón que utilizan las personas no videntes», Quito, 27 de marzo de 2017. Accedido: 26 de octubre de 2022. [En línea]. Disponible en: <https://www.elcomercio.com/afull/baston-personas-novidentes-historia-salud.html>
- [11] MEDECU, «Bastón Blanco (No videntes)», 2022. <https://www.medicacuador.com/productos/ayudas-tecnicas/baston-y-muletas/baston-blanco-no-videntes/> (accedido 27 de octubre de 2022).

- [12] Instituto de Tecnologías Educativas, «Educación Inclusiva: Discapacidad Visual», Secretaría de Estado de Educación y Formación Profesional, 2017.
- [13] M. Martínez Canelo, «¿Qué es la programación orientada a objetos?», *Marketing & Communications Manager en Profile. Inmersa en el mundo de la tecnología.*, 11 de febrero de 2020. <https://profile.es/blog/que-es-la-programacion-orientada-a-objetos/> (accedido 27 de octubre de 2022).
- [14] A. Mordvintsev y K. Abid, «OpenCV-Python Tutorials Documentation», " OpenCV Python Doc, Berlin, 2017.
- [15] L. Rouhiainen, «Inteligencia artificial; 101 cosas que debes saber hoy sobre nuestro futuro», Editorial Planeta, S.A., Madrid, 2018.
- [16] Atria Innovation, «El Machine Learning en la industria», 22 de febrero de 2021. <https://www.atriainnovation.com/el-machine-learning-en-la-industria/> (accedido 18 de enero de 2023).
- [17] J. L. Espinoza, «Machine Learning», BBVA México, Ciudad de México, 2019.
- [18] Atria Innovation, «Deep Learning y sus muchas aplicaciones», 3 de octubre de 2019. <https://www.atriainnovation.com/deep-learning-aplicaciones/> (accedido 18 de enero de 2023).
- [19] F. Joseph, «Clasificación de imágenes basada en RESNET», 8 de octubre de 2021. <https://iopscience.iop.org/article/10.1088/1742-6596/1634/1/012110> (accedido 17 de abril de 2023).
- [20] J. B. Picazo, K. T. Hemi, y M. M. Cabello, «Comparación de marcos de trabajo de aprendizaje profundo para la detección de objetos», Universidad de Málaga, Málaga, 2018.
- [21] J. Larkin Alonso, «¿Qué es TensorFlow y para qué sirve?», 15 de junio de 2022. <https://www.incentro.com/es-ES/blog/que-es-tensorflow> (accedido 11 de octubre de 2022).
- [22] K. Chandan, «Qué es una matriz de confusión en el aprendizaje automático?», *GeekFlare*, 27 de enero de 2023. <https://geekflare.com/es/confusion-matrix-in-machine-learning/> (accedido 25 de abril de 2023).
- [23] J. I. Bagnato, «Principales algoritmos usados en machine learning», 16 de junio de 2022. <https://www.aprendemachinelarning.com/principales-algoritmos-usados-en-machine-learning/> (accedido 15 de junio de 2022).
- [24] E. García Sánchez, «Introducción a las redes neuronales de convolución. Aplicación a la visión por ordenador.», Universidad Zaragoza, Zaragoza, 2019.
- [25] Redacción KeepCoding, «Arquitectura típica de una red neuronal convolucional», ene. 2023.
- [26] D. Calvo, «Red Neuronal Convolutacional CNN», DataScientest, Barcelona, 2017.

- [27] G. Edel y V. Kapustin, «Exploración de los MobileNet V1 y MobileNet V2 modelos en NVIDIA Jetson Nano microcomputadores», *J Phys Conf Ser*, vol. 2291, n.º 1, p. 012008, jul. 2022, doi: 10.1088/1742-6596/2291/1/012008.
- [28] A. Sarkar, «Comprender EfficientNet: la arquitectura CNN más poderosa», 28 de mayo de 2021. <https://medium.com/mllearning-ai/understanding-efficientnet-the-most-powerful-cnn-architecture-eaeb40386fad> (accedido 17 de abril de 2023).
- [29] M. Hassan, «ResNet (34, 50, 101): CNN residuales para tareas de clasificación de imágenes», 23 de enero de 2019. <https://neurohive.io/en/popular-networks/resnet/> (accedido 17 de abril de 2023).
- [30] Terence Shin, «Comprensión de la matriz de confusión y cómo implementarla en Python», *DataSource.AL*, 20 de abril de 2020. <https://www.datasource.ai/es/data-science-articles/comprencion-de-la-matriz-de-confusion-y-como-implementarla-en-python> (accedido 6 de enero de 2023).
- [31] Bobbitt Zach, «Cómo trazar una curva ROC en Python», 6 de abril de 2021. <https://www.statology.org/plot-roc-curve-python/> (accedido 20 de noviembre de 2022).
- [32] B. Jazmina Ramirez y J. C. Castillo Herrera, «Metodología de la investigación e investigación aplicada para ciencias económicas y administrativas», Universidad Nacional Autónoma de Nicaragua, Managua, 2018.
- [33] D. Terreros, «Qué es una encuesta, para qué sirve y qué tipos existen», *Hubspot*, 2021. <https://blog.hubspot.es/service/que-es-una-encuesta> (accedido 4 de noviembre de 2022).
- [34] E. Rus Arias, «Investigación documental», *Economipedia*, 2020. <https://economipedia.com/definiciones/investigacion-documental.html> (accedido 4 de noviembre de 2022).
- [35] P. Latam, «Métodos inductivo y deductivo: ¿cómo se utilizan en las empresas?», 18 de noviembre de 2021. <https://blog.pearsonlatam.com/talento-humano/metodos-inductivo-y-deductivo-en-las-empresas#:~:text=El%20pensamiento%20deductivo%20nos%20ayuda,a%20partir%20de%20f en%C3%B3menos%20individuales.> (accedido 4 de noviembre de 2022).
- [36] F. Heriyanti y A. Ishak, «Diseño de sistema de información logística en el producto terminado almacén con el método de la cascada: literatura de revisión», may 2020. doi: 10.1088/1757-899X/801/1/012100.
- [37] Culturación, «Raspberry Pi: Definición y características», 2017. Accedido: 27 de octubre de 2022. [En línea]. Disponible en: <http://culturacion.com/raspberry-pi-que-es-caracteristicas-y-precios/>
- [38] NayLamp Mechatronics, «Sensor de Distancia TOF VL53L0X», 2023. <https://naylampmechatronics.com/sensores-proximidad/325-sensor-vl53l0x-de-tof.html> (accedido 11 de mayo de 2023).

- [39] Forma Infancia, «Sistema Bliss: cómo comunicarse a través de los símbolos», 28 de noviembre de 2019. <https://formainfancia.com/que-es-sistema-bliss/> (accedido 11 de marzo de 2023).
- [40] B. Hehner, «Símbolos BLISS Diccionario Guia», *Conrecursos*, 16 de diciembre de 2018. <https://www.conrecursos.org/wp-content/uploads/2018/12/Bliss-Diccionario-Gui%CC%81a-de-Barbara-Hehner-Ministerio.pdf> (accedido 12 de marzo de 2023).
- [41] A. Agriam, *Matriz de confusión en Excel*, (8 de agosto de 2016). Accedido: 25 de mayo de 2023. [En línea Video]. Disponible en: <https://www.youtube.com/watch?v=U9oWprQhYh8>

## ANEXOS

### a) Entrevista para requerimientos



UNIVERSIDAD TÉCNICA DEL NORTE

INGENIERÍA EN MECATRÓNICA

La investigación que se realiza es para obtener información sobre como las personas con discapacidad visual puede movilizarse de un lugar a otro, como reconocen los objetos y la afectación que pueden sufrir el no reconocerlos, para ello se detalla los requerimientos que se desea implementar en un dispositivo para que el reconocimiento de objetos sea más eficiente.

Nombre: Wilo Pabón

Edad: 72 años

#### **¿Cómo obtuvo discapacidad visual?**

Desde temprana edad, ya tenía presencia de deterioro en la visión, la primera vez lo confirmé mediante una campaña de salud realizada en el parque principal de la ciudad, donde le comentaron sobre el principal riesgo de perder la visión por completo si no se sometía a una cirugía y un procedimiento de control. Pero sus padres al no tener los suficientes recursos económicos para seguir el tratamiento y como en esos momentos aún tenía visibilidad de su alrededor, optaron por no someterlo a dicha cirugía.

Pero pasaron los años y en la actualidad la discapacidad visual es evidente ya que menciona que solo puede visualizar sombras muy borrosas de aquellas cosas que se encuentren cerca, pero en escenarios donde exista luminosidad bien clara.

#### **¿Cuál es el rango de discapacidad que presenta?**

La discapacidad se fue haciendo más notoria y cada vez perdía más la visión hasta que en la última revisión médica realizado hace seis meses, el porcentaje de discapacidad visual supera el 92%.

### **¿Como puede moverse dentro del hogar?**

Es importante que todas las cosas que se encuentran dentro del hogar, estén en posiciones específicas y no sean movidas, ya que, mediante el contacto y el mapa mental creado en su memoria, sabe la ubicación exacta donde se encuentra cada objeto. Pero si estos objetos son movidos de lugar puede alterar la movilidad y tropezar, además las puertas que existen dentro de casa deben estar abiertas ya que no sabe la ubicación específica de la manivela para abrir, ya que sufre algún golpe en el rostro cuando una puerta se encuentre cerrada. Los años de vida de la persona han sido concurridos de un lugar a otro, no le ha dado mayor importancia a su discapacidad y se ha dado los métodos necesarios para moverse, pero ahora que mayor tiempo pasa dentro de casa es indispensable tener la localización de cada objeto para no tropezar, porque cada día siempre sufre algún altercado.

### **¿Como puede moverse en la ciudad?**

La mayor parte del tiempo pasa en el hogar, para moverse fuera del siempre necesita la ayuda de una persona que lo dirija y avise de aquellos objetos con los que se puede tropezar por lo que es un reto moverse en la ciudad.

### **¿Qué dispositivos utiliza para moverse?**

Solo utiliza un bastón que es de soporte y guía para moverse, ya que presenta amputación en su pierna derecha y utiliza prótesis para caminar, el bastón casi siempre lo lleva en el lado derecho para equilibrar el peso de su cuerpo.

¿Realiza alguna técnica de detección cuando no tiene algún dispositivo cerca?

En los últimos años ha recibido capacitaciones sobre técnicas de comunicación y técnicas de detección de objetos, por lo que, para movilizarse y no contar con el bastón, realiza la técnica de protección alta y baja, además de la de seguimiento al tacto.

**¿Utiliza dispositivos que utilizan nueva tecnología en la detección de obstáculos?**

No, porque tienen un alto valor económico, el único dispositivo que tiene, es un bastón que cuenta con sonidos e iluminación, que fue un obsequio que se ha puesto en práctica con muy poca eficiencia, porque no ayuda a la detección de objetos en específico, emite sonidos e iluminación para que las personas o vehículos puedan identificar a la persona no vidente, “para ser sincero, no es muy práctico en mi hogar”.

**¿Qué rango de alcance tienen el dispositivo que utiliza?**

El rango de alcance de dicho dispositivo en sonido es aproximado a los 5 metros y en iluminación tiene un aproximado de 10 metros y como bastón el rango de alcance es de un metro, el largo que tiene el bastón.

**¿Está de acuerdo que sea participe en la investigación?**

“Claro que sí, tengo muchas expectativas del proyecto”.

**¿Qué requerimientos son los que usted necesita en un dispositivo?**

El principal requisito es que sea de fácil utilización y mantenimiento, ya que al ser una persona de avanzada edad y con discapacidad visual, no tiene el suficiente conocimiento sobre el uso correcto de elementos electrónicos complejos, que sea montable y desmontable en un sector de fácil ubicación del cuerpo humano, con el que no se sienta complejo de utilizar el dispositivo.

Los principales objetos que se deben reconocer, deben estar en el hogar ya que es el ambiente en donde paso la mayor parte del día, siempre deben estar ubicados en lugares específicos, pero es de gran ayuda que me anticipe su ubicación actual, “así yo puedo dirigirme con total confianza hacia ese objeto”.

**b) Código de programación**

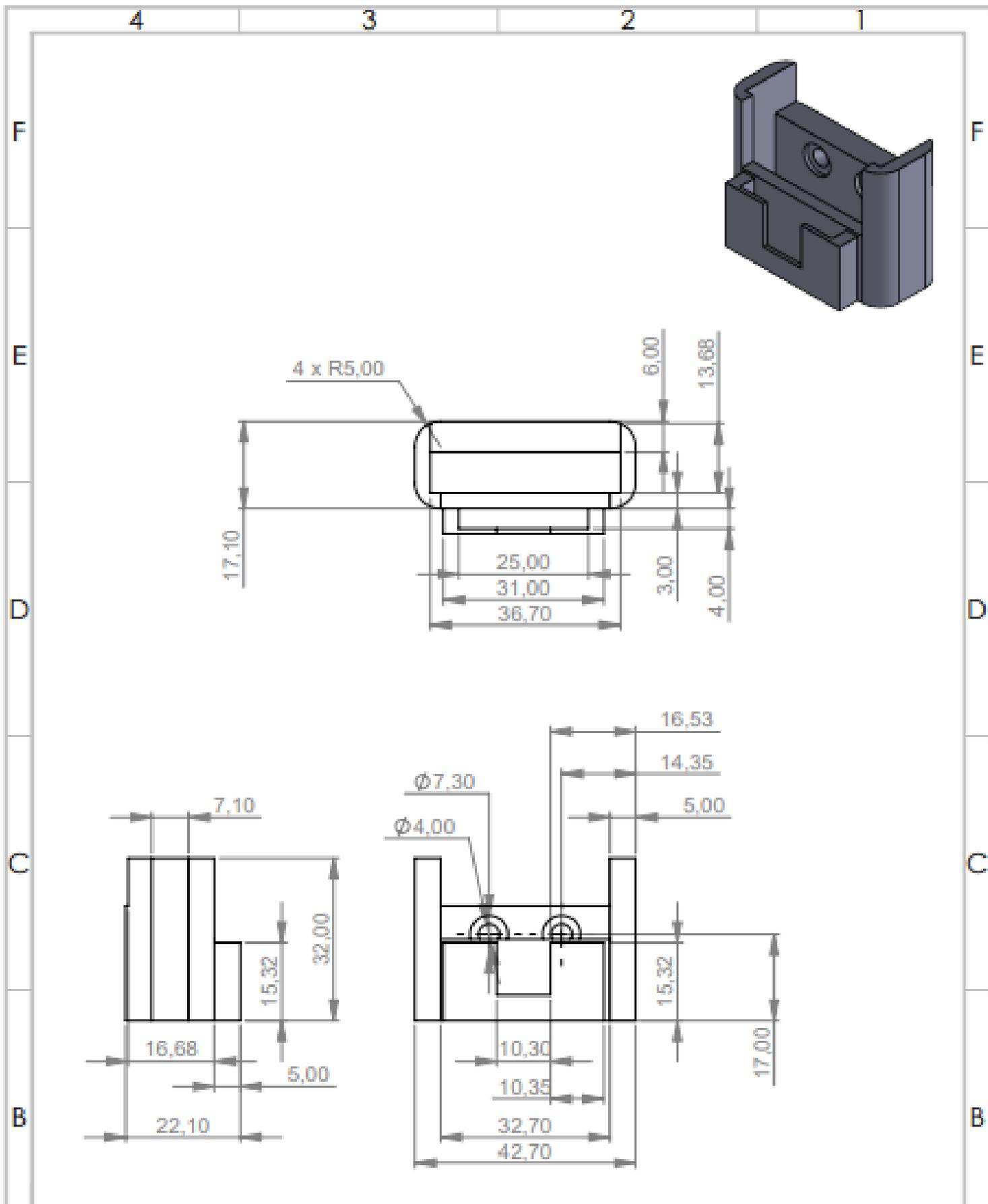
El código de programación del dispositivo detector de objetos se encuentre ubicado en el siguiente link:

[https://utneduec-my.sharepoint.com/:f:/g/personal/waortegab\\_utn\\_edu\\_ec/EnkCdN-ZhZxDqkRVUfbJ4WYBx2eqzCZXmvPE9T3xH2QQLA?e=EcZ4w7](https://utneduec-my.sharepoint.com/:f:/g/personal/waortegab_utn_edu_ec/EnkCdN-ZhZxDqkRVUfbJ4WYBx2eqzCZXmvPE9T3xH2QQLA?e=EcZ4w7)

### c) Pruebas de funcionamiento



### d) Planos



PROYECTO PORTA CAMARA

ESCALA



PARTE

CÓDIGO T-001

TOLERANCIA:

MATERIAL

PLA

DISEÑO

ORTEGA W

27/6/2023

HOJA

TRATAMIENTO

NINGUNO

DIBUJÓ

ORTEGA W

27/6/2023

1/1

RECUBRIMIENTO

NINGUNO

REVISÓ

GARCIA I

27/7/2023

TRATAMIENTO

NINGUNO

APROBÓ

FICA

CIME

4

3

2

1